Systems & Control: Foundations & Applications

Serdar Yüksel
Tamer Başar

Stochastic Networked Control Systems

Stabilization and Optimization under Information Constraints

Birkhäuser

Birkhäuser

Systems & Control: Foundations & Applications

For further volumes:
http://www.springer.com/series/4895

Serdar Yüksel • Tamer Başar

Stochastic Networked Control Systems

Stabilization and Optimization under Information Constraints

Birkhäuser

Serdar Yüksel
Department of Mathematics and Statistics
Queen's University
Kingston, Ontario
Canada

Tamer Başar
Department of Electrical
and Computer Engineering,
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

ISSN 2324-9749 ISSN 2324-9757 (electronic)
ISBN 978-1-4614-7084-7 ISBN 978-1-4614-7085-4 (eBook)
DOI 10.1007/978-1-4614-7085-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013936940

Mathematics Subject Classification (2010): 91B44, 93A14, 93E03, 93E15, 93E20, 94A15, 94A17, 94A24, 94A29, 60J05, 60J20, 60B10, 62B05, 62B10, 62C10

© Springer Science+Business Media New York 2013
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.birkhauser-science.com)

To
Kate, Sait, and Saibe Yüksel (S.Y.)
and
Tangül, Gözen, Elif, Altan, and Koray (T.B.)

Preface

Our goal in writing this book has been to provide a comprehensive, mathematically rigorous, but still accessible treatment of the interaction between information and control in multi-agent decision making in the context of networked control systems. These are systems where different decision units (or equivalently decision makers or agents, which could be sensors, controllers, encoders, or decoders) are connected over a real-time communication network, where the communication medium is heterogeneous, information is decentralized and distributed, and its acquisition is not instantaneous. The questions we address are all performance driven, and entail the issues of what data to pick and how to shape and transmit them for control purposes under various resource constraints as well as how to design optimal control policies with partial information. We deal specifically with the issues of quantization and encoding, design of optimum channels, effects of decentralization on control performance, stability, learning, signaling, and relationships between team performance (of a group of agents) and various information structures.

The book draws and utilizes a diverse set of tools (of both conceptual and analytical nature) from various disciplines, including stochastic control, stochastic teams, information theory, probability theory and stochastic processes, and source-coding and channel-coding theory, and amalgamates them into a unified, coherent, applicable theory. It could be used as a textbook or as an accompanying text in a graduate course on networked control or multi-agent decision making under informational constraints.

Acknowledgements We would like to thank several individuals whose work, either independently or through collaboration, has shaped the contents of the book. Of particular mention are Tamás Linder (Chaps. 4 and 10); Sean Meyn (Chaps. 6 and 8); Andrew Johnston (Chaps. 7 and 9); Sekhar Tatikonda (Chap. 11); Ali Zaidi, Tobias Oechtering, and Mikael Skoglund (Chap. 11); and Aditya Mahajan, Giacomo Como, and Nuno C. Martins, general collaborations with whom have been reflected in the book.

Abhishek Gupta, Andrew Johnston, Naci Saldi, Charalambos Charalambous, Maxim Raginsky, Emrah Akyol, and Daniel Quevedo read different versions of the book and provided substantial technical comments which this final version has greatly benefitted from.

The book has also benefitted from technical discussions with many of our colleagues and students, including Nevroz Şen, Marcos M. Vasconcelos, Debasish Chatterjee, Ather Gattami, Yücel Altuğ, Orhan Arikan, Demos Teneketzis, Alexander Condello, Adina Goldberg, Anant Sahai, Gürdal Arslan, R. Srikant, Chris Hadjicostis, Venu Veeravalli, Cédric Langbort, Todd Coleman, Orhan Çağri Imer, Atilla Eryilmaz, Akshay Kashyap, Tansu Alpcan, Abdol-Reza Mansouri, Andrew Lewis, and Fady Alajaji, to whom we also convey our thanks.

An earlier version of the book was used in an intensive course offered by one of us (S.Y.) at KTH, Stockholm, and the discussions there also helped shape this final version. We thank the hosts, Tobias Oechtering, Mikael Skoglund, and Ali Zaidi, for providing this opportunity.

Finally, we acknowledge the stimulating, conducive environment of the Coordinated Science Laboratory, University of Illinois, where the initial seeds of this book project were sown. The bulk of the work was carried out after the first author joined Queen's University, which provided another stimulating environment for the project to be completed. We thank both institutions and many of our colleagues there.

Kingston, ON Serdar Yüksel
Urbana, IL Tamer Başar

Contents

4 Topological Properties of Information Structures: Comparison, Convergence, and Optimization 119

Part II Stabilization of Networked Control Systems

Acronyms and Notations

trace(A)	Trace of a square matrix A
det(A)	Determinant of a square matrix A
1_E	Indicator function for event E
$\mathbb{X}$	A space of vectors
$\langle x, y \rangle$	Inner product between x and y on a Hilbert space
$\mathcal{B}(\mathbb{X})$	The Borel σ-field on $\mathbb{X}$
$\sigma(y)$	σ-field generated by a random variable y
$\mathcal{P}(\mathbb{X})$	Set of probability measures on $\mathcal{B}(\mathbb{X})$
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^n$	Vector space of n-dimensional real vectors
$\mathbb{Z}$	Set of integers
$\mathbb{Z}_+$	Set of nonnegative integers
$\mathbb{N}$	Set of positive integers
$\mathcal{Q}$	A space of quantizers
Q	Quantizer or channel depending on context
$\Pi^{comp,i}$	Composite quantization policy for encoder i
Q_t^i	Quantizer used by agent i at time t
DM i or **A** i	decision maker i or agent i
$\underline{\gamma}^i$	Policy of DM i, that is, $\{\gamma_t^i, t \geq 0\}$
$\underline{\gamma}$	Ensemble of policies for all decision makers, that is, $\{\underline{\gamma}^i\}$
I_t^i or $\mathcal{I}_t^i$	Information variable at agent i at time t
$\underline{\eta}$	Information structure inducing map $\{\eta^1, \ldots, \eta^N\}$
$E_P^{\underline{\gamma}}\{\cdot\}$	Expectation under policy $\underline{\gamma}$, with initial condition measure P
E_x	Expectation conditioned on an initial condition realization x, or with respect to a random variable x, depending on the context
$H(\cdot)$	Discrete entropy
$h(\cdot)$	Differential entropy
$I(\cdot;\cdot)$	Mutual information
$D(P_1 \| P_2)$	Kullback–Leibler divergence between P_1 and P_2

$\lvert x\rvert$	Euclidean norm of a finite-dimensional real vector x
A' or A^T	Transpose of matrix A
$\lvert S\rvert$	Cardinality of a set S
$A \setminus B$	Set difference: $\{x : x \in A, x \notin B\}$
$A \bigtriangleup B$	$(A \setminus B) \cup (B \setminus A)$
$\ln(x)$ or $\log(x)$	Natural logarithm of positive real x
$\underline{o}$	Zero vector
$\mathcal{T}$	Time/stage index set, $\{1, 2, \ldots, T\}$ or $\{0, 1, \ldots, T-1\}$
$\mathcal{N}$ ($\mathcal{L}$)	Decision maker (DM) index set, $\{1, 2, \ldots, N\}$ ($\{1, 2, \ldots, L\}$)
$u_{[k,s]}$	Action (decision) variables from $t = k$ to $t = s$ for $s > k$, $\{u_k, u_{k+1}, \cdots, u_s\}$
$\mathbf{u}$	Collection of actions in $\mathcal{N}$ (or $\mathcal{L}$): $\{u^1, u^2, \ldots, u^N\}$

Chapter 1
Introduction

This chapter provides an introduction to the field of networked control and thereby to this *book*. It highlights the main approaches taken to address issues unique to networked control and describes the scope of coverage and the contents of the book.

1.1 Information and Control

The *interaction between information and control* is a phenomenon that arises in every decision and control problem. On the one hand, any performance-driven controller requires information on the unknowns that affect the operation of the underlying system; on the other hand, the *quality* of the relevant information itself is typically affected by the choice of the control action in a closed-loop system. Further, the transmission of information over communication channels with high fidelity and the process of shaping the source output and recovering the transmitted signal at the other end can themselves be viewed as controller design problems. This book is a comprehensive undertaking aimed at furthering our understanding of this interaction in the context of decentralized and networked control systems.

Networked control refers to a decentralized control system in which the components are connected through real-time communication channels or a data network. Thus, there may be a data link between the sensors (which collect information), the controllers (which make decisions), and the actuators (which execute the controller commands); and the sensors, the controllers, and the plant themselves could be geographically separated.

With such a networked structure, many modern control systems are decentralized. Such systems feature multiple decision makers (e.g., sensors, controllers, and encoders) which have access to different and imperfect information, either cooperate with or compete against each other. Such systems are becoming ubiquitous, with applications ranging from automobile and inter-vehicle communications designs, control of surveillance and rescue robot teams for access to hazardous environments,

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_1,

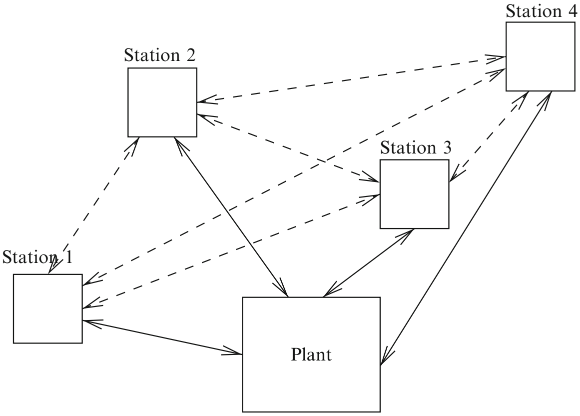

Fig. 1.1 A decentralized networked control system. *Solid lines* show the interaction between the control stations and the plant *Dashed lines* depict the possible communication links between the stations

remote surgery, space exploration and aircraft design, and control of economic systems, among many other fields of applications most of which involve *remote control*.

In such decentralized networked control problems, one major concern is the characterization of the *minimum* amount of information transfer needed for a satisfactory performance and particularly for *stability* of the overall system. This information transfer would be between various components of the networked control system. One necessity for satisfactory control performance is the ability for the controllers to track the plant state under various constraints on the communication (see Fig. 1.1). Another set of challenges is the determination of the data rate required for the transmission of control signals, and the construction of dynamic encoding, decoding, and control policies meeting some selected design criteria. Another important problem is the establishment of effective coordination among multiple sensors or multiple controllers/decision makers using minimum possible information exchange. Even in cases when communication resources are not scarce, a strong understanding of the fundamentals can be useful in constructing the system architecture, and finally, such an insight can help reduce the computation requirements and complexity.

Various forms of system architectures have been introduced and studied in the networked control literature. To be able to analyze different scenarios, it is important to identify and formalize the probabilistic description of the system and characterize the underlying information structure. Further, it is equally important to precisely pin down the objective in the system design, whether it is *stochastic stability* (in some appropriate sense) or *optimization*. Learning and identifying an unknown system through observations and actions is another important issue. Finally, the notion of *signaling*–that controllers could communicate through actions–is another aspect which has to be taken into account. In all this, it is essential to understand both the qualitative and the quantitative values of information for achieving different objectives in such networked settings.

The above description of a networked control system lies at the intersection of three disciplines of applied mathematics and engineering, namely, *decentralized control* (in view of information structures and decision making under measurability constraints), *stochastic control* (in view of decision making under uncertainty), and *communication, information, and quantization theories* (in view of information exchange among decision makers). Finally, *probabilistic analysis* (in view of being the ultimate mathematical tool needed to conduct a study in all of these disciplines) plays an essential role in the understanding, analysis, and synthesis of such systems.

Our aim in this book is to bridge these three disciplines in a precise and rigorous manner, while also conveying practical messages to systems designers and controller architects. The field of networked control has had accelerated growth during the past decade, but many of the problems and challenges that arise were actually obstacles identified already in the 1960s and the 1970s. What makes the situation different today is the accumulation (since that time) of an arsenal of more powerful results and tools from information theory, source coding theory, and the theory of Markov chains, directly applicable to the problems at hand. Furthermore, we have a richer pool of computational algorithms, and the computers of the current generation have significantly more processing power. The field has reached some state of maturity, and it seems timely to collect and present the main results in a book form. Furthermore, it is also useful to revisit many of the results of the 1970s and the 1980s and blend them in such a treatise with the current developments, as they have by no means lost their relevance. And this is another one of our goals here.

1.2 Coverage and the Intended Audience

Within the framework of networked control systems and in discrete time, five essential concepts are visited recurrently in the book:

- The characterization of information structures in team problems defined in terms of measurability relations in a given probability space. Comparison and topological properties of information structures for stochastic team optimization problems and identification of information structures which may lead to a systematic program for generation of optimal policies.
- Stochastic stability of systems (state, controller, and encoders). Converse results through information theoretic analysis and constructive algorithms via stochastic drift equations. Stochastic stability corresponds to the existence of an equilibrium distribution, ergodicity of a process, or existence of finite moments.

- The operational differences between information theoretic settings (which, in the classical sense, requires an infinite ensemble of messages to be transmitted or encoded) and real-time settings in control which do not tolerate delay. Use of information theory in establishing fundamental bounds on information requirements.
- Optimal information transmission under causality and delay constraints, and jointly optimal channel and controller design for real-time systems, under a variety of information structures. Structural results as well as existence results on optimal policies.
- The notion of signaling, its utilization in decentralized stabilization, and the technical issues that are associated with it, such as the lack of convexity, the dual effect of control, and non-neutrality.

These concepts and notions are interweaved throughout the book, constituting the backbone of a comprehensive theory of networked control. Specific results are built on that foundation and seamlessly presented throughout the book.

What is Not Covered

As indicated earlier, networked control systems, and multi-agent systems in general, entail multiple decision makers that provide input into the system using only local (decentralized) information. Throughout the book the underlying assumption is that the agents act in unison, toward a common goal, that is, as members of a team, even though they do not necessarily share the information they acquire. An extended framework would be one where the agents' goals are not aligned and may even be conflicting, which then cannot be cast as a decentralized team problem. Such problems belong to the realm of *dynamic noncooperative stochastic games*, where appropriate solution concepts are the Nash equilibrium or Stackelberg equilibrium, or a blend of the two, depending on whether there is a hierarchy in decision making or not [32]. Such problems entail other intricate issues and their analysis requires a different set of tools, beyond the scope of the coverage here.

Another direction in which the framework of this book can be broadened is as follows: The treatment in the book is restricted to settings where there is an underlying probability space and all the variables in the system are well-defined random variables on this probability space given the policies of the decision makers. Such a *Bayesian* setting does not include the probability-free settings occasionally used in control theory (as in robust control with distribution-free disturbances with norm constraints [28]) and in information theory and machine learning (as in coding and learning for individual sequences [90]). As objective or loss functionals, such settings typically admit a min-max type formulation instead of minimization of the expected value of a loss function with respect to a probability measure. This book does not consider such settings explicitly; however, the approaches presented here are applicable to many such scenarios.

Intended Audience for the Book

The intended audience is broad, including academic as well as industrial researchers interested in control theory, information theory, statistics, and applied mathematics. As indicated earlier, the book adopts a probability theory, information theory, and decentralized stochastic control theory view to networked control problems. To comfortably follow the material in the book, the reader should be familiar with linear systems (at the first-year graduate level), basics of information theory, and measure-theoretic stochastic processes (again at the first-year graduate level). The reader is also expected to have a basic understanding of Markov chains and martingale theory. For those who do not have the requisite background, appropriate references are provided throughout the development in the book, and four appendices are included, covering some of this material as well as others.

1.3 Contents of the Book

The book is comprised of twelve chapters, organized into three parts, as described below. It also has four appendices, providing background material.

1.3.1 Part I. Information Structures in a Networked Control System

This part is primarily concerned with the mathematical description of a networked control system as a stochastic dynamic team. It provides a treatise on stochastic dynamic teams and a detailed investigation of information structures. Comparison of different information structures is also covered.

In Chapter 2, a general probability theoretic framework for stochastic team decision problems is established, by defining and classifying information structures, interaction dynamics, policy spaces, and objective functions. A number of examples are included to provide a gentle introduction to the concepts. Team problems and information structures are classified according to various criteria. Solution methods for static teams are presented, with particular emphasis on convex cost functions. Dynamic teams are considered further in Chap. 3.

Chapter 3 focuses on comparison of information structures and solution approaches to a class of dynamic team problems. Under nonclassical information structures, the notion of signaling is introduced and thoroughly discussed. Witsenhausen's counterexample is studied, along with its generalizations and a class of dynamic team problems involving Gaussian sources and channels. Expansion of information structures is presented as a general recipe for studying dynamic teams with nonclassical information patterns. Witsenhausen's characterization of information structures is also presented in this chapter.

Chapter 4 investigates the optimal design of information structures and studies a number of topological properties of information structures modeled as observation channels under various topologies. Continuity, compactness, concavity, and existence properties are studied for single-stage and multistage optimal stochastic control problems. An introduction to quantizers is given. Quantizers are viewed as a special class of measurement channels, and existence of optimal quantizers is established. Furthermore, a partial ordering on the value of information channels for the minimization of cost functions is studied (known as Blackwell ordering). Applications to empirical consistency and learning are discussed.The results presented in this chapter are used extensively later in Part III.

1.3.2 Part II. Stabilization of Networked Control Systems

This part focuses on the stabilization of networked control systems, for both single-sensor/controller and multi-sensor/controller systems, and comprises five chapters.

The chapters in this part introduce fundamental criteria that need to be satisfied for stochastic stabilization. Constructive methods are presented which meet the fundamental (converse) bounds. The constructive proofs utilize a drift approach offered in a number of recent papers by us and our collaborators. Toward further understanding the value of information channels in stochastic control, it is shown that *Shannon capacity* provides a total ordering on the set of channels for the existence of policies for stochastic stability and ergodicity properties.

Chapter 5 introduces policies and actions regarding the selection of quantizers and controllers in networked control. It reviews fundamentals of information theoretic notions. The chapter exhibits the important differences between the real-time communication formulation and the traditional Shannon theoretic setup which allows for large blocks of data (with unbounded block length) to be encoded and transmitted. This distinction is highlighted in the context of distortion-constrained quantizer design and the rate-distortion theory. The chapter also establishes fundamental lower bounds on information rates needed for various forms of stochastic stabilization. These lower bounds are further studied in Chap. 9.

Chapter 6 is an important one for the general program of Part II, where random-time state-dependent stochastic drift criteria for stabilization of Markov chains are established together with a class of application areas in networked control systems. Criteria for transience and other forms of stochastic stability are presented. Related background material is reviewed in Appendix C.

In the context of stochastic stabilization of linear sources driven by noise with unbounded support, controlled over information channels, Chap. 7 focuses on finite-rate noiseless channels and provides the architectural setup for coding and control policies. Chapter 8 investigates stabilization over erasure channels, discrete memoryless channels with and without feedback as well as a class of continuous-alphabet channels (Gaussian channels are further discussed in Chap. 11). A common theme in these chapters is the relationship between Shannon capacity and the ergodicity of the controlled Markov process: For ergodicity (under additional technical assumptions), Shannon capacity (with feedback) provides a boundary condition in the space of communication channels for stochastic stabilization of unstable linear systems. For finite moment stability, however, further conditions are required both on the channels and on the tail distributions of the system noise. The results in these chapters also include extensions to multidimensional and partially observed settings.

Chapter 9 considers stabilization under a decentralized information structure for multi-sensor and multi-controller systems. Existence results on stabilizing policies under the decentralized information structure are obtained. In the absence of noise, it is shown that multi-controller systems, unlike multi-sensor systems with a centralized controller, entail a rate loss due to decentralization. The noisy cases are also investigated and rate conditions are established for multi-sensor systems.

1.3.3 Part III. Optimization in Networked Control: Design of Optimal Policies Under Information Constraints

The third, and final part of the book, comprising three chapters, studies simultaneous design of optimal encoding and control policies for networked control systems.

Chapter 10 establishes the structure of optimal quantization policies under various information structures for general cost functions. The coverage includes both single decision maker and multiple decision maker formulations, with partial as well as full observation. A dynamic programming approach is presented building on classical results by Witsenhausen, and Walrand and Varaiya. The chapter also presents optimal solutions for encoders and controllers under quadratic performance measure for linear Gaussian systems controlled over discrete noiseless channels.

Chapter 11 obtains optimal solutions for encoders and controllers under quadratic cost functions for linear Gaussian systems controlled over Gaussian channels, proving also the existence of optimal solutions. Furthermore, the chapter identifies conditions under which optimal coding and control policies are linear. Counterexamples on sub-optimality of linear policies are also presented.

Chapter 12 presents the notions of agreement and common knowledge and addresses the question of how to achieve common knowledge. The chapter presents a general framework for obtaining solutions to dynamic team problems under decentralized information structures based on dynamic programming and an evolving common knowledge, and applies this primarily in the context of the belief sharing information pattern. Information rates required for tractability of optimal solutions are also presented. Finally, the chapter introduces a team cost-rate function, which provides the minimum cost subject to a rate constraint on the information exchange among members of a team.

1.4 A Guide for the Reader or the Instructor

For students who have a background in stochastic processes at the first-year graduate level and who are also familiar with the basics of information theory, the book can be used as a textbook in a course on *networked control systems* or *stochastic control* or *multi-agent decision making*. It could also be used in a special topics course or as an independent resource.

For a graduate-level course on *decentralized control* where the students do not have any background on communication theory at the graduate level, Chaps. 2, 3, 4, 9, 10, 11, and 12 together with the Appendices can be used as primary or supporting material. A basic information theory background can be acquired from standard textbooks, such as [103] or [151]. For further advanced topics on information theory, the reader is referred to advanced texts such as [107] or [153].

Chapters 2, 3, 4, 6, 10, and 12 together with the Appendices can be used as primary or supporting material in a graduate-level course on *Stochastic Control* with limited information theory content.

For students who do not have a background at a graduate-level stochastic process course, all chapters except Chaps. 4, 6 should be accessible.

For students who are familiar with information theory, but not stochastic control and optimization, all chapters should be accessible with some further reading. The stochastic control and optimization fundamentals could be supplemented by resources such as [84, 194, 225, 269], and [14] on stochastic control and [55] and [242] on optimization. The material in Chap. 6 can be supplemented by [271]. A further useful reference is [189].

A useful related book in the literature on networked control systems is the text by Matveev and Savkin [266], which has partial overlap with the material in Chaps. 7, 8, and 9 of this book, even though the approaches are different. The two books can be used as complementary resources.

Part I
Information Structures in Networked Control

Chapter 2
Networked Control Systems as Stochastic Team Decision Problems: A General Introduction

2.1 Introduction

Networked control systems can be viewed as stochastic decision problems with dynamic decentralized information structures or as stochastic dynamic teams, with each subcontroller viewed as an *agent* in a dynamic team. The goal of this introductory chapter is accordingly to introduce the reader to a general mathematical formulation of stochastic teams, first with static and then with dynamic information structures, and to discuss some salient features of these decision problems and associated solution concepts through some simple but illustrative examples.

The chapter discusses both *static* stochastic teams (i.e., team decision problems where the information signals received by the decision makers are not affected by actions) and *dynamic* stochastic teams (where the information of at least one decision maker is affected by action). Sections 2.2, 2.3, and 2.6 deal with static teams, whereas Sects. 2.4 and 2.5 discuss dynamic teams. Section 2.2 provides a general formulation for static teams, which is followed by a complete analysis of a finite stochastic team problem under various information patterns, in Sect. 2.3. Section 2.6 provides some general explicit results on existence, uniqueness, and characterization of optimal solutions first for general static teams and then for special classes of teams with Gaussian statistics: those with quadratic and exponentiated quadratic costs.

Sections 2.4 and 2.5 can be viewed as the counterparts of Sects. 2.2 and 2.3 for dynamic teams. First a precise mathematical formulation for dynamic team decision problems is given, in Sect. 2.4, along with various dynamic information structures and appropriate solution concepts, and then an illustrative example of a finite dynamic team is provided in Sect. 2.5, within the framework of which some important features of optimal solutions in teams are discussed. The chapter concludes with Sect. 2.7 which provides some bibliographical notes and guidelines for further reading on the topics covered herein.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_2,

2.2 A Mathematical Framework For Static Decision Problems

Multiple person stochastic decision problems could be formulated with varying degrees of generality, abstraction, and rigor, depending on the types of problems to be solved (*i.e.*, the scope of coverage) and the level of mathematical sophistication to be expected from the reader. Common to all possible formulations, however, is the specification of **five** basic **ingredients** which are essential for a well-founded mathematical treatment of *decision making under uncertainty*. These are:

1. The number of **decision makers** (synonymously, *agents* or *controllers*) and the sets of alternative **actions** (synonymously, *decisions* or *controls*) available to them
2. The **uncertainty** and its probabilistic description
3. The **information** acquired by each decision maker on the uncertainty and the previous actions
4. The **payoff** (or *loss*) that accrues to each decision maker as a result of joint actions (over the decision period) and realization of uncertainty
5. A **solution concept** whereby "best" or "satisfactory" decision rules can be chosen

Before going into further specification of these entities, let us pause to introduce some terminology and notation which will be needed in the sequel. We will refer to a decision problem as **static** if the information available to each decision maker is independent of the actions of other decision makers (this statement will be made precise later in the section as well as in Sect. 3.8); otherwise, the decision problem is said to be **dynamic**. We will refer to decision makers interchangeably as *agents* or *controllers*, with the *i*th one denoted $\mathbf{A}i$, where i takes values in the set $\mathcal{N} := \{1, \ldots, N\}$ which is called the *agent* (*decision maker*) *set*. The variable under the control of each decision maker will be called the *action* (synonymously, *decision* or *control*) *variable* and will be denoted by u^i for $\mathbf{A}i$. Each u^i will take values in a given *action set* to be denoted by U^i. Finally, the N-tuple $(u^1, \ldots, u^N)$ will be denoted by $\mathbf{u}$ and the product action space $U^1 \times \cdots \times U^N$ by $\mathbf{U}$.

Basic Ingredients of Static Decision Problems

In the static framework we will initially study the class of problems where the action sets, $U^i, i \in \mathcal{N}$, are either (finitely or infinitely) countable or uncountable but finite dimensional. In the latter case, we take the action set (space) to be isomorphic to the Euclidean[1] space $\mathbb{R}^{m_i}$, for some integer $m_i, i \in \mathcal{N}$; furthermore, if there are any

---

[1] Some background material on sets and topological notions can be found in Appendix A.

constraints imposed on the action variable u^i, we introduce the *action constraint set* S^i, for $\mathbf{A}i$, as a proper subset of U^i.

The uncertainty in the decision problem is captured in the so-called random state of **nature**, ξ, which is a random variable (or vector) defined on a given probability space $(\Omega, \mathcal{F}, P_\Omega)$[2] and taking values in the Borel space $(\Xi, \mathcal{B}(\Xi))$ where either $\Xi \equiv \mathbb{R}^m$ for some positive integer m or Ξ is a countable set. Let P be the probability measure induced by ξ on $(\Xi, \mathcal{B}(\Xi))$, corresponding to P_Ω. To save from notation, the corresponding probability distribution function will also be denoted by P.

The decision makers do not, in general, have direct access to the true state of nature but instead observe the value of some other variable, known as the *measurement* (or *information*) *signal*. To define this quantity in precise mathematical terms, let us first introduce, for each $i \in \mathcal{N}$, the *information field*, $\mathbf{Y}^i$, for agent $\mathbf{A}i$ as a given sub σ-field of $\mathcal{B}(\Xi)$, generated by a measurable function η^i mapping $(\Xi, \mathcal{B}(\Xi))$ onto $(Y^i, \mathbf{B}^i)$. This is known as the *information function* for $\mathbf{A}i$, and the N-tuple $\underline{\eta} := (\eta^1, \ldots, \eta^N)$ is called the *information structure* (or *information pattern*) of the decision problem. The information function η^i induces a σ-field, $\mathbf{Y}^i$, of Ξ, and the information (measurement) signal y^i of $\mathbf{A}i$ (which lies in the *measurement set* Y^i) is generated according to η^i, which is symbolically written as

$$y^i = \eta^i(\xi) \equiv \tilde{\eta}^i(\omega)\,, \tag{2.1}$$

where the latter relates the measurement signal directly to the original probability space Ω, with elements ω. This is sometimes a more convenient representation to work with, especially if Ω is finite or countable. In that case one can consider Ω and Ξ to be essentially the same set and thereby view $\mathbf{Y}^i$ also as a partition of Ω, which is a convention we henceforth adopt. In the case of finite probability spaces we will also adopt the convention, perhaps by a slight abuse of notation and terminology, that the measurement signal y^i can be considered as an element of the partition set $\mathbf{Y}^i$.

The decision makers determine their actions using the measurement signals that they receive, under the *strategies* that they adopt for transforming measurements into actions. The *strategy* (synonymously, *decision rule* (function) or *control law*) of $\mathbf{A}i$ will be denoted by γ^i and is formally defined as a measurable mapping from $(\Xi, \mathbf{Y}^i)$ into the space $(U^i, \mathbf{B}_{U^i})$. This can also be written as a measurable mapping from $(Y^i, \mathbf{B}^i)$ to $(U^i, \mathbf{B}_{U^i})$, as we state explicitly below. We denote the set of all such mappings, which also satisfy the additional constraints that may have been imposed on u^i, by Γ^i, to be called the *strategy space* of $\mathbf{A}i$, and note the relationship

$$u^i = \gamma^i(y^i) = \gamma^i(\eta^i[\xi]),$$

---

[2] Necessary background material on probability theory, along with an explanation of the terminology and notation used here, can be found in Appendix B.

where the latter relates the action variable to the state of nature, ξ. We will denote the N-tuple $(\gamma^1, \ldots, \gamma^N)$ by $\underline{\gamma}$, and the product strategy space $\Gamma^1 \times \cdots \times \Gamma^N$ by $\mathbf{\Gamma}$. The individual strategy spaces Γ^i may also include the additional structural constraints that may have been imposed on the policies, such as linearity. What is not allowed in general, however, is for $\mathbf{\Gamma}$ to be *nonrectangular*, that is, for the choice out of Γ^i (for some i) to restrict the choice out of $\Gamma^j (j \neq i)$. For example, our formulation (at this point) does not cover "cross-constraints" of the type $f(u^i, u^j) \leq 0$, $i \neq j$, for some functional f.

Given an (N+1)-tuple $(\xi, \mathbf{u}) \in \Xi \times \mathbf{U}$, the loss incurred to the decision makers viewed collectively as a *team* will be denoted by $L(\xi, \mathbf{u})$, where the function $L : \Xi \times \mathbf{U} \to \mathbb{R}$ is known as the *loss function* for the team. Its negative, $-L(\xi, \mathbf{u}) =: \mathtt{U}(\xi, \mathbf{u})$, is known as the *payoff function*, which all agents collectively want to "maximize," in a sense to be defined shortly. Implicit here is the assumption that for the team there exists a unique (up to equivalence) utility function which numerically orders different outcomes corresponding to joint actions and realization of the state of nature, in a way consistent with the team's preference ordering among different alternatives.

The loss incurred is generally a random quantity, the randomness appearing through both ξ and $\mathbf{u}$, where the latter depends on ξ through the measurement signals and the strategies adopted by the decision makers. Therefore, one rather works with the expected value of this quantity, which we will be referring to as the *cost function*.[3] Other possible terminology would be *expected loss function*, *average risk*, or *expected cost*, all of which have been used in the literature, which we will also use interchangeably. The cost function, $J : \mathbf{\Gamma} \to \mathbb{R}$, is defined on the product strategy space $\mathbf{\Gamma}$ as[4]

$$J(\underline{\gamma}) = \int_{\Xi} L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))P(d\xi) = E[L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))] =: E_{\xi} L(\xi, \underline{\gamma}(\underline{\eta}[\xi])), \quad (2.2)$$

where

$$\underline{\gamma}(\underline{\eta}[\xi]) := (\gamma^1(\eta^1[\xi]), \ldots, \gamma^N(\eta^N[\xi])) \quad (2.3)$$

and E_{ξ} is the operator that takes the expected value of the quantity it precedes, over ξ. To show the explicit dependence of J on also the information structure $\underline{\eta}$, we will sometimes use the notation $J(\underline{\gamma}, \underline{\eta})$ and occasionally use $J(\underline{\gamma}, \underline{\eta}; L, P)$ to also indicate the dependence on the loss function L and the probability distribution P.

The specification of J, along with the product strategy space $\mathbf{\Gamma}$, provides a complete characterization (aside from the solution concept) of a stochastic multiple person decision problem and is known as the *normal form description*. Note that in

---

[3] There would be other ways of making the objective function deterministic, such as defining the cost function as the probability of the loss exceeding a given ceiling or taking it as the supremum of the loss function over $\omega \in \Omega$. We will not be devoting much discussion to such formulations in the book.

[4] See Appendix B for an explanation of the notation used here.

this description, the information structure is suppressed and it enters the problem formulation only through the strategy spaces Γ^i, $i \in \mathcal{N}$. The description which lays out explicitly the dependence of the measurement signals on the unknown state of nature is known as the *extensive form description* of the underlying (static) stochastic decision problem. The distinction between these two forms should be more transparent when we introduce dynamic decision problems, later in this chapter. We should note, however, that the two forms are in fact equivalent in the sense that they both uniquely characterize a given stochastic decision problem; the essential difference is that sometimes it is more convenient to work with one form than the other.

Notion of Optimality

In the framework laid out above, it would have been possible to endow each decision maker (agent) with a different loss function and also possibly a different subjective probability measure regarding the unknown state of nature. Either of these departures would take us outside the realm of team problems and necessitate consideration of the more general framework of stochastic (zero-sum or nonzero-sum) games, with associated solution concepts, such as *saddle-point equilibrium* or *Nash equilibrium* [32]. Covering this more general framework is outside the scope and the goals of this book, as here our interest is in problems originating in networked control systems, where decision makers have common objectives and act as a *team*, even though the information may not be centralized. More precisely:

> A **team** is a collection of individual decision makers who strive for the same goal, using the same (probabilistic) model of the underlying decision process, but not necessarily sharing the same online information (such as measurements) on the uncertainty.

For an N-person stochastic team problem, since all agents will be striving toward the same goal, with team preferences quantified in the given loss functional, the only reasonable solution that leads to optimal behavior is the global minimization of the team cost over the product strategy space. Hence, we have

Definition 2.2.1. For a given stochastic team problem with a fixed information structure, $\{J; \Gamma^i, i \in \mathcal{N}\}$, a strategy N-tuple $\underline{\gamma}^* := (\gamma^{1*}, \ldots, \gamma^{N*}) \in \mathbf{\Gamma}$ is an *optimal team decision rule* (synonymously, *team-optimal decision rule* or simply *team-optimal solution*) if

$$J(\underline{\gamma}^*) = \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}) =: J^*, \tag{2.4}$$

provided that such a strategy exists. The cost level achieved by this strategy, J^*, is the *minimum (or optimal) team cost.* ◇

In the above definition of "optimality in a team," we have taken the information structure as fixed and given *a priori*. Even though the class of systems one typically encounters are primarily of this type, it is worth mentioning that it is possible to consider the information structure of the problem as a variable, alongside the strategies of the agents. In fact, in the theory of organizations (as well as to a large extent in the design of networked systems), the prime goal is to obtain an optimal design for the pair $(\underline{\gamma}; \underline{\eta})$ which is known as the *organizational form* (Marschak and Radner [255]). Of course, to make the problem meaningful, we have to impose some restrictions on $\underline{\eta}$ (such as belonging to some prescribed class of comparable information structures, say $\underline{N}$) or attach some cost to it which would be directly proportional with its *value*.[5] In the absence of such realistic restraints on $\underline{\eta}$, the problem will admit the trivial solution where $\underline{\eta}^*$ (the optimal $\underline{\eta}$) allows the agents to acquire perfect information on the state of nature, ξ, and thereby $\underline{\gamma}^*$ to depend directly on ξ. Under realistic organizational constraints, however, say with $\underline{\eta} \in \underline{N}$, an optimal design $(\underline{\gamma}^*; \underline{\eta}^*) \in \mathbf{\Gamma} \times \underline{N}$ will have the property that there exists no $\underline{\eta} \in \underline{N}$ such that

$$\inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(\underline{\gamma}; \underline{\eta}) < J(\underline{\gamma}^*; \underline{\eta}^*) \tag{2.5}$$

where the cost function J may also include some additional (possibly additive) terms reflecting the costs associated with various $\underline{\eta}$'s. Furthermore, the policy space $\mathbf{\Gamma}$ implicitly depends on the choice out of $\underline{N}$, so that the product $\mathbf{\Gamma} \times \underline{N}$ is actually not *rectangular*. Note that a natural way of obtaining an optimal organizational form would be to minimize the function $J(\underline{\gamma}^*_{\underline{\eta}}; \underline{\eta})$ over $\underline{\eta} \in \underline{N}$, where $\underline{\gamma}^*_{\underline{\eta}}$ is the team-optimal solution corresponding to the fixed information structure $\underline{\eta}$. We use a subscript on $\underline{\gamma}^*$ here to explicitly point out the fact that the team-optimal solution depends on $\underline{\eta}$ structurally and in general in a fairly complicated manner, which makes the further optimization of J, as $\underline{\eta}$ varies over $\underline{N}$, a rather complex problem (not of the standard type), unless the cardinality of $\underline{N}$ is finite.

One important feature of the team-optimal solution that is worth mentioning at this point (perhaps as a cautionary remark) is that multiple solutions are *not* necessarily *interchangeable*. For a two-person team problem, for example, if the pairs of policies (γ^1, γ^2) and (β^1, β^2) are two team-optimal policy pairs, then it is not necessarily true that the pair (γ^1, β^2) will also constitute a team-optimal solution. Hence, in case of nonuniqueness of the solution, the agents need to have a common consistent rule as to which one of the possible solutions to adopt, in order to arrive at the optimum. This may require some pre-communication and pre-commitment to some protocols among the agents.

A weaker solution concept than that of team-optimality introduced in Definition 2.2.1 is that of *person-by-person optimality*, equivalently *Nash equilibrium*, introduced next.

---

[5] At this point this is a rather imprecise statement. The precise meaning of **value** of a given information structure and the notion of one information structure being more **valuable** (or **better**) than another one will be introduced and studied in the next chapter, particularly Sect. 3.2.

Definition 2.2.2. For a given N-person stochastic team with a fixed information structure, $\{J; \Gamma^i, i \in \mathcal{N}\}$, an N-tuple of strategies $\underline{\gamma}^* := ({\gamma^1}^*, \dots, {\gamma^N}^*)$ constitutes a *Nash equilibrium* (synonymously, a *person-by-person optimal* (pbp optimal) solution) if, for all $\beta \in \Gamma^i$ and all $i \in \mathcal{N}$, the following inequalities hold:

$$J^* := J(\underline{\gamma}^*) \leq J(\underline{\gamma}^{-i}*, \beta), \tag{2.6}$$

where we have adopted the notation

$$(\underline{\gamma}^{-i}*, \beta) := (\gamma^1*, \dots, \gamma^{i-1}*, \beta, \gamma^{i+1}*, \dots, \gamma^N*). \tag{2.7}$$

◇

Remark 2.2.1. Nash equilibrium is a weaker solution concept than team-optimality (*cf.* Definition 2.2.1), since satisfaction of the N inequalities (2.6) is clearly necessary but not sufficient for $\underline{\gamma}^*$ to be an optimal team decision rule. But, since every team-optimal solution is necessarily a *pbp* optimal solution, the latter plays an important role in the derivation of the former, as we will see later in the book, with the first demonstration being in Sect. 2.6. ◇

In the next section, we depart from the abstract formulation of the present section and provide an illustrative example which will aid in better understanding of the concepts introduced above.

2.3 An Illustrative Example of a Finite Stochastic Team

A stochastic team problem is said to be *finite* if both the action and the uncertainty sets are finite. In this case (as we have indicated earlier) there is no need to make any distinction between Ω and Ξ,[6] and one may as well work in the original probability space $(\Omega, \mathcal{F}, P_\Omega)$ where the probability measure will be replaced with the probability masses $\{p_j = P_\Omega(\{\omega_j\})\}_{j=1}^{\#}$ where ω_j is an element of Ω with positive probability, and $\# := |\Omega|$, the cardinality of Ω, with those elements of Ω receiving *zero* probability from P_Ω being irrelevant to the decision problem and therefore deleted. By a possible abuse of terminology, we will call the $\#$ elements of Ω the states of nature. We note that $\mathcal{F}$ is a collection of all subsets of Ω (hence it has $2^{\#}$ elements), and $\mathbf{Y}^i$ can be taken, without any loss of generality, as a partition of Ω, for each $i \in \mathcal{N}$.

Every two-person finite stochastic static team can be represented by a family of matrices, each matrix (and there will be $\#$ of them) corresponding to a different state of nature, ω. The rows of these matrices would correspond to action choices of one agent, say **A**1, the columns would correspond to action choices of the other agent, **A**2, and each entry would be the corresponding loss to the team for that particular ω.

---

[6] Actually here the only requirement is that the uncertainty set be finite.

This, together with a specification of the class of all possible information signals, $(\mathbf{Y}^1, \mathbf{Y}^2)$, would constitute the *extensive form* description for the team. Such a set-up can also naturally be extended to N-person finite static teams, where now the matrices are replaced by N-dimensional hypercubes.

One approach (and a universally applicable one) toward obtaining the team-optimal solution(s) of such finite static teams is to convert the above *extensive form* into a *normal form* by relating the strategies of the agents directly to the (expected) costs that accrue to the team. As we have indicated earlier, such a formulation would suppress the information signals as well as the role of nature in the decision problem, and it would involve only a single finite, albeit larger dimensional, matrix (or hypercube, if there are more than two agents) whose columns and rows are strategy choices of the agents and whose lowest entry (or entries) would yield the team-optimal solution. Note that for $\mathbf{A}i$ the number of alternative strategies (*i.e.,* $|\Gamma^i|$) would be $|U^i|^{|\mathbf{Y}^i|}$, and hence a derivation based solely on the normal form could easily get intractable if either the number of information signals or the cardinality of the action set for at least one agent is large. It is therefore necessary to look for alternative ways of obtaining the solution, by also exploiting the nature of the information available to the agents. Note that the solution to a finite static stochastic team problem always exists (but it may be nonunique), since it involves optimization over a finite set.

With this prelude, we consider in this section a two-person static stochastic team problem where $U^1 = \{U(\text{up}), D(\text{down})\}$, $U^2 = \{L(\text{left}), R(\text{right})\}$, $\Omega = \{\omega_1, \omega_2, \omega_3\}$, $p_1 = p_2 = 0.3$, $p_3 = 0.4$, and the loss matrices are given by

A2 **A2** **A2**

A1

	L	R
U	1	0
D	3	1

	L	R
U	2	3
D	2	1

	L	R
U	1	2
D	0	2

ω : $\omega_1 \leftrightarrow 0.3$ $\omega_2 \leftrightarrow 0.3$ $\omega_3 \leftrightarrow 0.4$

Under various information structures for the team, we now study the derivation of team-optimal decision rules and some of their properties.

1. *Perfect measurements*

Here both agents have access to the true state of nature, and hence $\mathbf{Y}^1 = \mathbf{Y}^2 = \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$, the $\sigma-$field generated by the singletons. The cardinality of the strategy spaces (Γ^1 and Γ^2) is $2^3 = 8$ each, and hence the normal form is an 8×8 matrix, requiring a comparison of 64 entries. The normal form of a decision problem could also be called a *pre-commitment model*, since the strategies of the agents tell them what to do under all possible realizations of the information signal, even before the actual state of nature is realized. If, however, the agents wait to make their decisions until after they receive the measurements (which we may call the *post-commitment scenario*), then the dimension of the problem could be reduced significantly. This is particularly true when the agents' measurements are identical, as in the present case, where intuition tells us that we may, without any loss of generality, obtain the minimum value of $L(\omega; u^1, u^2)$ for each $\omega \in \Omega$ and then construct the optimal decision rules from the solutions of these individual (deterministic) teams. A mathematical justification for this intuitively appealing approach follows from the identity

$$\begin{aligned} J^* := \min_{\gamma^1, \gamma^2} J(\gamma^1, \gamma^2) &= \min_{\gamma^1, \gamma^2} E_\omega L(\omega; \gamma^1(\omega), \gamma^2(\omega)) \\ &\equiv E_\omega \{ \min_{u^1, u^2} L(\omega; u^1, u^2) \}, \end{aligned} \tag{2.8}$$

which is true since the agents have perfect measurements on ω. Note that the inner minimization in (2.8) involves the minimization of 3 loss matrices with four elements each, while the normal form required the minimization of a cost matrix with 64 elements.

The individual minima of $L(\omega; u^1, u^2)$ are

$$\min L(\omega_1; u^1, u^2) = L(\omega_1; U, R) = 0,$$

$$\min L(\omega_2; u^1, u^2) = L(\omega_2; D, R) = 1,$$

$$\min L(\omega_3; u^1, u^2) = L(\omega_3; D, L) = 0,$$

which lead [from (2.8)] to $J^* = 0.3$ and the unique team-optimal decision rules:

$$\gamma^{1*}(\omega) = \begin{cases} U, \ \omega = \omega_1, \\ D, \text{ else}, \end{cases} \qquad \gamma^{2*}(\omega) = \begin{cases} L, \ \omega = \omega_3, \\ R, \text{ else}, \end{cases}$$

which we rewrite symbolically as

$$\underline{\gamma}^* = (UDD, RRL),$$

a convention we adopt (and will henceforth use) for representing strategies in finite spaces.

As a final note we point to the observation that even though the policy pair (UDD, RRL) is unique as a team-optimal solution (which is also, by definition, *pbp* optimal), it is not the unique *pbp* optimal solution. The policy pair (UUD, RLL) is also *pbp* optimal, but it carries the unfavorable cost of 0.6 which is significantly higher than J^*.

2. *Imperfect identical measurements*

Here we consider the situation where the agents can distinguish only between the pair (ω_1, ω_2) and the singleton ω_3, and hence $\mathbf{Y}^1 = \mathbf{Y}^2 =: \mathbf{Y} = \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$. The strategy spaces have four elements each, leading to a 4×4

matrix as the normal form. We write out this matrix, for instructional purposes, with the notation $\gamma^i(y^i) = (a, b)$ (with a, b denoting the possible actions of the agents) standing for

$$\gamma^i(y^i) = \begin{cases} a, \; y^i = \{\omega_1, \omega_2\}, \\ b, \; \text{else}, \end{cases}$$

A2

		LL	LR	RL	RR
	UU	1.3	1.7	1.3	1.7
A1	UD	0.9	1.7	0.9	1.7
	DU	1.9	2.3	1.0	1.4
	DD	1.5	2.3	0.6*	1.4

The matrix has a unique minimum entry, as indicated, and hence the team problem under the given information pattern admits the unique optimal solution $\underline{\gamma}^* = (DD, RL)$, yielding a cost level of $J^* = 0.6$. Note that this is twice the optimal cost level attained under the perfect state measurements, and we can refer to the difference between the two (informally) as the "value" of the additional measurement which enables the agents to distinguish between the two states ω_1 and ω_2. Note also that in addition to the team-optimal solution given above, the problem admits one other *pbp* optimal solution, which is (UD, LL), with a corresponding (unfavorable) cost level of 0.9.

An alternative derivation for the team-optimal solution, which would involve lower-dimensional matrices, follows from a reasoning similar to the one used for the perfect information case. Here the counterpart of (2.8) would be

$$\begin{aligned} J^* := \min_{\gamma^1, \gamma^2} J(\gamma^1, \gamma^2) &= \min_{\gamma^1, \gamma^2} E_\omega L(\omega; \gamma^1(y), \gamma^2(y)) \\ &\equiv E_y \{ \min_{u^1, u^2} E_{\omega|y} L(\omega; u^1, u^2) \}, \end{aligned} \tag{2.9}$$

where we have used the "iterated property" of the conditional expectation: $E_\omega = E_y E_{\omega|y}$ where $E_{\omega|y}$ is the conditional expectation of the random variable it precedes, given that $y \in \mathbf{Y}$ has been observed.[7] Also, since we are operating in finite spaces, expression (2.9) is well defined and thus we are allowed to interchange the operations of outer expectation (over $y \in \mathbf{Y}$) and minimization (over $\underline{\gamma} \in \Gamma$). Now, the inner minimization in (2.9) involves two matrices, corresponding to two different (and exhaustive) choices for y : $y_1 = \{\omega_1, \omega_2\}$ and $y_2 = \{\omega_3\}$. These

---

[7] For this and other properties of conditional expectation the reader is referred to Appendix B.

matrices, which we may call *conditional cost matrices*, are as follows, with the unique optimal solution indicated in each case[8]:

A2

		L	R
A1	U	1.5	1.5
	D	2.5	1.0*

A2

	L	R
U	1	2
D	0*	2

y : $y_1 \leftrightarrow 0.6$ $y_2 \leftrightarrow 0.4$

Since y_1 occurs with probability 0.6 and y_2 with probability 0.4, the (average) optimal team cost is $J^* = (0.6)(1) + (0.4)(0) = 0.6$, attained by the unique pair of decision rules (DD, RL). Note that the first matrix admits one other *pbp* optimal solution (U, L) which, together with the team-optimal solution of the second matrix, leads to a *pbp* optimal solution for the original team, (UD, LL), which is the one found earlier using the 4×4 normal form.

3. *No measurements*

When neither agent makes any measurements, $\mathbf{Y}^1$ and $\mathbf{Y}^2$ are trivial $\sigma-$fields $\{\emptyset, \Omega\}$, and hence all permissible decision rules are *constant* maps. The normal form is the 2×2 matrix

A2

		L	R
A1	U	1.3*	1.7
	D	1.5	1.4

from which we immediately read: $J^* = 1.3$ and $\underline{\gamma}^* = (U, L)$.

4. *Nonidentical measurements: Perfect for* **A**2 *and none for* **A**1

This is the first nonsymmetric information structure that we will be studying. The information sets are $\mathbf{Y}^2 = \sigma(\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\})$, $\mathbf{Y}^1 = \sigma(\{\omega_1, \omega_2, \omega_3\})$, leading to *eight* elements for Γ^2 and *two* for Γ^1. The normal form is given by the *two-by-eight* matrix

A2

		LLL	LLR	LRL	LRR	RLL	RLR	RRL	RRR
A1	U	1.3	1.7	1.6	2.0	1.0	1.4	1.3	1.7
	D	1.5	2.3	1.2	2.0	0.9	1.7	0.6*	1.4

---

[8] The entries of the first matrix are obtained from the relationship $E_{\omega|y_1} L(\omega; u^1, u^2) = L(\omega; u^1, u^2)p_{1|1} + L(\omega; u^1, u^2)p_{2|1}$, where $p_{1|1}$ and $p_{2|1}$ are the conditional probabilities, each equal to $1/2$.

and the unique team-optimal solution is, as indicated, $\underline{\gamma}^* = (D, RRL)$, and the value is $J^* = 0.6$. Note that the optimal cost is the same here as in case 2, even though the information structures are incomparable. (As compared with case 2, here $\mathbf{A}1$ has **worse** and $\mathbf{A}2$ has **better** information, in the sense that $\mathbf{Y}^1_{(4)} \subset \mathbf{Y}^1_{(2)}$ and $\mathbf{Y}^2_{(4)} \supset \mathbf{Y}^2_{(2)}$, where the subscripts on $\mathbf{Y}$ refer to the two different cases and inclusion is a strict one.[9])

The question arises now as to whether a procedure similar to those used in cases 1 and 2 could also be used here to simplify the derivation (i.e., to avoid working with a large dimensional matrix, each entry of which has to be computed). Clearly, an identity such as (2.9) cannot be used since the agents do not make identical measurements. However, for each fixed decision rule γ^1 of $\mathbf{A}1$ (and there are only *two*), one can obtain the best response (minimizing solution) $T(\gamma^1)$ for $\mathbf{A}2$ by using the original matrices, since $\mathbf{A}2$ has perfect information:

$$\gamma^1 = U \Rightarrow \gamma^2 = T(U) = (RLL) \Rightarrow J(U, T(U)) = 1.0,$$
$$\gamma^1 = D \Rightarrow \gamma^2 = T(D) = (RRL) \Rightarrow J(D, T(U)) = 0.6.$$

The best choice for $\mathbf{A}1$, then, is $\gamma^{1^*} = D$, and the corresponding best response for $\mathbf{A}2$ is $\gamma^{2^*} = T(D) = (RRL)$, thus agreeing with what we had obtained earlier.

The above is yet another procedure for obtaining the team-optimal solution in two-person stochastic teams: Fix the policies of one of the agents (preferably the one whose strategy space has fewer elements), obtain the best response of the other agent to each such policy, and compute the corresponding (average) team cost in each case. The lowest such cost is the optimal team cost, and the corresponding policies are the team-optimal decision rules. Such a procedure is always justified because of the following sequence of identities (where we have taken $\mathbf{A}1$ as the starting agent):

$$\begin{aligned} J^* &= \min_{\gamma^1} \min_{\gamma^2} E_\omega L(\omega; \gamma^1(y^1), \gamma^2(y^2)) \\ &\equiv \min_{\gamma^1} E_{y^2} \left\{ \min_{u^2} E_{\omega|y^2} L(\omega; \gamma^1(y^1), u^2) \right\} \\ &\equiv \min_{\gamma^1} E_\omega L(\omega; \gamma^1(y^1), T(\gamma^1)(y^2)). \end{aligned}$$

This would be applicable even if $\mathbf{Y}^1$ and $\mathbf{Y}^2$ do not satisfy an inclusion relationship (in the particular case above we had $\mathbf{Y}^1 \subset \mathbf{Y}^2$), but then one has to construct new *conditional cost matrices* ($|\mathbf{Y}^2|$ of them, each of dimension $|\Gamma^1|$-by-$|U^2|$) in order to obtain the optimal response of $\mathbf{A}2$. [See the next case for an information structure of the type where the inclusion does not hold.]

---

[9] An equivalent statement would be $\Gamma^1_{(4)} \subset \Gamma^1_{(2)}$ and $\Gamma^2_{(4)} \supset \Gamma^2_{(2)}$. A more formal treatment of comparison of two information structures will be done in the next chapter, in Sect. 3.2.

5. *Nonidentical imperfect measurements*

This case will serve to illustrate a point which is sometimes very useful in the derivation of team-optimal solutions. Consider the information structure given by $\mathbf{Y}^1 = \sigma(\{\{\omega_1\}, \{\omega_2, \omega_3\}\})$ and $\mathbf{Y}^2 = \sigma(\{\{\omega_1, \omega_2\}, \{\omega_3\}\})$, where an inclusion property does not hold between $\mathbf{Y}^1$ and $\mathbf{Y}^2$. This, therefore, does not fall into any of the categories of information structures considered so far in this section (for the specific example). The two methods of derivation here would be:

(a) The direct solution based on the normal form (which is a 4×4 matrix)
(b) The sequential approach (which involves *two* 4×2 matrices and hence does not offer any savings (and thereby advantage) over the normal form)

These are the two general methods which would be applicable to this class of problems; however, in the present case a simple (but useful) observation yields the solution immediately: The team-optimal decision rules γ^{1^*} and γ^{2^*} for case 1 (i.e., under perfect measurements) are also well-defined functions on the signal spaces $\mathbf{Y}^1$ and $\mathbf{Y}^2$ above, and hence under the right kind of interpretation, they belong to the strategy spaces Γ^1 and Γ^2 of the present problem. The information structure in case 1 being *richer* (in fact, the *richest* possible),[10] this observation directly implies that the pair $\{\gamma^{1^*} = UD, \gamma^{2^*} = RL\}$ is the unique team-optimal solution of the new problem with “coarser” information. Note that the pair (UD, RL) here is indeed the pair (UDD, RRL) of case 1, simply rewritten using the adopted convention, on the restricted information space. If we write them out, they both correspond to

$$\gamma^{1^*}(y^1) = \begin{cases} U, \ y^1 = \{\omega_1\}, \\ \\ D, \text{ else}, \end{cases} \qquad \gamma^{2^*}(y^2) = \begin{cases} R, \ y^2 = \{\omega_1, \omega_2\}, \\ \\ L, \text{ else}. \end{cases}$$

A mathematically precise statement of the property (of the team-optimal solution) used here will be given later in the chapter.

6. *“Noisy” measurements*

For reasons which will become clear later, it is useful to distinguish between “imperfect” and “noisy” measurements. The information signals of cases 2 and 5, studied above, belong to the former category because they do not bring in additional uncertainty into the problem formulation, other than what exists already in the complete description of the cost matrices. In a sense, an imperfect measurement brings in a refinement on the information available to an agent under the

---

[10] At this point, this statement should be interpreted as saying “there is no other information structure which provides the agents with more information on the state of nature.” The underlying notion will be made precise later.

no-measurement scenario (such as case 3) **without** bringing in additional elements of uncertainty. In the "noisy measurement" case, however, the sample space has some additional elements which are not needed in the complete description of the loss (payoff) functions. To further elaborate on this point, consider the scenario depicted below, which uses essentially the same team problem as before, but with a different type of information.

Agent **A**2 makes no measurements, while **A**1 observes the value of a random variable z, taking two possible (distinct) values, y_1 and y_2. The loss matrices are the same as before, where we now adopt a different symbol, x, to replace ω, with $x_i = \omega_i$. To complete the description of the team problem, we now specify, in the following table, the joint probability mass function (*pmf*) of the pair (x, z), which has to be consistent with the marginal *pmf* of x:

	x_1	x_2	x_3
y_1	0.12	0.21	0.12
y_2	0.18	0.09	0.28

Note that

$$\mathit{Prob}(x_i \mid y_1) = \begin{cases} 4/15, \, i = 1, 3, \\ 7/15, \, i = 2, \end{cases} \qquad \mathit{Prob}(x_i \mid y_2) = \begin{cases} 18/55, \, i = 1, \\ 9/55, \quad i = 2, \\ 28/55, \, i = 3, \end{cases}$$

and hence after observing y_1 or y_2 it is not possible for **A**1 to tell, with certainty, the true value of x. We refer to the measurement signal as "noisy" because

(a) It does not transmit the true value of x (which, along with the action variables, completely determines the loss).
(b) It introduces additional elements of uncertainty into the problem.

The problem can now be cast in the framework of the general formulation of Sect. 2.2 by constructing an appropriate sample space. Toward this end, let Ω be a set of cardinality 6, with elements ω_{ij} $(i = 1, 2, 3; j = 1, 2)$, where ω_{ij} corresponds to the pair (x_i, y_j) and hence $\mathit{Prob}(\omega = \omega_{ij}) = \mathit{Prob}(x = x_i, y = y_j)$. The two possible measurement signals of **A**1 are $y_1^1 = \{\omega_{11}, \omega_{21}, \omega_{31}\}$ and $y_2^1 = \{\omega_{12}, \omega_{22}, \omega_{32}\}$ which together determine the partition $\mathbf{Y}^1$ introduced in Sect. 2.2.[11] We thus have a team problem of the standard type, for which the normal form is

A1

		UU	UD	DU	DD
A2	L	1.30*	1.38	1.42	1.50
	R	1.70	1.70	1.40	1.40

---

[11] Here, since we have a finite decision problem, we do not distinguish between Ω and Ξ, and hence consider $\mathbf{Y}^1$ as a partition of the sample space Ω.

which admits the unique team-optimal solution $\underline{\gamma}^* = (UU, L)$, with a corresponding value of $J^* = 1.30$. An immediate observation here is that this is the same value as that obtained in case 3, and hence the additional (noisy) information to **A**1 is of no value to the team. We leave it to the reader to verify that if, instead, agent **A**2 had received this measurement signal, then the team-optimal solution would again be unique and be given by $\underline{\gamma}^* = (D, RL)$, yielding this time a value of $J^* = 1.29$. Hence the same measurement is of some (positive) value to the team, if received by the second agent. As a final scenario, let us consider the information structure under which **both** agents have access to the realization of z (i.e., they have a complete sharing of information, which makes the problem essentially no different from a single agent stochastic decision problem). In view of the discussion for case 2, and especially the relation (2.9), we first form the *conditional cost matrices* corresponding to the two realizations of the measurement signal, y_1 and y_2:

		A2			**A2**	
		L	R		L	R
A1	U	22/15	29/15	U	64/55*	83/55
	D	26/15	19/15*	D	72/55	83/55
z :		$y_1 \leftrightarrow 0.45$			$y_2 \leftrightarrow 0.55$	

Then we can readily read, from the above matrices, the unique team-optimal strategy pair: (DU, RL), with a corresponding cost value of 1.21. Note that here, to determine the optimal strategies, all we need are the six conditional probabilities, *Prob*$(x_i \mid y_j)$, $i = 1, 2, 3;\ j = 1, 2$, and not the individual probabilities for y_1 and y_2.[12] The latter are, of course, needed in the computation of the corresponding cost value.

It is worth noting that the main feature of this last case, which distinguishes it from the earlier ones, is that the random quantity ω (or, equivalently here, the state of nature ξ) has two identifiable components: the "payoff relevant" part, x, and the information signal, y, with some correlation between them. The role of y is to carry information regarding the true value of x, and it affects the value of the loss function **only** through the strategy of the agent who receives this information. The advantage of splitting ξ into two components, as above, may not be that obvious at this point, but we will later observe the versatility of such a formulation, especially in the context of infinite decision problems.

---

[12] This would not have been true if the agents had made nonidentical measurements.

2.4 A Mathematical Framework for Dynamic Decision Problems

As mentioned earlier in Sect. 2.2, a decision problem is said to be *dynamic* if the measurements of at least one of the agents involve past actions (of that particular agent or some other agent(s)). In the literature, the connotation "dynamic" is also used to characterize decision problems where an agent acts more than once, even if the measurements do not depend on past actions (the case of *open-loop* information structure). In principle such problems can be converted into static decision problems by essentially working in higher-dimensional spaces, but it is generally found convenient to treat them also in the context of truly dynamic problems because of the similarities in the derivation of the optimal solutions. We will have occasions to use both approaches in this book. We describe below an appropriate setup for the study of truly dynamic decision problems, restricting the exposition to discrete time.

For a truly dynamic problem, we follow the formulation of Sect. 2.2, prior to (2.1) but now replace the static relationship (2.1) with the dynamic equation

$$y^i = \eta^i(\xi; \mathbf{u}), \quad i \in \mathcal{N}, \tag{2.10}$$

where the dependence on $\mathbf{u}$ is assumed to be *strictly causal*, which means that under a given *fixed clock* the information received by each agent can depend only on actions taken in the past. To give this statement a more precise mathematical meaning, let us consider a *discrete-time* framework where actions are taken at discrete instants of time, $1, 2, \ldots, T$. Let t stand for the generic time variable and $\mathcal{T}$ denote the (discrete) time set

$$\mathcal{T} := \{1, \ldots, T\}. \tag{2.11}$$

Let u_t^i and y_t^i denote, respectively, the action (decision) variable and the information variable of agent $\mathbf{A}i$ at the time instant $t \in \mathcal{T}$. Furthermore, introduce the notation:

$$\mathbf{u}_t := \{u_t^1, \ldots, u_t^N\}, \quad \mathbf{y}_t := \{y_t^1, \ldots, y_t^N\}, \tag{2.12}$$

$$\mathbf{u}_{[t_0,t_1)} \equiv \mathbf{u}_{[t_0,t_1-1]} := \{\mathbf{u}_{t_0}, \mathbf{u}_{t_0+1}, \ldots, \mathbf{u}_{t_1-1}\} \equiv \{u^1_{[t_0,t_1)}, \ldots, u^N_{[t_0,t_1)}\}. \tag{2.13}$$

Then, under the strict causality assumption, (2.10) becomes equivalent to

$$y_t^i = \eta_t^i(\xi; \mathbf{u}_{[1,t)}), \qquad t \in \mathcal{T}, i \in \mathcal{N} \tag{2.14}$$

for some "information functions" η_t^i, $t \in \mathcal{T}$, $i \in \mathcal{N}$. The stochastic variable y_t^i, taking values in Y_t^i, is the *online* information available to $\mathbf{A}i$ which he can use in the construction of the decision u_t^i at time t, through an appropriate policy variable $\gamma_t^i : Y_t^i \to U_t^i$

$$u_t^i = \gamma_t^i(y_t^i) \equiv \gamma_t^i(\eta_t^i[\xi; \mathbf{u}_{[1,t)}]), \qquad t \in \mathcal{T}, i \in \mathcal{N}. \tag{2.15}$$

A permissible policy γ_t^i is one under which u_t^i becomes a well-defined random variable, defined on the original probability space, and taking values in $S_t^i \subset U_t^i$, where S_t^i is the *action constraint* set for $\mathbf{A}i$ at time t. Let us denote the set of all such maps by Γ_t^i, which is the *policy space* of $\mathbf{A}i$ at time t. The construction of such a policy space will depend on the problem at hand, and we will see several such constructions throughout the book. At this point let us simply assume that such a construction is given, and rewrite (2.15) in the following compact form:

$$\mathbf{u} = \underline{\gamma}(\underline{\eta}[\xi; \mathbf{u}]), \quad \underline{\gamma} \in \boldsymbol{\Gamma} := \Gamma^1 \times \cdots \times \Gamma^N, \tag{2.16}$$

where Γ^i is the *composite* (over-time) *policy space* of $\mathbf{A}i$:

$$\Gamma^i := (\Gamma_1^i, \Gamma_2^i, \ldots, \Gamma_T^i), \qquad i \in \mathcal{N}.$$

Note that the right-hand side of (2.16) also depends on $\mathbf{u}$, which is what makes dynamic decision problems intrinsically different from the static ones introduced in Sect. 2.2. Equation (2.16) is called a *loop equation*, and a dynamic decision problem is well defined only if this loop equation has a unique solution for every ξ, that is, for some $\underline{\tilde{\gamma}} : \Xi \to \mathbf{U}$,

$$\mathbf{u} = \underline{\tilde{\gamma}}(\xi) \tag{2.17}$$

uniquely solves (2.16). The strict causality condition, or equivalently the structural assumption (2.15), is precisely the condition that guarantees this.

It is possible to relax the strict causality condition and the fixed clock assumption and replace them by some other conditions under which the loop equations (2.16) still admit a unique solution. A precise study of these conditions is beyond the level of our treatment here; for this more general treatment the reader is referred to Witsenhausen [393] and Teneketzis [360] (see also Sect. 3.7 for further discussions). To just provide a flavor of these extensions here, let us note that in (2.15) we can allow u_t^i to depend on $u_t^j, j \neq i$, provided that u_t^j is not allowed to depend on u_t^i either directly or through the actions of other agents. In Fig. 2.1 we have depicted two such scenarios for a four-agent problem with time step t isolated. A pointed arrow indicates that information flows at this stage in the direction of the arrow. The first (upper) configuration of Fig. 2.1 does not lead to a well-defined (physically realizable) decision problem because u_t^4 depends on u_t^1 while at the same time u_t^1 depends on u_t^4; hence there is a deadlock. (Clearly closed directed graphs should not be allowed for unique solvability of the loop equations.) The second (lower) configuration of Fig. 2.1, on the other hand, depicts an acceptable information exchange[13] since it does not contain any closed direct graph. Hence,

---

[13] Here, and throughout, we are using a finer partition of the time interval in between the two (discrete) time points t and $t+1$, so that agents can select u_t^i, $i \in \mathcal{N}$, in a (partially) sequential order. A possible strict time order of the configuration of Fig. 2.1 (lower) is $(u_t^1, u_t^4, u_t^2, u_t^3)$, but not all configurations have to have a strict time order.

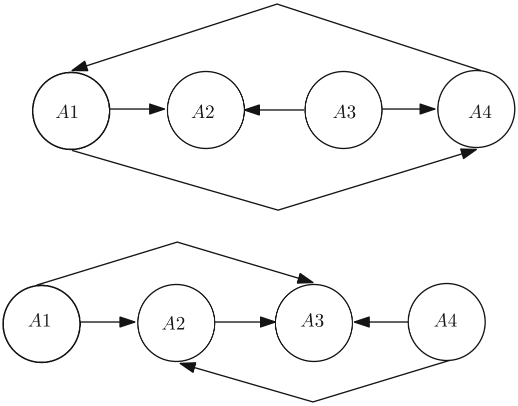

Fig. 2.1 Two scenarios on available action information at stage t

the strictly causal relation (2.15) can accommodate such permissible informational relationships among the u_t^i's, and the configuration could be different for each t.

In such a team, if there is a prespecified order in which the agents act, then such a team is said to be a *sequential team*. However, one can allow these "permissible configurations" to be sample-path dependent (i.e., dependent on ξ), provided that certain measurability conditions are satisfied. This is the situation where the order in which the agents act is determined (partially) by a chance mechanism. Such a dynamic team model is said to be *a nonsequential* team.

Sequential Dynamic Teams in State-Space Form

Now, coming back to (2.14) and (2.15), it is generally convenient to introduce an intermediate variable, called the *state variable*, recursively defined by

$$x_{t+1} = f_t(x_t, u_t^1, \ldots, u_t^N; w_t^0), \quad t \in \mathcal{T}, \tag{2.18}$$

where x_1 and $w^0 := \{w_1^0, \ldots, w_T^0\}$ are random exogenous variables with given probability distributions (with the latter being the *system noise*); furthermore, x_{t+1} takes values in a given topological space X_{t+1}. We further introduce the equation

$$y_t^i = g_t^i(x_t, u_{t-1}^1, \ldots, u_{t-1}^N; w_t^i), \qquad i \in \mathcal{N}, \qquad t \in \mathcal{T}, \tag{2.19}$$

where y_t^i is again the measurement of $\mathbf{A}i$ at stage t and $\{w_t^i, t \in \mathcal{T}\}$ is known as the measurement noise of $\mathbf{A}i$, which has a given probability distribution, possibly different for different agents as to be elucidated below. Note that for $t = 1$, g_t^i would have as argument only x_t and w_t, as controls have not yet been applied. Further note that the state variable x_t, $t \in \mathcal{T}$, can easily be eliminated (for $t > 1$) in (2.19) by recursive substitution, so that y_t^i can be expressed solely in terms of the action variables and the primitive random variables:

$$y_t^i = \tilde{g}_t^i(x_1, w_{[1,t-1]}^0, \mathbf{u}_{[1,t-1]}, w_t^i). \tag{2.20}$$

Here, the primitive random variables, or the "states of nature," are

$$\xi := \{x_1, w_{[1,T]}^0, \mathbf{w}_{[1,T]}\}, \tag{2.21}$$

where $\mathbf{w}_{[1,T]}$ is defined similar to $\mathbf{u}_{[1,T]}$, with w replacing u.

Now, the operation of a decision process described by (2.18) and (2.19) would proceed chronologically as follows:

- Generation of an initial random state x_1 with distribution P_{x_1}
- Observation of measurements $\mathbf{y}_1 := \{y_1^1, \ldots, y_1^N\}$, where the composite measurement noise $\mathbf{w}_1$ has a given conditional distribution $P_{\mathbf{w}_1|x_1}, i \in \mathcal{N}$
- Application of controls $\mathbf{u}_1$
- Generation of the "system noise" w_1^0, with conditional distribution $P_{w_1^0|x_1,\mathbf{u}_1}$, and transition to the next state x_2

……………………

- $\cdots$ transition to state x_t
- Observation of measurements $\mathbf{y}_t$, with $\mathbf{w}_t \sim P_{\mathbf{w}_t|x_t}, u_{t-1}$
- Application of controls $\mathbf{u}_t$
- Generation of the "system noise" $w_t^0 \sim P_{w_t^0|x_t,\mathbf{u}_t}$

……………………

- $\cdots$ transition to state x_{T+1}

The above evolution does not completely describe the dynamic decision process, because the construction of the controls and the allowable dependence of the controls on the past measurements and/or actions have not yet been specified. Toward this end, let $\tilde{y}_t^i$ denote some prespecified subset of the collection $\{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}$, possibly a different subcollection for different $i \in \mathcal{N}$. Note that it is possible to find an η_t^i, so that

$$\tilde{y}_t^i = \eta_t^i(\xi; \mathbf{u}_{[1,t-1]}), \quad t \in \mathcal{T}, i \in \mathcal{N},$$

where ξ was introduced earlier by (2.21). The $\tilde{y}_t^i$ defined above can definitely be viewed as a (high-dimensional) vector, and it is precisely the information variable (2.14) where we have not used *tilde* on y simply not to clutter the notation. But the distinction should be clear from context. In static decision problems, there is, of course, no difference between the information and measurement variables, and indeed in Sect. 2.2 we have called y^i as both measurement variable and information variable. In dynamic problems, however, there is a distinction between the two, and this has to be recognized in the derivation of optimal solutions to dynamic teams, as we will see later.

The collection of individual information functions $\{\eta_t^i, t \in \mathcal{T}, i \in \mathcal{N}\}$ constitutes the *information structure* of the dynamic decision problem. Perhaps by a slight abuse of notation and terminology, we introduce, for each $t \in \mathcal{T}$ and $i \in \mathcal{N}$, a finite set $\mathcal{I}_t^i$ which specifies precisely which elements of the set of vectors $\{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}$ will be used in the construction of the control u_t^i, and we call the collection $\mathcal{I} := \{\mathcal{I}_t^i, i \in \mathcal{N}, t \in \mathcal{T}\}$ again as the *information structure* of the decision problem.

We list below some important information structures which will be used throughout the book.

1. *Sole prior information* (SPI): $\mathbf{A}i$ is said to have *SP* information if she makes no measurements and the only information she works with is the prior statistics on the random variables. A decision problem has *SP* information if all agents have *SP* information.
2. *Open-loop* (*OL*) *information*: $\mathbf{A}i$ is said to have *OL* information if $\mathcal{I}_t^i = \mathcal{I}_1^i$ for all $t \in \mathcal{T}$. A decision problem has *OL information structure* if all agents have OL information (which are not necessarily the same). Note that OL information is different from SPI.
3. *Complete information sharing* (*CIS*):
$$\mathcal{I}_t^i = \{\mathbf{y}_{[1,t]}, \mathbf{u}_{[1,t-1]}\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$
Here there is a complete exchange of present and past measurements as well as past actions.
4. *Complete measurement sharing* (*CMS*):
$$\mathcal{I}_t^i = \{\mathbf{y}_{[1,t]}\}, \quad i \in \mathcal{N}.$$
Here the past actions are not shared.
5. n*-step delayed information sharing* (n*DIS*):
$$\mathcal{I}_t^i = \begin{cases} \{y_{[t-n+1,t]}^i, \mathbf{y}_{[1,t-n]}, \mathbf{u}_{[1,t-n]}\}, & t > n, \\ \\ \{y_{[1,t]}^i\}, & t \leq n, \end{cases} \qquad i \in \mathcal{N}.$$
6. n*-step delayed measurement sharing* (n*DMS*):
$$\mathcal{I}_t^i = \begin{cases} \{y_{[t-n+1,t]}^i, \mathbf{y}_{[1,t-n]}\}, & t > n, \\ \\ \{y_{[1,t]}^i\}, & t \leq n, \end{cases} \qquad i \in \mathcal{N}.$$

7. *n-step delayed control sharing* (*nDCS*):

$$\mathcal{I}_t^i = \begin{cases} \{y_{[1,t]}^i, \mathbf{u}_{[1,t-n]}\}, & t > n, \\ \\ \{y_{[1,t]}^i\}, & t \leq n, \end{cases} \qquad i \in \mathcal{N}.$$

8. *k-step periodic information sharing* (*kPIS*):

$$\mathcal{I}_t^i = \begin{cases} \{y_{[\lfloor t/k \rfloor k, t]}^i, \mathbf{y}_{[1, \lfloor t/k \rfloor k]}, \mathbf{u}_{[1, \lfloor t/k \rfloor k]}\}, & t \geq k, \\ \\ \{y_{[1,t]}^i\}, & t < k, \end{cases} \qquad i \in \mathcal{N}.$$

9. *Completely decentralized information* (*CDI*):

$$\mathcal{I}_t^i = \{y_{[1,t]}^i\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

Note that this corresponds to n*DMS* with $n = T$.

All the information structures given above are of the *perfect recall* (*PR*) type, in the sense that the agents have full memory of their information in the past. An example of an information structure (*IS*) which is not of the *PR* type is

$$\mathcal{I}_t^i = \{y_t^i\}, \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

Stochastic decision problems whose *IS*s are not of the *PR* type are relatively more difficult to analyze than those with *PR* type *IS*, as we shall see later. Another class of challenging decision problems are those with so-called nonclassical ISs. Under such *IS*s an agent sees the action variable of another agent in her information set, or her information is indirectly affected by it, but she does not have access to the measurements/information based on which that action was taken; n*DCS* introduced above is one such *IS*, so could n*DMS*, *CDI*, or information structures which are not *PR*. We will say more on such nonclassical *IS*s in the next chapter, Sect. 3.2. Furthermore, we will observe that not all nonclassical information structures lead to computational difficulties. Examples will be considered further in the next chapter, as well as in Chap. 12.

Now, fixing the *IS* of a dynamic decision problem also fixes the strategy (policy) spaces of the agents, as in (2.16). To complete the description as a team problem, we have to specify the cost structure, which we do as follows:

Adopting the description (2.18) and (2.19), also known as the *state-space model*, we associate with the team the loss function[14]

$$L(x_{[1,T+1]}, \mathbf{u}_{[1,T]}) = \sum_{t \in \mathcal{T}} c_t(x_{t+1}, \mathbf{u}_t), \tag{2.22}$$

---

[14] Here, an alternative form can be $L(x_{[1,T+1]}, \mathbf{u}_{[1,T]}) = \sum_{t \in \mathcal{T}} c_t(x_t, \mathbf{u}_t) + c_{T+1}(x_{T+1})$.

where each term in the summation is known as the *incremental* (*stagewise*) *loss*. Since $x_{[1,t+1]}$ can be expressed in terms of the primitive random variables and the action variables, replacing u^i_t in (2.22) by $\gamma^i_t(y^i_t)$, $\gamma^i_t \in \Gamma^i_t$, constructed under the given IS, L becomes a function of only ξ (for each fixed $\underline{\gamma} \in \mathbf{\Gamma}$), whose expectation with respect to the subjective probability distribution function of ξ leads as in (2.2) to the cost function:

$$J(\underline{\gamma}) = E_\xi L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))\,. \tag{2.23}$$

Here L is given by (2.22), with the intermediate variables eliminated by using (2.18) and (2.19).

The function J, along with the product strategy space $\mathbf{\Gamma}$, constitutes the *normal* (*strategic*) form of the dynamic decision problem, and as such is no different (in abstract form) from the normal form introduced in Sect. 2.2 for static multiple person decision problems. Hence, all the solution concepts introduced there, viz., *team-optimality* and *person-by-person optimality* (or *Nash equilibrium*), are equally valid (and relevant) here, which we do not give to avoid repetition. In addition, however, some new features emerge due to the dynamic nature of the information pattern, which use particularly the sequential (extensive form) description of the decision problem. We introduce below two such general features associated with the team-optimal or *pbp* optimal solutions of dynamic team problems.

Definition 2.4.1. Let $D := \{J, \mathbf{\Gamma}, \mathcal{T}\}$ be a dynamic team problem which admits a solution $\underline{\gamma}^* \in \mathbf{\Gamma}$. Let $t > 1$ be an arbitrary point in $\mathcal{T}$ and consider the decision problem $D^{\beta}_{[t,T]}$ which is derived from D by setting $\underline{\gamma}_{[1,t-1]} = \underline{\beta}_{[1,t-1]}$, for an arbitrary $\underline{\beta}_{[1,t-1]} \in \mathbf{\Gamma}_{[1,t-1]}$. Then:

(i) The solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ is *strongly time consistent* (STC) if the subpolicy $\underline{\gamma}^*_{[t,T]}$ constitutes a *solution* to the dynamic team $D^{\beta}_{[t,T]}$, this being so for every $t \in \mathcal{T}, t > 1$, and every permissible $\underline{\beta}_{[1,t-1]} \in \mathbf{\Gamma}_{[1,t-1]}$.

(ii) The solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ is *weakly time consistent* (WTC) if the subpolicy $\underline{\gamma}^*_{[t,T]}$ constitutes a *solution* to the dynamic team $D^{\beta}_{[t,T]}$ when $\underline{\beta}_{[1,t-1]} = \underline{\gamma}^*_{[1,t-1]}$.

◇

Note that if an equilibrium solution is *STC*, then the past actions do not rein in the present and future actions of the agents under the same solution concept, i.e., the agents have no reason to renege (and deviate from the equilibrium policy or the course of action) even if some inadvertent deviations have taken place in the past. With the WTC solution, however, there is no incentive to renege only if the declared course of action has been followed in the past.

The Intrinsic Model and the Markov Transition Model

Before concluding this section, we should mention that in our general formulation of a dynamic team decision problem, we have allowed an agent to act multiple times, at different time instants, using possibly different information, that is, $\mathbf{A}i$ has $u^i_{[1,T]}$ as her action variable. An alternative (but equivalent) formulation would be to have an agent $\mathbf{A}i$ be split into T agents, with $\mathbf{A}$*i(t)* (t'th agent in this split, where $t = 1, \dots, T$) controlling only u^i_t. This would then transform the original N-agent team to an NT-agent team problem, but other than a difference in semantics, the two formulations are essentially the same. More details on these different viewpoints to dynamic decision problems can be found in Witsenhausen [399–401] (see Sect. 3.7, where Witsenhausen's *intrinsic model* as well as other models for dynamic teams are reviewed).

Another point worth mentioning is that an alternative to the state-space model (2.18) and (2.19) exists, especially if the probability and action spaces are finite. This so-called Markov transition model involves $N + 1$ conditional probability laws at each stage $t \in \mathcal{T}$, to replace (2.18) and (2.19). The state equation (2.18) is replaced by a controlled probability transition:

$$P_{x_{t+1|x_t,w^0_t}}(\mathbf{u}_t), \quad i \in \mathcal{N}, t \in \mathcal{T}.$$

If all the variables belong to finite spaces, then the model is completely described by a finite number of finite-dimensional (probability) matrices.

2.5 An Illustrative Example of a Finite Dynamic Team

To illustrate some salient aspects of the formulation of dynamic decision problems, we consider in this section a finite dynamic team problem with two agents and two stages and with the agents having the same subjective prior probabilities on the random variables. We will study the derivation of the team-optimal solution under several different *IS*s of the type introduced in the previous section. Since the underlying team is finite (with a finite probability space), a team-optimal solution will exist under all *IS*s.

Now, the description of the stochastic dynamic team follows: At each stage, the control (decision) spaces of the agents have two elements, as in the static team of Sect. 2.3.

$$U^1_1 = U^1_2 = \{U(\text{up}), D(\text{down})\}, \quad U^2_1 = U^2_2 = \{L(\text{left},)R(\text{right})\}.$$

The initial state, x_1, is a discrete random variable, taking two values, x_{11} and x_{12}, with respective probabilities 0.4 and 0.6. If $x_1 = x_{1i}$ and $u^1_1 = U$ or D, and $u^2_1 = L$ or R, the loss to the team (i.e., the stagewise loss, $c_1(x_1, u^1_1, u^2_1)$) is given by the "loss matrices"

A2

		L	R
A1	U	1	0
	D	3	1

$x_1 = x_{11}$

A2

	L	R
U	1	2
D	0	2

$x_1 = x_{12}$

We denote the first matrix above by LM1 and the second matrix by LM2. The transition to the second stage and the associated cost is now described as follows: Let x_2 take three distinct values, x_{21}, x_{22} and x_{23}, with the rule

$$x_2 = \begin{cases} x_{21}, \text{ if } (u_1^1, u_1^2) = (U, L) \text{ and } x_1 = x_{11}, \\ \\ x_{22}, \text{ if } (u_1^1, u_1^2) = (U, L) \text{ and } x_1 = x_{12}, \\ \\ x_{23}, \text{ otherwise.} \end{cases}$$

The corresponding cost is determined by x_2, u_2^1, u_2^2, and an independent random variable, in terms of the loss matrices LM1 and LM2 at stage 1, according to the following table.

x_2	Loss at stage 2 (c_2)
x_{21}	LM1 *w.p.* 0.4, LM2 *w.p.* 0.6
x_{22}	LM1 *w.p.* 0.5, LM2 *w.p.* 0.5
x_{23}	LM1 *w.p.* 0.3, LM2 *w.p.* 0.7

All random mechanisms are assumed to be independent and the total loss to the team is the arithmetic sum of the stagewise losses, as in (2.22). Note that here we basically have two random variables, x_1 and w_2, say, with the statistics of the latter governing the loss structure in the table above, i.e.,

$$\textit{Prob}(w_2 = w_{21} \mid x_2) = 1 - \textit{Prob}\ (w_2 = w_{22} \mid x_2) = \begin{cases} 0.4,\ x_2 = x_{21}, \\ 0.5,\ x_2 = x_{22}, \\ 0.3,\ x_2 = x_{23}, \end{cases}$$

where w_{21} corresponds to LM1 and w_{22} to LM2.

Now we specify the measurements available to the agents: It is assumed that **A**1 knows exactly the value of x_1 at stage 1, and does not make any further measurements (at stage 2). **A**2, on the other hand, makes no measurements at stage 1 but knows precisely the value of w_2 at stage 2; hence

$$y_1^1 = x_1, \quad y_2^1 : \text{void}; \quad y_1^2 : \text{void}, \quad y_2^2 = w_2.$$

For any given permissible policy, the (expected) cost to the team is given by (2.23), and with $\xi := (x_1, w_2)$. The precise form, of course, will depend on the *IS* to be adopted, as delineated below:

1. *SPI*. This will lead to the worst performance for the team, because both agents work under only the given prior information. The policy spaces of the agents are

$$\Gamma^1 = \{Uu, Ud, Du, Dd\}; \quad \Gamma^2 = \{Ll, Lr, Rl, Rr\}$$

where we have used “lower case letters” for the second-stage decisions. The only way to solve this team problem is to convert it to *normal form*, which is a 4×4 matrix

		A2			
		Ll	Lr	Rl	Rr
A1	Uu	2	2.08	2.2	1.8
	Ud	2.38	3.5	3.3	2.5
	Du	2.2	2.2	2.6	2.2
	Dd	3.3	2.9	3.7	3.2

The unique team-optimal solution is (Uu, Rr), with a cost level of $J_{SP} = 1.8$.

2. *CMS*. This is the other extreme case, when the agents know exactly what loss matrix is being optimized at each stage. Since the lowest entries of LM1 and LM2 are both *zero*, the optimal team cost is $J_{CMS} = 0$. The solution would have been the same if, instead, we had the *CIS IS*.
3. *1-step-delayed measurement sharing* (*1DMS*). Here the information available to the agents at each stage are

$$\mathcal{I}_2^1 = \mathcal{I}_1^1 = \{x_1\}, \quad \mathcal{I}_1^2 = \phi, \mathcal{I}_2^2 = \{x_1, w_2\},$$

and hence $|\Gamma_1^1| = |\Gamma_2^1| = 4$, $|\Gamma_1^2| = 2$, and $|\Gamma_2^2| = 8$. The cardinality of the composite policy spaces are $|\Gamma^1| = 8$,[15] $|\Gamma^2| = 16$, which means that the normal form would be an 8×16 dimensional matrix. One possible approach to the problem would be to compute the entries of this matrix and choose the smallest one as the solution. An alternative approach is a sequential derivation, which makes use of the fact that measurements are shared with a delay of *one* time unit, which is what we discuss below.

Suppose that the actions at stage 1 have been taken, and the agents are facing the decision problem (at the second stage) where the value of x_1 is now *common knowledge*. **A**1 has no other information, and hence his possible actions (policies) are u and d. **A**2, on the other hand, has the additional information, the precise value of w_2, and hence he has *four* possible policies: ll, lr, rl, and rr, where lr stands for $u_1^2 = 1$, if $w_2 = w_{21}$, and $u_1^2 = r$, otherwise. Now, conditioned on the value of x_1 (which is common knowledge) and the actions taken by the agents at stage 1 (which can also be considered to be common knowledge since we have a (cooperative) team problem and the information based on which these actions were taken is common knowledge), we have the following eight *total cost matrices*.

---

[15] This one is not 16 because for each value of x_1, **A**1 has four choices, which makes the total 8. Hence, one can replace $\Gamma^1 := \Gamma_1^1 \times \Gamma_2^1$ with a smaller set, without any loss in performance.

$x_1 = x_{11}$:

$(u_1^1, u_1^2) = (U, L)$

	ll	lr	rl	rr
u	2	2.6	1.6	2.2
d	2.2	3.4	1.4*	2.6

$(u_1^1, u_1^2) = (U, R)$

	ll	lr	rl	rr
u	1	1.3	0.3*	0.6
d	2.1	3.3	0.7	1.3

$(u_1^1, u_1^2) = (D, L)$

	ll	lr	rl	rr
u	4	4.3	3.3*	3.6
d	5.1	6.3	3.7	4.3

$(u_1^1, u_1^2) = (D, R)$

	ll	lr	rl	rr
u	2	2.3	1.3*	1.6
d	3.1	4.3	1.7	2.3

$x_1 = x_{12}$:

$(u_1^1, u_1^2) = (U, L)$

	ll	lr	rl	rr
u	2	2.5	1.5*	2
d	2.5	3.5	1.5*	2.5

$(u_1^1, u_1^2) = (U, R)$

	ll	lr	rl	rr
u	3	3.3	2.3*	2.6
d	4.1	5.3	2.7	3.3

$(u_1^1, u_1^2) = (U, L)$

	ll	lr	rl	rr
u	1	1.3	0.3*	0.6
d	2.1	3.3	0.7	1.3

$(u_1^1, u_1^2) = (U, R)$

	ll	lr	rl	rr
u	3	3.3	2.3*	2.6
d	4.1	5.3	2.7	3.3

In each case, the "starred" entry(ies) denotes the minimum entries of the corresponding matrices, which will be carried over to the *first* stage to determine the optimal policies there. Now, at stage 1 **A**1 knows the value of x_1, but **A**2 does not, so that possible policies for **A**1 are UU, UD, DU, and DD, while the permissible policies for **A**2 are L and R. Using the optimum entries above, we can construct an equivalent cost matrix at stage 1 through an appropriate averaging process:

A2

		L	R
	UU	1.46	1.5
A1	UD	0.74*	1.5
	DU	2.22	1.9
	DD	1.5	1.9

This is known as the *optimum cost-to-go matrix* at stage 1, because of the following interpretation that the entries admit. Consider, for example, the entry with the numerical value 2.22: If **A**1 chooses D when $x_1 = x_{11}$ and U when $x_1 = x_{12}$, and **A**2 chooses L, all at stage 1, then whatever choices are made at stage 2 (under the given information) the total (expected) cost can never be lower than 2.22. To arrive at this numerical value, we note that if $x_1 = x_{11}$, $u_1^1 = D$ and $u_1^2 = L$, the cost resulting from an optimum choice of policies at stage 2 would be 3.3 (the lowest entry of the *third* conditional total cost matrix), whereas if $x_1 = x_{12}$, $u_1^1 = U$, $u_1^2 = L$, the optimum total cost would be 1.5 (the lowest entry of the *fifth* matrix). Since the value of x_1 is not available to **A**2 at stage 1, we average these values of x_1 to obtain

$$(0.4)(3.3) + (0.6)(1.5) = 2.22.$$

All other entries of the 4×2 *cost-to-go matrix* can be computed analogously.

Clearly, the minimum cost is $J_{1DMS} = 0.74$, with the unique team-optimal policy being

$$(\gamma_1^{1*}(x_1), \gamma_2^{1*}(x_1)) = \begin{cases} Ud, \ x_1 = x_{11}, \\ Du, \ x_1 = x_{12}, \end{cases}$$

$$\gamma_1^{2*} = L, \quad \gamma_2^{2*}(x_1, w_2) = \begin{cases} r, \ w_2 = w_{21}, \\ l, \ w_1 = w_{22}, \end{cases}$$

Note that γ_2^{2*} is actually independent of the value of x_1 (which means that even if the measurement y_1^1 had not been shared, the team-optimal solution would still be the same) and that γ_2^{1*} does depend on x_1 (which means that if **A**1 were not allowed to recall the value of x_1 at stage 2, the optimal team cost would have been higher—see the next case).

4. *No sharing*, *no recall* (*NSR*). Here we have

$$\mathcal{I}_1^1 = \{x_1\}, \mathcal{I}_2^1 = \phi = \mathcal{I}_1^2, \quad \mathcal{I}_2^2 = \{w_2\}.$$

If we had allowed perfect recall (i.e., $\mathcal{I}_2^1 = \mathcal{I}_1^1$), then the solution would be the one obtained in case 3, as discussed there[16]; however, without perfect recall the solution does not follow from the one in case 3. This information structure is *nonclassical* and hence a recursive derivation as in case 3 is also not possible. The only possibility is to construct the normal form for the team, which is characterized by the 8×8 matrix given below.

A2

		Lll	Llr	Lrl	Lrr	Rll	Rlr	Rrl	Rrr
A1	UUu	2	2.54	1.54	2.08	2.2	2.5	1.5	1.8
	UDu	1.4	1.82	0.82*	1.24	2.2	2.5	1.5	1.8
	DUu	2.8	3.22	2.22	2.64	2.6	2.9	1.9	2.2
	DDu	2.2	2.5	1.5	1.8	2.6	2.9	1.9	2.2
	UUd	2.38	3.46	1.46	2.54	2.86	3.44	1.9	2.5
	UDd	2.14	3.34	0.98	1.82	3.3	4.5	1.9	2.5
	DUd	3.54	4.62	2.38	3.22	3.7	4.9	2.3	2.9
	DDd	3.3	4.5	1.9	2.50	3.7	4.9	2.3	2.9

---

[16] We should note that this is specific to the problem at hand and is not a general rule. In general, optimal team cost in case 4 and with perfect recall will be higher than the one at case 3 where some sharing of measurements is allowed.

The unique solution is (UDu, Lrl), with a cost level of $J_{NSR} = 0.82$. Note that UDu stands for

$$\gamma_1^{1*}(x_1) = \begin{cases} U, \; x_1 = x_{11}, \\ D, \; x_1 = x_{12}, \end{cases} \quad \gamma_2^{1*} = u,$$

and Lrl denotes the policy

$$\gamma_1^{2*} = L, \quad \gamma_2^{2*}(w_2) \begin{cases} r, \; w_1 = w_{21}, \\ l, \; w_1 = w_{22}, \end{cases}$$

5. *Open-loop information* (*OLI*). Here the agents use only the measurements they have obtained at stage 1, which is x_1 for **A**1 and no measurement for **A**2. The normal form here is, in fact, a submatrix that can be obtained from the 8×8 matrix of case 4. For **A**1 the permissible policies are still the same. For **A**2, however, the permissible ones are Ll, Lr, Rl, Rr, which correspond in case 4 to Lll, Lrr, Rll, Rrr. Hence, we retain only the *first*, *fourth*, *fifth*, and *eighth* columns of the matrix of case 4, and the result is the unique team-optimal solution (UDu, Lr), with a cost level of $J_{OL} = 1.24$. We should note in passing that the normal form of case 1 can also be recovered from the normal form of case 4, this time by also eliminating the *second*, *third*, *sixth*, and *seventh* rows of the matrix of case 4.

Cost Comparisons

Clearly, more information to any one agent in a team will never result in higher optimal team cost and in fact could lead to a strictly lower value. In the latter case, we say that the extra information is *useful* (or *worth receiving*). In the context of this specific example, we have the optimum cost comparisons

$$\begin{aligned} &J_{SP} = 1.8 > J_{OL} = 1.24 > J_{NSR} = 0.82 > J_{1DMS} = J_{1DIS} = 0.74 \\ &\quad > J_{CMS} = J_{CIS} = 0, \end{aligned}$$

and hence in each case the extra information to one or more agents has been worth receiving (with the exception of pure action information in the cases of *CIS* and *1DIS ISs*). The equalities $J_{1DIS} = J_{1DMS}$, $J_{CMS} = J_{CIS}$ hold not only for this specific example but also for the general stochastic team problems.

2.6 Team-Optimal Solutions for Static Teams

We present, in this section, a theory for static N-person stochastic teams, by focusing on fundamental issues such as the existence, uniqueness and derivation of team-optimal solutions, establishing conditions under which person-by-person (*pbp*) optimal solutions are also team-optimal, and studying the relationships between achievable optimal team costs and information structures in static teams.

Using the terminology and notation introduced in Sect. 2.2, we represent a general team by the $(N+1)$-tuple $\{J; \Gamma^i, i \in \mathcal{N}\}$, where the cost J is derived from a loss function, using a probability measure P on the states of nature, common to all agents. We consider the cases where the action spaces $(U^i, i \in \mathcal{N})$ are either finite or infinite but finite dimensional, and for the latter class we also include the possibility that some hard constraints may be imposed on the decision variables, in which case the action constraint sets $(S^i, i \in \mathcal{N})$ are taken as appropriate closed subsets of the corresponding action spaces.

In the first subsection (Sect. 2.6.1), we consider the class of teams which are either finite or have finite measurement spaces for all agents. For this class, we provide general existence and uniqueness results for team-optimality and discuss the relationship with *pbp* optimality and the notion of *stationarity* (which is to be defined shortly). In Sect. 2.6.2, we extend this study to teams where the measurement spaces are infinite (but finite dimensional) and develop conditions under which stationarity implies team-optimality. Section 2.6.3 discusses two special, but important, classes of teams: (1) those with quadratic loss functions, first under general probability distributions and then under the Gaussian distribution, and (2) static teams with exponentiated quadratic loss functions. We also discuss recursive algorithms for the computation of the team-optimal solution in each case.

2.6.1 Teams with Finite Measurement Spaces

We have already identified, in Sect. 2.3, one class of team problems for which an optimal solution always exists, namely, *static finite teams*, *i.e.,* teams where both the action and the measurement spaces are finite or, equivalently, the product strategy space ($\mathbf{\Gamma}$) is finite. This conclusion would also be valid for dynamic teams where the product strategy space ($\mathbf{\Gamma}$) is finite, since we would be doing a comparison among only a finite number of choices. It would be appropriate first to present this trivial, but useful, result as a **fact**.

Fact 2.6.1. *Every finite stochastic team admits at least one team-optimal solution.* ◇

Two other key observations we made in Sect. 2.3 were that multiple team-optimal solutions are not necessarily interchangeable (respecting the order) and that a *pbp* optimal solution is not necessarily team-optimal, both of which we summarize below for future reference. These two *facts* are naturally valid not only for finite teams but for infinite teams as well and further not only for static teams but also for dynamic teams in normal (strategic) form.

Fact 2.6.2. *Neither multiple team-optimal solutions nor multiple pbp optimal solutions are necessarily interchangeable.* ◇

Fact 2.6.3. *Every team-optimal solution is pbp optimal, but not vice versa.* ◇

If a team problem is not finite, then one has to bring additional structure into the formulation in order to guarantee the existence, as well as the uniqueness, of the solution. Viewing a stochastic team in normal form as one of minimization[17] of a functional, J, over a set, $\mathbf{\Gamma}$, an optimum may fail to exist (in infinite teams) for basically one of two reasons:

1. The (cost) functional J is unbounded below.
2. There exists an infimizing sequence in $\mathbf{\Gamma}$, without any limit in $\mathbf{\Gamma}$.

The former basically says that J^*, defined by the *RHS* of (2.4), is $-\infty$, implying that a sequence can be found in $\mathbf{\Gamma}$ which makes the value of J arbitrarily small (negative). The only way to avoid this difficulty is to formulate, from the beginning, a well-defined team problem whose cost is bounded away from $-\infty$. The latter reason, however, cannot be dispensed with that easily since it places some nontrivial restrictions on the topology of the product policy space $\mathbf{\Gamma}$ as well as on the structure of J. In this case J^*, defined by the *RHS* of (2.4), is a finite quantity, but one can only achieve values arbitrarily close to (but larger than) J^*, and never equal to it. This could arise if, for example, the function J has some discontinuities on $\mathbf{\Gamma}$ or $\mathbf{\Gamma}$ has some "holes" in it so that the infimizing sequence cannot have a limit in $\mathbf{\Gamma}$. The most general condition that ensures that these two things do not happen is the celebrated *Weierstrass theorem* (see Appendix A, Sect. A.5), rephrased below as a fact using the team framework.

Fact 2.6.4. *The team problem* $\{J; \Gamma^i, i \in \mathcal{N}\}$ *admits a team-optimal solution if the product policy space* $\mathbf{\Gamma}$ *is a compact subset of a normed linear vector space, and the cost function* J *is lower semicontinuous (lsc) on* $\mathbf{\Gamma}$. ◇

As one useful application of the above result, consider the class of stochastic team problems which satisfy the following four hypotheses:

$(c.1)$ Each action constraint set S^i $(i \in \mathcal{N})$ is a closed and bounded subset of the action space U^i $(i \in \mathcal{N})$ which is itself a finite-dimensional vector space.

$(c.2)$ $L(\xi; u^1, \ldots, u^N)$ is *almost surely* (*a.s.*) jointly *lsc* in $(u^1, \ldots, u^N) =: \mathbf{u}$, on $\mathbf{U} := U^1 \times \cdots \times U^N$.

$(c.3)$ Each measurement set Y^i $(i \in \mathcal{N})$ is finite, with no element receiving *zero* probability from the probability measure P, or equivalently, for each $i \in \mathcal{N}$, the

---

[17] For some background material on the optimization of functionals, see Sect. A.5 of Appendix A.

partition set $\mathbf{Y}^i$ has a finite number of elements, with each element receiving positive probability from P.

$(c.4)$ $E_{\xi|y^i} L(\xi; u^1, \ldots, u^N)$ is finite for every $y^i \in Y^i$, $u^j \in U^j$, $i, j \in \mathcal{N}$.

Then we have the following theorem:

Theorem 2.6.1. *For an N-agent static stochastic team problem satisfying $(c.1)$–$(c.4)$ above, there exists at least one team-optimal solution.* ◇

Proof. The result follows from Fact 2.6.4, once we observe that, under the given specifications, the normal form has a *lsc* cost function J on a compact policy space $\mathbf{\Gamma}$. We first show the latter, which is equivalent to showing that, for each $i \in \mathcal{N}$, Γ^i is a closed and bounded subset of a finite-dimensional space. Toward this end, let Y^i be generated (without any loss of generality) by the n_i-tuple $\{y_1^i, \ldots, y_{n_i}^i\}$, where $n_i := |Y^i|$ is finite by $(c.3)$. Then, every permissible strategy γ^i for $\mathbf{A}i$ (*i.e.,* every element of Γ^i) can be written as

$$\gamma^i(y^i) = u_j^i, \qquad \text{if} \quad y^i = y_j^i, \quad j = 1, \ldots, n_i,$$

where each u_j^i lies in S^i. Hence, each strategy can be viewed as an n_i-tuple of vectors $(u_1^i, \ldots, u_{n_i}^i)$ belonging to

$$\mathbf{S}^i := \underbrace{S^i \times \cdots \times S^i}_{n_i \text{times}} \subset \underbrace{U^i \times \cdots \times U^i}_{n_i \text{times}} =: \mathbf{U}^i,$$

which makes Γ^i isomorphic to $\mathbf{S}^i$ which is closed and bounded, and finite dimensional, since it is a finite product of S^i which itself is closed and bounded [by $(c.1)$].

We now show that J is *lsc* on $\mathbf{U} := \mathbf{U}^1 \times \cdots \times \mathbf{U}^N$. To obtain a description of J on $\mathbf{U}$, let us first introduce the notation $\mathbf{y}$ to denote an N-tuple of scalars

$$\mathbf{y} := (y_{t_1}^1, \ldots, y_{t_N}^N), \quad t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N},$$

where $y_{t_i}^i$ denotes one possible (generic) measurement of $\mathbf{A}i$. By a possible abuse of terminology, we will consider $\mathbf{y}$ as a random quantity, which has $\mathbf{N} := \prod_{i \in \mathcal{N}} n_i$ different realizations. Then, we have the following sequence of equalities:

$$\begin{aligned} J(\underline{\gamma}) &= E_\xi L(\xi; \gamma^1(\eta^1(\xi)), \ldots, \gamma^N(\eta^N(\xi))) \\ &= E_{\mathbf{y}} \underbrace{E_{\xi|\mathbf{y}} L(\xi; u_{t_1}^1, \ldots, u_{t_N}^N)}_{L_{\text{av}}(\mathbf{y}; u_{t_1}^1, \ldots, u_{t_N}^N)} \\ &\equiv \sum_{t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N}} L_{\text{av}}(\mathbf{y}; u_{t_1}^1, \ldots, u_{t_N}^N) \mathit{Prob}(\mathbf{y}), \end{aligned} \qquad (\star)$$

where the second line follows from the "iterated property" of conditional expectations (see Appendix B) and the adopted convention that $\gamma^i(y^i) = u^i_{t_i}$, when $y^i = y^i_{t_i}$.

Now, under $(c.2)$ and $(c.4)$, $L_{\mathrm{av}}(\mathbf{y}_t; u^1_{t_1}, \ldots, u^N_{t_N})$ is a *lsc* function on $\mathbf{u}$ (as well as on U) for each $\mathbf{y}_t$, since it is the integral of a *lsc* function (L) under the conditional measure *Prob*$(\xi|\mathbf{y})$, which is finite by $(c.4)$. In view of this, the last line of $(\star)$ (which is a finite weighted sum of individual *lsc* functions) provides a representation for J on U, which is *lsc*. This then completes the proof of the theorem. It is worth noting at this point that the result would be true even if condition $(c.2)$ is relaxed somewhat, requiring instead that the function $L_{\mathrm{av}}(\mathbf{y}; \mathbf{u})$ be *lsc* on $\mathbf{u}$, where $L_{\mathrm{av}}(\mathbf{y}; \mathbf{u}) := E_{\xi|\mathbf{y},\mathbf{u}} L(\xi; \mathbf{u})$. ☐

For the general result of Theorem 2.6.1 to be valid, conditions $(c.2)$ and $(c.3)$ cannot be relaxed any further, because the relaxation of $(c.2)$ (with the provision above) would lead to violation of the *lsc* part of Fact 2.6.4, and the relaxation of $(c.3)$ (meaning that some of the Y^i's might be infinite sets) leads to policy spaces that are no longer finite dimensional, in which case $(c.1)$ does not imply the compactness of $\mathbf{\Gamma}$. Relaxation of $(c.4)$, on the other hand, would lead to a J that is not *lsc* at those points of $\mathbf{u}$ where it is unbounded. Of course, this condition is automatically satisfied if the underlying probability space is finite. The only condition that can be relaxed, without affecting the basic result of the theorem is $(c.1)$. To accommodate in our formulation the situation where some (or all) of the decision variables do not have hard constraints imposed on them, we have the following substitute condition:

$(c.1')$ Let $\mathcal{N}_h$ and $\mathcal{N}_s$ be two complementary subsets of $\mathcal{N}$ (*i.e.*, $\mathcal{N}_h \cup \mathcal{N}_s = \mathcal{N}$, and $\mathcal{N}_h \cap \mathcal{N}_s = \emptyset$) such that S^i is compact for all $i \in \mathcal{N}_h$, and $S^j \equiv U^j$ for all $j \in \mathcal{N}_s$. Then, as $\sum_{j\in\mathcal{N}_s} |u^j| \to \infty$, $L(\xi; u^1, \ldots, u^{\mathcal{N}}) \to \infty$ *a.s.*, for every fixed $u^i \in S^i, i \in \mathcal{N}_h$.

This condition ensures that $u^j, j \in \mathcal{N}_s$, can be restricted to a (possibly sufficiently large) compact set, thus making the result of Theorem 2.6.1 still valid. Hence we have

Corollary 2.6.1. *An N-agent static stochastic team problem satisfying* $(c.1')$, $(c.2)$, $(c.3)$, *and* $(c.4)$ *admits at least one team-optimal solution.* ◇

Uniqueness

In view of Fact 2.6.2, it may be important to determine the conditions under which a team-optimal solution is unique, since as we have discussed earlier in Sect. 2.3, multiple optima may lead to an inferior outcome if the agents do not have a consistent protocol to resolve the dilemma. Once the existence of an optimum has been established, there would be two ways to verify uniqueness of the solution. One would be to write down a set of necessary conditions to be satisfied by the team-optimal solution and show that these conditions admit at most one solution—as to be discussed later in this subsection. A second way to verify unicity would be to use the normal form for the team and show strict convexity[18] of J over the product policy space $\mathbf{\Gamma}$, which has to be a convex set. The following theorem now does precisely that, by relating the (strict) convexity of L to the convexity of J under the hypotheses of Theorem 2.6.1.

Theorem 2.6.2. *In addition to the four hypotheses of Theorem 2.6.1 or of Corollary 2.6.1, let S^i be a convex set for each $i \in \mathcal{N}$ and $L(\xi; \cdot)$ be strictly convex on $\mathbf{U}$ a.s.*[19] *Then, the stochastic team problem admits a unique team-optimal solution.* $\diamond$

Proof of Theorem 2.6.2. First note that for X a finite-dimensional vector space and I a finite index set, if f_i, $i \in I$, is a convex (respectively, strictly convex) functional defined on X, then the functional f: $f = \sum_{i \in I} f_i$ is also convex (respectively, strictly convex) on X. Now, the construction given for J, in the proof of Theorem 2.6.1, satisfies the hypotheses of this result with $f_t(\cdot) = L_{\mathrm{av}}(\mathbf{y}_t; \cdot) \mathit{Prob}(\mathbf{y}_t)$, since for each $\mathbf{y}_t$, $L_{\mathrm{av}}(\mathbf{y}_t; \cdot)$ is strictly convex on $\mathbf{u}$ (being the conditional average of an *a.s.* strictly convex functional), and every $u^i_{t_i}$, $t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N}$, appears in at least one of the additive terms in the representation $(\star)$ for J. Note, in passing, that $\mathit{Prob}(\mathbf{y}_t)$ may not be positive for every possible N-tuple $(t_1, \ldots, t_N)$, $t_i \in \{1, \ldots, n_i\}, i \in \mathcal{N}$, but for any $j \in \mathbf{N}$, and $t_j \in \{1, \ldots, n_j\}$, $y^j_{t_j}$ will receive positive probability in at least one such sequence, since otherwise this would imply that $\mathit{Prob}(y^j = y^j_{t_j}) = 0$, a contradiction to our initial hypothesis. $\square$

The following example serves to illustrate some of the fine points of the results of Theorems 2.6.1 and 2.6.2 and the analyses that led to these results, including the construction $(\star)$ used in the proof of Theorem 2.6.1.

Example 2.6.1. Let $N = 2$, $\Xi = U^1 = U^2 \equiv \mathbb{R}$, ξ be a (continuous) random variable *uniformly distributed* on the open interval $(0, 2)$, and the loss functional L be given by

$$L(\xi; u^1, u^2) = (u^1)^2 + (u^2)^2 + \xi u^1 u^2 - u^1 - 2u^2.$$

Suppose that $\mathbf{A}1$ can tell (through his measurements and with certainty) whether the realized value of ξ belongs to the open interval $(0, 1)$ or not and $\mathbf{A}2$ can similarly tell whether it belongs to the subinterval $(\frac{1}{2}, \frac{3}{2})$ or not. The question is whether this (static) stochastic team problem admits a team-optimal solution or not and, if it does, whether it is unique and how it can be computed.

Let us first check the conditions of Theorem 2.6.1 and Corollary 2.6.1. Clearly $(c.2)$ and $(c.3)$ are satisfied, where in the latter we choose $\mathbf{y}^1 = \{(0, 1), [1, 2)\}$, $\mathbf{y}^2 = \{(\frac{1}{2}, \frac{3}{2}), (0, \frac{1}{2}] \cup [\frac{3}{2}, 2)\}$. We associate the measurement y^i_1 with the first subinterval (in each corresponding partition) and the measurement y^i_2 with the

---

[18] See Appendix A, Sect. A.4, for a definition.

[19] A random function $L(\xi; \mathbf{u})$ is *a.s. strictly convex* in $\mathbf{u}$ if the set of ξ for which L is not strictly convex in $\mathbf{u}$ is of P_ξ-measure zero.

complement (*i.e.*, the second) set. Condition $(c.1)$ is not satisfied, but $(c.1')$ is (with $\mathcal{N}_s = \mathcal{N}$), because, for each $\xi \in (0, 2)$, the *Hessian matrix* of L (see Appendix A, Sect. A.4),

$$\bigtriangledown^2 L(\xi, \mathbf{u}) = \begin{pmatrix} 2 & \xi \\ \xi & 2 \end{pmatrix},$$

is *positive definite* (*p.d.*), implying that $L(\xi, \mathbf{u}) \to \infty$ as $|u^1| + |u^2| \to \infty$, for every fixed $\xi \in (0, 2)$. Note that even if the open interval $(0, 2)$ is replaced with the closed interval $[0, 2]$ (the distribution still being the same), $(c.1')$ would still be satisfied because, even though $\bigtriangledown^2 L(\xi, \mathbf{u})$ is no longer *p.d.* at $\xi = 2$ (in fact, then choosing $u^1 = -u^2$ and letting $u^2 \to \infty$, one can drive L to $-\infty$), the singleton event $\{\xi = 2\}$ receives *zero* probability under the given continuous distribution, and hence the condition holds in the *a.s.* sense. If, however, we had a probability distribution with a *jump* at the point $\xi = 2$, assigning, say, a weight of $\frac{1}{3}$ to that single value, then $(c.1')$ would have been violated.

Now, in the course of the discussion above, we have also established the validity of the two additional hypotheses of Theorem 2.6.2 (the first one trivially and the second one because of the reason that the Hessian matrix of L is *p.d. a.s.*), from which it follows that the problem indeed admits a unique team-optimal solution.

To obtain a characterization of the solution, let us first construct the normal form, following the steps outlined in the proof of Theorem 2.6.1. Noting that the summation in $(\star)$ has *four* terms, with *Prob*$((y_i^1, y_j^2)) = \frac{1}{4}, i, j = 1, 2$, some algebra leads to the expression

$$J(\underline{\gamma}) = \frac{1}{2} \sum_{i,j=1}^{2} (u_j^i)^2 - \frac{1}{2} \sum_{j=1}^{2} u_j^1 - \sum_{j=1}^{2} u_j^2 + \frac{1}{16} [u_1^1 u_2^2 + 3u_1^1 u_1^2 + 5u_2^1 u_1^2 + 7u_2^1 u_2^2],$$

which is to be minimized with respect to $(u_1^1, u_2^1, u_1^2, u_2^2)$ over $\mathbb{R}^4$. This is a strictly convex functional and is differentiable, which means that the unique solution should satisfy (uniquely) the *stationarity* conditions:

$$\partial J / \partial u_j^i = 0, \quad i, j = 1, 2.$$

These conditions reduce to the set of *four* linear equations:

$$\left.\begin{aligned} 16u_1^1 + 3u_1^2 + u_2^2 &= 8, \\ 16u_2^1 + 5u_1^2 + 7u_2^2 &= 8, \\ 3u_1^1 + 5u_2^1 + 16u_1^2 &= 16, \\ u_1^1 + 7u_2^1 + 16u_2^2 &= 16, \end{aligned}\right\} \qquad (\circ)$$

which admits the unique solution (to the nearest 6 decimal places):

$$u^{1*} := (u_1^{1^*}, u_2^{1^*}) = (0.231214, -0.323699),$$

$$u^{2^*} := (u_1^{2^*}, u_2^{2^*}) = (1.057803, 1.127168),$$

with the minimum value being

$$J^* \approx -1.141618.$$

It would be instructive to compare this value for the team cost with what would have been achieved if the agents had not made any measurements (*i.e.*, operated in an "open-loop" fashion with no measurements). In such a case, the normal form for the team would be given by the cost functional J_{OL} (where the subscript $_{\mathrm{OL}}$ stands for "open-loop");

$$J_{\mathrm{OL}} = (u^1)^2 + (u^2)^2 - u^1 - 2u^2 + u^1 u^2,$$

since $E[\xi] = 1$. A straightforward minimization of this quadratic (and strictly convex) functional leads to the unique solution

$$u^{1^*} = 0, u^{2^*} = 1 \quad \Rightarrow \quad J_{\mathrm{OL}}^* = -1.$$

Hence, we observe that the presence of the measurements (which bring the uncertainty in the true value of ξ to intervals of length 1, instead of the original interval of length 2) leads to an improvement of (approximately) $14\,\%$ in the performance attained by the team.

Another extreme case to consider would be the information structure that provides the agents with the "maximum" information regarding the true value of ξ, which, unquestionably, is the measurement signal $y^i = \xi, i = 1, 2$ (*i.e.*, perfect measurement), unless some restrictions are imposed on the information structure. Our results, so far, as embodied in Theorems 2.6.1 and 2.6.2 and Corollary 2.6.1, are not (strictly speaking) applicable to problems of this type, since the measurements belong to infinite sets—this is the topic of the next subsection. However, because of the fact that both agents make perfect measurements here, the problem would be easy to analyze, since it is no different from a deterministic (quadratic) optimization problem. As such, this particular (team) problem is not well defined, since the objective functional is not strictly convex at $\xi = 2$, meaning (in this case) that for $\xi = 2$ the loss functional can be driven to $-\infty$, which in turn implies (in that case) that under the uniform distribution on $(0, 2)$ the cost (average loss) can be made arbitrarily small (negative).[20] The message here is that not only the characterization

---

[20] One can determine the optimal decision rules for each value of $\xi \in (0, 2)$ (this can be done analytically), substitute these unique rules into the given loss functional, and see that its integral over the interval $(0, 2)$ does not exist—which shows that J^* is unbounded under the perfect measurement information structure.

but also the existence of a solution in a stochastic team problem could very much depend on the underlying information structure. ◇

We have thus seen, in the preceding example, a constructive procedure for obtaining closed-form solutions to a stochastic static team problem with finite measurement spaces. The question now is whether there are other (alternative) ways of obtaining the solution and also of verifying team-optimality of a given candidate solution without going through the derivation. Such tools would be provided by the *necessary conditions* satisfied by a team-optimal solution, one of which is *person-by-person* (*pbp*) optimality, the necessity of which has already been given in Fact 2.6.3. Recalling Remark 2.2.1, and particularly inequality (2.7), a *pbp* solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ for a static team problem $(J, \mathbf{\Gamma})$ would be given by

$$\min_{\beta \in \Gamma^i} J(\underline{\gamma}^{-i}*, \beta) = J(\underline{\gamma}^*), \quad i \in \mathcal{N}, \tag{2.24}$$

which can equivalently be written as

$$\min_{u \in S^i} E_{\xi|y^i} L(\xi; \underline{\gamma}^{-i} * (\mathbf{y}^{-i}), u) = E_{\xi|y^i} L(\xi; \underline{\gamma} * (\mathbf{y})), \quad i \in \mathcal{N}. \tag{2.25}$$

In other words, we have N separate optimization problems, one for each agent, and in each case the remaining agents' policies frozen at their *pbp* optimal choices. Note that (2.24) is an optimization (of total expectation) in the policy space, whereas (2) is optimization (of conditional expectation) in the action space, for every value of the conditioning variable. If (2.25) admits a unique solution and if the original problem is known to have a team-optimal solution (as in the case of Theorem 2.6.1 or Corollary 2.6.1), then (2.25) provides an alternative way of obtaining that solution. If, however, (2.25) admits more than one solution, then one would like to determine whether all or some of these are team-optimal. Hence, derivation of conditions under which *pbp* optimality implies team-optimality is of natural interest. At the outset, one would expect *a.s.* convexity of $L(\xi; \mathbf{u})$ over $\mathbf{u} \in \mathbf{U}$ to play a role here. This is indeed the case, but convexity in itself is not a sufficient condition, as the following example demonstrates:

Consider the purely deterministic loss function $L : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$L(u^1, u^2) = \begin{cases} (u^1)^2 + (1 - u^2)^2, \, u^1 \geq u^2, \\ \\ (u^2)^2 + (1 - u^1)^2, \, u^1 < u^2, \end{cases}$$

which is strictly convex on $\mathbb{R}^2$. For any fixed $u^2 \in \mathbb{R}$, $\arg\min_{u^1} L(u^1, u^2) = u^2$, and likewise for fixed $u^1 \in \mathbb{R}$, $\arg\min_{u^2} L(u^1, u^2) = u^1$. Hence, there exist infinitely many *pbp* optimal solutions $(u^1 = u^2 = u, u \in \mathbb{R})$, but only one of these, namely, $(u^1 = u^2 = \frac{1}{2})$ is team-optimal.

The function L above is nondifferentiable at the *pbp* optimal points $(u^1 = u^2)$, and in view of this observation one might wonder whether a similar "negative" result can be obtained if the loss function were continuously differentiable. The following lemma outrules this possibility for deterministic problems and provides a set of (tight) sufficient conditions for a *pbp* optimal solution to be globally optimal.

Lemma 2.6.1. *Let* $L : \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N} \to \mathbb{R}$ *be a convex (deterministic) loss function, with a pbp optimal solution* $\mathbf{u}^\circ := (u^{1^\circ}, \ldots, u^{N^\circ})$. *If* L *is continuously differentiable*[21] *at* $\mathbf{u}^\circ$*, then* $\mathbf{u}^\circ$ *is globally (team) optimal.* ◇

Proof. From the definition of convexity, we have the inequality

$$L(\alpha\mathbf{v} + (1-\alpha)\mathbf{u}^\circ) \le \alpha L(\mathbf{v}) + (1-\alpha)L(\mathbf{u}^\circ)$$

for any $\mathbf{v} = (v^1, \ldots, v^N) \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$ and every $\alpha \in [0,1]$. Rearranging this inequality, we obtain, for $0 < \alpha \le 1$,

$$\frac{1}{\alpha}[L(\mathbf{u}^\circ + \alpha(\mathbf{v} - \mathbf{u}^\circ)) - L(\mathbf{u}^\circ)] \le L(\mathbf{v}) - L(\mathbf{u}^\circ),$$

and letting $\alpha \downarrow 0$, we arrive at

$$\sum_{i=1}^{N} \nabla_{u^i} L(\mathbf{u}^\circ)(v^i - u^{oi}) \le L(\mathbf{v}) - L(\mathbf{u}^\circ), \quad \forall \mathbf{v} \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N},$$

where the required derivatives exist and the chain rule applies since L is continuously differentiable at the given point. Furthermore, by the *pbp* optimality of $\mathbf{u}^\circ$, all these partial derivatives vanish, leading to

$$L(\mathbf{u}^\circ) \le L(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N},$$

which proves global optimality of $\mathbf{u}^\circ$. □

This lemma now finds a natural generalization to static stochastic team problems with finite measurement spaces. First, we formally introduce the notion of a "stationary policy."

Definition 2.6.1. Given a static stochastic team problem $\{J; \Gamma^i, i \in \mathcal{N}\}$, a policy N-tuple $\underline{\gamma} \in \mathbf{\Gamma}$ is *stationary* if (i) $J(\underline{\gamma})$ is finite, (ii) the N partial derivatives in the following equations are well defined (locally), and (iii) $\underline{\gamma}$ satisfies these equations:

$$\left[\nabla_{u^i} E_{\xi|y^i} L(\xi; \underline{\gamma}^{-i}(\mathbf{y}^{-i}), u^i)\right] |_{u^i = \gamma^i(y^i)} = 0, a.s. \quad i \in \mathcal{N}. \tag{2.26}$$

◇

---

[21] Here, if one is to generalize the space on which L is defined, such as an infinite-dimensional space, Fréchet differentiability would be a sufficient condition. In fact continuous differentiability, and thus continuity of partial derivatives for a finite-dimensional function (see Appendix A.4), implies Fréchet differentiability [140]. The key aspect required is that the chain rule in differentiation applies, which is the case for Fréchet differentiable functions, and not necessarily the case for weaker forms of differentiability. See also Radner [316] for a related discussion.

Clearly, the *stationarity condition* (2.26) is a necessary condition for (2.25) if $L(\xi; \mathbf{u})$ is continuously differentiable in each agent's action variable (not necessarily jointly) for every $\xi \in \Xi$, and $S^i, i \in \mathcal{N}$, are open subsets of finite-dimensional vector spaces. It is equivalent to (2.25) if furthermore $S^i, i \in \mathcal{N}$, are convex sets, and $L(\xi; \mathbf{u})$ is convex in $u^i, i \in \mathcal{N}$, for every $\xi \in \Xi$. The following theorem now basically says that if the convexity and continuous differentiability of L is jointly in all the agents' action variables, then a stationary policy is necessarily team-optimal.

Theorem 2.6.3. *For an N-agent static stochastic team problem, let the hypotheses $(c.3)$ and $(c.4)$ be satisfied, S^i be an open convex subset of a finite-dimensional vector space, for each $i \in \mathcal{N}$, and $L(\xi, \cdot)$ be convex and continuously differentiable on $\mathbf{S} := S^1 \times \cdots \times S^N$. Under these conditions, if the policy $\underline{\gamma}^\circ$, taking values in $\mathbf{S}$, is stationary, it is team-optimal.* ◇

Proof. Using the construction given in the proof of Theorem 2.6.2, J admits a representation on the space U, which is convex and continuously differentiable (this last property follows because by $(c.4)$ the function $L_{av}(\mathbf{y}_t; \cdot)$ is continuously differentiable, and the representation for J is a finite weighted sum of such functions). Then, the result follows by a direct application of Lemma 2.6.1. □

Example 2.6.1. continued. Returning to the static team of Example 2.6.1, so as to apply Theorem 2.6.3, first by the "monotone convergence theorem" (see the proof of Theorem 2.6.4), the conditional expectation and differentiation in (2.26) can be interchanged, leading to the equivalent stationarity conditions

$$E_{\xi|y^i}\{(\partial/\partial u^i)L(\xi; \gamma^j(y^j), u^i)\}|_{u^i=\gamma^i(y^i)} = 0, \quad i \neq j, i, j = 1, 2$$

⇔

$$2\gamma^1(y^1) + E[\xi\gamma^2(y^2)|y^1] - 1 = 0,$$

$$2\gamma^2(y^2) + E[\xi\gamma^1(y^1)|y^2] - 2 = 0.$$

Since y^1 and y^2 each take two different values, this pair of equations is in fact a set of *four* linear equations, identical with the equations (○) encountered earlier—as we would have expected. Note that to further simplify the pair of equations above, we can substitute for γ^2 from the second into the first, to arrive at a single equation in terms of γ^1,

$$4\gamma^1(y^1) + 2E[\xi|y^1] - 1 - E[\xi E[\gamma^1(y^1)|y^2]|y^1] = 0,$$

which can be solved uniquely for $u_1^1 = \gamma^1(y_1^1)$ and $u_2^1 = \gamma^1(y_2^1)$, to yield $u_1^{1^*} = 0.231214, u_2^{1^*} = -0.323699$. The stationary policies of agent 2 can likewise be obtained. Since the loss function L satisfies all the hypotheses of Theorem 2.6.3, these stationary policies are indeed team-optimal. ◇

2.6.2 Teams on Finite-Dimensional Spaces

We now extend the theory of the previous subsection from finite spaces to a class of uncountable (but finite-dimensional) measurement spaces, with the action spaces again taken as finite-dimensional vector spaces. Given such a team problem $\{J; \Gamma^i, i \in \mathcal{N}\}$, at least one of the policy spaces (Γ^i) will be infinite dimensional, which means that condition $(c.1)$ will no longer imply that the policy space $\mathbf{\Gamma}$ is compact. Hence, even though Fact 2.6.1 would still be applicable in this case, a counterpart of Theorem 2.6.1 (on the existence of a team solution) will not follow from the given conditions. If $\underline{\gamma}^* \in \mathbf{\Gamma}$ is a team-optimal solution, then it will necessarily satisfy the *pbp* optimality condition (2.25) where now $Y^i = \mathbb{R}^{r_i}, i \in \mathcal{N}$. If furthermore, the action constraint sets are open, and the function to be minimized in (2.25) is continuously differentiable in the minimizing argument, this being so for all $i \in \mathcal{N}$, then the team solution should satisfy the stationarity conditions (2.26). The question now is whether there exists a counterpart of Theorem 2.6.3, to ensure that every stationary solution is also team-optimal. We first have the following theorem which provides a set of sufficient conditions for a policy N-tuple to be team-optimal.

Theorem 2.6.4. *Let $\{J; \Gamma^i, i \in \mathcal{N}\}$ be a static stochastic team problem where $U^i \equiv \mathbb{R}^{m_i}, i \in \mathcal{N}$, the loss function $L(\xi, \mathbf{u})$ is convex and continuously differentiable in $\mathbf{u}$ a.s., and $J(\underline{\gamma})$ is bounded from below on $\mathbf{\Gamma}$. Let $\underline{\gamma}^* \in \mathbf{\Gamma}$ be a policy N-tuple with a finite cost $(J(\underline{\gamma}^*) < \infty)$, and suppose that for every $\underline{\gamma} \in \mathbf{\Gamma}$ such that $J(\underline{\gamma}) < \infty$, the following N inequalities hold:*

$$E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i*}(y^i)]\} \geq 0, \quad i \in \mathcal{N}, \tag{2.27}$$

where $E\{\cdot\}$ denotes the total expectation. Then, $\underline{\gamma}^*$ is a team-optimal policy, and it is unique if L is strictly convex in $\mathbf{u}$. ◇

Proof. First, by the convexity of L, we obtain (as in the proof of Lemma 2.6.1)

$$\frac{1}{\alpha}\left[L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})]) - L(\xi; \underline{\gamma}^*(\mathbf{y}))\right] \leq L(\xi; \underline{\gamma}(\mathbf{y})) - L(\xi; \underline{\gamma}^*(\mathbf{y})),$$

for all $\alpha \in (0, 1]$. Using the definition of J, this inequality can equivalently be written as (by taking the total expectation):

$$h(\alpha) := \frac{1}{\alpha}[E\{L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})])\} - J(\underline{\gamma}^*)] \qquad \leq J(\underline{\gamma}) - J(\underline{\gamma}^*),$$

where $\alpha \in (0, 1]$. Note that both $J(\underline{\gamma})$ and $J(\underline{\gamma}^*)$ are finite, by hypothesis, and the first random variable (i.e., the first loss function) also has a finite expectation for every $\alpha \in (0, 1]$ because of the bound provided by the inequality. Now, due to the convexity of L, its finite integral, $E\{L(\xi; \underline{\gamma}^*(\mathbf{y}) + \alpha[\underline{\gamma}(\mathbf{y}) - \underline{\gamma}^*(\mathbf{y})])\}$ is also convex in α. This leads to the conclusion that (by a property of convex functionals, given in Appendix A, Sect. A.4) $h(\alpha)$ is a monotonically nonincreasing function as $\alpha \downarrow 0$, and furthermore $h(1) \equiv J(\underline{\gamma}) - J(\underline{\gamma}^*)$ is bounded (by hypothesis). It then follows from the *monotone convergence theorem* (see Appendix B) that $\lim_{\alpha \downarrow 0} h(\alpha)$ exists, and the limit and expectation operations can be interchanged. As a consequence of continuous differentiability, this then leads to the inequality

$$\sum_{i=1}^{N} E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i^*}(y^i)]\} \quad \leq J(\underline{\gamma}) - J(\underline{\gamma}^*)$$

from which team-optimality of $\underline{\gamma}^*$ follows, since the left-hand side is nonnegative, by (2.27).

If L were strictly convex in $\mathbf{u}$, a.s., then all the inequalities above would be strict, for $\underline{\gamma} \neq \underline{\gamma}^*$, thus leading to

$$J(\underline{\gamma}^*) < J(\underline{\gamma}),$$

which says that $\underline{\gamma}^*$ is the unique team-optimal solution. □

Note that the conditions of Theorem 2.6.4 above do not include the stationarity of $\underline{\gamma}^*$, and furthermore inequalities (2.27) may not generally be easy to check, since they involve all permissible policies $\underline{\gamma}$ (with finite cost)—generally an uncountable set. It is therefore important to obtain more readily checkable conditions to replace (2.27) and to relate team-optimality to stationarity. Either one of the following two conditions will accomplish this goal:

(c.5) For all $\underline{\gamma} \in \mathit{\Gamma}$ such that $J(\underline{\gamma}) < \infty$, the following random variables have well-defined (finite) expectations (i.e., mean values):

$$\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))[\gamma^i(y^i) - \gamma^{i*}(y^i)], \quad i \in \mathcal{N}$$

(c.6) $\mathit{\Gamma}^i$ is a Hilbert space for each $i \in \mathcal{N}$ and $J(\underline{\gamma}) < \infty$ for all $\underline{\gamma} \in \mathit{\Gamma}$. Furthermore,

$$E_{\xi|y^i}\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y})\} \in \mathit{\Gamma}^i, \quad i \in \mathcal{N}.$$

Of course, (c.6) can be obtained from (c.5) if $\mathit{\Gamma}^i$, $i \in \mathcal{N}$, are taken as Hilbert spaces. Here we give it as a separate condition because in some problems (such as linear quadratic—as we shall see shortly) (c.6) follows quite readily from the problem formulations.

Theorem 2.6.5. *Let* $\{J; \mathit{\Gamma}^i, i \in \mathcal{N}\}$ *be a static stochastic team problem which satisfies all the hypotheses of Theorem 2.6.4, with the exception of the set of inequalities (2.27). Instead of (2.27), let either (c.5) or (c.6) be satisfied. Then, if* $\underline{\gamma}^* \in \mathbf{\Gamma}$ *is a stationary policy, it is also team-optimal. Such a policy is unique if* $L(\xi; \mathbf{u})$ *is strictly convex in* $\mathbf{u}$*, a.s.* ◇

Proof. We prove the result under condition (c.6) and leave its verification under (c.5) as an exercise. Clearly, what we need to show is that stationarity of $\underline{\gamma}^*$ implies [under (c.6)] the set of inequalities (2.27). Firstly note that since $\mathit{\Gamma}^i$ is a vector space, $\gamma^i - \gamma^{i*} \in \mathit{\Gamma}^i$ for every $\gamma^i \in \mathit{\Gamma}^i$, and for every $\beta^i \in \mathit{\Gamma}^i$, there exists a $\gamma^i \in \mathit{\Gamma}^i$

such that $\beta^i = \gamma^i - \gamma^{i*}$. Since $\beta^i \in \Gamma^i \Rightarrow -\beta^i \in \Gamma^i$, the set of inequalities (2.27) become equivalent to

$$E\{\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))\beta^i(y^i)\} = 0, \quad \forall \beta^i \in \Gamma^i, \quad i \in \mathcal{N}$$

$$\Leftrightarrow$$

$$E_{y^i}\{E_{\xi|y^i}[\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))]\beta^i(y^i)\} = 0, \quad \forall \beta^i \in \Gamma^i, \quad i \in \mathcal{N},$$

where the second line follows from the iterated property of conditional expectation, under condition (c.6). Since both product terms above belong to Γ^i which is a Hilbert space, and the equality is required to hold for every element of Γ^i, $i \in \mathcal{N}$, the last line becomes equivalent to

$$E_{\xi|y^i}[\nabla_{u^i} L(\xi; \underline{\gamma}^*(\mathbf{y}))] = 0, \text{a.s. } y^i, \quad i \in \mathcal{N}.$$

To complete the proof, we now have to show that the stationarity condition (2.26) implies the above, which would be true if we were able to interchange the derivative (which is a limit) and conditional expectation operations. This however is justified (using again the monotone convergence theorem, as in the proof of Theorem 2.6.4), since $J(\underline{\gamma})$ is finite for all $\underline{\gamma} \in \mathbf{\Gamma}$ and the conditional expectation above is well defined (as an element of a Hilbert space). ☐

Theorem 2.6.5 above thus provides an extension of the result of Theorem 2.6.3 from finite to infinite measurement sets. To appreciate some of the fine points of Theorems 2.6.4 and 2.6.5, let us now consider the following example, which was discussed by Radner [316] and Krainak et al. [218].

Example 2.6.2. Let $N = 2$, $\Xi = U^1 = U^2 = \mathbb{R}$, $\xi = x$ be a Gaussian random variable with zero mean and unit variance ($\sim N(0,1)$), and the loss functional be given by

$$L(x; u^1, u^2) = (u^1 - u^2)^2 e^{x^2} + 2u^1u^2.$$

Note that L is strictly convex and continuously differentiable in (u^1, u^2) for every value of x. Hence, if the true value of x were known to both agents, the problem would admit a unique team-optimal solution: $u^1 = u^2 = 0$, which is also stationary. Since this team-optimal solution does not use the precise value of x, it is certainly optimal also under "no-measurement" information (the other extreme scenario). Note, however, that in this case the only pairs that make $J(\gamma)$ finite are $u^1 = u^2 = u \in \mathbb{R}$, since

$$E[e^{x^2}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{+\frac{x^2}{2}} dx = \infty.$$

With the set of permissible policies not being an open set, clearly we cannot talk about stationarity in this case. Theorem 2.6.4 (which does not involve stationarity) is applicable here, where inequality (2.27) is satisfied trivially. Note also that for every $u \in \mathbb{R}$, $u^1 = u^2 = u$ is a *pbp optimal* solution, but only one of these is team-optimal.

Now, as a more interesting case, consider the measurement scheme:

$$y^1 = x + w^1; \qquad y^2 = x + w^2,$$

where w^1 and w^2 are independent random variables uniformly distributed on the interval $[-1, 1]$, which are also independent of x.[22] Clearly, $u^1 = u^2 = 0$ is team-optimal for this case also, but it is not obvious at the outset whether it is stationary or not. Toward this end, let us evaluate (2.26) for $i = 1$ and with $\gamma^2(y^2) = 0$:

$$(\partial/\partial u^1)E_{x,y^2|y^1}\{(u^1)^2 e^{\xi^2}\} = (\partial/\partial u^1)[(u^1)^2 E_{x|y^1}\{e^{\xi^2}\}] = 2u^1 E_{x|y^1}\{e^{\xi^2}\}$$

where the last step follows because the conditional probability density of x given y^1 is nonzero only in a finite interval (thus making the conditional expectation finite). By symmetry, it follows that both derivatives in (2.26) vanish at $u^1 = u^2 = 0$, and hence the team-optimal solution is stationary. It is not difficult to see that in fact this is the only pair of stationary policies. Note that all the hypotheses of Theorem 2.6.5 are satisfied here, under condition $(c.5)$. ◇

2.6.3 Two Special Cost Structures

We now specialize the above general result to two classes of teams with special cost structures, namely, quadratic and exponentiated quadratic loss functions. In both cases the team loss function will be strictly convex and continuously differentiable, so that (2.27) provides a sufficient condition for a policy $\underline{\gamma}^* \in \mathbf{\Gamma}$ to be team-optimal. We will further observe that the conditions of Theorem 2.6.5 are satisfied, so that stationary policies are also team-optimal.

Static Teams with Quadratic Loss

Given a probability space $(\Omega, \mathbf{F}, P_\Omega)$, and an associated vector-valued random variable ξ, let $\{J; \Gamma^i, i \in \mathcal{N}\}$ be a static stochastic team problem with the following specifications:

(i) $U^i \equiv \mathbb{R}^{m_i}, i \in \mathcal{N}$, i.e., the action spaces are unconstrained Euclidean spaces.

(ii) The loss function is a quadratic function of $\mathbf{u}$ for every ξ:

$$L(\xi; \mathbf{u}) = \sum_{i,j \in \mathcal{N}} u^{i'} R_{ij}(\xi) u^j + 2 \sum_{i \in \mathcal{N}} u^{i'} r_i(\xi) + c(\xi), \tag{2.28}$$

---

[22] Note that here the random state of nature, ξ, is chosen as $(x, w^1, w^2)'$.

where $R_{ij}(\xi)$ is a matrix-valued random variable (with $R_{ij} \equiv R'_{ji}$), $r_i(\xi)$ is a vector-valued random variable, and $c(\xi)$ is a random variable, all generated by measurable mappings on the random state of nature, ξ.

(iii) $L(\xi; \mathbf{u})$ is strictly (and uniformly) convex in $\mathbf{u}$ a.s., i.e., there exists a positive scalar α such that, with $R(\xi)$ defined as a matrix comprised of N blocks, with the *ij*'th block given by $R_{ij}(\xi)$, the matrix $R(\xi) - \alpha I$ is positive definite a.s., where I is the appropriate dimensional identity matrix.

(iv) $R(\xi)$ is uniformly bounded above, i.e., there exists a positive scalar β such that the matrix $\beta I - R(\xi)$ is positive definite a.s.

(v) $Y^i \equiv \mathbb{R}^{r_i}$, $i \in \mathcal{N}$, i.e., the measurement spaces are unconstrained Euclidean spaces.

(vi) $y^i = \eta^i(\xi), i \in \mathcal{N}$, for some appropriate Borel measurable functions η^i, $i \in \mathcal{N}$.

(vii) Γ^i is the (Hilbert) space of all Borel measurable mappings of $\gamma^i : \mathbb{R}^{r_i} \to \mathbb{R}^{m_i}$, which have bounded second moments, i.e., $E_{y^i}\{\gamma^{i'}(y^i)\gamma^i(y^i)\} < \infty$.

(viii) $E_\xi[r'_i(\xi) r_i(\xi)] < \infty, i \in \mathcal{N}; \quad E_\xi[c(\xi)] < \infty$.

Definition 2.6.2. A static stochastic team is *quadratic* if it satisfies (i)–(viii) above. It is a *standard quadratic team* if furthermore the matrix R is constant for all ξ (i.e., it is deterministic). If, in addition, ξ is a Gaussian distributed random vector, and $r_i(\xi) = Q_i\xi, \eta^i(\xi) = H^i\xi$, $i \in \mathcal{N}$, for some deterministic matrices Q_i, H^i, $i \in \mathcal{N}$, the decision problem is a *quadratic-Gaussian team* (more widely known as a linear-quadratic-Gaussian (LQG) team under some further structure on Q_i and H^i). ◇

We now first show that the cost function of a quadratic team is bounded and strictly convex on $\boldsymbol{\Gamma}$.

Proposition 2.6.1. *For a quadratic team,*

(i) $|J(\underline{\gamma})| < \infty$ *for all* $\underline{\gamma} \in \boldsymbol{\Gamma}$.
(ii) $J(\underline{\gamma})$ *is strictly convex on* $\boldsymbol{\Gamma}$. ◇

Proof. For each $\underline{\gamma} \in \boldsymbol{\Gamma}$, each component of $u^i = \gamma^i(y^i)$ is a random variable on $(\Omega, \mathcal{F}, P_\Omega)$ with a bounded second moment (i.e., it is a *second-order* random variable), this being true for all $i \in \mathcal{N}$. Now using the fact that the product of any two second-order random variables defined on the same probability space is a well-defined random variable (on the same probability space) with a finite mean value (see Appendix B), it follows that the expected value of the second term of (2.28) is finite. Furthermore, since $R(\xi)$ is uniformly bounded, the expected value of the first term satisfies the bound

$$0 \leq E\{\sum_{i,j} u^{i'} R_{ij}(\xi) u^j\} \equiv E\{\mathbf{u}' R(\xi) \mathbf{u}\} \leq \beta E\{\mathbf{u}'\mathbf{u}\},$$

where $E\{\mathbf{u}'\mathbf{u}\}$ is finite by the same reasoning as above. Then, it readily follows that $L(\xi; \mathbf{u})$, with $u^i = \gamma^i(y^i)$, $i \in \mathcal{N}$, is a well-defined random variable with a *finite* expectation. Now, since $L(\xi, \mathbf{u})$ is strictly convex in $\mathbf{u}$ for every ξ, we have the strict inequality

$$L(\xi; \tilde{\alpha}\underline{\gamma}(\mathbf{y}) + (1 - \tilde{\alpha})\underline{\hat{\gamma}}(\mathbf{y})) < \tilde{\alpha} L(\xi; \underline{\gamma}(\mathbf{y})) + (1 - \tilde{\alpha}) L(\xi; \underline{\hat{\gamma}}(\mathbf{y}))$$

for all $\tilde{\alpha} \in (0,1)$, and every $\underline{\gamma}$, $\underline{\hat{\gamma}} \in \mathbf{\Gamma}$, $\underline{\gamma} \neq \underline{\hat{\gamma}}$. Taking the expected values of both sides, which are finite as shown above, we arrive at

$$J(\tilde{\alpha}\underline{\gamma} + (1-\tilde{\alpha})\underline{\hat{\gamma}}) < \tilde{\alpha}J(\underline{\gamma}) + (1-\tilde{\alpha})J(\underline{\hat{\gamma}}),$$

which shows that J is strictly convex. □

Now, the stationarity conditions (2.26) associated with the loss functional (2.28) can be evaluated:

$$[\nabla_{u^i}\left\{E_{\xi|y^i}\sum_{k,j\in\mathcal{N}} u^{k'}R_{kj}(\xi)u^j + 2u^{i'}E_{\xi|y^i}r_i(\xi) + 2E_{\xi|y^i}\sum_{j\in\mathcal{N},j\neq i} u^{j'}r_j(\xi) + E_{\xi|y^i}c(\xi)\right\}]|_{u^i=\gamma^i(y^i)} = 0, \quad i\in\mathcal{N}$$

$$\Leftrightarrow [E_{\xi|y^i}[R_{ii}(\xi)]u^i + \sum_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}R_{ij}(\xi)u^j + E_{\xi|y^i}r_i(\xi)]|_{u^i=\gamma^i(y^i)} = 0, \quad i\in\mathcal{N}$$

$$\Leftrightarrow E_{\xi|y^i}[R_{ii}(\xi)]\gamma^i(y^i) + \sum_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j(y^j)] + E_{\xi|y^i}r_i(\xi) = 0, \quad i\in\mathcal{N}, \tag{2.29}$$

where in going from the first to the second line of the equation we have simply performed vector differentiation with respect to u^i which is outside the conditional expectation and have also used the fact that $R_{ij} \equiv R'_{ji}$.

Hence, (2.29) constitutes the set of stationarity conditions for the quadratic team. The following theorem, due to Radner [316], now says that the solution is unique and is team-optimal.

Theorem 2.6.6. *A quadratic static team (á la Definition 2.6.2) admits a unique team-optimal solution $\underline{\gamma}^* \in \mathbf{\Gamma}$, which is also the unique stationary solution satisfying (2.29).* ◇

Proof. Assuming that there exists a stationary solution [i.e., a solution to (2.29)], the uniqueness and team-optimality follow from Theorem 2.6.5, since all its hypotheses are satisfied along with condition (c.6). Hence the proof will be completed if we can show that there exists at least one solution $\underline{\gamma}^* \in \mathbf{\Gamma}$ to (2.29). Here the verification is somewhat technical and requires some results from functional analysis and particularly Hilbert spaces (which are summarized in Appendix A, Sect. A.2). We outline here the crucial steps in this verification; the approach is essentially due to Radner [316].

Let us first note that the quadratic loss function (2.28) can equivalently be written as

$$L(\xi;\mathbf{u}) = \mathbf{u}'R(\xi)\mathbf{u} + 2\mathbf{u}'\mathbf{r}(\xi) + c(\xi)$$

$$\equiv [\mathbf{u} + R^{-1}(\xi)\mathbf{r}(\xi)]'R(\xi)[\mathbf{u} + R^{-1}(\xi)\mathbf{r}(\xi)] + c(\xi) - \mathbf{r}(\xi)'R^{-1}(\xi)\mathbf{r}(\xi),$$

where $R(\xi)$ is the matrix whose ijth block is $R_{ij}(\xi)$, and has an inverse (pointwise) by the a.s. strict convexity of L. Hence, if all agents had perfect access to the precise value of ξ, the minimum value of J would be given by the expected value of the last two terms above. Since this is not the case, the actual minimum value of J will be higher, the difference being due to the error in "approximating the vector $-R^{-1}(\xi)\mathbf{r}(\xi)$ using policies out of $\mathbf{\Gamma}$." An equivalent problem, therefore, is

$$\min_{\underline{\gamma}\in\mathbf{\Gamma}} \tilde{J}(\underline{\gamma}), \quad \tilde{J}(\underline{\gamma}) := E\{\|\underline{\gamma}(\mathbf{y}) + R^{-1}(\xi)\mathbf{r}(\xi)\|^2_{R(\xi)}\}, \tag{2.30}$$

and the statement just made (in "inverted commas") can be given a precise mathematical meaning as follows.

First note that the policy space $\mathbf{\Gamma}$ is the product space $\Gamma^1 \times \cdots \times \Gamma^N$, where each Γ^i is in fact a Hilbert space, with the inner product

$$\langle \alpha^i, \beta^i \rangle_i := E[\alpha^i(y^i)'\beta^i(y^i)]$$

(see Appendix A, Sect. A.2, and Appendix B, Sect. B.1). This makes $\mathbf{\Gamma}$ also a Hilbert space, with the inner product

$$\langle \underline{\alpha}, \underline{\beta} \rangle := E\left\{\sum_{i\in\mathcal{N}} \alpha^i(y^i)'\beta^i(y^i)\right\} \equiv E[\underline{\alpha}(\mathbf{y})'\underline{\beta}(\mathbf{y})]. \tag{$\circ$}$$

Note the important restriction that $\underline{\alpha}$ and $\underline{\beta}$ are not allowed to depend on all components of $\mathbf{y}$, because different agents do not have access to the same set of measurements. Now, in order to be able to use (2.30) as a norm compatible with the given inner product, we have to change the ($\circ$) somewhat by weighting it with $R(\xi)$:

$$\langle \underline{\alpha}, \underline{\beta} \rangle = E[\underline{\alpha}(\mathbf{y})'R(\xi)\underline{\beta}(\mathbf{y})]. \tag{$\circ\circ$}$$

This actually changes $\mathbf{\Gamma}$, the space where $\underline{\alpha}$ and $\underline{\beta}$ belong, but because of the given properties of $R(\xi)$, we have an isometry between the two spaces and therefore can denote the one under the new inner product ($\circ\circ$) also by $\mathbf{\Gamma}$. If every component of $\underline{\gamma}$ were allowed to depend on the entire measurement vector $\mathbf{y}$ (which would be the case if all the agents were to share their measurements), then the set of all permissible $\underline{\gamma}'$s bounded under the norm induced by ($\circ\circ$) would be a much larger (than $\mathbf{\Gamma}$) space. Let us denote this space by $\mathbf{H}$, and note that it is also a Hilbert space, under the inner product ($\circ\circ$). An important observation now is that $\mathbf{\Gamma}$ is a closed linear subspace of $\mathbf{H}$, *closed* because every convergent sequence in $\mathbf{\Gamma}$ with a limit point will have the limit point in $\mathbf{\Gamma}$. Hence, the team-minimization problem (2.30) is in fact an orthogonal projection problem, one of orthogonally projecting the random vector $\mathbf{x}(\xi) := R^{-1}(\xi)\mathbf{r}(\xi)$ from $\mathbf{H}$ onto $\mathbf{\Gamma}$. The conditions of the *orthogonal projection theorem* given in Appendix A, Sect. A.2, are satisfied, and therefore there exists a *unique* element of $\mathbf{\Gamma}$ that solves (2.30). Furthermore, this unique element, say $\underline{\gamma}^*$, has the property that

$$\underline{\gamma}^* + \mathbf{x} \perp \gamma, \qquad \forall \gamma \in \mathbf{\Gamma}$$

(see Appendix A, Sect. A.2, for notation and terminology). Using the inner product (∘∘), this orthogonality relationship can be written as

$$\begin{aligned}
&\langle \underline{\gamma}^* + \mathbf{x}, \underline{\gamma} \rangle = E\left\{[\underline{\gamma}^*(\mathbf{y}) + \mathbf{x}(\xi)]' R(\xi) \underline{\gamma}(\mathbf{y})\right\} = 0 \\
\Leftrightarrow \\
&E\left\{\sum_{i\in\mathcal{N}} \gamma^i(y^i)' \left[\sum_{j\in\mathcal{N}} R_{ij}(\xi)\gamma^{*j}(y^j) + r_i(\xi)\right]\right\} = 0 \\
\Leftrightarrow \\
&E\Bigg\{\sum_{i\in\mathcal{N}} \gamma^i(y^i)' \Bigg[E_{\xi|y^i}\{R_{ii}(\xi)\}\gamma^{*i}(y^i) + \sum_{j\in\mathcal{N}, j\neq i} E_{\xi|y^i}\left\{R_{ij}(\xi)\gamma^{*j}(y^j)\right\} \\
&\qquad\qquad + E_{\xi|y^i} r_i(\xi)\Bigg]\Bigg\} = 0,
\end{aligned}$$

where in arriving at the last line we have used the iterative property of conditional expectations. Now, since this equality has to hold for all $\underline{\gamma} \in \mathbf{\Gamma}$ and since $\mathbf{\Gamma}$ is a Hilbert space, it follows that the expression in brackets should vanish for every $i \in \mathcal{N}$[23] which is precisely (2.29). □

The proof of the theorem, as presented above, provides us with a new interpretation to the stationarity conditions (2.29). Note that they can be rewritten (compactly) as

$$PR\underline{\gamma} + P\mathbf{r} = 0, \tag{2.31}$$

where P is a linear operator, block diagonal, with the *ii*'th block defined through

$$P_{ii}\beta^i(\xi) = E_{\xi|y^i}\beta^i(\xi), \quad i \in \mathcal{N},$$

where $\beta^i(\xi)$ is an m_i-dimensional measurable function of ξ, satisfying the boundedness condition $E\{\beta^i(\xi)'\beta^i(\xi)\} < \infty$. As such, the linear operator P is a *projection operator* defined on a Hilbert space, whose operator norm is one (see Appendix A, Sect. A.2). Note that if the agents had full access to the value of ξ, then the stationarity condition would be

$$R(\xi)\underline{\gamma}(\xi) + \mathbf{r}(\xi) = 0, \tag{2.32}$$

---

[23] Here we have used the following property of Hilbert spaces: if $\langle \alpha, \beta \rangle = 0$ for all $\beta \in \mathbf{H}$, then $\alpha \equiv 0$.

which of course admits a unique solution since the matrix $R(\xi)$ is invertible for all ξ. Hence, the stationarity equation in the decentralized measurement case is a "projected" version of the one in the centralized full information case, but note, however, that the unique (decentralized) team-optimal solution is *not* a projected version of the centralized one $(-R^{-1}\mathbf{r})$.

The unique team-optimal solution can be obtained using some approximation schemes. Viewing (2.29) (or equivalently (2.31) as a fixed-point equation (see Appendix A, Sect. A.6, for details), one approach would be to use successive approximations:

$$\begin{cases} \gamma^i_{(k+1)}(y^i) = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}\left\{\sum\limits_{j\in\mathcal{N},j\neq i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k)}(y^j)] + E_{\xi|y^i}r_i(\xi)\right\}, \\ \gamma^i_{(0)}(y^i) \quad\ \equiv 0, \quad i\in\mathcal{N}, \quad k=0,1,\dots \end{cases} \tag{2.33}$$

which is called the *parallel update scheme*, where we have taken the starting points of the iteration as the *zero* function, as an arbitrary choice. This iteration models a dynamic decision process where the agents exchange policy information at every (discrete) point in time, and at the $k+1$'th instant agent i solves the (stochastic) optimization problem

$$\begin{aligned} \min\nolimits_{\gamma^i\in\Gamma^i} J(\gamma^1_{(k)},\dots,\gamma^{i-1}_{(k)},\gamma^i,\gamma^{i+1}_{(k)},\dots,\gamma^N_{(k)}) \\ = J(\gamma^1_{(k)},\dots,\gamma^{i-1}_{(k)},\gamma^i_{(k+1)},\gamma^{i+1}_{(k)},\dots,\gamma^N_{(k)}),\ i\in\mathcal{N}. \end{aligned}$$

Clearly, if the parallel scheme converges, it will yield the (unique) team-optimal solution in the limit. However, there is generally no guarantee that it will converge, unless some conditions are imposed on the matrix R and the probabilistic structure of the problem. To state two such conditions, let us first write (2.33) in compact form:

$$\underline{\gamma}_{(k+1)} = \mathbf{F}\underline{\gamma}_{(k)} + \hat{\mathbf{r}}, \tag{2.34}$$

where $\mathbf{F}$ is a linear operator mapping $\mathbf{\Gamma}$ into itself and composed of block operators with the diagonal blocks being zero and off-diagonal blocks given by

$$[\mathbf{F}_{ij}\gamma^j](y^i) = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}E_{\xi|y^i}[R_{ij}(\xi)\gamma^j(y^j)], \quad j\neq i, j\in\mathcal{N}. \tag{2.35}$$

Furthermore, $\hat{\mathbf{r}}\in\mathbf{\Gamma}$, with the ith block vector given by

$$[\hat{\mathbf{r}}(\mathbf{y})]_i = -[E_{\xi|y^i}R_{ii}(\xi)]^{-1}E_{\xi|y^i}r_i(\xi), \quad i\in\mathcal{N}. \tag{2.36}$$

Using the notation introduced in Appendix A, Sect. A.2, let $\ll \mathbf{F} \gg$ denote the *operator norm* of $\mathbf{F}$ and $\rho(\mathbf{F})$ denote its *spectral radius*; furthermore, note the inequality

$$\rho(\mathbf{F}) \leq \quad \ll \mathbf{F} \gg .$$

The *Banach* and *successive approximation* theorems of Appendix A, Sect. A.6, now readily lead to the following result.

Proposition 2.6.2. *Consider the parallel update scheme (2.33) [equivalently (2.34)] for the general stochastic static team problem:*

(i) The iteration converges for all starting points $\underline{\gamma}_{(0)} \in \boldsymbol{\Gamma}$ *if, and only if,*

$$\rho(\mathbf{F}) < 1. \tag{2.37}$$

(ii) The iteration converges for all starting points $\underline{\gamma}_{(0)} \in \boldsymbol{\Gamma}$ *if*

$$\ll \mathbf{F} \gg \quad < 1, \tag{2.38}$$

which is therefore a sufficient condition for (2.37).

◇

It is important to note that nonsatisfaction of (2.37) does not necessarily imply that there is no recursive scheme which would compute $\underline{\gamma}^*$; in fact, there may exist nonparallel schemes or schemes that use relaxation (i.e., higher-order memory), which will have better convergence properties. As an example of a nonparallel scheme consider the so-called sequential scheme where the agents take their turns, one at a time and in strict order, to re-optimize their policies, i.e., $\gamma^i_{(k+1)}$ is determined through the minimization of

$$\min_{\gamma^i \in \Gamma^i} J(\gamma^1_{(k+1)}, \ldots, \gamma^{i-1}_{(k+1)}, \gamma^i, \gamma^{i+1}_{(k)}, \ldots, \gamma^N_{(k)}).$$

This then leads to the following counterpart of (2.33):

$$\begin{cases} \gamma^i_{(k+1)} = [E_{\xi|y^i} R_{ii}(\xi)]^{-1} \left\{ \sum_{j \in \mathcal{N}, j<i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k+1)}(y^i)] \right. \\ \qquad \left. + \sum_{j \in \mathcal{N}, j>i} E_{\xi|y^i}[R_{ij}(\xi)\gamma^j_{(k)}(y^j)] + E_{\xi|y^i} r_i(\xi) \right\}, i \in \mathcal{N}, k = 0, 1 \ldots, \\ \gamma^i_{(0)} \quad \equiv 0, \quad i \in \mathcal{N}, i \neq 1. \end{cases} \tag{2.39}$$

Note that this recursion cannot be written in a compact form as in (2.34). However, for such convex team problems, sequential schemes have more desirable convergence properties since the sequence of minimizations above leads to a monotone non-increasing sequence of positive real numbers (associated with the team cost) which has a limit (unlike the general parallel scheme in (2.33)).

Standard Quadratic Teams

We now study the class of quadratic teams where the matrix $R(\xi)$ is a constant in (ξ), i.e., R is deterministic. The basic equation of stationarity, (2.29), simplifies to

$$\gamma^i(y^i) + \sum_{j\in\mathcal{N},j\neq i} \tilde{R}_{ij} E_{\xi|y^i}[\gamma^j(y^j)] + E_{\xi|y^i}\tilde{r}_i(\xi) = 0, \quad i \in \mathcal{N}, \tag{2.40}$$

where

$$\tilde{R}_{ij} := R_{ii}^{-1}R_{ij}; \quad \tilde{r}_i(\xi) = R_{ii}^{-1}r_i(\xi). \tag{2.41}$$

Clearly, by Theorem 2.6.6, this equation admits a unique solution $\underline{\gamma}^* \in \mathbf{\Gamma}$, whenever the loss function is strictly convex (equivalently, if the constant matrix R is positive definite). The counterpart of the parallel update scheme (2.33) is

$$\gamma^i_{(k+1)}(y^i) = -\sum_{j\in\mathcal{N},j\neq i} \tilde{R}_{ij} E_{y^j|y^i}[\gamma^j_{(k)}(y^j)] - E_{\xi|y^i}\tilde{r}_i(\xi) \quad i \in \mathcal{N}, k = 0,1,\ldots, \tag{2.42}$$

which we now further study for the case $N = 2$ (i.e., with only two agents). Substituting $\gamma^2_{(k+1)}$ obtained from (2.42) into the same for $i = 1$, we obtain

$$\gamma^1_{(k+2)}(y^1) = \tilde{R}_{12}\tilde{R}_{21}E_{y^2|y^1}E_{y^1|y^2}[\gamma^1_{(k)}(y^1)] + c^1(y^1), \tag{2.43}$$

where

$$c^1(y^1) := -E_{\xi|y^1}\tilde{r}_1(\xi) + \tilde{R}_{12}E_{y^2|y^1}E_{\xi|y^2}\tilde{r}_2(\xi). \tag{2.44}$$

Note that if we instead had the sequential update, (2.39), the resulting equation for $i = 1$ would be exactly (2.43) with simply $\gamma^1_{(k)}$ replaced by $\gamma^1_{(k+1)}$. Hence the parallel and sequential update schemes are essentially identical in the case of a two-agent team problem. The following proposition states this result, along with two other useful observations.

Proposition 2.6.3. *For the standard quadratic team with $N = 2$:*

(i) The parallel update schemes (2.42) converge (to a limiting policy pair $\underline{\gamma}^* \in \mathbf{\Gamma}$, which is a team-optimal solution) if, and only if, the single update scheme (2.43) converges.

(ii) If (2.43) converges to a limiting policy $\underline{\gamma}^{1*} \in \Gamma^1$, then $\underline{\gamma}^{2*}$ is the unique team-optimal policy of agent $\mathbf{A}1$, and

$$\gamma^{2*}(y^2) = -\tilde{R}_{21}E_{y^1|y^2}[\gamma^{1*}(y^1)] - E_{\xi|y^2}\tilde{r}_2(\xi)$$

is the unique team-optimal policy of agent $\mathbf{A}2$.

(iii) The parallel and sequential update schemes require the same condition of convergence, which is

$$\rho(\tilde{R}_{12}\tilde{R}_{21}E_{\xi|y^1}E_{\xi|y^2}) < 1\,,$$

where $\rho(\cdot)$ is the spectral radius of its linear operator argument mapping Γ^1 into itself.

◇

Proof. Parts (i) and (ii) are mere observations and require no proof. Part (iii) follows from the *successive approximation theorem* of Appendix A, Sect. A.6, since in (2.43) $\tilde{R}_{12}\tilde{R}_{21}E_{y^2|y^1}E_{y^1|y^2}$, which can equivalently be written as $\tilde{R}_{12}\tilde{R}_{21}E_{\xi|y^1}E_{\xi|y^2}$ is a linear bounded operator mapping Γ^1 (a Hilbert space) into itself. See also Proposition 2.6.2, and compare (2.43) with (2.31). □

The following lemma now paves the way toward showing that the condition of Proposition 2.6.3(iii) is satisfied for all standard quadratic teams.

Lemma 2.6.2. *The loss function (2.28), with $N = 2$ and R_{ij} constant matrices, is strictly convex if, and only if, R_{22} is positive definite and*

$$\rho(\tilde{R}_{12}\tilde{R}_{21}) < 1.$$

◇

Proof. Strict convexity of L is equivalent to the positive definiteness of the matrix

$$R := \begin{pmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{pmatrix},$$

which is further equivalent to (by definition)

$$\begin{pmatrix} x \\ y \end{pmatrix}' \begin{pmatrix} R_{11} & R_{12} \\ R_{12}' & R_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} > 0, \qquad \forall \begin{pmatrix} x \\ y \end{pmatrix} \neq \underline{o},$$

where $\underline{o}$ is the zero vector in $\mathbb{R}^m$, $m := m_1 + m_2$, and x, y have dimensions compatible with the dimensions of the blocks of R. The above can be rewritten as

$$x'R_{11}x + 2x'R_{12}y + y'R_{22}y > 0$$

from which it follows that $R_{11} > 0$, $R_{22} > 0$ are necessary conditions for strict convexity. Now, minimizing this expression with respect to y, we have, by differentiation,

$$y = -R_{22}^{-1}R_{21}x$$

as the unique solution, substitution of which into the original expression leads to

$$\begin{aligned} \min_y \begin{pmatrix} x \\ y \end{pmatrix}' R \begin{pmatrix} x \\ y \end{pmatrix} &= x'(R_{11} - R_{12}R_{22}^{-1}R_{12}')x \\ &\equiv (R_{11}^{\frac{1}{2}}x)'(I - (R_{11}^{\frac{1}{2}})^{-1}R_{12}R_{22}^{-1}R_{12}'(R_{11}^{\frac{1}{2}})^{-1})(R_{11}^{\frac{1}{2}}x) > 0, \end{aligned}$$

for $x \neq 0$, where $R_{11}^{\frac{1}{2}}$ is the unique square root of R_{11}.

The strict inequality holds, for all nonzero x, if, and only if, the matrix

$$I - (R_{11}^{\frac{1}{2}})^{-1} R_{12} R_{22}^{-1} R'_{12} (R_{11}^{\frac{1}{2}})^{-1}$$

is positive definite, which is equivalent to all eigenvalues of the second (nonnegative definite) matrix to be less than one. Hence,

$$\rho((R_{11}^{\frac{1}{2}})^{-1} R_{12} R_{22}^{-1} R'_{12} (R_{11}^{\frac{1}{2}})^{-1}) \equiv \rho(R_{11}^{-1} R_{12} R_{22}^{-1} R'_{12}) < 1,$$

where we used the fact that for two square matrices A and B, $\rho(AB) = \rho(BA)$. Since $\tilde{R}_{12} = R_{11}^{-1} R_{12}$, $\tilde{R}_{21} = R_{22}^{-1} R_{21} \equiv R_{22}^{-1} R'_{12}$, this completes the proof of the lemma. □

This brings us to the following strengthened version of Theorem 2.6.6 for standard quadratic teams with $N = 2$.

Theorem 2.6.7. *For the two-agent standard quadratic team,*

(i) There exists a unique team-optimal solution $\underline{\gamma}^* \in \mathbf{\Gamma}$*, which is also the unique solution of (2.40) with* $N = 2$*.*

(ii) Both the parallel and sequential update schemes converge for all starting points in $\mathbf{\Gamma}$*.*

(iii) Agent $\mathbf{A}i$*'s optimal policy is given by the infinite sum*

$$\gamma^{i*}(y^i) = \sum_{k=0}^{\infty} (\tilde{R}_{ij} \tilde{R}_{ji} E_{y^j|y^i} E_{y^i|y^j})^k c^i(y^i), \quad i, j = 1, 2; j \neq i, \qquad (2.45)$$

where c^1 *is given by (2.44) and* c^2 *is defined by the same with 1's and 2's interchanged.* ◇

Proof. Of course, (i) follows from Theorem 2.6.6, but since (ii) implies (i) in view of Proposition 2.6.3, the independent proof that we will give for (ii) will also provide an alternative proof to this special case of Theorem 2.6.6.

To prove part (ii), it will be sufficient to verify the condition of Proposition 2.6.3(iii). Toward this end, let us first introduce a (Hilbert) space $\hat{\Gamma}^1$ of all m_1-dimensional measurable functions $\hat{\gamma}(y^1, y^2)$ with bounded second moments: $E_{\mathbf{y}}\{|\hat{\gamma}(y^1, y^2)|^2\} < \infty$. Clearly, Γ^1 is a subspace of $\hat{\Gamma}^1$. Now, the conditional expectation operator $E_{\xi|y^i} =: P^i$ is a projection operator on $\hat{\Gamma}^1$ (see Appendix B) and hence has operator norm one, for both $i = 1$ and $i = 2$. Since the product $P^1 P^2$ is also a linear bounded operator on $\hat{\Gamma}^1$, its norm is bounded by

$$\ll P^1 P^2 \gg \quad \leq \quad \ll P^1 \gg \ll P^2 \gg = 1.$$

An important observation here is that for any $\hat{\gamma} \in \hat{\Gamma}'$, $P^1 P^2 \hat{\gamma} \in \Gamma^1 \subset \hat{\Gamma}^1$, and hence we can also view the product operator $P^1 P^2$ as a bounded linear operator mapping Γ^1 into itself. Since $\tilde{R}_{12} \tilde{R}_{21}$ also maps Γ^1 into itself, we have

$$\rho(\tilde{R}_{12} \tilde{R}_{21} P^1 P^2) \leq \rho(\tilde{R}_{12} \tilde{R}_{21}) \rho(P^1 P^2),$$

which is the spectral radius inequality on Hilbert spaces (see Appendix A, Sect. A.2). The first product term above is strictly less than *one* by Lemma 2.6.2, and the second term is no greater than *one*, by

$$\rho(P^1P^2) \leq \quad \ll P^1P^2 \gg \quad \leq 1.$$

This completes the proof of (ii), in view of Proposition 2.6.3(iii).

For part (iii) simply note that for $i = 1$, (2.45) is the infinite summation obtained from (2.43) by taking $\gamma^1_{(0)} \equiv 0$, with the limit being a valid (well-defined) element of Γ^1 by part (ii). Clearly the same result holds for $i = 2$. □

Remark 2.6.1. The iteration (2.43), or more generally (2.42), is sometimes called the *infinite second guessing* scheme. If the agents had known each other's (optimal) policies, then the iteration would halt after one step. Since this knowledge is not there, they have to estimate (or guess) each other's actions, which would also involve the estimates of each other's estimates, etc., leading in general to an infinite, albeit convergent, sequence. ◇

Even though iteration (2.42) converges for $N = 2$, it does not necessarily converge for $N > 2$. This is mainly due to the fact that strict convexity of L (equivalently, positive definiteness of R) does not imply that[24] $\rho(\tilde{R}) < 1$, unless $N = 2$.

Also one intuitive explanation for this discrepancy is that for $N = 2$ the iteration (2.42) corresponds to the sequence of minimizations

$$\begin{aligned} J(\gamma^1_{(k+1)}, \gamma^2_{(k)}) \quad &= \min_{\gamma^1 \in \Gamma^2} J(\gamma^1, \gamma^2_{(k)}) =: J_{(k+1)}, \\ J(\gamma^1_{(k+1)}, \gamma^2_{(k+2)}) &= \min_{\gamma^2 \in \Gamma^2} J(\gamma^1_{(k+1)}, \gamma^2) =: J_{(k+2)}, \quad k = 0, 1, \ldots, \end{aligned}$$

with the property

$$J_{(1)} \geq J_{(2)} \geq \ldots \geq J_{(k)} \geq J_{(k+1)} \geq \ldots,$$

Hence, at every step of the iteration, the value of J cannot increase, thus generating a nonincreasing convergent sequence (of costs). This, of course, could also converge to a *pbp*-optimal solution, but we know in this case (as already shown) that a *pbp*-optimal solution (which is also stationary because of the special quadratic structure of the loss function) is also team-optimal.

For $N > 2$, however, the iteration (2.42) does not necessarily generate a monotonic cost sequence nor a subsequence that is monotonic, which is a reason for the failure of (2.42) to converge.

---

[24] Here $\tilde{R}$ is the $(m \times m)$ matrix whose diagonal blocks are zero, and off-diagonal blocks are given by $[\tilde{R}]_{ij} = \tilde{R}_{ij}$, as defined by (2.41).

In view of the above, a natural question that arises is whether there exists some other (computational) algorithm that would yield the unique solution of (2.40) (which is known to exist by Theorem 2.6.6). Toward studying this question, let us first rewrite (2.40) as follows, using the compact notation of (2.31):

$$PR\underline{\gamma} + P\mathbf{r} = 0. \tag{$*$}$$

Let us add $-\epsilon\underline{\gamma}$ to both sides (where $\epsilon > 0$), and divide throughout by ϵ, to obtain

$$\underline{\gamma} = P(I - \frac{1}{\epsilon}R)\underline{\gamma} - \frac{1}{\epsilon}P\mathbf{r}, \tag{$**$}$$

where we have used the fact that the projection operator P and the identity operator (matrix) I commute. Note that $(*)$ and $(**)$ are in fact identical equations. Now, we associate the following iteration with $(**)$:

$$\underline{\gamma}_{(k+1)} = P(I - \frac{1}{\epsilon}R)\underline{\gamma}_{(k)} - \frac{1}{\epsilon}P\mathbf{r}, \quad k = 0, 1, \tag{2.46}$$

which, in component form, is, for $i \in \mathcal{N}$, $k = 0, 1, \ldots$,

$$\gamma^i_{(k+1)}(y^i) = (I - \frac{1}{\epsilon}R_{ii})\gamma^i_{(k)}(y^i) - \frac{1}{\epsilon}\sum_{j\in\mathcal{N}, j\neq i} R_{ij}E_{\xi|y^i}\gamma^j(y^j) - \frac{1}{\epsilon}E_{\xi|y^i}r_i(\xi). \tag{2.47}$$

Clearly, if the sequence generated by (2.46) converges to a limit in $\mathbf{\Gamma}$, this also solves $(**)$ and equivalently $(*)$. Furthermore, (2.40) being a linear iteration, we know from Proposition 2.6.3(i) that the sequence $\{\underline{\gamma}_{(k)}\}$ converges if, and only if,

$$\rho(P(I - \frac{1}{\epsilon}R)) < 1.$$

Since both P and $(I - \frac{1}{\epsilon}R)$ map $\mathbf{\Gamma}$ into itself and since P has operator norm equal to *one*, this inequality will be satisfied if

$$\rho(I - \frac{1}{\epsilon}R) < 1.$$

The matrix R being positive definite, this inequality can always be met by choosing $\epsilon > 0$ sufficiently large. If $\lambda_{\max}(R)$ denotes the maximum eigenvalue of R, choosing $\epsilon > \frac{1}{2}\lambda_{\max}(R)$ will in fact do the job. The specific choice of ϵ within this region may be dictated by other considerations, such as the speed of convergence. Since the smaller the spectral radium $\rho(I - \frac{1}{\epsilon}R)$ is, the "faster" the algorithm will (in general) converge, a reasonable choice of ϵ, with this in mind, is

$$\epsilon = \frac{1}{2}[\lambda_{\max}(R) + \lambda_{\min}(R)], \tag{2.48}$$

where $\lambda_{\min}(R)$ denotes the minimum eigenvalue of R. These results are now summarized in the following proposition.

Proposition 2.6.4. *For the standard quadratic team with N agents, the parallel update scheme (31b) converges to the unique team-optimal solution whenever $\epsilon > \frac{1}{2}\lambda_{\max}(R)$. A particular value of ϵ which leads to relatively fast convergence is given by (2.48).* ◇

Proof. The result has already been verified prior to the statement of the proposition. Note that this also provides an alternative proof for Theorem 2.6.6 for the special case of standard quadratic teams. □

Remark 2.6.2. The algorithm (2.46) should be viewed (at this point) only as a computational tool, and not carry any significant interpretation in terms of the original team decision problem. There are also other variations of this algorithm which lead to convergence, but further discussion is beyond the scope and goal of this book. ◇

Quadratic-Gaussian Teams

One class of quadratic teams for which the team-optimal solution can be obtained in closed form are those where the random state of nature ξ is a Gaussian random vector. Let us decompose ξ into $N+1$ block vectors

$$\xi = (x', y^{1'}, y^{2'}, \ldots, y^{N'})' \tag{2.49}$$

of dimensions r_0, r_1, $r_2, \ldots, r_N$, respectively. Being a Gaussian random vector, ξ is completely described in terms of its mean value and covariance matrix, which we specify below:

$$E[\xi] =: \bar{\xi} = (\bar{x}', \bar{y}^{1'}, \ldots, \bar{y}^{N'}), \tag{2.50}$$

$$\text{cov}\ (\xi) =: \Sigma,\ \text{with}\ [\Sigma]_{ij} =: \Sigma_{ij}, \quad i,j = 0,1,\ldots,N, \tag{2.51}$$

$[\Sigma]_{ij}$ denotes the *ij*th block of the matrix Σ of dimension $r_i \times r_j$, which stands for the cross-variance between the *i*th and *j*th block components of ξ. We further assume (in addition to the natural condition $\Sigma \geq 0$) that $\Sigma_{ii} > 0$ for $i \in \mathcal{N}$, which means that the measurement vectors y^i's have nonsingular distributions. To complete the description of the quadratic-Gaussian team, we finally take the linear terms $r_i(\xi)$ in the loss function (2.28) to be linear in x, which makes x the "payoff relevant" part of the state of nature (recall the earlier discussion in Sect. 2.4 on the use of this terminology):

$$r_i(\xi) = D_i x, \quad i \in \mathcal{N}, \tag{2.52}$$

where D_i is an $(r_i \times r_0)$ dimensional constant matrix. Note that in Definition 2.6.2 we simply have $Q_i = (D_i, 0, 0)$.

In the characterization of the team-optimal solution for the quadratic-Gaussian team we will need the following important result on the conditional distributions of Gaussian random vectors, which we will have occasion to use also in other chapters in the book. A proof of this result can be found in any standard text on probability theory.

Lemma 2.6.3. *Let z and y be jointly Gaussian distributed random vectors with mean values $\bar{z}$, $\bar{y}$, and covariance*

$$cov\,(z,y) = \begin{pmatrix} \Sigma_{zz} & \Sigma_{zy} \\ \Sigma_{zy}' & \Sigma_{yy} \end{pmatrix} \geq 0, \quad \Sigma_{yy} > 0. \tag{2.53}$$

Then, the conditional distribution of z given y is Gaussian, with mean

$$E[z|y] = \bar{z} + \Sigma_{zy}\Sigma_{yy}^{-1}(y - \bar{y}) \tag{2.54}$$

and covariance

$$cov(z|y) = \Sigma_{zz} - \Sigma_{zy}\Sigma_{yy}^{-1}\Sigma_{zy}' \tag{2.55}$$

◇

The complete solution to the quadratic-Gaussian team is now given in the following theorem:

Theorem 2.6.8. *The quadratic-Gaussian team decision problem as formulated above admits a unique team-optimal solution that is affine in the measurement of each agent:*

$$\gamma^{i*}(y^i) = \Pi^i(y^i - \bar{y}^i) + M^i\bar{x}, \quad i \in \mathcal{N}. \tag{2.56}$$

Here, Π^i is an $(m_i \times r_i)$ matrix $(i \in \mathcal{N})$, uniquely solving the set of linear matrix equations:

$$R_{ii}\Pi^i\Sigma_{ii} + \sum_{j \in \mathcal{N}, j \neq i} R_{ij}\Pi^j\Sigma_{ji} + D_i\Sigma_{0i} = 0, \tag{2.57}$$

and M^i is an $(m_i \times r_0)$ matrix for each $i \in \mathcal{N}$, obtained as the unique solution of

$$\sum_{j \in \mathcal{N}} R_{ij}M^j + D_i = 0, \quad i \in \mathcal{N}. \tag{2.58}$$

◇

Proof. Referring back to iteration (2.47), and initializing it with $\gamma^i_{(0)} \equiv 0$, $i \in \mathcal{N}$, it follows from repeated application of Lemma 2.6.3 that $\gamma^i_{(k)}(y^i)$ that is generated by (2.47) is necessarily affine in y^i, for all $k = 1, 2, \ldots$, with the structure given by

$$\gamma^i_{(k)}(y^i) = \Pi^i_{(k)}(y^i - \bar{y}^i) + M^i_{(k)}\bar{x}.$$

By Proposition 2.6.4, this sequence converges, and the limiting solution is necessarily in the form (2.56). Further, by Theorem 2.6.6, this limiting policy should uniquely solve the stationarity equations (2.40). Therefore, all that remains to be done is to substitute (2.56) into (2.40), to arrive at (in view of Lemma 2.6.3):

$$[R_{ii}\Pi^i + \sum_{j\in\mathcal{N}, j\neq i} R_{ij}\Pi^j \Sigma_{ji}\Sigma_{ii}^{1} + D_i\Sigma_{oi}\Sigma_{ii}^{-1}](y^i - \bar{y}^i) + [\Sigma_{j\in\mathcal{N}} R_{ij}M^j + D_i]\bar{x} \equiv 0,$$

which is an identity for each $i \in \mathcal{N}$. Since $y^i - \bar{y}^i$ and $\bar{x}$ are independent, (2.57) and (2.58) readily follow. Clearly, in view of our reasoning above, the solutions to (2.57) and (2.58) have to be unique. This algebraic result can in fact also be proven directly. For (2.58), it trivially follows because it can be rewritten as

$$RM + D = 0,$$

where $M := (M^{1'}, M^{2'}, \ldots, M^{N'})'$; $D := (D^{1'}, D^{2'}, \ldots, D^{N'})'$ and hence the unique solution is

$$M = -R^{-1}D.$$

□

A quadratic-Gaussian team is known as a LQG team, if furthermore the measurements have the special structure

$$y^i = H^i x + w^i, \quad i \in \mathcal{N}, \tag{2.59}$$

where w^i, $i \in \mathcal{N}$, constitutes an independent sequence of zero-mean Gaussian random vectors, also independent of x. Let us denote the covariance of w^i, known as the measurement noise for agent $\mathbf{A}i$, by $N^i > 0$, $i \in \mathcal{N}$. Note that in this setup the state of nature is given as

$$\xi = (x', w^{1'}, \ldots, w^{N'})',$$

which is again an $r := \sum_{i=0}^{N} r_i$-dimensional Gaussian random vector.

Now, in view of (2.59), and the independence of the noise sequence, we have

$$\bar{y}^i = H^i\bar{x}, \Sigma_{0i} = \Sigma_{00}H^{i'}, \Sigma_{ij} = H^i\Sigma_{00}H^{j'}, \Sigma_{ii} = H^i\Sigma_{00}H^{i'} + N^i, i \in \mathcal{N}.$$

Clearly, by the positive definiteness of N^i's, Σ_{ii}'s are positive definite, which means that all the hypotheses of Theorem 2.6.8 are satisfied. The following corollary then follows as a special case.

Corollary 2.6.2. *The LQG team decision problem as formulated above admits a unique team-optimal solution given by*

$$\gamma^{i*}(y^i) = \Pi^i(y^i - H^i\bar{x}) + M^i\bar{x}, \quad i \in \mathcal{N}, \tag{2.60}$$

where M^i, $i \in \mathcal{N}$, *is the unique solution of (2.58) and* Π^i *solves uniquely the following version of (2.57):*

$$R_{ii}\Pi^i + \left(\sum_{j\in\mathcal{N},j\neq i} R_{ij}\Pi^j H^j \Sigma_{00} H^{i'} + D_i \Sigma_{00} H^{i'}\right)(H^i \Sigma_{00} H^{i'} + N^i)^{-1} = 0\,. \tag{2.61}$$

◇

Example 2.6.3. To illustrate the preceding results, consider the two-agent scalar LQG team with loss function

$$L(x,\mathbf{u}) = (u^1 + u^2 + x)^2 + (u^1)^2 + (u^2)^2$$

and measurements

$$y^1 = x + w^1, \quad y^2 = x + w^2$$

under the independent statistics

$$x \sim N(1,2), w^1 \sim N(0,2), w^2 \sim N(0,1).$$

Direct application of Corollary 2.6.2 leads to the unique team-optimal solution

$$\left.\begin{aligned} \gamma^{1*}(y^1) &= -\tfrac{2}{11}(y^1 - 1) - \tfrac{1}{3}, \\ \gamma^{2*}(y^2) &= -\tfrac{3}{11}(y^2 - 1) - \tfrac{1}{3}. \end{aligned}\right\} \tag{$*$}$$

The corresponding minimum team cost can be computed to be

$$J^* := J(\gamma^{1*}, \gamma^{2*}) \simeq 1.424.$$

Note that this is a symmetric team as far as the loss function goes (i.e., $L(u^1, u^2; x) = L(u^2, u^1, x)$), but as far as the measurements go agent $\mathbf{A}1$ has "higher" measurement noise than agent $\mathbf{A}2$. This is reflected in the team-optimal policies, with the measurement of $\mathbf{A}1$ weighted less than the measurement of $\mathbf{A}2$ (compare the gain $\frac{2}{11}$ against the gain $\frac{3}{11}$).

If the agents did not have access to any measurements, and thus optimize in the class of constant policies, the unique solution can easily be read off from $(*)$ to be

$$\gamma^1_{OL} = \gamma^2_{OL} = -\frac{1}{3}$$

with the corresponding cost being

$$J_{OL} := J(\gamma^1_{OL}, \gamma^2_{OL}) = \frac{7}{3} \cong 2.3333.$$

Hence, the decentralized measurements lead to about 39% improvement (reduction) in the team cost, as compared with the no-measurement (*open-loop*) case.

If, on the other hand, the agents shared their measurements, with the team's common measurement now being $\mathbf{y} = (y^1, y^2)'$, the optimum cost should be lower than J^*. To study this specific model, we first note that Theorem 2.6.8 is directly applicable here, with

$$\Sigma_{11} = \Sigma_{12} = \Sigma_{22} = \text{cov}(\mathbf{y}) = \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix}; \Sigma_{01} = \Sigma_{02} = (2, 2).$$

The unique team-optimal solution can readily be obtained to be

$$\gamma^1_{sh}(\mathbf{y}) = \gamma^2_{sh}(\mathbf{y}) = -\frac{1}{4}[\frac{1}{3}(y^1 - 1) + \frac{2}{3}(y^2 - 1)] - \frac{1}{3}$$

with the corresponding cost being

$$J_{sh} := J(\gamma^1_{sh}, \gamma^2_{sh}) = \frac{4}{3} \cong 1.3333.$$

The improvement here, over the open-loop cost, is 43 %, and over the decentralized case is about 6%.

Finally, if both agents had perfect access to the true value of x (the case of perfect measurements), the unique optimal solution would be

$$\gamma^1_{pr}(x) = \gamma^2_{pr}(x) = -\frac{1}{3}x$$

with a cost level of

$$J_{pr} := J(\gamma^1_{pr}, \gamma^2_{pr}) = \frac{1}{3}E[x^2] = 1,$$

which is the lowest possible value for J, under any measurement scheme. ◇

Positively Exponentiated Quadratic Loss

Consider again the formulation of the quadratic-Gaussian team (á la Definition 2.6.2) but with the loss function being a positively exponentiated quadratic function, i.e.,

$$L(\xi; \mathbf{u}) = \theta e^{\frac{\theta}{2}C(\xi;\mathbf{u})}, \quad \theta > 0, \tag{2.62}$$

where C is a strictly convex (in $\mathbf{u}$) function:

$$C(\xi; \mathbf{u}) = \sum_{i,j\in\mathcal{N}} u^{i'} R_{ij} u^j + 2 \sum_{i\in\mathcal{N}} u^{i'} D_i x + x'Qx. \tag{2.63}$$

The state of nature, ξ, is a Gaussian vector, as specified earlier by (2.49)–(2.51).

A static team problem with the structure above is known as an *exponential-Gaussian team* or a *linear-exponential-Gaussian team* (LEGT), the latter used especially if the measurements are given in the form (2.59). An exponential (of a quadratic) loss function captures phenomena not obtainable from a quadratic loss function and is preferred in situations where higher (than second) order moments of the statistical quantities should also be taken into consideration. A team using an exponential quadratic loss function in the construction of policies is called *risk averse* if $\theta > 0$ and *risk preferring* if $\theta < 0$. Here we will discuss only the case $\theta > 0$ because it is only in this case that L is convex in $\mathbf{u}$, which will enable us to apply some of the results of Sect. 2.6.2. The "optimistic" case $\theta < 0$ does not lead to a convex loss function, and hence it is not possible to obtain a general theory to cover this case as well. However, this should not be construed as the $\theta < 0$ case not being well defined or interesting. In fact, the stationarity conditions could hold in this case also, but one has to study each problem individually before concluding global (team) optimality.

Returning to the LEGT problem with $\theta > 0$, the first step toward studying its solution would be to obtain a characterization of the stationarity conditions (2.26). To avoid some unnecessary complexity in the analysis to follow, let us take the mean value of ξ to be zero, and furthermore let us restrict ourselves at the outset to linear decision rules (policies) for the agents—the latter will actually create no loss of generality as we shall see later.

Accordingly, let the decision rules be given as

$$\gamma^j(y^j) = A^j y^j, \quad j \in \mathcal{N}. \tag{2.64}$$

Let us fix all but one (say ith one) as above, and substitute them into (2.63) to obtain

$$C(\xi; u, \{\gamma^j\}_{j \neq i}) = u' R_{ii} u + 2u' T_i' \xi_i + \xi_i' S_i \xi_i =: C_i(\xi_i, u),$$

where u stands for u^i and T_i, S_i are defined as follows:

$$\xi_i := (x', y^{1'} \ldots y^{i-1'} y^{i+1'} \ldots y^{N'})',$$

$$T_i' \xi_i = D_i x + \sum_{j \in \mathcal{N}, j \neq i} R_{ij} A^j y^j, \tag{2.65}$$

$$\xi_i' S_i \xi_i = x' Q x + \sum_{j,k \in \mathcal{N}; j,k \neq i} y^{j'} A^{j'} R_{jk} A^k y^k + 2 \sum_{j \in \mathcal{N}, j \neq i} y^{j'} A^{j'} D_j x. \tag{2.66}$$

The important point here is that T_i and S_i are constant matrices (not dependent on ξ), but they depend on the policy gain matrices A^j, for all $j \in \mathcal{N}$, except $j = i$.

We now evaluate

$$E_{\xi | y^i} L(\xi; u, \{\gamma^j\}_{j \neq i}) =: J_i(u; y),$$

where, for convenience, we have dropped the superscript from y^i. Using Lemma 2.6.3, the distribution of ξ conditioned on $y^i = y$ is Gaussian, with mean and covariance given by

$$E[\xi|y] = (\Sigma'_{0i}\Sigma'_{1i}\dots\Sigma'_{Ni})'\Sigma_{ii}^{-1}y =: \hat{\xi}_i, \tag{2.67}$$

$$\text{cov}(\xi_i|y) = \Sigma^{(i)} - (\Sigma'_{0i}\dots\Sigma'_{Ni})'\Sigma_{ii}^{-1}(\Sigma'_{0i}\dots\Sigma'_{Ni}) =: \hat{\Sigma}_i, \tag{2.68}$$

which we assume to be positive definite. In (2.67) the matrix Σ_{ii} does not appear in $(\dots)$, and likewise in (2.68). Furthermore $\Sigma^{(i)}$ in (2.68) is the covariance of ξ_i, which is the Σ of (2.51) with the $(i+1)$th row and column block deleted.

Now, aside from a *pdf* normalization constant,

$$J_i(u;y) = \int \theta e^{\frac{\theta}{2}C_i(\xi_i;u)} e^{-\frac{1}{2}(\xi_i-\hat{\xi}_i)'\hat{\Sigma}_i^{-1}(\xi_i-\hat{\xi}_i)} d\xi_i,$$

where the integration is over the vector ξ_i belonging to an appropriate dimensional Euclidean space. This integral will have a finite value if, and only if, the quadratic term in ξ_i is negative definite, which brings in the condition

$$M_i := \hat{\Sigma}_i > 0. \tag{2.69}$$

Under this condition, the integral can be evaluated (using a property of Gaussian *pdf*'s) to yield (again aside from a positive multiplying constant)

$$J_i(u;y) = \theta e^{\tilde{C}_i(\hat{\xi}_i,u)},$$

where

$$\tilde{C}_i(\hat{\xi}_i,u) := \frac{\theta}{2}u'R_{ii}u + \frac{1}{2}(\theta T_i u + \hat{\Sigma}_i^{-1}\hat{\xi}_i)'M_i^{-1}(\theta T_i u + \hat{\Sigma}_i^{-1}\hat{\xi}_i) - \frac{1}{2}\hat{\xi}_i^{-1}\hat{\Sigma}_i^{-1}\hat{\xi}_i.$$

Note that for $\theta > 0$, $\tilde{C}_i$ is strictly convex in u, which implies that $J_i(u;y)$ is also strictly convex in u. Minimization of J_i is equivalent to minimization of $\tilde{C}_i$ which, being quadratic, immediately leads to

$$u = \gamma^i(y^i) = -(R_{ii} + \theta T_i' M_i^{-1} T_i)^{-1} T_i M_i^{-1} \hat{\Sigma}_i^{-1}\hat{\xi}_i =: A^i y^i. \tag{2.70}$$

Clearly this solution also satisfies the stationarity condition (2.26), with all other (than ith) agents' policies fixed as in (2.64).

Note that A^i determined as in (2.70) depends on the fixed gain matrices A^j's, for $j \neq i$, this dependence being through M_i and T_i. Let us denote this relationship by

$$A^i = f^i(A^i,\dots,A^{i-1},A^{i+1},\dots,A^N) \tag{2.71}$$

where f^i is a nonlinear but continuous function, determined uniquely by (2.70). Since agent $\mathbf{A}i$ was selected arbitrarily in the preceding analysis, a similar function will exist for each agent, so that (2.71) will hold for all $i \in \mathcal{N}$. This readily brings us to the following proposition.

Proposition 2.6.5. *The positively exponentiated Gaussian team problem admits a linear stationary solution if, and only if, there exists a set of matrices A^i, $i \in \mathcal{N}$, mutually satisfying (2.71), and under which (2.69) holds for all $i \in \mathcal{N}$, and $J(\gamma)$ remains finite.* ◇

Proof. In view of Definition 2.6.1, the result follows from the analysis that led to the proposition. □

Remark 2.6.3. If the random state of nature, ξ, has nonzero mean, as in (2.50), then the policies (2.64) will have to be replaced by the affine structure

$$\gamma^j(y^j) = A^j(y^j - \bar{y}^j) + b^j, \quad i \in \mathcal{N}.$$

Within this structure, one can again proceed through the preceding analysis and arrive at a counterpart of Proposition 2.6.5. ◇

Remark 2.6.4. The boundedness of the cost corresponding to the linear stationary solution can be checked by evaluating the quantity

$$E_{y^i} J_i(A^i y^i; y^i), \tag{$*$}$$

where J_i was defined in the analysis leading to the proposition. We first substitute (2.70) in $C_i(\hat{\xi}_i, u)$ to obtain

$$\tilde{C}_i(\hat{\xi}_i, \gamma^i(y^i)) = -\frac{1}{2}\hat{\xi}_i' N_i \hat{\xi}_i,$$

where

$$N_i := \hat{\Sigma}_i^{-1} + \hat{\Sigma}_i^{-1} M_i^{-1} T_i (\frac{1}{\theta} R_{ii} + T_i' M_i^{-1} T_i)^{-1} T_i' M_i^{-1} \hat{\Sigma}_i^{-1} - \hat{\Sigma}_i^{-1} M_i^{-1} \hat{\Sigma}_i^{-1}.$$

Then, we observe that $(*)$ is finite if, and only if, the integral

$$\int \theta e^{-\frac{1}{2}\hat{\xi}_i' N_i \hat{\xi}_i} e^{-\frac{1}{2} y' \Sigma_{ii}^{-1} y} dy$$

is finite, where $\hat{\xi}_i$ is related to y through (2.67). This condition is equivalent to the exponent being negative definite, that, is

$$\hat{\xi}_i' N_i \hat{\xi}_i + y' \Sigma_{ii}^{-1} y > 0 \qquad \forall y \in \mathbb{R}^{m_i}, y \neq 0. \tag{$**$}$$

◇

Proposition 2.6.5 above leaves a number of questions unanswered. First, we would like to know whether a linear stationary solution, whenever it exists, is team-optimal and secondly whether there would be other team-optimal solutions if the linear stationary solution (*lss*) ceases to exist. Clearly, we would not generally expect the *lss* to exist for all (especially arbitrarily large) values of θ, because of condition (2.69). The theorem below now provides an answer to the first question raised above; the second question is a most difficult one for which no general answer is known as yet.

Theorem 2.6.9. *Let* $\underline{\gamma}^* \in \mathbf{\Gamma}$ *be the linear stationary solution of Proposition 2.6.5, and let there exist some other linear policy* $\underline{\beta} \in \mathbf{\Gamma}$ *such that* $J(\underline{\beta}) < \infty$. *Then,* $\underline{\gamma}^*$ *is the unique team-optimal solution of the (positively) exponentiated-Gaussian team problem.* ◇

Proof. Here we resort to Theorem 2.6.5, which delineates the conditions under which stationarity implies team-optimality. Clearly $L(\xi; \mathbf{u})$ is strictly convex and continuously differentiable (in $\mathbf{u}$), and $J(\underline{\gamma})$ is bounded from below (by zero) for all $\underline{\gamma} \in \mathbf{\Gamma}$. The stationary solution $\underline{\gamma}^*$ has finite cost by hypothesis, and the subset (say, $\hat{\mathbf{\Gamma}}$) of $\mathbf{\Gamma}$ on which J is finite is not a singleton, again by hypothesis. Hence, to apply Theorem 2.6.5, one has to show that condition (*c.5*) holds for this problem. This is indeed the case and follows from the fact that the subset $\hat{\mathbf{\Gamma}}$ referred to above is not a singleton. The proof of this result is quite technical and will not be given here; it can be found in [218]. ☐

Remark 2.6.5. A sufficient condition for the second hypothesis of Theorem 2.6.9 is the following: Choose $\beta \equiv 0$, which is clearly a linear policy. Then,

$$J(\beta) = E\{\theta e^{\frac{\theta}{2}x'Qx}\},$$

which is finite if, and only if,

$$\Sigma_{00}^{-1} - \theta Q > 0.$$

Hence, if θ is chosen to be smaller than $1/[\lambda_{\max}(\Sigma_{00})\lambda_{\max}(Q)]$, the second hypothesis is satisfied. Of course, this condition (on θ) can be made less stringent by choosing some other (nonzero) β. ◇

Remark 2.6.6. For the negatively exponentiated Gaussian team problem, Proposition 2.6.5 remains equally valid (now in fact condition (2.69) would be satisfied with a bigger margin on θ), but we do not have the counterpart of Theorem 2.6.9 because of lack of convexity. ◇

Example 2.6.4. To illustrate the main result of Theorem 2.6.9 and to study the restrictions imposed on the parameters of the problem by the various conditions stated there, let us reconsider the static two-agent team problem of Example 2.6.3, with two differences: Now, the loss functional is a positive exponential of the one given there, i.e.,

$$L(x,u) = e^{\theta C(x,\mathbf{u})},$$
$$C(x,\mathbf{u}) = (u^1 + u^2 + x)^2 + (u^1)^2 + (u^2)^2,$$

and the random variable x has zero mean, i.e.,

$$x \sim N(0,2), \quad w^1 \sim N(0,2), \quad w^2 \sim N(0,1).$$

The measurements are still given by

$$y^1 = x + w^1, \quad y^2 = x + w^2.$$

Writing out the stationarity conditions (2.71), we obtain, after some algebra,

$$A^1 = -\frac{1}{2}[1+A^2-\theta(A^2)^2]/c_1(\theta,A^2), \quad A^2 = -[1+A^1-\theta(A^1)^2]/c_2(\theta,A^1) \quad (*)$$

where

$$c_1(\theta,A^2) := 2 - \theta[1 + 2A^2 + 6(A^2)^2] + \theta^2(A^2)^2,$$
$$c_2(\theta,A^1) := 3 - \theta[1 + 12(A^1)^2 + 2A^1] + 2\theta^2(A^1)^2.$$

The matrices M_1 and M_2, defined by (2.69), are given by

$$M_1 = \begin{pmatrix} 2-\theta & -1-\theta A^2 \\ -1-\theta A^2 & 1-2\theta(A^2)^2 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 2-\theta & -\frac{1}{2}-\theta A^1 \\ -\frac{1}{2}-\theta A^1 & \frac{1}{2}-2\theta(A^1)^2 \end{pmatrix},$$

so that condition (2.69) reads

$$0 < \theta < 2, \quad c_1(\theta,A^2) > 0, \quad c_2(\theta,A^1) > 0. \quad (**)$$

Trying out two different values of θ, namely, $\theta = 1$ and $\theta = \frac{1}{3}$, we find that for the former there is no solution to $(*)$ that also satisfies $(**)$; for the latter, however, there exists a unique solution $(*)$ that also meets $(**)$, which is

$$A^{1^*} = -0.236375, \quad A^{2*} = -0.345398 \quad (\theta = \frac{1}{3}).$$

This solution and the associated value of θ also satisfy the conditions of Remarks 2.6.4 and 2.6.5, and hence by Theorem 2.6.9 there exists a unique team-optimal solution to the LEGT problem:

$$\gamma^{1*}(y^1) = -0.236375y^1, \quad \gamma^{2*}(y^2) = -0.345398y^2.$$

It is important to note that, as $\theta \to 0$ in $(*)$, the nonlinear equations reduce to linear equations:

$$A^1 = -\frac{1}{4}(1 + A^2), \quad A^2 = -\frac{1}{3}(1 + A^1),$$

whose unique solution $(-\frac{2}{11}, -\frac{3}{11})$ is precisely the gain coefficients in the team-optimal solution of Example 2.6.3. Hence in the limit as $\theta \to 0$ we recover the solution of the corresponding LQG team (with loss function $C(x, \mathbf{u})$). This is to be expected because for any positive function C,

$$\lim_{\theta \to 0} \frac{1}{\theta}(e^{\theta C} - 1) = C.$$

Remark 2.6.7. A method for solving the set of coupled nonlinear equations $(*)$ of Example 2.6.4 for a particular value of θ, or more generally equations (2.71), is provided by the parallel update scheme

$$A^i_{(k+1)} = f^i(A^1_{(k)}, \ldots, A^{i-1}_{(k)}, A^{i+1}_{(k)}, \ldots, A^N_{(k)}), \quad i \in \mathcal{N}, \quad k = 0, 1, \ldots,$$

where the starting point is arbitrary. This set of recursive equations admits exactly the same interpretation as in the LQ team case, and as was the case there, this recursion may not converge (even if the LEG team problem admits a solution) for $N \geq 3$. For $N = 2$, however, the recursion will converge whenever the LEQ team problem admits a solution; this is because at each step this corresponds to an agent's minimization of J by fixing the other agent's policy at its most recently updated value. Since one is basically minimizing a strictly convex functional (in the LEGT problem), the unique minimum, whenever it exists, should be reachable by such a unilaterally cost-minimizing update scheme. ◇

2.7 Concluding Remarks

This chapter has provided a general introduction to stochastic team decision problems and associated solution concepts. Static and dynamic teams have been identified, and in the context of static teams conditions for existence of team-optimal solutions and for person-by-person optimality to imply team-optimality have been obtained. The chapter has also discussed iterative methods for obtaining team-optimal solutions and illustrated the theory presented with numerical examples.

2.8 Bibliographic Notes

Team decision theory has its roots in both control theory and economics. Jacob Marschak [254] was perhaps the first to introduce the basic elements of teams and to provide the first steps toward the development of a *team theory*. Roy Radner [316] followed with a precise mathematical formulation and provided conclusive results to some classes of static teams, establishing precise connections between person-by-person optimality, stationarity, and team-optimality. Marschak's and Radner's collaborative work culminated in the publication of their influential 1972 book [255]. At the time when such developments were being made, significant progress in the theory of statistical decision theory was also taking place: Bahadur's characterization of information fields and sufficient statistics [35, 36]; Blackwell's sufficient statistics and comparison of experiments results [61]; and Wald's [383], Savage's [333], and Chernoff's [94] contributions to statistical decision theory, among other major developments in probability theory, contributed to the rapid development of team decision theory.

Contributions of Hans Witsenhausen [393, 394, 399–401] on dynamic teams and characterization of information structures have been crucial in the progress of our understanding of dynamic teams; see Sect. 3.7, where Witsenhausen's *intrinsic model* as well as other models for dynamic teams are discussed in detail. This section also includes a brief discussion for nonsequential teams where important contributions in the literature have been due to Andersland and Teneketzis [9, 10] and Teneketzis [360], in addition to Witsenhausen [393].

Considerations of risk sensitivity motivated researchers to look into team problems with exponentiated loss function, with substantial results in this domain obtained (for teams) by Krainak et al. [219]. De Waal and van Schuppen [114] considered extensions to discrete action spaces. Bagchi and Başar [34] studied teams in continuous time as well as non-Gaussian settings.

Başar [24] studied team problems and more general nonzero-sum stochastic games when agents do not agree on a common *a priori* probability measure on the primitive random variable and work under their own subjective views of the environment, with team models in this context necessarily leading to game formulations. A more detailed discussion in this context of inconsistent probability models among a group of decision makers is presented in Chap. 12.

Further discussion on design of information structures in the context of team theory and economics applications is available in [15, 372], among a rich collection of other contributions.

In the next chapter, Chap. 3, we will see extensions of the static team theory of this chapter to dynamic teams, where information structures are of paramount importance. We will also consider Witsenhausen's intrinsic model more explicitly in Sect. 3.7. We refer the reader to also Teneketzis [360], in addition to [400], in this context.

Part of the chapter uses results from [219,316], however, with somewhat different proofs for some of the key results. The update schemes considered in Sect. 2.6 are based on [24].

Chapter 3
Characterization and Comparison of Information Structures

3.1 Introduction

In Chap. 2, we introduced a general framework for stochastic decision problems with static and dynamic information patterns, which include stochastic teams, stochastic control problems, or stochastic optimization problems with distributed decision makers (DMs) and decentralized information. We have discussed general existence results for finite stochastic teams as well as infinite but static teams, with some specific results for those with quadratic or exponentiated quadratic loss functions and when the underlying statistics of the primitive random variables are Gaussian. In this chapter we concentrate on information structures (*IS*s) in stochastic teams, and discuss them from various angles: comparison of two (or more) information structures from the point of view of *informativeness*, classification of different *IS*s and the notion of *signaling*, *IS*s that allow for sequential decomposition of a multi-stage problem into static ones, and the difficulties involved in solving dynamic stochastic teams with nonclassical information.

The first section, Sect. 3.2, discusses comparison of information structures, and makes precise the notion of one information structure being "better" than another. It also introduces some examples on the design of "optimal" information structures for static stochastic teams under some "hard" or "soft" costs attached to the acquisition of information.

This is then followed by Sect. 3.3, which introduces nonclassical information structures, and highlights the challenges involved in solving stochastic teams which feature such patterns. It discusses the notion of *signaling* and the celebrated *Witsenhausen's counterexample* and several variations on it. Section 3.4 discusses stochastic dynamic team problems with nested or partially nested information structures (classical and quasi-classical), mostly from a structural angle. Section 3.5 focuses on probability and cost-dependent aspects of information structures. The section identifies dynamic teams with nonclassical information which displays irrelevant signaling and also presents classes of problems with nonclassical information which can be solved by conversion into an equivalent one

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_3,

(as far as optimality goes) with classical or quasi-classical information; this one displays versatility of embedding a given dynamic team into a broader class of teams with expanded (such as quasi-classical) information. Section 3.6 discusses information transmission through signaling for a class of information patterns. Finally, Sect. 3.7 summarizes a precise characterization of information structures in view of Witsenhausen's intrinsic model and presents further dynamic team decision models due to Witsenhausen. The chapter concludes with Sect. 3.8 which provides some historical notes and guidelines for further reading on the topics covered herein.

3.2 Comparison of Information Structures

The issues addressed in the previous chapter all relate to the questions of existence, uniqueness, and characterization (derivation) of *team-optimal* solutions in static team problems for a given *fixed* set of measurements for the agents (or, equivalently, a fixed information structure (*IS*) for the team). In this section, we allow the flexibility of having more than one information structure for the team, in which context a natural question that arises is "when is one information structure better than another?" To shed some light on this question, let us first introduce the notation $R(\underline{\gamma};\underline{\eta})$ (as in Sect. 2.2) as the cost function of the team (to replace "J" which we had used all along), to recognize the dependence of the team performance (or cost) also on the information structure $\underline{\eta}$.[1] This function R may also include some additional terms quantifying costs due to the acquisition of information, i.e., cost directly associated with various $\underline{\eta}$'s, as in

$$R(\underline{\gamma};\underline{\eta}) := \tilde{J}(\underline{\gamma};\underline{\eta}) + c(\underline{\eta}) \equiv E_{\xi} L(\xi, \underline{\gamma}(\underline{\eta}[\xi])) + c(\underline{\eta}), \tag{3.1}$$

where $\tilde{J}$ is the team (expected) cost under $\underline{\gamma}$ and $\underline{\eta}$, without any cost on information, and $c(\underline{\eta}) \geq 0$ is the cost due to the choice of the particular information structure $\underline{\eta}$. We should note that as $\underline{\eta}$ changes, the policy space of the team, $\mathbf{\Gamma}$, may also change, and hence to make this dependence explicit, we will write the policy space as $\mathbf{\Gamma}(\underline{\eta})$.

Clearly, being a scalar valued function, (3.1) provides a strict ordering of all permissible information structures for a given team problem (with a fixed loss function L and a fixed cost of information c) according to the best that can be achieved. This leads us to the following first comparison (strict ordering) of *IS*s for a given team problem.

Definition 3.2.1. For a stochastic team problem with a fixed cost structure as given by (3.1) but a variable information structure (*IS*), an *IS* $\underline{\eta}$ is *better* (or *more valuable*) than another *IS* $\underline{\eta}'$ if

---

[1] Throughout this section, we adopt the framework of Sect. 2.2, unless otherwise stated; in particular, all random variables take values either in finite sets or finite-dimensional Euclidean spaces, and all information structures are static, as in Sect. 2.2.

$$\inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta})} R(\underline{\gamma};\underline{\eta}) < \inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}')} R(\underline{\gamma};\underline{\eta}').$$

Two *IS*s $\underline{\eta}$ and $\underline{\eta}'$ are of *equal value* if the inequality above is an equality. ◇

The strict ordering provided by Definition 3.2.1 is unfortunately dependent on the underlying cost structure (i.e., L and c) and it would be desirable to obtain an ordering that is valid for all reasonable loss functions (e.g., for all strictly convex loss functions, all continuous and bounded costs functions, or all measurable and bounded cost functions). The following definition which provides only a partial ordering serves this purpose[2]:

Definition 3.2.2. Let $\mathcal{L}$ be a given class of loss functions for a given stochastic team problem with fixed information cost and variable *IS*. Then, an *IS* $\underline{\eta}$ is *uniformly better* (*uniformly more valuable, more informative*) than another *IS* $\underline{\eta}'$ (with respect to the class $\mathcal{L}$) if, under the cost structure (3.1),

$$\inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta})} R(\underline{\gamma};\underline{\eta}) \leq \inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}')} R(\underline{\gamma};\underline{\eta}')$$

for all $L \in \mathcal{L}$, with strict inequality for at least one $L \in \mathcal{L}$. A somewhat weaker notion which allows the possibility for the inequality above to be an equality for all $L \in \mathcal{L}$ is "$\underline{\eta}$ is *uniformly at least as valuable as (least as informative as)* $\underline{\eta}'$" (with respect to the class $\mathcal{L}$). ◇

Now, to obtain a sufficient condition for an *IS* to be *uniformly at least as valuable as* another *IS* , let us recall [from Sect. 2.2, Eq. (2.1)] that given an *IS* $\underline{\eta} = (\eta^1, \ldots, \eta^N)$, there is an associated set of measurements $\mathbf{y} = (y^1, \ldots, y^N)$, which can also be viewed as random variables (or vectors) defined directly on the original probability space $(\Omega, \mathbf{F}, P)$. Let $\sigma(y^i)$ be the smallest sigma-field (on Ω) with respect to which y^i is measurable. Clearly, $\sigma(y^i) \subseteq \mathbf{F}$. To indicate the dependence of this sigma-field on η^i, we will use a subscript on σ: $\sigma_\eta(y^i)$. We are now in a position to state our first result.

Proposition 3.2.1. *Suppose we are given a static stochastic N-agent team problem with two possible ISs $\underline{\eta}$ and $\underline{\eta}'$. Let $\mathcal{L}$ be a class of loss functions with the property that for $L \in \mathcal{L}$, the random variable $L(\xi, \underline{\gamma}(\underline{\eta}[\xi]))$ has a well-defined (possibly infinite) expected value for every $\underline{\gamma} \in \mathbf{\Gamma}(\underline{\eta})$ and every $\underline{\gamma} \in \mathbf{\Gamma}(\underline{\eta}')$. Then, if $c(\eta) = c(\eta')$ and*

$$\sigma_\eta(y^i) \supseteq \sigma_{\eta'}(y^{i'}) \quad \textit{for all } i \in N, \textit{then}$$

IS $\underline{\eta}$ is uniformly at least as valuable as $\underline{\eta}'$. ◇

---

[2] If the objective is not minimization of a cost function, but stabilization of a dynamical system, then obtaining a total order is possible for certain criteria, as information-theoretic notions may be applicable. More on this will be presented later in the book.

Proof. Suppose that the result is not true. Then, there exists a loss function L such that

$$\inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta})} \tilde{J}(\underline{\gamma};\underline{\eta}) > \inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}')} \tilde{J}(\underline{\gamma};\underline{\eta}'). \qquad (*)$$

But under the given sigma-field inclusion, $\Gamma^i(\eta^i) \supseteq \Gamma^i(\eta^{i'})$, $i \in \mathcal{N}$, which implies that $\mathbf{\Gamma}(\underline{\eta}) \supseteq \mathbf{\Gamma}(\underline{\eta}')$. The infimum over a larger set cannot lead to a higher value (note that the function $\tilde{J}$ does not depend on $\underline{\eta}$ or $\underline{\eta}'$ *explicitly*—but *only* through $\underline{\gamma}$), and hence the inequality $(*)$ leads to a contradiction. □

The sigma-field inclusion condition of Proposition 3.2.1 is also known as the *fineness condition* (i.e., the sigma-field under one *IS* being finer (strictly speaking, *not coarser*) than under another *IS*), and clearly it is necessary that $c(\eta) = c(\eta')$ for the result to hold in this generality. It is quite possible to produce simple examples which show that the result could fail if some nonconstant cost is attached to the *IS*, with the following being one such illustration:

Example 3.2.1. Consider Example 2.6.3 of the previous chapter, this time with a cost on the acquisition of information. The loss function is

$$L(x,\mathbf{u}) = (u^1 + u^2 + x)^2 + (u^1)^2 + (u^2)^2$$

and the individual measurements are

$$y^1 = x + w^1, \quad y^2 = x + w^2,$$

where $x \sim N(1,2)$, $w^1 \sim N(0,2)$, $w^2 \sim N(0,1)$ are independent. Consider three possible *IS*s with the following associated costs:

(i) No measurement: $\sigma(y^1) = \sigma(y^2) = \{\phi, \Omega\}, \quad c(\underline{\eta}_{(i)}) = 0$

(ii) Individual measurements: $y^1 = y^1, y^2 = y^2, \quad c(\underline{\eta}_{(ii)}) = 0.25$

(iii) Sharing of measurements: $y^1 = y^2 = (y^1, y^2)', \quad c(\underline{\eta}_{(iii)}) = 0.5$

Let $R^*_{(\cdot)}$ denote the total team cost when the *IS* corresponding to case $(\cdot)$ is used, i.e.,

$$R^*_{(\cdot)} := \inf_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}_{(\cdot)})} R(\underline{\gamma};\underline{\eta}_{(\cdot)}) \equiv \min_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}_{(\cdot)})} R(\underline{\gamma};\underline{\eta}_{(\cdot)}).$$

Then, it readily follows from the results of Example 2.6.3 that

$$R^*_{(i)} \cong 2.3333, \quad R^*_{(ii)} \cong 1.6742, \quad R^*_{(iii)} \cong 1.83333,$$

and hence under the given cost structure the one that has the best trade-off between performance and cost of information is the *second IS*, even though the third one is uniformly at least as valuable as (in fact more valuable than) the second one. Clearly, by selecting the information cost function $c(\cdot)$ appropriately, even the “no measurement case” $(\eta_{(i)})$ can be made overall more desirable. ◇

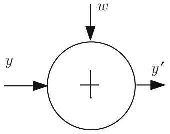

Fig. 3.1 Garbled information y', generated from y by additive independent noise

An easy way of checking the sigma-field inclusion condition of Proposition 3.2.1 is to find N-functions $h^i : Y^i \to Y^{i'}$, so that

$$y^{i'} = h^i(y^i) = h^i(\eta^i[\xi]) \equiv \eta^{i'}(\xi), \quad i \in \mathcal{N}.$$

As an illustration of this consider the *IS* of the Linear-Quadratic-Gaussian (LQG) team introduced by (2.59):

$$\underline{\eta} : y^i = H^i x + w^i, \quad i \in \mathcal{N},$$

where y^i is a (Gaussian) random vector of dimension m_i. Let Λ^i be a matrix of dimension $m_i' \times m_i$ (with $m_i' \leq m_i$), and define

$$\underline{\eta}' : y^{i'} = \Lambda^i y^i, \quad i \in \mathcal{N}.$$

Then, for the LQG team with no communication costs, the *IS* $\underline{\eta}$ is uniformly at least as valuable as $\underline{\eta}'$, which means that for any team with an arbitrary (but fixed) loss function, the optimum performance obtained under $\underline{\eta}$ can be no worse than the optimum performance under $\underline{\eta}'$.

The result of Proposition 3.2.1, although useful, is typically too restrictive and is not always directly applicable to situations where our intuition tells us that one particular *IS* should be *uniformly better* than another. As a case in point, consider the standard formulation of a stochastic team with measurement vectors $y^1, \ldots, y^N$ [defined on the common probability space $(\Omega, \mathbf{F}, \mathbf{P})$], generating an information structure $\underline{\eta}$. Suppose that the measurement of one of the agents, say $\mathbf{A}\mathit{1}$, is corrupted by additive noise (independent of $y^1, \ldots, y^N$, as well as of the "payoff-relevant" part of the random state of nature, x), thus leading to the new information structure (see Fig. 3.1)

$$\eta' : \quad y^{1'} = y^1 + w^{1'}, \quad y^{i'} = y^i, \quad i = 2, \ldots, N.$$

Here we say that $y^{1'}$ is a *garbled* version of the measurement (or information) y^1 of $\mathbf{A}\mathit{1}$, and as a consequence $\underline{\eta}'$ is a garbled version of the *IS* $\underline{\eta}$. Now, since $w^{1'}$ is independent noise, we would expect $\underline{\eta}'$ to be less valuable (less informative) to the team (and, particularly, $y^{1'}$ to be less informative to $\mathbf{A}\mathit{1}$) than $\underline{\eta}$, and hence the team should not be able to achieve better performance under $\underline{\eta}'$ than under $\underline{\eta}$. To prove this, however, Proposition 3.2.1 cannot be directly used, because we do not necessarily have an inclusion relationship between $\sigma(y^1)$ and $\sigma(y^{1'})$. Even if not directly, the proposition can still be used (indirectly), however, to put the above intuition on the same solid ground, as shown in the sequel. We first make precise what it means for a random vector to be a garbled version of another random vector.

Definition 3.2.3. Let x, y, y' be three real-valued random vectors defined on a common probability space $(\Omega, \mathbf{F}, P)$. Let $P_{y'|y,x}$ be the conditional probability distribution of y' given (y, x), and the same convention applies to $P_{y,y'|x}, P_{x|y,y'}$, etc. Then, y' is a *garbled version* of y *with respect to* x, if any one of the following three equivalent conditions hold:

(1) $P_{y'|y,x} = P_{y'|y}$
(2) $P_{x|y,y'} = P_{x|y}$
(3) $P_{y,y'|x} = P_{y|x}P_{y'|y}$

◇

Remark 3.2.1. The definition given above is more general than what Fig. 3.1 depicts, because the noise w' does not necessarily have to be additive. The equivalence of conditions (1)–(3) follows from a simple application of the Bayes theorem; for example, to show that $(1) \to (2)$, we note the sequence of equalities

$$P_{x|y,y'} = \frac{P_{x,y,y'}}{P_{y,y'}} = \frac{P_{y'|y,x} \cdot P_{y,x}}{P_{y'|y} \cdot P_y} = \frac{P_{y,x}}{P_y} = P_{x|y}$$

where at the next-to-the-last step we have used (1). Verification of (2) $\to$ (1), and the other implications are left to the reader. Note that all these conditions basically say that the whole information on x in y' is the information received through y. ◇

It is useful first to consider the one-agent stochastic optimization problem where the roles of the three random vectors x, y, y' used in Definition 3.2.3 are more transparent.

Proposition 3.2.2. *Consider a one-agent stochastic optimization problem with loss function $L(x, u)$, two separate ISs η and η' corresponding to static measurements y and y', respectively, and no cost of information. Let $\xi := (x, y, y')$ be a random vector defined on the probability space $(\Omega, \mathbf{F}, P)$, and let $\Gamma(\eta)$ and $\Gamma(\eta')$ be two policy spaces for the agent, corresponding to the ISs η and η', which are consistent with any constraint that might have been imposed on the action variable $u \in U$. Further assume that for every $\gamma \in \Gamma(\eta)$ and $\gamma' \in \Gamma(\eta')$, the loss functions $L(x, \gamma(y))$ and $L(x, \gamma'(y'))$ are integrable (i.e., have well-defined expected values). Then, if y' is a garbled version of y with respect to x, η is uniformly at least as valuable as η'.* ◇

Proof. Let $y'' := (y, y')$, with corresponding *IS* denoted by η''. Let $\Gamma(\eta'')$ be the class of all permissible policies for the agent under the *IS* η''. Clearly, $\Gamma(\eta) \subseteq \Gamma(\eta'')$ and $\Gamma(\eta') \subseteq \Gamma(\eta'')$, and furthermore $\sigma(y'')$ includes both $\sigma(y)$ and $\sigma(y')$. Hence, by Proposition 3.2.1, η'' is uniformly at least as valuable as both η' and η. Particularly

$$\inf_{\gamma'' \in \Gamma(\eta'')} E_\xi L(x, \gamma''(y, y')) \le \inf_{\gamma' \in \Gamma(\eta')} E_\xi L(x, \gamma'(y')) \,.$$

Now, given any $\gamma'' \in \Gamma(\eta'')$, note the following set of equalities:

$$\begin{aligned} E_\xi L(x, \gamma''(y, y')) &= E_{y,y'} E_{x|y,y'} L(x, \gamma''(y, y')) = E_{y,y'} E_{x|y} L(x, \gamma''(y, y')) \\ &=: E_{y,y'} \tilde{L}(y, \gamma''(y, y')) \,, \end{aligned}$$

where the *first* equality follows from the Bayes rule: $E_{x,y,y'} \equiv E_{y,y'} \cdot E_{x|y,y'}$, the *second* equality follows from condition (2) of Definition 3.2.3 under the "garbling" hypothesis of the proposition, and the *third* equality defines $\tilde{L}$ [as a measurable function of (y, y')] through its inner conditional expectation (of x given y). Likewise, for any $\gamma \in \Gamma(\eta)$, $E_{x|y} L(x, \gamma(y)) = \tilde{L}(y, \gamma(y))$. Now, concentrating on $\tilde{L}(y, \gamma(y))$ and $\tilde{L}(y, \gamma''(y, y'))$, it follows from Blackwell's *Irrelevant Information Theorem* (see Theorem D.1.1) that for any given $\gamma'' \in \Gamma(\eta'')$, there exists a $\gamma \in \Gamma(\eta)$ such that

$$E_y \tilde{L}(y, \gamma(y)) \leq E_{y,y'} \tilde{L}(y, \gamma''(y, y')) \,,$$

and hence that

$$\inf_{\gamma \in \Gamma(\eta)} E_{x,y} L(x, \gamma(y)) \leq \inf_{\gamma \in \Gamma(\eta'')} E_\xi L(x, \gamma(y, y')) \,,$$

and since $\Gamma(\eta) \subseteq \Gamma(\eta'')$, we have to have an equality here. Therefore,

$$\inf_{\gamma \in \Gamma(\eta)} E_{x,y} L(x, \gamma(y)) = \inf_{\gamma'' \in \Gamma(\eta'')} E_\xi L(x, \gamma''(y, y')) \leq \inf_{\gamma' \in \Gamma(\eta')} E_\xi L(x, \gamma'(y')) \,,$$

which establishes the desired result. □

Note the role x plays in the result of Proposition 3.2.2: it is the *payoff-relevant* portion of the random state of nature. If we have a genuine team problem, with $N (\geq 2)$ agents, then this result does not immediately extend because the payoff-relevant portion of the random state of nature will be a different random vector to different agents; specifically, with **A***i* isolated and under a fixed *IS* $\underline{\eta}$, the vector $\xi^i := (x, y^1, \ldots, y^{i-1}, y^{i+1}, \ldots, y^N)$ will be the payoff-relevant portion of ξ from the point of view of **A***i*. On the positive side, this says that in an N-agent static team problem, if *only one* of the agents' information changes, so that it is a garbled version of the old one, then the new (garbled) *IS* cannot lead to improved performance (which follows directly from Proposition 3.2.2). An iterative use of this feature leads to the following more general result where the measurements of all or a subset of the agents are allowed to change.

Proposition 3.2.3. *Consider an N-agent stochastic team problem, under the hypothesis of Proposition 3.2.2, naturally extended to N agents, and with $\underline{\eta} := \{y^1, \ldots, y^N\}$ and $\underline{\eta}' := \{y^{1'}, \ldots, y^{N'}\}$, where $y^{i'}$ is a garbled version of y^i with* *respect to* $\xi_i' := (x, y^{1'}, y^{2'}, \ldots, y^{i+1'}, \ldots, y^{N'}), \quad i \in \mathcal{N}$.[3] *Then, with no cost of information for the team,* $\underline{\eta}$ *is uniformly at least as valuable as* $\underline{\eta}'$. ◇

Proof. For the sake of simplicity and without much loss of generality, we assume that all the minima in the following development exist. If a minimum does not exist, then it is possible to consider an infimizing sequence, and the main result again goes through.

Now, consider the following sequence of equalities and inequalities, where we let $\underline{\gamma}' \in \mathbf{\Gamma}(\underline{\eta}')$ denote a team-optimal solution under the *IS* $\underline{\eta}'$:

$$\min_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta}')} E_\xi L(x, \underline{\gamma}(\mathbf{y}')) = \min_{\beta\in \mathit{\Gamma}^N(\eta^{N'})} E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-1}, \beta(y^{N'}))$$

$$\geq \min_{\beta\in \mathit{\Gamma}^N(\eta^{N})} E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-1}, \beta(y^{N}))$$

$$\equiv E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-1}, \gamma^{N*}(y^{N}))$$

$$\geq \min_{\beta\in \mathit{\Gamma}^{N-1'}(\eta^{N-1'})} E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-2}, \beta(y^{N-1'})), \gamma^{N*}(y^{N}))$$

$$\geq \min_{\beta\in \mathit{\Gamma}^{N-1}(\eta^{N-1})} E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-2}, \beta(y^{N-1})), \gamma^{N*}(y^{N}))$$

$$\equiv E_\xi L(x, \{\gamma^{j'}(y^{j'})\}_{j=1}^{N-2}, \gamma^{N-1*}(y^{N-1}), \gamma^{N*}(y^{N}))$$

$$\geq \ldots\ldots \geq \min_{\beta\in \mathit{\Gamma}^1(\eta^1)} E_\xi L(x, \beta(y^1), \{\gamma^{i*}(y^i)\}_{i=2}^{N})$$

$$\geq \min_{\underline{\gamma}\in\mathbf{\Gamma}(\underline{\eta})} E_\xi L(x, \underline{\gamma}(\mathbf{y})).$$

In the above, the first equality is a property of sequential minimization; the next inequality follows from Proposition 3.2.2 by identifying x there with $\xi_N' := (x, y^{1'}, \ldots, y^{N-1'})$, y with y^N, and y' with $y^{N'}$ and by making use of the given hypothesis that $y^{N'}$ is a garbled version of y^N with respect to ξ_N'. The next identity defines $\gamma^{N*} \in \mathit{\Gamma}^N(\eta^N)$, and the following inequality says that there may be some other element of $\mathit{\Gamma}^{N-1}(\eta^{N-1'})$ (other than $\gamma^{N-1'}$) that provides a lower value for the given cost. The next inequality again follows from Proposition 3.2.2 since $y^{N-1'}$ is a garbled version of y^{N-1} with respect to ξ_{N-1}', and the identity again defines $\gamma^{N-1*} \in \mathit{\Gamma}^{N-1}(\eta^{N-1})$. An induction argument then shows that the minimum of $J(\underline{\gamma})$ over $\mathbf{\Gamma}(\underline{\eta})$ cannot be higher than the minimum over $\mathbf{\Gamma}(\underline{\eta}')$, thus completing the proof. □

---

[3] Here, we allow for the possibility that y^i and $y^{i'}$ are identical for some $i \in \mathcal{N}$.

Remark 3.2.2. As we will see in the next chapter, garbling is not a necessary condition for one information structure to be *uniformly at least as valuable* as another one. A somewhat more relaxed condition (which is necessary and sufficient for a large class of setups) is presented in Theorem 4.3.2. ◇

Remark 3.2.3. It should be clear from the proof given above that in the hypotheses of Proposition 3.2.3 the order of indices in $\mathcal{N}$ under which ξ_i' has been defined could be arbitrary. ◇

Remark 3.2.4. An important special case under which the "garbling" condition of Proposition 3.2.3 is satisfied occurs when $\underline{\eta}$ and $\underline{\eta}'$ are related by

$$y^{i'} = y^i + w^{i'}, \quad i \in \mathcal{N},$$

where $w^{i'}$, $i \in N$, is a sequence of random vectors that are independent of the $N+1$ tuple $(x, y^1, \ldots, y^N)$. Then, clearly for every $i \in \mathcal{N}$,

$$P_{y^{i'}|y^i,\xi_i'} = P^{w^{i'}}_{y^{i'}-y^i}$$

where $P^{w^{i'}}$ is the probability distribution function of $w^{i'}$, which shows that $y^{i'}$ is a garbled version of y^i with respect to ξ_i'. ◇

In view of Remark 3.2.4 above, we have the following useful result, which we state as a corollary to Proposition 3.2.3.

Corollary 3.2.1. *Consider the LQG team problem of Sect. 2.6.3, under the measurement scheme (2.59), where the measurement noises $w^i \sim N(0, N^i)$, $i \in \mathcal{N}$, are independent. Let the optimum team cost be given by $J^*(N^1, \ldots, N^N)$ as a function of the noise covariances. Then, J^* is a nondecreasing function of the N^i's, under the matrix partial ordering.*[4] ◇

So far in this section, we have shed some light on the question "under what conditions is one *IS* better than another?" The answer is definitely context-dependent, but it has also been possible to obtain some fairly general results, which lead to only a partial ordering in the class of all (permissible) *IS*s. Two other related questions here are:

(i) What is the value of a given extra information (measurement)?

(ii) What is the "best" *IS* for a given team problem?

The answer to the first question will definitely depend on the level of information already at hand, but even under this provision there is the further issue of which agent(s) should receive the extra available measurement(s) assuming that a sharing of information is not possible. Then, the value of the extra measurement to the team

---

[4] In other words, $N^{i'} - N^i \geq 0$ (nonnegative definite), $i \in N$, implies $J^*(N^{1'}, \ldots, N^{N'}) \geq J^*(N^1, \ldots, N^N)$. In words, the more "noisy" the measurements are, the higher is the team cost.

will depend very much on which agent receives (and utilizes) it. This is clearly a well-posed finite optimization problem which in general requires an exhaustive search (to determine the agent who will "make the most out of the measurement"). The following simple example will illustrate this point.

Example 3.2.2. Consider the two-agent LQG team problem with loss function

$$L(x, u^1, u^2) = (u^1 + u^2 + x)^2 + (u^1)^2 + 2(u^2)^2,$$

where all variables are scalar, and $x \sim N(0, 1)$. Note that the two agents enter the loss function quite symmetrically, with the exception of the "soft constraint" on the action variable of $\mathbf{A}$*2*, which shows that his action is costlier than that of $\mathbf{A}$*1*.

In the absence of any measurements, the unique team-optimal solution is $u^{1*} = u^{2*} = 0$, leading to the cost level of $J^*_\emptyset = E[x^2] = 1$.

Let us now assume that a single measurement

$$z = x + w\,, \quad w \sim N(0, 1)\,, \quad \text{independent of } x\,,$$

becomes available and only one agent is allowed to use it. If $\mathbf{A}$*1* uses it, the unique team-optimal solution is

$$\gamma^{1*}(z) = -\frac{1}{4}z\,, \quad u^{2*} = 0\,,$$

with a cost level of $J^*_{11} = 0.75$. If, on the other hand, $\mathbf{A}$*2* uses it, the unique team-optimal solution is

$$u^{1*} = 0\,, \quad \gamma^{2*}(z) = -\frac{1}{6}z\,,$$

with a cost level of $J^*_{12} = 0.833$. Clearly, it is to the team's advantage for $\mathbf{A}$*1* to receive it, and hence the value of this measurement to the team is $1 - 0.75$ = 0.25 units (with the comparison made with the *no measurement* case).

Next, suppose that an additional measurement is made available

$$y = x + v, \quad v \sim N(0, 2)\,,$$

where v is independent of both x and w. Again, the decision as to which agent should receive it will be made based on its value to the team. The following table summarizes the four possible scenarios (Table 3.1).

Again it is optimal for the team for both measurements to go to $\mathbf{A}$*1*, the agent with a lower cost on action "effort." The value of the second channel to the team is $0.75 - 0.7 = 0.05$ units, on the top of the optimum one-channel performance. ◇

In summary, for a general team problem,

$$\begin{aligned} \textit{Value} \text{ of extra information} = U(&\text{optimum team cost without the information} \\ &- \text{optimum team cost with the extra information}), \end{aligned}$$

*IS*s	Team-optimal policies $(\gamma^{1*}, \gamma^{2*})$	Optimal cost J^*
$\eta^1 = \{z\}, \eta^2 = \{y\}$	$\big(-\frac{8}{35}y, -\frac{3}{35}z\big)$	0.76735
$\eta^1 = \{y\}, \eta^2 = \{z\}$	$\big(-\frac{1}{7}y, -\frac{1}{7}z\big)$	0.71428
$\eta^1 = \{z, y\}, \eta^2 = \phi$	$\big(-\frac{1}{10}(y + 2z), 0\big)$	0.7
$\eta^1 = \phi, \eta^2 = \{z, y\}$	$\big(0, -\frac{1}{15}(y + 2z)\big)$	0.8

Table 3.1 Four possible scenarios for the two-channel two-agent team of Example 3.2.2

where U is some appropriate utility function. In determining whether a given measurement should be accepted or not, this *value* should be weighted against the cost associated with the acquisition (and utilization) of this extra information. Of course, in an "information cost-free world," it is never detrimental for the team to acquire the additional measurement, regardless of which agent utilizes it. We will make use of this observation throughout the book, in particular in Chap. 10.

The second question we raised above, viz., the construction of the best *IS* for a given team problem, can be addressed by formulating it as an optimization problem. If there are only a finite number of *IS*s from which to select, the procedure is basically one of exhaustive search, requiring the computation of the optimal team cost under each *IS*; clearly costs associated with different *IS*s can also be accommodated in this formulation. There are also cases when the class of *IS*s is not finite, in which case one has to be more precise in the formulation of the underlying optimization problem and the appropriate topology on the space of information structures. We will provide further discussion along this direction in the next chapter. Below, we provide one illustration in the context of LQG team problems.

Consider the LQG team framework of Sect. 2.6.3 (or equivalently of Corollary 3.2.1), but with the observation matrices H^i, $i \in N$, in (2.59) yet to be designed—a situation depicted in Fig. 3.2. We have N Gaussian independent vector channels, each of dimension m_i, $i \in N$, with the covariances of the zero-mean channel noises (w^i's) fixed, and the outputs of the channels (y^i's) available as measurements to individual agents. The *IS* is determined by the H^i, $i \in N$, where H^i is of dimension $n \times m_i$, and the *IS* design problem is one of coming up with the best selection of these matrices (under some constraints) so that the team cost will be minimized, provided that the γ^i's are chosen optimally. Note that if no constraints (bounds) are imposed on these matrices, then the optimal solution may not exist, because there would be a tendency to choose elements of the observation matrices as large as possible so as to combat the channel noise. Hence it is reasonable to have individual channel output energy constraints, such as

$$E[y^{i'} y^i] \leq c_i, \quad i \in \mathcal{N},$$

which translate into the following constraints on the observation matrices:

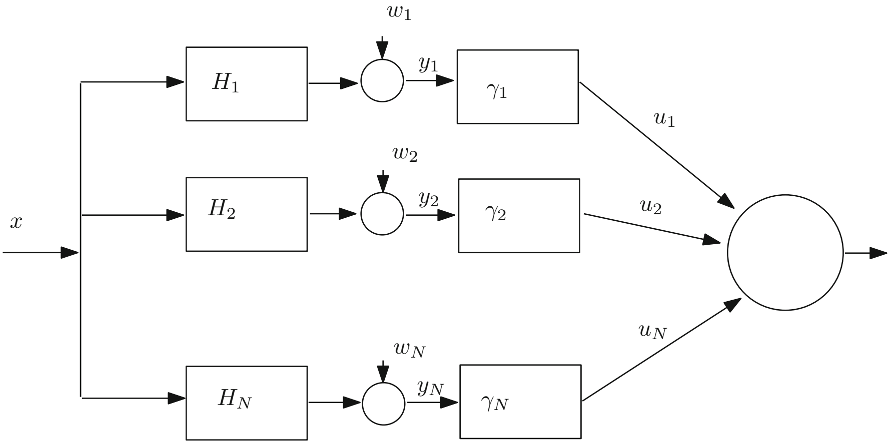

Fig. 3.2 An LQG team decision problem with a variable *IS*

$$E[y^{i'} y^i] = Tr[\Sigma_{00} H^{i'} H^i] + Tr[N^i] \le c_i, \quad i \in \mathcal{N},$$

where $\Sigma_{00} := \mathrm{cov}(x)$, $N^i := \mathrm{cov}(w^i)$. Denote the class of $(n \times m_i)$-dimensional matrices that satisfy the energy constraint above by $\mathcal{M}_i$, $i \in \mathcal{N}$. Now, adopting a constant parameter strictly convex quadratic loss function as in Sect. 2.6.3, we know from Corollary 2.6.2 (with $\bar{x} = 0$) that for each $H^i \in \mathcal{M}_i$, $i \in \mathcal{N}$, the LQG team admits a unique team-optimal solution

$$\gamma^{1*}(y^i) = \Pi^i y^i, \quad i \in \mathcal{N}$$

where Π^i, $i \in \mathcal{N}$, satisfy the set of equations (2.61) and each Π^i depends not only on H^i but also on H^j, $j \neq i$. The optimal team cost $J(\underline{\gamma}^*) =: \tilde{J}(H^1, \ldots, H^N)$ will depend on the H^i's continuously, and furthermore since the sets $\mathcal{M}_i$, $i \in N$, are closed and bounded subsets of finite-dimensional (Euclidean) spaces, the minimum of $\tilde{J}$ will exist on $\mathcal{M} := \mathcal{M}_1 \times \ldots \times \mathcal{M}_N$, by the *Weierstrass theorem* (Appendix A.5). We summarize this result in the following proposition.

Proposition 3.2.4. *Consider the LQG team depicted in Fig. 3.2, where L is a strictly convex quadratic loss function and the observation matrices H^i, $i \in \mathcal{N}$, satisfy the constraints $E[y^{i'} y^i] \le c_i$, with $c_i > Tr[N^i]$, $i \in \mathcal{N}$. The problem of jointly designing these observation matrices and the decision rule γ^i, $i \in \mathcal{N}$, so that $E_\xi L(x, \mathbf{u})$ is minimized, admits a solution, with the optimum γ^i's being linear in the respective channel outputs, y^i.* ◇

Remark 3.2.5. The optimum choice(s) out of $\mathcal{M}$ cannot be obtained in general in analytic form, but since $\tilde{J}$ is a differentiable function, where differentiability can be established by a variational argument, some of the existing numerical optimization algorithms [55, 243] can be used. Note that, even in a single channel setup, the problem is in general not convex [302]. Nonetheless, if the information is centralized, however, such as the case of a single-agent problem, the optimum H can be obtained analytically, as shown in [23]. In the terminology of *information theory*, this can also viewed as a *linear encoding-decoding* problem, for Gaussian channels (see, e.g., [410]), under a further *decentralized structure*. If the encoding part is taken to be a general mapping (not necessarily linear), then the optimal measurement structure will generally be nonlinear, but the form of this nonlinear encoding scheme is currently not known. We will say more on this in the next section and in Chap. 11. $\diamond$

3.3 Dynamic Teams with Nonclassical Information: Importance of Signaling

In a dynamic stochastic decision problem (or dynamic stochastic team, or stochastic control problem), if the action to be taken at some point in (discrete) time, say k, by an agent **A***i* relies on a measurement (information) which is affected by an action taken at some previous point in time, say k' ($k' < k$), by the same agent or by some other agent and **A***i* does not have access to the measurement (information) used in the construction of that previous action, then the underlying decision problem is said to be one with *nonclassical information*. Such problems are inherently difficult to solve, and there is no general theory which will aid in the construction of optimal policies that use nonclassical information. The goal of this section is to discuss the source of these difficulties, by studying what is perhaps the simplest such system, first introduced by Witsenhausen in a 1968 paper [398]. We discuss here not only that infamous counterexample but also some variants of it, including ones where the underlying distributions are discrete.

Consider the following two-stage stochastic control or equivalently two-agent dynamic stochastic team problem, where all quantities are scalar (see Fig. 3.3 for flow of information and relationships between different variables).

A random variable, x, with a given distribution is to be transformed into another random variable, $u_0 = \gamma_0(x)$, which is transmitted over a channel, $y = u_0 + w$, with additive noise w of known distribution, the output, y, of which is to be further transformed into another random variable, $u_1 = \gamma_1(y)$. The objective is to choose the transformations γ_0 and γ_1 in such a way that a given performance index,

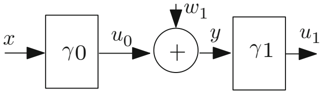

Fig. 3.3 Flow of information in Witsenhausen's counterexample

$Q(x, u_0, u_1)$, quadratic in x, u_0, and u_1, is minimized in the average sense. That is, we seek the pair $\gamma^* := (\gamma_0^*, \gamma_1^*)$, if exists, such that

$$J(\gamma^*) = \inf_{\gamma} J(\gamma) =: J^*, \tag{3.2}$$

where

$$J(\gamma) = E\left[Q(x, \gamma_0(x), \gamma_1(y))\right] \tag{3.3}$$

with expectation, $E[\cdot]$, taken over the statistics of x and w, which are assumed to be independent. Furthermore, the minimization is over the space of all Borel-measurable maps, that is, both policies (decision rules) γ_0 and γ_1 are taken to be Borel-measurable maps of the real line into itself.

This is a stochastic decision problem with *nonclassical information*, because the information to be used by the decision rule, γ_1, of the second agent depends on the action, u_0, of the first agent (and thereby on the decision rule of the first agent), but the second agent does not have access to the information of the first agent (i.e., x). If we view it as a single-agent problem where the agent acts twice, then it is one where the agent is *memoryless*, that is, she does not remember what she had observed at the earlier stage. As such, these problems belong to the realm of inherently difficult decision problems for which a systematic solution process generally does not exist, one of the main reasons being that due to loss in memory, a sequential decomposition or a dynamic programming approach is not possible [399, 400]. We now consider different instances of this class of problems, corresponding to different choices of the performance index Q, and different choices for the distributions for x and w, some of which admit explicit, relatively simpler solutions, while some others do not. Further, in some cases an optimal solution exists (even though not available in closed form), and in other cases only an ϵ-optimal solution exists. Hence, part of the message here is that even though the nonclassical nature of the information is generally responsible for the difficulty in obtaining the optimal solution, in some cases the structure of the loss function (performance index) also contributes to the difficulty in solving these problems.

3.3.1 Witsenhausen's Counterexample with Discrete Distributions

Let the quadratic performance index Q be picked as

$$Q_W(x, u_0, u_1) = k(u_0 - x)^2 + (u_1 - u_0)^2, \tag{3.4}$$

where $k > 0$ is a given parameter. Note that here the first agent wants to stay as close to x as possible, while the second agent wants to stay as close to the action of the first agent, u_0, as possible.

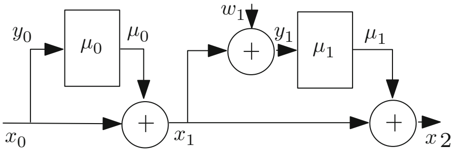

Fig. 3.4 Witsenhausen's counterexample in two-stage state-space form

This can also be viewed as a standard discrete-time two-stage stochastic optimal control problem, with state equations (see Fig. 3.4)

$$x_1 = x_0 + v_0\,, \quad x_2 = x_1 - v_1\,,$$

measurement equations

$$y_0 = x_0\,, \quad y_1 = x_1 + w_1\,,$$

and memoryless controls

$$v_0 = \mu_0(y_0)\,, \quad v_1 = \mu_1(y_1)\,,$$

where μ_0 and μ_1 are the instantaneous measurement output control policies at stages 0 and 1, respectively. This becomes equivalent to the earlier formulation in view of the correspondences

$$u_0 = x_0 + v_0\,, \quad u_1 = v_1\,, \quad x = x_0\,, \quad w = w_1\,, \quad y = y_1\,,$$

if we pick the cost function as

$$\tilde{Q}(x_2, v_0) = (x_2)^2 + k(v_0)^2 \equiv Q_W(x_1 - v_0, x_1, x_1 - x_2)\,.$$

Now let both x_0 and w_1 admit two-point probability mass functions, each taking values 1 and -1 with equal probability $\frac{1}{2}$. Consider, for each $\epsilon > 0$ sufficiently small, the construction

$$\gamma_0(x) = x_0 + \epsilon\,\text{sgn}(x) \qquad \gamma_1(y) = \begin{cases} 1+\epsilon, & \text{if } y = 2+\epsilon \text{ or } \epsilon, \\ \\ -1-\epsilon, & \text{if } y = -2-\epsilon \text{ or } -\epsilon, \end{cases}$$

and note that the corresponding value of J is $k\epsilon^2$. This shows that J, which is a nonnegative quantity, can be made sufficiently close to *zero* by picking ϵ sufficiently small.

Note that if u_1 also had access to the true value of x, then the optimum solution would be $\gamma_0(x) = \gamma_1(x) = x$, resulting in *zero cost*, thus achieving the lowest possible value for J. Now, since u_1 does not have access to the value of x, this being a team problem the true value of x has to be transmitted to u_1 through the action u_0, which is a realization of a policy γ_0; this in turn provides u_1 with the true value of u_0 (even though it is received over a noisy channel) and thus enables the second agent to make the second term of the cost zero. The channel noise being discrete and

u_0 being a continuous variable (taking values on the real line), transmission without any error is possible, albeit at some (arbitrarily small) cost to u_0. It should be clear here that the infimum of J is zero, but it cannot be achieved because in the limit as $\epsilon \to 0$, it is no longer possible to transmit x (and thus u_0, which inevitably depends on it to keep the first term in J small) perfectly.

The phenomenon we have observed above—the true value of a random variable being transmitted by an agent who observes it to another agent who observes only the action of the first agent—is called *signaling through control actions*. Stochastic team problems where the information structure is nonclassical generally entail *signaling through control actions* by at least one agent to other agents who act later, and signaling generally involves a trade-off or a compromise from an otherwise optimal control action, and this trade-off cannot generally be formalized in a precise mathematical form, or an explicit characterization cannot always be provided as in this example. This is what contributes to the challenges underlying derivation of team-optimal solutions in dynamic stochastic teams with nonclassical information. Furthermore, as we will see in Sect. 4.10, team problems with such an information structure are typically non-convex problems due to the signaling aspect.

Hence, the message conveyed by the analysis of the two-agent dynamic team problem of this subsection is that in problems with nonclassical information a minimum may not exist, and *signaling* may be essential to achieve a cost arbitrarily close to the infimum. Note that the conclusions would have been the same if instead of (3.4) the loss function was

$$Q'_W(x, u_0, u_1) = k(u_0 - x)^2 + (u_1 - x)^2 \, , \tag{3.5}$$

which has no cross terms between the two action variables (but IS is still nonclassical). Further, if the distributions had been m-point, instead of 2-point, where m is any positive integer, again the conclusion would essentially be the same.

In the next subsection we discuss the original counterexample by Witsenhausen, where the distributions are not discrete, in which case the minimum exists, but the policies that achieve it are not known.

3.3.2 Witsenhausen's Counterexample

Now assume that the distributions of the two random variables are Gaussian: x is a Gaussian random variable with mean zero and variance σ_x^2, and w is also a zero-mean Gaussian random variable, with variance σ_w^2, the two again being independent.

Witsenhausen has shown in a 1968 paper [398] that the optimal solution to this problem exists, but there are instances of the problem where the optimum decision rules (μ_0 and μ_1, or equivalently γ_0 and γ_1) are not linear. For the latter, he has shown that for some values of the parameters defining the problem, there exist nonlinear policies which outperform the best linear policies.[5] A class of such nonlinear policies introduced by Witsenhausen [398] and further improved upon in [40] is

$$u_0 = \gamma_o(x) = \epsilon\,\mathrm{sgn}(x) + \lambda x\,,$$
$$u_1 = \gamma_1(y) = E[\epsilon\,\mathrm{sgn}(x) + \lambda x|y],$$

where ϵ and λ are parameters to be optimized over (in [398] the values are picked as $\lambda = 0$ and $\epsilon = \sigma_x$, and some asymptotics are studied). Clearly, if $\epsilon = 0$, this class of decision rules will be linear, since $E[\lambda x|y]$ will be linear for each λ however when $\epsilon \neq 0$, the decision rules at both stages will be nonlinear. To give some indication of how much can be gained by taking $\epsilon \neq 0$, let us consider the case with parameter values $k = 0.1, \sigma_x^2 = 10, \sigma_w^2 = 1$; then the best linear policy at stage *zero* has the gain $\lambda_{\text{opt}} = -0.1127$, with the corresponding value of J being -0.100. If however ϵ is picked to be 2, the corresponding value of J (for the same choice of λ which is clearly not optimal and can be further improved upon) is -0.4797, which registers a substantial improvement over the best linear solution. For another scenario, let us take $k = 0.01, \sigma_x^2 = 80, \sigma_w^2 = 1$; in this case the best linear policy at stage *zero* has the gain $\lambda_{\text{opt}} = 0.01006$, with the corresponding value of J being -7.98×10^{-3}, whereas for the same value of λ, picking $\epsilon = 5$ leads to a value of $J = -0.4691$. Further numerical results can be found in [40], which also shows that if $\lambda = 0$, $\epsilon = \sqrt{2/\pi}$ and $k\sigma_x^2 = 1$, as $k \to 0$ the bound on asymptotic performance becomes $(1 - (2/\pi)) = 0.363$.

Note that the policy γ_0 above is of the type "linear plus 2-point quantized," which begs the question whether one can improve upon that by using a finer quantizer. For example, one could start with the class of policies

$$u_0 = \gamma_0(x) = \epsilon\,\mathrm{Quant}(x) + \lambda x\,,$$
$$u_1 = \gamma_1(y) = E[\epsilon\,\mathrm{Quant}(x) + \lambda x|y],$$

where *Quant* denotes the quantization operator (which will be defined precisely later in the book; see Definition 4.7.1), and optimize over the parameter values ϵ and λ, as well as over all quantizers.[6] Since the 2-point quantizer is a special case, one could naturally achieve an improvement in performance as a result of this optimization, and in fact one can show [274] that there are instances of the problem where the best performance in the linear class (i.e., with $\epsilon = 0$, and optimized over λ) could be arbitrarily bad against the best performance in this class; in precise mathematical terms, within the structure above,

---

[5] We will shortly provide a proof of existence of the team-optimal solution, not based on the original proof by Witsenhausen, but following a recent one given in [409] which is more direct. As of today, closed-form expressions for the optimal nonlinear policies are not available, and their characterization is not known.

[6] Note that γ_1 above is optimal against the γ_0 picked.

$$\sup_{k>0,\sigma>0} \frac{\inf_{\epsilon=0,\lambda} E[Q_W(x,\gamma_0(x),\gamma_1(y))]}{\inf_{\epsilon,\lambda,\text{Quant}} E[Q_W(x,\gamma_0(x),\gamma_1(y))]} = \infty.$$

In spite of this improvement, however, the optimal solution to the problem (which exists as mentioned earlier and to be proved below) is not of the quantized type.

Proof of Existence. We now proceed with the proof of existence of the team-optimal solution, following [409]. First note that given any measurable γ_0, the optimal γ_1 is uniquely given by the conditional mean

$$\gamma_1(y) = E[\gamma_0(x)|y]\,,$$

and hence the problem becomes one of minimization with respect to all measurable functions $\gamma_0 = f : \mathbb{R} \to \mathbb{R}$ of the function

$$J'(f) = E[k(f(x)-x)^2+[f(x)]^2-[E[f(x)|y]]^2] \equiv kE[(z-x)^2]+\text{mmse}(\Phi), \tag{3.6}$$

where Φ is the probability distribution of $z = f(x)$ and mmse(Φ) stands for the minimum mean square (MMS) error of estimating a random variable z with probability distribution Φ using measurement $y = z+w$, where w is as given before and is independent of z. Note that the second term depends only on Φ, whereas the first term depends on the joint distribution of x and z. Hence, in the process of minimization of (3.6) with respect to f, we could first hold Φ fixed, minimize the first term over the joint distribution of x and z, say $P_{x,z}$, with the marginals fixed at N and Φ, respectively, and then minimize the resulting expression with respect to the unknown marginal Q. The outcome of the first minimization (infimization) is

$$k[W_2(N,\Phi)]^2 + \text{mmse}(\Phi)\,, \tag{3.7}$$

where

$$W_2(N,\Phi) := \inf_{\{P_{x,z}:P_x=N,P_z=\Phi\}} \sqrt{E[(z-x)^2]}$$

is the quadratic *Wasserstein* distance between the two probability distributions N and Φ. Hence,

$$\inf_f J'(f) = \inf_\Phi [k[W_2(N,\Phi)]^2 + \text{mmse}(\Phi)] =: \inf_\Phi F(\Phi)\,.$$

For each fixed N, $W_2(N,\Phi)$ is weakly lower semicontinuous in Φ [7]. Furthermore, if Φ has a bounded second moment $m_2(\Phi)$, with the bound say $\alpha^2 > 0$, then $W_2(N,\Phi)$ is minimized by linearly correlating z with x, more precisely $z = (\alpha/\sigma_x)x$, leading to the value of

$$\min_{\{\Phi:m_2(\Phi)\le\alpha^2\}} W_2(N,\Phi) = |\alpha - \sigma_x|\,,$$

which is an increasing function of α for $\alpha > \sigma_x$. Likewise, mmse($\varPhi$) is an increasing function of α. Hence in the minimization of $F(\varPhi)$, $\varPhi$ can be restricted to distributions with a uniformly bounded second moment and thus to a weakly compact set.[7] Then, $\inf_{\varPhi} F(\varPhi)$ involves infimization of a weakly lower-semicontinuous function on a weakly compact set, and hence $F(\varPhi)$, and consequently $J'(f)$, has a minimum (by extended Weierstrass theorem [242]).[8]

3.3.3 Generalized Gaussian Test Channel

Now consider a different choice for Q:

$$Q_{\text{TC}}(x, u_0, u_1) = k(u_0)^2 + (u_1 - x)^2 \tag{3.8}$$

where again $k > 0$. Note that here the second agent's objective is to estimate the random variable x in the MMS sense, using a measurement that is transmitted over a Gaussian channel where the input to the channel is *shaped* by the first agent who has access to x and has a soft constraint ($kE[(u_0)^2]$) on its action. The version of this problem where the soft constraint is replaced by a hard power constraint, $E[(u_0)^2] \leq k$, is known as the *Gaussian test channel* (GTC), and in this context γ_0 is the *encoder* and γ_1 the *decoder*, where the latter's optimal choice is clearly the conditional mean of x given y, that is, $E[x|y]$. The best encoder for the GTC can be shown to be linear (a scaled version of the source output, x), which in turn leads to a linear optimal decoder. The approach here (as we will discuss further below for a more general, soft-constrained version), which is in fact the only approach known to apply here, is to obtain bounds on the performance using an inequality from information theory involving channel capacity [410] and rate distortion function [53] and then to show that the bound can be achieved using linear policies.[9]

Now, consider the more general version of (3.8):

$$Q_{\text{GTC}}(x, u_0, u_1) = k(u_0)^2 + (u_1 - x)^2 + b_0 u_0 x\,, \tag{3.9}$$

where b_0 is a parameter. Let

$$E[(u_0)^2] =: \alpha \quad \text{and} \quad E[(u_1 - x)^2] =: \beta\,.$$

---

[7] The next chapter, Chap. 4, provides a detailed coverage of some of the topological notions used here, such as the space of probability distributions and weak topology on such spaces.

[8] We further note that the term $W_2(N, \varPhi)$ is convex in $\varPhi$, but mmse($\varPhi$) is concave in $\varPhi$, which makes the function to be minimized, $F(\varPhi)$, in general non-convex. Some results and general discussion on concavity of optimization problems in information structures can be found in Chap. 4.

[9] A detailed coverage of information theoretic notions can be found in Chap. 5. See also Theorem 11.2.2.

Then, with J defined as before, by (3.3), and with γ_0 and γ_1 constrained as above, we have the inequalities

$$\begin{aligned}\inf_{\gamma} J(\gamma) &\geq k\alpha + \beta + \inf_{\gamma_0} b_0 E[\gamma_0(x)x] \\ &\geq k\alpha + \beta - |b_0|\sigma_x\sqrt{\alpha}, \end{aligned} \tag{3.10}$$

where the second one follows from the Cauchy–Schwarz inequality.

Now, by the data processing theorem [410] (see Lemma 5.3.1), in a linear configuration the mutual information[10] between two random variables closer to each other is no smaller than the mutual information between two random variables farther apart. In our case, this translates to

$$I(x;y) \geq I(x;u_1), \tag{3.11}$$

where $I(\cdot;\cdot)$ stands for mutual information. For each fixed $\alpha > 0$, $I(x;y)$ is bounded from above by the capacity of the channel, $C(\alpha)$, which is known for the Gaussian channel to be [151]

$$C(\alpha) = \frac{1}{2}\log(1 + (\alpha/\sigma_w^2)\,)\,.$$

Further, for each fixed β, the quantity $I(x;u_1)$ is bounded from below by the rate distortion function, $R(\beta)$, for which the expression, when $\beta \leq \sigma_x^2$, is [53]

$$R(\beta) = \frac{1}{2}\log(\sigma_x^2/\beta)\,.$$

In view of (3.11), we have

$$\frac{1}{2}\log(1 + (\alpha/\sigma_w^2)\,) = C(\alpha) \geq R(\beta) = \frac{1}{2}\log(\sigma_x^2/\beta)$$

leading to the following bound on β: $\beta \geq \sigma_w^2\sigma_x^2/(\alpha + \sigma_w^2)$, which is tight with

$$\gamma_0(x) = -\text{sgn}(b_0)\frac{\sqrt{\alpha}}{\sigma_x}x, \tag{3.12}$$

Substitution of this in (3.10) leads to

$$\inf_{\gamma} J(\gamma) \geq k\alpha + \sigma_w^2\sigma_x^2/(\alpha + \sigma_w^2) - |b_0|\sigma_x\sqrt{\alpha}. \tag{3.13}$$

Let α^* be the positive value of α that minimizes the bound in (3.13), which exists and is unique. It is a solution of the polynomial equation

---

[10] See Definition 5.3.3 given in Chap. 5.

$$[2k\sqrt{\alpha} - |b_0|\sigma_x]\,[\alpha + \sigma_x^2]^2 = 2\sigma_w^2\sigma_x^2\sqrt{\alpha}\,. \tag{3.14}$$

Then, when Q is in the structural form (3.9), the solution to (3.3) exists, is linear, and is given by

$$\gamma_0^*(x) = -\mathrm{sgn}(b_0)\frac{\sqrt{\alpha^*}}{\sigma_x}x\,, \tag{3.15}$$

$$\gamma_1^*(y) = E[x|y] = -\frac{\mathrm{sgn}(b_0)\sigma_x\sqrt{\alpha^*}}{\alpha^* + \sigma_w^2}y\,. \tag{3.16}$$

Remark 3.3.1. The main difference between the two problems of Sects. 3.3.2 and 3.3.3 is that Q in the former has a product term between the decision rules of the two agents while in the latter it does not. Hence, it is not only the nonclassical nature of the information structure but also the structure of the performance index that determines whether linear policies are optimal in these quadratic dynamic decision problems with Gaussian statistics and nonclassical information. Another point to note here is that for a similar formulation, but with discrete distributions, both the generalized test channel model [note that (3.5) is a special case of (3.9)] and the Witsenhausen counterexample system exhibit similar features (not admitting a minimum, but allowing computation of the infimum and construction of explicit policies that achieve a cost arbitrarily close to that value), thus indicating that the nature of the distribution of the primitive random variables also makes a difference in the level of complexity of the solution. More on this can be found in the following sections. ◇

3.4 Dynamic Teams with Classical or Quasi-classical Information Patterns

We have seen in the previous section the formidable difficulties associated with solving stochastic dynamic team problems with nonclassical information. Again nonclassical information structure (*IS*) arises if an agent $\mathbf{A}i$'s action affects the information available to another agent $\mathbf{A}j$, who however does not have access to the information available to $\mathbf{A}i$ based on which her action was constructed. Another perhaps mathematically more precise way of stating this is that the information sigma-field (over Ω) of agent $\mathbf{A}j$ is dependent explicitly on the policy (decision rule or control law) of $\mathbf{A}i$. The difficulties in developing a comprehensive and broadly applicable theory for such decision problems with nonclassical information stem from the fact that actions (controls) exhibit *triple* roles: (i) the control effort of reducing the cost, (ii) improvement of future knowledge of uncertainty, and (iii) *signaling* to agents acting in the future some useful information on relevant random variables, which they do not necessarily acquire—and these roles are generally conflicting.

Stochastic dynamic teams which do not entail nonclassical information are generally simpler, since there is no possibility of affecting the information content of measurements or uncertainty at future stages through actions at the present. Such dynamic teams are also known as *neutral*—a term coined by Feldbaum [136]—where [assuming that there is a state-space representation, as in (2.18) and (2.19)] the conditional probability distribution of the state vector given past and present measurements, past control actions, and past control laws (or policies) does not depend on the control laws. *IS*s which are not of the nonclassical type are generally grouped into two categories: *classical* and *quasi-classical*.

*Classical IS*s include deterministic patterns and centralized information patterns. Referring back to the formulation (2.18) and (2.19) where the state-variable-based description was introduced, deterministic patterns arise when the information is not noise-corrupted and may be of the *open-loop* type in which only the initial value of the state is available and no dynamic information is acquired or of the *closed-loop* type where perfect information concerning the current value of the state is also acquired, that is, in (2.19) $y_t^i = x_t$, for all i and t, and there is perfect recall. Centralized patterns arise when all agents exchange their measurements without any delay and also recall the past information.

Under the deterministic or centralized stochastic *IS*s, stochastic team problems become equivalent to stochastic control problems and the solution techniques for these reviewed in Appendix D (see, e.g., [56]) aredirectly available. In particular, for stochastic teams with classical *IS*s, when everything is expressed in terms of action variables and primitive random variables, as in (2.20), if

(i) the measurements are linear in the primitive random variables and past controls (or actions of agents),
(ii) the primitive random variables are jointly Gaussian,
(iii) the loss function is quadratic jointly in the action variables and the primitive random variables, and
(iv) the loss function is strictly convex in the action variables,

then there exists a unique team-optimal solution, which is affine in the available information and can readily be computed by solving a set of minimization problems. A special case of this is the so-called *LQG* control problem (or LQG dynamic team with centralized information) which can be formulated using state variables [as in (2.18) and (2.19)] where we require the state equation (2.18) and the measurements (2.19) to be linear in all the variables and the loss function to be quadratic and jointly strictly convex in the state and control (action) variables, which then admits a unique optimal solution where the optimal control laws are linear in the minimum-mean-square estimate of the state using the measurements, where this estimate is generated recursively (in time) using the Kalman filter (conditional mean and error covariance, independent of the past actions) [56]. The problem features a *separation* of estimation and control as well as *certainty equivalence*, which means that the control gain multiplying the output of the Kalman filter does not depend on the statistics of the primitive random variables, and hence is the same as in the deterministic version of the problem where all random quantities are replaced by their mean values. *Separation* (in a weaker sense) holds even for nonlinear, non-quadratic, and non-Gaussian systems (again under centralized *IS*s) where again the "quality" of the information carried to future stages cannot be affected by the choice of the control policies in the past; this allows for a two-step derivation of the optimum controller: First determine the conditional probability distribution (*cpd*) of the state, express the (expected) cost in terms of this quantity and the control (yet to be determined), and subsequently minimize the new expected cost function over all control laws as functions of the *cpd*, which provides *sufficient statistics* for the stochastic control problem. See Appendix D for further discussions.

The second type of *IS*, which is not nonclassical, is of the *quasi-classical* type, also known as *partially nested*. Assuming again the existence of a common clock for all agents, an *IS* is partially nested if an agent's information at a particular stage t can depend on the action of some other agent at some stage $t' \leq t$ only if she also has access to the information of that agent at stage t'.[11] This would allow, for example, for two agents acting at the same stage t not to share their current measurements, but sharing their past measurements, or two agents who are on different parts of the decision tree, whose actions and information are completely decoupled not to share any information. As a special case, the static team problems considered in Chap. 2, Sect. 2.6, are partially nested, but not with classical *IS*, unless there is complete sharing of (the static) measurements. The one-step delayed information sharing pattern or the one-step delayed measurement sharing pattern introduced in Sect. 2.4 of Chap. 2 is of the quasi-classical type, and it is also important to note that any performance (for a team) that can be achieved under the former can also be achieved under the latter, and *vice versa*, because they generate the same information sigma-fields for each agent, which are further policy invariant (see [29]). In team problems with partially nested information, one talks about *precedence relationships* among agents: an agent $\mathbf{A}i$ is *precedent* to another agent $\mathbf{A}j$ (or $\mathbf{A}i$ *communicates* to $\mathbf{A}j$), if the former agent's actions affect the information of the latter, in which case (to be partially nested) $\mathbf{A}j$ has to have the information based on which the action-generating policy of $\mathbf{A}i$ was constructed.

Under quasi-classical information, LQG stochastic team problems are tractable by conversion into equivalent static team problems of the type discussed extensively in Sect. 2.6.3 of Chap. 2. As an example, consider the following dynamic team with N agents, where each agent acts only once, with $\mathbf{A}k$, $k \in \mathcal{N}$, having the following measurement:

$$y^k = C^k \xi + \sum_{i:i \to k} D_{ik} u^i, \tag{3.17}$$

---

[11] Here we can assume without any loss of generality that each agent acts only *once* in the decision process, or equivalently appears only once on the decision tree. If an agent acts more than once, then (as discussed earlier) she can be split into multiple agents, with each one again acting only once; see also Witsenhausen's *intrinsic model* [400, 401] discussed in Sect. 3.7.

where ξ is an exogenous random variable picked by nature and $i \to k$ denotes the precedence relation that the action of $\mathbf{A}i$ affects the information of $\mathbf{A}k$ and u^i is the action of $\mathbf{A}i$.

If the information structure is quasi-classical, then

$$\mathcal{I}^k = \{y^k, \{\mathcal{I}^i, i \to k\}\}.$$

That is, $\mathbf{A}k$ has access to the information available to all the signaling agents. Such an IS is equivalent to the IS $\mathcal{I}^k = \{\tilde{y}^k\}$, where $\tilde{y}^k$ is a static measurement given by

$$\tilde{y}^k = \left\{C^k\xi, \{C^i\xi, i \to k\}\right\}. \tag{3.18}$$

Such a conversion can be done provided that the policies adopted by the agents are deterministic, with the equivalence to be interpreted in the sense that any deterministic policy measurable under the original *IS* being measurable also under the new (static) *IS* and vice versa, since the actions are determined by the measurements. The restriction of using only deterministic policies is, however, without any loss of optimality: with policies of all other agents fixed (possibly randomized) no agent can benefit from randomized decisions in such team problems. We will discuss this property of irrelevance of random information/actions in optimal stochastic control further in the following chapter in view of Blackwell's Irrelevant Information Theorem (see Theorem D.1.1 in Appendix D).

If the underlying optimization problem is quadratic in all variables and the random variables are all Gaussian, by such a reduction, the optimization problem can be converted into the class of quadratic Gaussian static team problems considered in Sect. 2.6.3. The team-optimal solution under this new (static) *IS* can then be reexpressed in terms of the original *IS*. Examples of such an indirect derivation for dynamic teams with quasi-classical information have been given in several papers [34, 198]. Team-optimal solutions have been constructed under the one-step delayed information sharing pattern [223, 224, 332, 414] as well as under the one-step delayed observation sharing [22] pattern. These do not exhibit *certainty equivalence* or even *separation* as in the case of classical *IS*s, but are recursively computable, involving solutions of coupled linear matrix equations of the type given in Theorem 2.6.8. Another class of dynamic team problems that can be converted into solvable dynamic optimization problems are those where even though the information structure is nonclassical, there is no incentive for signaling because any signaling from say agent $\mathbf{A}i$ to agent $\mathbf{A}j$ conveys information to the latter which is "cost irrelevant," that is, it does not lead to any improvement in performance. If, for example, given a dynamic stochastic team with information structure $\underline{\eta} = (\eta^1, \ldots, \eta^N)$, possibly nonclassical, and the corresponding strategy space $\mathbf{\Gamma} = \Gamma^1 \times \cdots \times \Gamma^N$ and normal form description $(J, \mathbf{\Gamma})$, there is another information structure $\underline{\eta}' = (\eta^{1'}, \ldots, \eta^{N'})$ (quasi-classical or classical), with corresponding composite strategy space $\tilde{\mathbf{\Gamma}} = \tilde{\Gamma}^1 \times \cdots \times \tilde{\Gamma}^N$, where $\mathbf{\Gamma} \supset \tilde{\mathbf{\Gamma}}$, such that

$$\inf_{\underline{\gamma}\in\mathbf{\Gamma}} J(\underline{\gamma}) = \inf_{\underline{\gamma}\in\tilde{\mathbf{\Gamma}}} J(\underline{\gamma}) =: J(\underline{\tilde{\gamma}}^*)\,, \tag{3.19}$$

then $\underline{\tilde{\gamma}}^*$ provides an optimal team solution to the original problem even though it utilizes less information. This statement may seem to be a trivial one (and indeed it is an obvious fact), but it does play an important role in expanding the class of solvable stochastic dynamic teams by converting some with nonclassical information and specific structures into ones with classical information with equality holding in (3.19), which is because the *nonclassical* part of the information structure is *performance irrelevant*. We now make this statement concrete in the next section by considering a class of dynamic teams which fit into this framework.

3.5 Probability and Cost-dependent Properties and Expansion of Information Structures

In this section, we further discuss the role of information structures in dynamic stochastic teams and identify useful refinements that depend on the underlying probability measure and the cost structure.

3.5.1 Performance-irrevelant Signaling and a Stochastic Interpretation of Nestedness

Let us start with a specific 2-agent LQG team with nonclassical information, within the framework of (2.18), (2.19), (2.22), and (2.23), using the notation introduced there:

$$\begin{aligned} x_{t+1}^1 &= x_t^1 + u_t^1 + w_t^{01}, \\ x_{t+1}^2 &= x_t^2 + u_t^2 + w_t^{02}, \end{aligned} \tag{3.20}$$

$$\begin{aligned} y_t^1 &= x_t^1 + w_t^1\,, \quad y_t^2 = x_t^2 + w_t^2, \\ y_t^{12} &= y_t^2 + w_t^{12}\,, \quad y_t^{21} = y_t^1 + w_t^{21}\,, \end{aligned} \tag{3.21}$$

where $t \in \mathcal{T} := \{1, \ldots, T\}$, all variables are scalar and take values on the real line, and $\{x_1^i, w_t^{0i}, w_t^i, w_t^{ij}, i, j = 1, 2, j \neq i, t \in \mathcal{T}\}$ are independent zero-mean Gaussian random variables with specified variances. We assume that at time t, agent $\mathbf{A}i$ has access to the present and past values of y_t^i and y_t^{ij}, $j \neq i$, $i, j = 1, 2$. Hence, $\mathcal{I}_t^i = \{y_{[1,t]}^i, y_{[1,t]}^{ij}\}$, $i, j = 1, 2, j \neq i$. A permissible policy for $\mathbf{A}i$ at time t is a measurable function of $\mathcal{I}_t^i$ into the real line; denote the set of all such policies for $\mathbf{A}i$ over the time interval $\mathcal{T}$ by $\mathbf{\Gamma}$. Note that the information pattern here is neither classical nor partially nested. Note further that there is no sharing of information (even with delay), and the second measurement stream, y_t^{ij}, an agent $\mathbf{A}i$ receives, contains noisy information on the first component of measurement of the other agent $\mathbf{A}j$, but she does not have access to the control actions of that agent that influences her ($\mathbf{A}j$'s) state. It is therefore plausible that $\mathbf{A}j$ could signal to $\mathbf{A}i$ her past measurements by appropriate control actions, through this second channel $\mathbf{A}i$ has access to.

Now, let us consider, as a special case of (2.22), the following loss function for the team:

$$L(x_{[1,T+1]}, \mathbf{u}_{[1,T]}) = L^1(x^1_{[1,T+1]}, u^1_{[1,T]}) + L^2(x^2_{[1,T+1]}, u^2_{[1,T]}), \tag{3.22}$$

where

$$L^i(x^i_{[1,T+1]}, u^i_{[1,T]}) = \sum_{t\in\mathcal{T}} (x^i_{t+1})^2 + r^i(u^i_t)^2\,, \quad i = 1, 2\,, \tag{3.23}$$

where r^i is some positive parameter, for $i = 1, 2$. Let J denote the corresponding cost (expected loss) function for the team [as counterpart of (2.23)], defined over the policy space $\mathbf{\Gamma}$, additively decomposed as in (3.22):

$$J(\gamma^1_{[1,T]}, \gamma^2_{[1,T]}) = J^1(\gamma^1_{[1,T]}, \gamma^2_{[1,T]}) + J^2(\gamma^1_{[1,T]}, \gamma^2_{[1,T]}), \tag{3.24}$$

Note that the loss function (3.22) and (3.23) is *additively decoupled* as far as the agents' state variables and action variables go, and hence if the agents had access to perfect state measurements, each would use only her own state in minimizing the expected value of the individual loss function, and independent minimizations of these would lead to the unique optimal policies for the team (which would be linear for each agent in the current value of her own state). Such a decomposition would hold also if the agents' information structures were instead

$$\tilde{\mathcal{I}}^i_t = \{y^i_{[1,t]}\}\,, \; i = 1, 2\,, \; t \in \mathcal{T}\,, \tag{3.25}$$

which eliminates the possibility of any signaling (note that all random variables are independent), in which case the agents solve independently their scalar LQG problems. Now, with the original information structure, this reasoning does not immediately hold, because even if the loss functions are decoupled, the agents would be coupled through their second measurement channels, which is why in (3.24) we have kept both γ^1 and γ^2 in the arguments of J^1 and J^2, even though L^i in (3.23) depends only on u^i. In spite of this possibility of coupling through information, however, we will show below that this coupling is in fact irrelevant as far as the underlying optimization problem goes, and there is no loss of performance if one works instead with the *IS* (3.25). In a nutshell, the main reason for this is that (in addition to L being additively decoupled) $y^{ij}_{[1,T]}$ is a *garbled* version of $y^j_{[1,T]}$, and hence agent $\mathbf{A}i$ cannot provide any useful information to agent $\mathbf{A}j$ on the second component of her measurement, and also signaling the first component is not useful either because L is additively decoupled.

For a mathematically precise verification of this intuitive result, let us introduce $\tilde{\Gamma}^i$ as the strategy space of $\mathbf{A}i$ corresponding to the information structure (3.25), with the corresponding composite strategy space denoted by $\tilde{\mathbf{\Gamma}}$. Let $\tilde{\mathbf{\Gamma}}_c$ and $\mathbf{\Gamma}_c$ denote the strategy spaces corresponding to *centralized* information structures

$$\tilde{\mathcal{I}}_t^c := \{y^1_{[1,t]}, y^2_{[1,t]}\} \text{ and } \mathcal{I}^c{}_t := \{y^i_{[1,t]}, y^{ij}_{[1,t]}, j \neq i,\ i,j = 1,2\}\,,$$

respectively. Note that

$$\mathbf{\Gamma}_c \supset \tilde{\mathbf{\Gamma}}_c \supset \tilde{\mathbf{\Gamma}} \quad \text{and} \quad \mathbf{\Gamma}_c \supset \mathbf{\Gamma} \supset \tilde{\mathbf{\Gamma}}.$$

Now note the following set of inequalities and equalities:

$$\begin{aligned}
\inf_{\underline{\gamma}\in\mathbf{\Gamma}} J(\gamma^1,\gamma^2) &\geq \inf_{\underline{\gamma}\in\mathbf{\Gamma}_c} J(\gamma^1,\gamma^2) = \inf_{\underline{\gamma}\in\tilde{\mathbf{\Gamma}}_c} J(\gamma^1,\gamma^2) \\
&\geq \inf_{\underline{\gamma}\in\tilde{\mathbf{\Gamma}}_c} J^1(\gamma^1,\gamma^2) + \inf_{\underline{\gamma}\in\tilde{\mathbf{\Gamma}}_c} J^2(\gamma^1,\gamma^2) \qquad (3.26) \\
&= \inf_{\gamma^1\in\tilde{\Gamma}^1} J^1(\gamma^1) + \inf_{\gamma^2\in\tilde{\Gamma}^2} J^2(\gamma^2)\,,
\end{aligned}$$

where the first inequality follows because the second infimization is over a larger set; the next equality follows because under centralized information the collection of y^{ij}'s is a garbled version (cf. Definition 3.2.3) of the collection of y^i's with respect to the state vector $\mathbf{x} = (x^1, x^2)$ and hence is superfluous (not contributing to improvement in performance)[12]; the next inequality follows because the sum of two infima cannot be larger than the infimum of the sum; and the last equality follows because under the reduced centralized information structure we have both informational decoupling and cost decoupling (and hence J^i does not depend on $\gamma^j, j \neq i$). Each of these single-agent optimization problems is a strictly convex LQG problem and hence admits a unique solution where $\tilde{\gamma}^i*$ is a linear function of the conditional mean of x^i at time t, using $y^i_{[1,t]}$ (given by the Kalman filter). Since $\tilde{\gamma}^i*$ is also an element of Γ^i, we have equality throughout in (3.26), and hence (3.19) holds.

We should note that the structure of the dynamic team problem above was picked for convenience in conveying the message succinctly, and the same result applies to more general, linear as well as nonlinear problems, as long as there is decoupling in both the state equation and the loss function, and the garbling condition holds. For example, we could havean N-agent dynamic stochastic team, with (3.20) replaced by

$$x^i_{t+1} = f^i_t(x^i_t, u^i_t, w^{0i}_t)\,,\ t \in \mathcal{T}, i = 1, \ldots, N\,,$$

---

[12] This follows from Proposition 3.2.2, naturally extended to the dynamic case.

(3.21) by

$$y_t^i = g_t^i(x_t^i, w_t^i),$$

$$y_t^{ij} = y_t^j + w_t^{ij}\ ,\ j \neq i,\ i,j = 1,\ldots,N\,,$$

and (3.23) by

$$L^i(x^i_{[1,T+1]}, u^i_{[1,T]}) = \sum_{t\in\mathcal{T}} c_t^i(x_{t+1}^i, u_t^i)\ ,\ i = 1,\ldots,N\,,$$

with

$$L := \sum_{i=1}^{N} L^i\,,$$

where f^i's, g^i's, and c^i's are measurable functions, the variables are all possibly non-scalar, and all random variables are statistically independent of each other and over time. Again, even though this team problem is one with nonclassical information, the measurements $y_t^{ij}, j \neq i$, do not contribute toward improving the performance and hence can be discarded. The resulting problem then is equivalent to N stochastic control problems with classical information and is hence readily solvable.

We should also note that if the loss function had not been additively decoupled, then signaling would come into play because an agent would be able to contribute improvement to the performance by receiving some information on the measurement of some other agent with whom she is coupled through the loss function.

Remark 3.5.1. In the discussion of this section, there is an inherent nestedness condition: A signaling DM's private information is not informative given the signaled DM's information even though the signaled DM does not have access to the information at the signaling DM. The nestedness notion here is termed as *stochastic nestedness* in [417]. The difference between partial nestedness and the setting here is the following. Partial nestedness (or nestedness) is defined independent of the probability measure on the system variables and strictly depends on the information field relationships of the DMs. However, when probabilistic aspects are also considered, a weaker condition identifying the irrelevance of private information at a signaling DM can be established leading to a reduction to a static team. Further aspects of such a stochastic nestedness interpretation are presented in the next subsection where a signaling DM's private information is not informative given the signaled DM's information, provided that the DM's actions are made available to the signaled DM. ◇

3.5.2 Expansion of Information Structures: A Recipe for Identifying Sufficient Information

We start with a general result on *optimum-performance equivalence* of two stochastic dynamic teams with different information structures. This is in fact a result which has a very simple proof, but it is quite effective as we will see shortly.

Proposition 3.5.1. *Let D_1 and D_2 be two stochastic dynamic teams with the same loss function and differing only in their information structures, $\underline{\eta}_1$ and $\underline{\eta}_2$, respectively, with corresponding composite strategy spaces $\mathbf{\Gamma}_1$ and $\mathbf{\Gamma}_2$, such that $\mathbf{\Gamma}_2 \subseteq \mathbf{\Gamma}_1$. Let D_1 admit a team-optimal solution, denoted by $\underline{\gamma}_1^* \in \mathbf{\Gamma}_1$, with the further property that $\underline{\gamma}_1^* \in \mathbf{\Gamma}_2$. Then $\underline{\gamma}_1^*$ also solves D_2.* ◇

Proof. Let J_1 and J_2 be the cost functions of the two teams D_1 and D_2 when expressed in normal form. Note that $\mathbf{\Gamma}_2 \subseteq \mathbf{\Gamma}_1$ says that the strategy space for D_1 is richer than that for D_2. Consider now the following set of equalities and inequalities:

$$J_1(\underline{\gamma}_1^*) = \inf_{\underline{\gamma}_1 \in \mathbf{\Gamma}_1} J_1(\underline{\gamma}_1) \leq \inf_{\underline{\gamma}_1 \in \mathbf{\Gamma}_2} J_1(\underline{\gamma}_1) = \inf_{\underline{\gamma}_2 \in \mathbf{\Gamma}_2} J_2(\underline{\gamma}_2) \leq J_2(\underline{\gamma}_1^*),$$

where the first equality is the definition of $\underline{\gamma}_1^*$; the first inequality follows because infimization over a smaller set cannot lead to a smaller value; the next equality follows because J_1 and J_2 agree on $\mathbf{\Gamma}_2$; and the last inequality follows because as an element of $\mathbf{\Gamma}_2$, $\underline{\gamma}_1^*$ cannot lead to a lower value than the infimum of J_2 on $\mathbf{\Gamma}_2$. □

A recipe for utilizing the result above would be:

> Given a team problem, say D_2, with *IS* $\underline{\eta}_2$, which is presumably difficult to solve, obtain a *finer IS* $\underline{\eta}_1$ and solve the team problem under this expanded *IS* (assuming that this new team problem is easier to solve). Then, if the team-optimal solution here is adapted to the sigma-field generated by the original coarser *IS*, it solves also the original problem D_2.

To see such a recipe at work, consider the following class of two-agent LQG team problems[13] with nonclassical information considered in [417]: The state equation is given by

$$x_{t+1} = Ax_t + B^1 u_t^1 + B^2 u_t^2 + w_t^0 \, , \; t \in \mathcal{T} \, ,$$

the measurement of agent **A**1 by

$$y_t^1 = C^1 x_t + w_t^1 \, , \; t \in \mathcal{T},$$

---

[13] As it will be clear later, restriction to two agents and to the LQG framework is only for convenience in presentation; the result of Proposition 3.5.1 is applicable to a broader class of N-agent teams with nonlinear state equation and non-quadratic loss function, provided that they have an *IS* similar to that of the two-agent team problem considered here.

and that of **A**2 by

$$y_1^2 = C^2 x_1 + w_1^2,$$

where the initial state x_1 and noise sequences $w^i_{[1,T]}$, $i = 0, 1$, are independent zero-mean Gaussian random vectors with specified positive-definite covariance matrices, and w_1^2 is also Gaussian zero-mean with a specified positive-definite covariance matrix. The only restriction is that the statistics of w_1^2 are picked, along with the matrices C^1 and C^2, such that y_1^2 is a *garbled* version of y_1^1 with respect to x_1; in other words (in view of Definition 3.2.3) the conditional probability density function (cpdf) of x_1 given (y_1^1, y_1^2) does not depend on y_1^2. One way of satisfying this requirement would be if y_1^2 was a noisy version of y_1^1, obtained, for example, through

$$y_1^2 = C_3 y_1^1 + v \equiv C_3 C_1 x_1 + C_3 w_1^1 + v =: C_2 x_1 + w_1^2 \,,$$

where v is another independent zero-mean Gaussian random vector with positive-definite covariance.

Now given all of the above, let us specify the information structure of the problem as

$$I_t^1 = \{y^1_{[1,t]}, u^2_{[1,T]}\} \,, \; I_t^2 = \{y_1^2\} \,, \; t \in \mathcal{T},$$

where agent **A**1 keeps her noisy measurement of the current value of the state and has access to the past, present, and future control actions of the other agent, both with perfect recall. **A**2, on the other hand, has open-loop information: she has access to imperfect noisy measurement of the initial value of the state and uses this in the construction of her policies **for all time**, which she does in a *pre-commitment mode*, that is, she decides on $u^2_{[1,T]} = \gamma^2_{[1,T]}(y_1^2)$ at time $t = 1$ and sticks to it for all t. Note that this is a problem which exhibits nonclassical information, because **A**1 observes the actions of **A**2 but does not have access to the measurement (y_1^2) used in the construction of these controls, which opens the possibility of signaling, with **A**2 signaling the value of y_1^2 to **A**1 through his actions.

It is worth noting that an *IS* equivalent to the one above would be the one where the agents' self control actions are also included in the information sets, namely,

$$I_t^1 = \{y^1_{[1,t]}, u^1_{[1,t]}, u^2_{[1,T]}\} \,, \; I_t^2 = \{y_1^2, u^2_{[1,t]}\} \,, \; t \in \mathcal{T},$$

which is sometimes more convenient to work with. Now, an *IS* which is more informative (finer) than the one above would be the one where **A**1 also has access to y_1^2:

$$\tilde{I}_t^1 = \{y^1_{[1,t]}, y_1^2, u^1_{[1,t]}, u^2_{[1,T]}\} \,, \; \tilde{I}_t^2 = \{y_1^2, u^2_{[1,t]}\} \,, \; t \in \mathcal{T} \,.$$

This is a partially nested *IS* because **A**1 now knows everything **A**2 knows but not *vice versa*.

To proceed further, let us adopt a loss function for the team:

$$L(x_{[1,T+1]}, u_{[1,T]}) = \sum_{t\in\mathcal{T}} |x_{t+1}|_Q^2 + |u_t|_R^2 ,$$

where $u_t := (u_t^{1'}\ u_t^{2'})'$, $Q \geq 0$ and $R > 0$ are symmetric weighting matrices of appropriate dimensions, and $|\cdot|_Q$ denotes the Euclidean (semi-)norm weighted by Q. The team problem then is minimization of the expected value of L, expressed in terms of the policy variables using the given information. Now, if instead of the original information structure, we take the expanded one $\tilde{I}$, then the problem is readily solvable and in closed form, as to be shown below, because it is a partially nested LQG team.

We first note that if the weighting matrix R in the loss function is not block diagonal, it can be made one by redefining u_t^1 through a translation that is linear in u_t^2, that is, there exists a matrix M of appropriate dimensions, such that with $\tilde{u}_t^1 := u_t^1 + Mu_t^2$, $|u_t|_R^2 = |\tilde{u}_t^1|_{R^1}^2 + |u_t^2|_{R^2}^2$, where both R^1 and R^2 are positive-definite matrices. Note that such a transformation is compatible with the original information structure, because **A**1 is allowed to have access to the control actions of **A**2. To keep the notation simple, we will assume henceforth (and clearly without any loss of generality) that the original formulation has the weightings on u^1 and u^2 decoupled, that is, that $|u_t|_R^2 = |u_t^1|_{R^1}^2 + |u_t^2|_{R^2}^2$.

Denote the composite strategy spaces corresponding to the original and expanded information structures (i.e., I and $\tilde{I}$) by $\mathbf{\Gamma}_2$ and $\mathbf{\Gamma}_1$, respectively. Let the corresponding ones for an individual agent be superscripted by the identity of that agent (1 or 2). Recall that a team-optimal solution $\underline{\gamma}^{1^*}, \underline{\gamma}^{2^*}$ is one that is defined through the relationship (this one under the expanded *IS*):

$$J(\underline{\gamma}^{1^*}, \underline{\gamma}^{2^*}) = \inf_{\underline{\gamma}^1\in\mathbf{\Gamma}_2^1, \underline{\gamma}^2\in\mathbf{\Gamma}_2^2} J(\underline{\gamma}^1, \underline{\gamma}^2) \equiv \inf_{\underline{\gamma}^2\in\mathbf{\Gamma}_2^2} \inf_{\underline{\gamma}^1\in\mathbf{\Gamma}_2^1} J(\underline{\gamma}^1, \underline{\gamma}^2) , \qquad (\star)$$

where the last expression says that the composite infimization can be carried out sequentially; we can first hold $\underline{\gamma}^2$ fixed as an arbitrary element of $\mathbf{\Gamma}_2^2$, perform minimization with respect to $\underline{\gamma}^1$ over $\mathbf{\Gamma}_2^1$ (let us call this *inner minimization*), and then come back and infimize the resulting expression over $\mathbf{\Gamma}_2^2$. The inner minimization is a discrete-time stochastic control problem with state dynamics

$$x_{t+1} = Ax_t + B^1u_t^1 + B^2\gamma_t^2(y_1^2) + w_t^0 \ , \ t \in \mathcal{T},$$

and loss function

$$L(x_{[1,T+1]}, u_{[1,T]}) = \sum_{t\in\mathcal{T}} |x_{t+1}|_Q^2 + |u_t^1|_{R^1}^2 + |u_t^2|_{R^2}^2 , \qquad (\star\star)$$

where all terms have been defined as before. The controller has access to (y_1^1, y_1^2) at $t = 1$ and to y_t^1 for $t > 1$, and also has access to the realized values of $\gamma_{[1,T]}^2(y_1^2)$ at $t = 1$, and has no restrictions on memory. As such, this is a standard LQG control problem, but with an additional input $B^2\gamma_t^2(y_1^2)$ in the state equation, which however is measurable with respect to the sigma-field generated by y_1^2, and hence can be treated as a deterministic quantity (and in fact it *is* given as $B^2 u_t^2$) when the expectation of L is conditioned on y_1^2.[14] Note that the third term in the expression for L does not enter into the minimization at this stage.

Now, using standard LQG theory with an additional deterministic driving term in the state equation, given also that $Q \geq 0$ and $R^1 > 0$, the problem of minimization of the conditional expected loss function

$$E[L(x_{[1,T+1]}, \gamma^1_{[1,T]}(\tilde{I}^1_{[1,T]}), \gamma^2_{[1,T]}(y_1^2))|y_1^2]$$

over $\mathbf{\Gamma}_2^1$ for fixed $\gamma^2_{[1,T]}$ admits an optimal solution that is of the form

$$\gamma_t^1(y^1_{[1,t]}, u^1_{[1,t-1]}, y_1^2, u^2_{[1,T]}) = \hat{\gamma}_t^1(\hat{x}_t, u^2_{[1,T]})\,, \qquad (o)$$

where

$$\hat{\gamma}_t^1(\hat{x}_t, u^2_{[1,T]}) = -P_t S_{t+1} A \hat{x}_t - P_t(s_{t+1} + S_{t+1} B^2 u_t^2)\,,\ t \in \mathcal{T}\,, \qquad (i)$$

$P_t,\ t \in \mathcal{T}$, is given by

$$P_t = [R^1 + B^{1\prime} S_{t+1} B^1]^{-1} B^{1\prime}\,, \qquad (ii)$$

$S_t,\ t \in \mathcal{T}$, is an appropriate-dimensional matrix sequence generated by the matrix Riccati difference equation

$$S_t = Q + A' S_{t+1}[I - B^1 P_t S_{t+1}]A; \quad S_{T+1} = Q\,, \qquad (iii)$$

and $s_t,\ t \in \mathcal{T}$, is a vector sequence generated by a linear difference equation in retrograde time (where s_t depends linearly on $u^2_{[t+1,T]}$):

$$s_t = A'[I - B^1 P_t S_{t+1}]'[s_{t+1} + S_{t+1} B^2 u_t^2]; \quad s_{T+1} = 0\,. \qquad (iv)$$

Finally,

$$\hat{x}_t = E[x_t \,|\, y^1_{[1,t]}, u^1_{[1,t-1]}, y_1^2, u^2_{[1,t-1]}]\,, \qquad (v)$$

that is, it is the conditional mean of x at time t, given all the past and present measurements and past control actions, and as such is given by the Kalman filter. Note that the solution $\hat{\gamma}_t^1,\ t \in \mathcal{T}$, in (i) does not depend on $\gamma^2_{[1,T]}$ but depends on its realized value $u^2_{[1,T]}$ and also possibly on y_1^2. We next argue that it actually does not depend on the latter. Let $\tilde{x}_t,\ t \in \mathcal{T}$, be generated by

$$\tilde{x}_{t+1} = A\tilde{x}_t + w_t^0\,,\ \tilde{x}_1 = x_1,\ t \in \mathcal{T}\,,$$

---

[14] For the argument here it is not necessary that the controller has access to $u^2_{[1,T]}$ since with the expanded *IS* it has access to y_1^2, but shortly we will see that accessibility to $u^2_{[1,T]}$ is crucial to ensure that the optimal solution to the inner minimization is independent of the policy $\underline{\gamma}^2$.

and

$$\tilde{y}_t^1 := C^1 \tilde{x}_t + w_t^1 \;,\; t \in \mathcal{T} \,.$$

Then, the cpdf of the random variable x_t given the set of random variables $\{y^1_{[1,t]}$, $u^1_{[1,t-1]}$, y^2_1, and $u^2_{[1,t-1]}\}$ is the same as that of $\tilde{x}_t$ given $\{\tilde{y}^1_{[1,t]}, y^2_1\}$, that is, $p(\tilde{x}_t|\tilde{y}^1_{[1,t]}, y^2_1)$. For $t = 1$, this is the cpdf of x_1 given $\{y^1_1, y^2_1\}$, which is independent of y^2_1 since y^2_1 is a garbled version of y^1_1. For $t > 1$, we have

$$p(\tilde{x}_t|\tilde{y}^1_{[2,t]}, y^1_1, y^2_1) = \int_{x_1} p(\tilde{x}_t|x_1, \tilde{y}^1_{[2,t]}, y^1_1, y^2_1)\, p(x_1|y^1_1, y^2_1)\, dx^1 \;/\; p(\tilde{y}^1_{[2,t]}|y^1_1, y^2_1)\,.$$

The first cpdf above is the cpdf of $w^0_{[1,t-1]}$ which is independent of y^2_1; the second cpdf is independent of y^2_1 because of the garbling condition; and the denominator cpdf is also independent of y^2_1 by the same reasoning as in the first cpdf, because $\tilde{y}^1_t := C^1\tilde{x}_t + w^1_t$ and w^1_t is independent of y^2_1. Hence, the Kalman filter does not use the extra measurement y^2_1 when it is a garbled version of y^1_1 with respect to x_1. And consequently the inner minimization problem in $(\star)$ would admit the same solution if the information structure for **A**1 did not include y^2_1.

We now focus on the outer minimization problem in $(\star)$. First, because of the linear structure of the optimal controller (i), the minimum value of the inner minimization in $(\star)$ consists of three general terms: a term which is quadratic in $u^2_{[1,T]}$, a second (bilinear) term which is a product of $u^2_{[1,T]}$ and the primitive random variables, and a third term which depends only on the primitive random variables but not on $u^2_{[1,T]}$. Further, since $(\star\star)$ is positive for all $u^2_{[1,T]}$ whenever x_1 has a positive-definite covariance, the minimum value of the inner minimization in $(\star)$ is strictly convex in $u^2_{[1,T]}$. Hence, the outer minimization in $(\star)$ is a standard static quadratic stochastic optimization problem, which (being strictly convex in the decision variables) admits a unique optimal solution $u^2_t = \gamma^2_t(y^2_1)$, $t \in \mathcal{T}$, which is linear in y^2_1.

Hence, problem $(\star)$ admits a globally optimal solution, $(\underline{\gamma}^{1^*} \in \mathbf{\Gamma}^1_2, \underline{\gamma}^{2^*} \in \mathbf{\Gamma}^2_2)$, which is linear in the available information to the two agents, and furthermore $\underline{\gamma}^{1^*} \in \mathbf{\Gamma}^1_1$ (and naturally also $\underline{\gamma}^{2^*} \in \mathbf{\Gamma}^1_2$). In view of Proposition 3.5.1, $(\underline{\gamma}^{1^*}, \underline{\gamma}^{2^*})$ also solves the original problem with nonclassical information.

Before closing the discussion on the solution to the problem, we note that linearity of the state equation and quadratic structure of the loss function did not play any role in the applicability of Proposition 3.5.1; other than that the optimal solution was linear because of the linear-quadratic structure. Hence, provided the basic assumptions of garbling and the specific *IS* structure hold, a similar result can be obtained for nonlinear, non-quadratic problems as well. We remark, however, that both the garbling assumption and the one that allows **A**1 to have full access not only to the past and current values of the actions of **A**2 but also their future values (which entails a pre-commitment mode of operation) are essential for linearity of the optimal solution. The former is needed for the problem with expanded (partially nested) information to have the same minimum as the one with the original nonclassical information. The latter is needed because otherwise the recursion (iv) cannot be implemented. If any of these two assumptions do not hold, then there emerges the possibility of $\mathbf{A}2$ signaling the value of y_1^2 to $\mathbf{A}1$ through her control actions. We discuss below one such class of problems where signaling through control actions becomes instrumental in improving the performance of a team.

3.6 Signaling Through Control Actions

Consider the following 2-agent stochastic dynamic team problem: u^1 and u^2 are real-valued control actions of agents $\mathbf{A}1$ and $\mathbf{A}2$, respectively, who have access to information $I^1 = \{u^2\}$ and $I^2 = \{\alpha, x\}$, where α takes two values, 1 and 2, each with equal probability $\frac{1}{2}$, x is a uniform random variable with support set the interval $[-\sqrt{3}, \sqrt{3}]$, α and x are statistically independent, and all this is common information to both agents. Hence, $u^1 = \gamma^1(u^2)$ and $u^2 = \gamma^2(\alpha, x)$, where the policies $\gamma^1 \in \Gamma^1$ and $\gamma^2 \in \Gamma^2$ are measurable maps. Note that this is a problem with nonclassical information, because $\mathbf{A}1$ has access to the action variable of $\mathbf{A}2$, but does not have access to the information used by $\mathbf{A}2$. Let the loss function for the team be given by

$$L(\alpha, x, u^1, u^2) = \alpha(u^1 + u^2)^2 + (u^2 - x)^2 + (u^1)^2 \,, \tag{\#}$$

with the corresponding expected loss (or cost) given by $J(\gamma^1, \gamma^2)$. The goal is to minimize J over $\Gamma^1 \times \Gamma^2$.

Now, if $\mathbf{A}1$ instead had the expanded *IS* $\tilde{I}^1 = \{u^2, \alpha, x\}$, then this would be a problem with classical information, and the unique pair of policies would be the ones that minimize $(\#)$ for each α and x:

$$\tilde{\gamma}^1(u^2, \alpha, x) = -\frac{1}{1+\alpha} u^2 \,, \quad \tilde{\gamma}^2(\alpha, x) = -\frac{1}{2+\alpha} x \,, \tag{\#\#}$$

with the minimum value of L being

$$\min_{u^1, u^2} L(\alpha, x, u^1, u^2) = \frac{1}{2+\alpha} x^2 \,,$$

whose expected value, $\frac{7}{24}$, is the minimum of J over the expanded policy spaces.

The question now is whether this expected value can be achieved, exactly or approximately, under the original information structure. Note that $\mathbf{A}1$ does not use the additional information on x in $(\#\#)$, but needs the value of α. Hence, if it were possible for $\mathbf{A}1$ to decipher the true value of α through her observation of u^2 or, put it another way, if $\mathbf{A}2$ were able to signal to $\mathbf{A}1$ the true value of α through her action u^2 without compromising performance attained under $\tilde{\gamma}^2$ in $(\#\#)$, then the same performance would be achieved under the original *IS*. But a closer look will reveal that since $\tilde{\gamma}^2$ is unique as given, *signaling at zero cost* is not possible. Then, is perfect signaling at some infinitesimal cost possible? We will argue below that yes, given any $\epsilon > 0$, it is possible to come up with a policy for $\mathbf{A}2$ which will lead to a value smaller than $\frac{7}{24} + \epsilon$.

First note that any realized value of $u^2 = \tilde{\gamma}^2(\alpha, x)$ in $(\#\#)$ will be in the interval $[-\frac{\sqrt{3}}{3}, \frac{\sqrt{3}}{3}]$, and signaling the true value of α will require just *one* bit of information. Hence, if $\mathbf{A}2$ adopts a strategy of truncating the realized value of $\tilde{\gamma}^2(\alpha, x)$ to n decimal points and uses an extra $n+1$'-th decimal point to convey the true value of α, 1 or 2,[15] then $\mathbf{A}1$ will be able to decipher the true value of α from her observation of u^2 and be able to implement the policy $\tilde{\gamma}^1(u^2, \alpha)$. There will be a small deviation from the ultimate performance level of $\frac{7}{24}$, with the *smallness* determined by the value of n. Clearly, given any small $\epsilon > 0$, one can always find a large enough integer n such that the strategy above for $\mathbf{A}2$ will result in a cost no larger than $\frac{7}{24} + \epsilon$. But clearly, ϵ cannot be driven to zero, and hence this problem admits only an ϵ-optimal solution, as in the case of the team problem of Sect. 3.3.1, but for a totally different reason.

This type of a result, considered by Bismut [60], where part of the action variable in a decimal expansion carries information from one agent to another on an observed random variable through appropriate encoding, resulting in ϵ-optimality, applies to more general problems as long as one agent has perfect access to the action variable of another agent. To put what we have observed above in more precise terms, the ϵ term arises due to the fact that the control policy is to encode information on both the control action and the observation, with as minimum damage as possible to the control action; and this is possible due to the fact that a real number carries infinite amount of information (when information is measured in Shannon information theoretic bits). One way to achieve this is as follows: Since rational numbers are dense in $\mathbb{R}$, for any ϵ', there exists an $n \in \mathbb{Z}$ such that an n-decimal representation which is at most at an ϵ' distance from any real number in a compact set is possible. Therefore, if one is to represent a finite-dimensional (say, r-dimensional) control variable in a compact set $\mathbb{U} \subset \mathbb{R}^r$ and a finite-dimensional (say, m-dimensional) observation variable in a compact set $\mathbb{Y} \subset \mathbb{R}^m$, these signals can be represented with an arbitrarily small error by a single rational number. Thus, one may embed in this number, the ϵ'-approximate decimal expansions of the numbers to be represented, thus leading to a total of $n(m+r)$ decimal letters, by allocating the most significant nr letters for the control variable. Clearly, the ϵ used in the performance is directly related to the ϵ' used in the truncation argument above.

If the control and observation variables take values in a non-compact set, then, by separability, a countable representation is possible, but a uniform number of decimal letters will then not be sufficient, making the coding design impractical.

---

[15] For example, let a particular realization be $x = 1.7$ and $\alpha = 1$, and n be picked as 7. Then, $\tilde{\gamma}^2(\alpha, x) = -1.7/3$, which would be, using a 7-decimal point truncation, -0.5666666. Then, $\mathbf{A}2$'s action would be $u^2 = -0.56666661$.

3.7 Revisiting Witsenhausen's Characterization of Information Structures

Heretofore in this chapter, as well as in Chap. 2, we have presented a comprehensive study of the characterization of information structures for stochastic team decision problems. In this section, we revisit these and present some further characterizations as laid out by Witsenhausen, termed as *the intrinsic model* [400]. In this model (described in discrete time), any action applied at any given time is regarded as applied by an individual decision maker/agent, who acts only once. One advantage of this model, in addition to its generality, is that the definitions regarding information structures can be compactly described.

Suppose that in the decentralized system considered below, there is a predefined order in which the decision makers act. Such systems are called *sequential systems*. In the context of a sequential system, the *intrinsic model* has three components:

- A collection of *measurable spaces* $\mathcal{I} := \{\Omega, \mathcal{F}, (U^i, \mathcal{U}^i), (Y^i, \mathcal{Y}^i), i \in \mathcal{N}\}$, specifying the system's distinguishable events, and the control and measurement spaces. Here $N = |\mathcal{N}|$ is the number of control actions taken, and each of these actions is taken by an individual (different) DM (hence, even a DM with perfect recall can be regarded as a separate decision maker every time it acts). The pair $(\Omega, \mathcal{F})$ is a measurable space (on which an underlying probability may be defined). The pair $(U^i, \mathcal{U}^i)$ denotes the measurable space from which the action of decision maker i, u^i is selected. The pair $(Y^i, \mathcal{Y}^i)$ denotes the measurable observation/measurement space.
- A *measurement constraint* which establishes the connection between the observation variables and the system's distinguishable events. The Y^i-valued observation variables are given by $\mathcal{I}^i = \eta^i(\omega, \mathbf{u}^{-i})$, $\mathbf{u}^{-i} = \{u^k, k \leq i-1\}$, η^i measurable functions, and u^k denotes the action of DM k.
- A *design constraint*, which restricts the set of admissible N-tuple control laws $\underline{\gamma} = \{\gamma^1, \gamma^2, \ldots, \gamma^N\}$, also called *designs* or *policies*, to the set of all measurable control functions, so that $u^i = \gamma^i(\mathcal{I}^i)$, with $\mathcal{I}^i = \eta^i(\omega, \mathbf{u}^{-i})$, and γ^i, η^i measurable functions.

Hence, the information variable $\mathcal{I}^i$ induces a σ-field, $\sigma(\mathcal{I}^i)$ over $\Omega \times \prod_{k=1}^{i-1} U^k$.

One can also introduce a fourth component:

- A *probability measure* P defined on $(\Omega, \mathcal{F})$. To accommodate randomizations in individual decisions, this can be expanded both in $\mathcal{F}$ as well as $\mathcal{I}^i$, by including randomization events.

Under this intrinsic model, a team problem is *dynamic* if the information available to at least one DM is affected by the action of at least one other DM. A decentralized problem is *static*, if the information available at every decision maker is only affected by exogenous disturbances (nature); that is, no other decision maker can affect the information at any given decision maker.

As before, information structures can also be classified as classical, quasi-classical, or nonclassical. An IS $\{\mathcal{I}^i, 1 \leq i \leq N\}$ is *classical* if $\mathcal{I}^i$ contains all of the information available to DM k for $k < i$. An IS is *quasi-classical* or *partially nested*, if whenever u^k, for some $k < i$, affects $\mathcal{I}^i$, $\mathcal{I}^i$ contains $\mathcal{I}^k$ (i.e., $\sigma(\mathcal{I}^k) \subset \sigma(\mathcal{I}^i)$). An IS which is not partially nested is *nonclassical*.

The Nonsequential Case

The order of actions can also be nature-dependent in some settings. If there is a pre-defined order in which the decision makers act, as above, then we say that a system is sequential. Otherwise, the system is nonsequential. Such systems are substantially more difficult to study, since the ambiguities in the order of actions lead to challenges on the interpretation of local information.

On a conceptual level, the intrinsic model described above captures such dynamic teams as well. In particular, the measurement constraint described above needs to be adjusted such that the observation variables are given by $\mathcal{I}^i = \eta^i(\omega, \mathbf{u}^{-i})$, $\mathbf{u}^{-i} = \{u^k, k \neq i, k \in \mathcal{N}\}$, where η^i is measurable and u^k denotes the action of DM k.

Design of such nonsequential systems requires a careful construction since the systems should be deadlock-free, that is, the actions of a given DM should not depend on the actions of DMs acting in the future, for any realized random ordering (see Fig. 2.1 for an example of a system with deadlock). Furthermore, the optimization problem for such systems should be well posed, since for some designs the expected cost may not be well defined. To gain further insight into these intricacies, consider the following two examples (see Teneketzis [360]):

Example 3.7.1. Let $\Omega = U^1 = U^2 = U^3 = \{0, 1\}$ and

$$\sigma(\mathcal{I}^1) = \Big\{\emptyset, \Omega \times U^1 \times U^2 \times U^3, \{(\omega, u^1, u^2, u^3) : \omega(1-u^2)u^3 = 1\},$$
$$\{(\omega, u^1, u^2, u^3) : \omega(1-u^2)u^3 = 0\}\Big\},$$

$$\sigma(\mathcal{I}^2) = \Big\{\emptyset, \Omega \times U^1 \times U^2 \times U^3, \{(\omega, u^1, u^2, u^3) : \omega(1-u^3)u^1 = 1\},$$
$$\{(\omega, u^1, u^2, u^3) : \omega(1-u^3)u^1 = 0\}\Big\},$$

$$\sigma(\mathcal{I}^3) = \Big\{\emptyset, \Omega \times U^1 \times U^2 \times U^3, \{(\omega, u^1, u^2, u^3) : \omega(1-u^1)u^2 = 1\},$$
$$\{(\omega, u^1, u^2, u^3) : \omega(1-u^1)u^2 = 0\}\Big\}.$$

This system has a deadlock, since no DM can act as in the top figure of Fig. 2.1. ◇

Example 3.7.2. Let $\Omega = U^1 = U^2 = \{0,1\}$ and $\sigma(\mathcal{I}^1) = 2^{\Omega \times U^2}$, $\sigma(\mathcal{I}^2) = 2^{\Omega \times U^1}$ (where the notation 2^U denotes the power set, i.e., the collection of all subsets of U). Consider the following team policy:

$$\gamma^1(\omega, u^2) = 0 \times 1_{\{u^2=0\}} + 1 \times 1_{\{u^2=1\}},$$

$$\gamma^2(\omega, u^1) = 0 \times 1_{\{u^1=0\}} + 1 \times 1_{\{u^1=1\}},$$

where 1_E denotes the indicator function for event E. For this design, consider the realization $\omega = 0$. In this case, $(\omega, u^1, u^2) = (0,0,0)$ as well as $(0,1,1)$ are acceptable realizations given the policy stated above. A similar setting occurs for $\omega = 1$, since $(1,0,0)$ and $(1,1,1)$ are acceptable realizations. Hence, for a given cost function c, there does not exist, in general, a well-defined (measurable) cost realization variable $c(\omega, u^1, u^2)$ under this policy, and the expectation $E[c(\omega, u^1, u^2)]$ is not well defined given the policy (γ^1, γ^2). ◇

These two examples exhibit the difficulties arising in non-sequential systems. For such systems, one needs to consider the sequences of possible events to ensure that these issues do not arise. We refer the reader to Witsenhausen [393], Andersland and Teneketzis [9, 10], and Teneketzis [360] for a comprehensive study of nonsequential systems.

Witsenhausen's Equivalent Model and Static Reduction of Sequential Dynamic Teams

Another equivalence between sequential dynamics teams and their static reduction is as follows (termed as *the equivalent model* [401]).

Consider a dynamic team setting according to the intrinsic model where there are N time stages, and each DM observes, for some t, $y^t = g_t(\omega_0, \omega_t, u^1, u^2, \cdots, u^{t-1})$, and the decisions are generated by $u^t = \gamma_t(\{y^\tau, \tau \in K_t\})$, where K_t is the set of observations available at DM t. Here $\omega_0, \omega_1, \cdots, \omega_N$ are primitive (exogenous) variables. We assume that all variables take values in real Euclidean spaces (or complete separable metric spaces). The resulting cost under a given team policy is

$$J(\underline{\gamma}) = E[c(\omega_0, \mathbf{y}, \mathbf{u})],$$

where, as before, we have the notation $\mathbf{y} = \{y^k, k \in \mathcal{N}\}$. This dynamic team can be converted to a static team provided that the following absolute continuity condition holds: For every $t \in \mathcal{N}$, there exists a function f_t such that for all Borel S

$$P(y^t \in S | \omega_0, u^1, u^2, \cdots, u^{t-1}) = \int_S f_t(y_t, \omega_0, u^1, u^2, \cdots, u^{t-1}) Q_t(dy^t).$$

We can then write (since the action of each DM is determined by the measurement variables under a policy)

$$P(d\omega_0, d\mathbf{y}) = P(d\omega_0) \prod_{t=1}^{N} f_t(y_t, \omega_0, u^1, u^2, \cdots, u^{t-1}) Q_t(dy^t).$$

The cost function $J(\underline{\gamma})$ can then be written as

$$J(\underline{\gamma}) = \int P(d\omega_0) \prod_{t=1}^{N} (f_t(y_t, \omega_0, u^1, u^2, \cdots, u^{t-1}) Q_t(dy^t)) c(\omega_0, \mathbf{y}, \mathbf{u}),$$

where now the measurement variables can be regarded as independent and by incorporating the $\{f_t\}$ terms into c, we can obtain an equivalent *static team* problem. Hence, the essential step is to appropriately adjust the probability space and the cost function. For the (Witsenhausen's) counterexample considered in Sect. 3.3.2, this reduction leads to the following equivalent static problem:

$$\int P(dy_0) P(dy_1) \Big((y_0 + u_0 - u_1)^2 + k^2 u_0^2 \Big) e^{(y_0+u_0)(2y_1-y_0-u_0)/2},$$

where $u_0 = \gamma_0(y_0), u_1 = \gamma_1(y_1)$ and y_0, y_1 are independent, zero-mean Gaussian variables with variances σ^2 and 1, respectively. One immediately notices that this static reduction leads to a loss function which is not convex in the control actions further explaining the source of difficulties in solving the counterexample.

Such a reduction is both conceptually and computationally useful. The computational benefit is particularly evident in finite state-space decision problems, since it amounts to a reformulation of the underlying optimization problem.

Standard Form for Sequential Dynamic Teams

According to another model for sequential teams, known as *Witsenhausen's standard form* [394], for optimization of sequential dynamic teams with finite horizons, a dynamic programming principle can be applied which essentially expresses the optimization problem as a terminal-stage cost function. Here, every DM acts given the policies of the previous DMs optimally. When the cost function is stagewise additive and further assumptions are placed on the primitive variables, the analysis reduces to the dynamic programming formulation for state-space models (see Appendix D for a review of dynamic programming).

3.8 Concluding Remarks

The main theme of this chapter has been information structures and the role they play in stochastic dynamic teams. We have seen the complexity and challenges involved in obtaining team-optimal solutions for dynamic teams when the information structure is nonclassical, which would arise when, for example, the agents do not share information and/or have limited memory. We have also seen that the structures of the cost function and the system dynamics (describing the evolution of the decision process) as well as the overall probabilistic description of the team problem also play important roles in the existence, complexity, solvability, and characterization of team-optimal solutions. We have obtained a method for solving nonclassical information structures through expanding the information set, obtaining a solution and checking if the solution is realizable according to the original information structure. We also discussed performance-irrelevant signaling for a class of nonclassical information structures.

3.9 Bibliographic Notes

In addition to the extensive references provided in the main body of the chapter, we refer the reader to [29, 197, 250, 332, 417] for further discussion on and examples of information structures. Partially nested information structures also include the cases where explicit information exchange in a decentralized system among decision makers is faster than information propagation through system dynamics; see [87,320,382]. Related to the notion of partial nestedness, Bamieh and Voulgaris, Rotkowitz and Lall, and Voulgaris [38,324,382] have studied sufficiency conditions for tractability and convexity in optimal decentralized control problems.

Further examples on information patterns and structural results for optimal team policies have been considered in [4,13,22,224,286,332,374,414,417]. In Chap. 12, further discussions on the generation of optimal team policies are presented.

In the context of dynamic quadratic team problems, if one poses the problem not as an expectation minimization but as a min-max optimization where nature acts as the maximizer and the encoders/decoders (or the controllers) act as the minimizer, then linear policies are optimal for a class of settings; see, for example, [26, 30, 31, 157, 323]; [26] also provides a review on LQG problems under nonclassical information including Witsenhausen's counterexample. Also regarding this counterexample, connections with information-theoretic concepts such as binning have been made in [177] to obtain approximation bounds.

Approximately optimal solutions for weakly coupled teams with nonclassical information structures have been considered in [345]. The discussion in the chapter on performance-irrelevant signaling builds on and extends the *stochastically nested information structure* in [417]. This generalizes earlier characterizations based on information-field inclusion based characterizations of nestedness by Witsenhausen

[399] and Ho and Chu [198]. Blackwell's [61] comparison of information structures is a related concept in a single decision maker setting—a topic to be covered further in the next chapter. Mahajan and Yüksel [251] generalized this view to Witsenhausen's intrinsic model.

We recall that an important property for nonclassical information structures is that the information fields at the decision makers are not nested. A result is that the conditional expectation operations of each decision maker require careful constructions (and in particular, these operations do not satisfy the law of the iterated expectations; see (B.2) in Appendix B, also known as the smoothing property of conditional expectation). With this observation and the motivation that the information fields generated by local measurements lead to nontrivial constructions for nonsequential systems, an alternative algebraic model for describing team decision problems has been proposed in [44, 45].

Game theory and economics literatures have also considered *signaling* extensively in the context when decision makers (players) have different objective functions to be optimized (see, e.g., [25, 33, 105]). In these cases, the design of the information structure leads to further subtle intricacies.

Section 3.5 builds, in part, on [417]. The material on garbling in the chapter builds, in part, on [61, 255].

Chapter 4
Topological Properties of Information Structures: Comparison, Convergence, and Optimization

4.1 Introduction

In Chaps. 2 and 3, we introduced the notion of an information structure and studied various properties associated with information structures in the contexts of static and dynamic stochastic teams, identifying classes of teams (in terms of their information structures) which are tractable and others which present major challenges because of the nonclassical nature of the information.

In this chapter, we will continue with our discussion of information structures but from a different perspective: studying their topological properties. In particular, we study topological and structural properties of the functional $\inf_{\underline{\gamma}} R(\underline{\gamma}; \underline{\eta})$, where $R(\cdot;\cdot)$ is defined in (3.1), on a space of information structures. We investigate values of measurement channels and the problem of channel optimization for a class of stochastic control problems. In addition to establishing existence and structural results for optimal policies, the topological properties developed in the chapter will also be useful in assessing robustness of designs when a probabilistic description of a model is not fully accurate. We will also introduce quantizers as a special class of measurement channels and obtain several useful topological properties and existence results for optimal quantizers. Quantizers will be studied further in the book, in Parts II and III.

We start the chapter by first defining measurement channels as information structures and stating the problems considered in this chapter, in Sect. 4.2. We then consider comparison of measurement channels in Sect. 4.3. Topological characterization of information structures is presented in Sect. 4.4. Continuity of optimal solutions in channels for single-stage optimization problems is considered in Sect. 4.5 and existence analysis is provided in Sect. 4.6. Quantizers, viewed as a subclass of measurement channels, are studied in detail in view of continuity and existence properties in Sect. 4.7. Multistage problems are considered in Sect. 4.8 and multi-agent setting is investigated in Sect. 4.9. Applications of the results in

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_4,

the chapter to nonclassical team problems in view of lack of convexity in control policies are discussed in Sect. 4.10. Finally, relaxations of the continuity properties under setwise and weak convergence notions are discussed in Sect. 4.11.

4.2 Measurement Channels as Information Structures

Consider a single-agent dynamical system described by the discrete-time equations [as single-agent, time-invariant counterparts of (2.18) and (2.19)]

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, w_t), \\ y_t &= g(x_t, v_t), t \in \mathcal{T}, \end{aligned}$$

for some measurable functions f, g, with $\{w_t\}$ being an independent and identically distributed (i.i.d) system noise process and $\{v_t\}$ an i.i.d. measurement disturbance process, which are independent of x_0 and each other. Here, $x_t \in \mathbb{X}$, $y_t \in \mathbb{Y}$, $u_t \in \mathbb{U}$, where we assume that these spaces are Borel subsets of finite-dimensional Euclidean spaces.

In the above, we can view g as inducing a measurement channel Q, which is a stochastic kernel or a regular conditional probability measure from $\mathbb{X}$ to $\mathbb{Y}$ in the sense that $Q(\,\cdot\,|x)$ is a probability measure on the (Borel) σ-algebra $\mathcal{B}(\mathbb{Y})$ on $\mathbb{Y}$ for every $x \in \mathbb{X}$, and $Q(A|\,\cdot\,) : \mathbb{X} \to [0, 1]$ is a Borel-measurable function for every $A \in \mathcal{B}(\mathbb{Y})$.

As before, an *admissible policy* is a sequence of control functions $\{\gamma_t,\ t \in \mathbb{Z}_+\}$ such that γ_t is measurable with respect to the σ-algebra generated by the information variables

$$I_t = \{y_{[0,t]}, u_{[0,t-1]}\}, \quad t \in \mathbb{N}, \qquad I_0 = \{y_0\},$$

where

$$u_t = \gamma_t(I_t), \quad t \in \mathbb{Z}_+ \tag{4.1}$$

are the $\mathbb{U}$-valued control actions. Let $\mathbf{\Gamma}$ be the space of all admissible policies.

With the above setup, let the objective be one of minimization of the cost

$$J(P, Q, \underline{\gamma}) = E_P^{Q,\underline{\gamma}}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{4.2}$$

over the set of all admissible policies $\underline{\gamma}$, where $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is a Borel-measurable stagewise cost function and $E_P^{Q,\underline{\gamma}}$ denotes the expectation with initial state probability measure given by P, under policy $\underline{\gamma}$ and given channel Q.

For $P \in \mathcal{P}(\mathbb{X})$ and $Q \in \mathcal{Q}$ we let PQ denote the joint distribution induced on $(\mathbb{X} \times \mathbb{Y}, \mathcal{B}(\mathbb{X} \times \mathbb{Y}))$ by channel Q with input distribution P:

$$PQ(A) = \int_A Q(dy|x)P(dx), \quad A \in \mathcal{B}(\mathbb{X} \times \mathbb{Y}).$$

We adopt the convention that given a probability measure μ, the notation $z \sim \mu$ means that z is a random variable with distribution μ.

Comparison of Measurement Channels (Stochastic Kernels): The first question we ask is when can one compare two measurement channels Q^1, Q^2 such that

$$\inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(P, Q^1, \underline{\gamma}) \leq \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(P, Q^2, \underline{\gamma}),$$

for all measurable and bounded cost functions c in (4.2)?

Continuity on the Space of Measurement Channels (Stochastic Kernels): Suppose that $\{Q_n, n \in \mathbb{N}\}$ is a sequence of communication channels converging in some sense to a channel Q. Then the question we ask is, when does

$$Q_n \to Q$$

imply

$$\inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(P, Q_n, \underline{\gamma}) \to \inf_{\underline{\gamma} \in \mathbf{\Gamma}} J(P, Q, \underline{\gamma})?$$

Existence of Optimal Measurement Channels and Quantizers: Let $\mathcal{Q}$ be a set of communication channels. A second question we ask is, when do there exist minimizing and maximizing channels for the optimization problems

$$\inf_{Q \in \mathcal{Q}} \inf_{\underline{\gamma}} E_P^{Q, \underline{\gamma}} \left[\sum_{t=0}^{T-1} c(x_t, u_t) \right]$$

and

$$\sup_{Q \in \mathcal{Q}} \inf_{\underline{\gamma}} E_P^{Q, \underline{\gamma}} \left[\sum_{t=0}^{T-1} c(x_t, u_t) \right]?$$

If solutions to these problems exist, are they unique?

4.3 Concavity on the Space of Channels and Blackwell's Comparison of Information Structures

The following result has important consequences in decentralized stochastic control problems as will be elaborated on later. It is also related to the discussion in Chap. 3 on garbling.

Theorem 4.3.1 ([436]). *Let $T = 1$ and let the integral $\int c(x, \gamma(y)) PQ(dx, dy)$ exist for all $\gamma \in \mathbf{\Gamma}$ and $Q \in \mathcal{Q}$. Then, the function*

$$J(P, Q) = \inf_{\gamma \in \mathbf{\Gamma}} E_P^{Q, \gamma}[c(x, u)]$$

is concave in Q. ◇

Proof. For $\alpha \in [0, 1]$ and $Q', Q'' \in \mathcal{Q}$, let $Q = \alpha Q' + (1 - \alpha)Q'' \in \mathcal{Q}$, i.e.,

$$Q(A|x) = \alpha Q'(A|x) + (1 - \alpha)Q''(A|x)$$

for all $A \in \mathcal{B}(\mathbb{Y})$ and $x \in \mathbb{X}$. Noting that $PQ = \alpha PQ' + (1 - \alpha)PQ''$, we have

$$\begin{aligned}
J(P, Q) &= J(P, \alpha Q' + (1 - \alpha)Q'') \\
&= \inf_{\gamma \in \mathbf{\Gamma}} E_P^{Q,\gamma}[c(x, u)] \\
&= \inf_{\gamma \in \mathbf{\Gamma}} \int c(x, \gamma(y))PQ(dx, dy) \\
&= \inf_{\gamma \in \mathbf{\Gamma}} \left(\alpha \int c(x, \gamma(y))PQ'(dx, dy) + (1 - \alpha) \int c(x, \gamma(y))PQ''(dx, dy)\right) \\
&\geq \inf_{\gamma \in \mathbf{\Gamma}} \left(\alpha \int c(x, \gamma(y))PQ'(dx, dy)\right) \\
&\qquad + \inf_{\gamma \in \mathbf{\Gamma}} \left((1 - \alpha) \int c(x, \gamma(y))PQ''(dx, dy)\right) \\
&= \alpha J(P, Q') + (1 - \alpha)J(P, Q'') \qquad (4.3)
\end{aligned}$$

proving that $J(P, Q)$ is concave in Q. □

The following result is a folk theorem in statistical decision theory whose proof is similar to that of Theorem 4.3.1.

Proposition 4.3.1 ([436]). *The function*

$$V(P) := \inf_{u \in \mathbb{U}} \int c(x, u)P(dx),$$

is concave in P, under the assumption that c is measurable and bounded. ◇

We will use the preceding observation to revisit a classical result in statistical decision theory and comparison of experiments, by David Blackwell [61], who considered a finite $\mathbb{X}$. In a single decision maker setup, we refer to the probability space induced on $\mathbb{X} \times \mathbb{Y}$ as an information structure.

As we observed in Remark 3.2.2 in the previous chapter, *garbling* can be further weakened for comparison of information structures.

Definition 4.3.1. An information structure induced by some channel Q_2 is weakly stochastically degraded with respect to another one, Q_1, if there exists a channel Q' on $\mathbb{Y} \times \mathbb{Y}$ such that

$$Q_2(B|x) = \int_{\mathbb{Y}} Q'(B|y)Q_1(dy|x), \quad B \in \mathcal{B}(\mathbb{Y}), \ x \in \mathbb{X}.$$

◇

In view of Proposition 4.3.1, we have the following.

Theorem 4.3.2 (Blackwell [61]). *If Q_2 is weakly stochastically degraded with respect to Q_1, then the information structure induced by channel Q_1 is more informative with respect to the one induced by channel Q_2 in the sense that*

$$\inf_{\gamma} E_P^{Q_2,\gamma}[c(x,u)] \geq \inf_{\gamma} E_P^{Q_1,\gamma}[c(x,u)],$$

for all measurable and bounded cost functions c. ◇

Proof. Let $(x, y^1) \sim PQ_1$, y^2 be such that $\Pr(y^2 \in B | x = x, y^1 = y) = Q'(B|y)$ for all $B \in \mathcal{B}(\mathbb{Y})$, $y^1 \in \mathbb{Y}$, and $x \in \mathbb{X}$. Then x, y^1, and y^2 form a Markov chain in this order and therefore $P(dy^2|y^1, x) = P(dy^2|y^1)$ and $P(x|dy^2, y^1) = P(x|y^1)$.[1] Thus we have

$$\begin{aligned} J(P, Q_2) &= \int V\big(P(\,\cdot\,|y^2)\big)P(dy^2) \\ &= \int V\Big(\int P(\,\cdot\,|y^1)P(dy^1|y^2)\Big)P(dy^2) \\ &\geq \int \Big(\int P(dy^1|y^2)V\big(P(\,\cdot\,|y^1)\big)\Big)P(dy^2) \\ &= \int V\big(P(\,\cdot\,|y^1)\big)\Big(\int P(dy^1|y^2)P(dy^2)\Big) \\ &= \int V\big(P(\,\cdot\,|y^1)\big)P(dy^1) \\ &= J(P, Q_1), \end{aligned}$$

where in arriving at the inequality, we used Proposition 4.3.1 and Jensen's inequality. □

Remark 4.3.1. When $\mathbb{X}$ is finite, Blackwell showed that the above condition also has a converse theorem if P has positive measure on each element of $\mathbb{X}$: For an information structure to be more informative, weak stochastic degradedness is a necessary condition. For Polish $\mathbb{X}$ and $\mathbb{Y}$, the converse result holds under further technical conditions on the stochastic kernels (information structures); see [67, 86]. ◇

The comparison argument applies also for multistage settings with $T > 1$.

Theorem 4.3.3. *For the multistage problem (4.2), if Q_2 is weakly stochastically degraded with respect to Q_1, then the information structure induced by channel Q_1 is more informative with respect to the one induced by channel Q_2 in the sense that for all measurable and bounded cost functions c in (4.2),*

---

[1] We have slightly abused notation here, using, for example, $P(dy^2|y^1, x)$ instead of $P_{y^2|y^1,x}(dy^2|y^1, x)$.

$$J(P,Q_1) \leq J(P,Q_2).$$

◇

Proof. We will follow an argument based on *simulation*. As in the proof of Theorem 4.3.2, let for $t \geq 0$, $(x_t, y_t^1) \sim PQ_1$, y_t^2 be such that $\Pr(y_t^2 \in B|x = x, y_t^1 = y) = Q'(B|y)$ for all $B \in \mathcal{B}(\mathbb{Y})$, $y_t^1 \in \mathbb{Y}$, and $x_t \in \mathbb{X}$. Then x_t, y_t^1, and y_t^2 form a Markov chain in this order. Now, the cost value achieved under any given policy $\underline{\gamma}^2$ which is measurable under the channel Q_2 can also be achieved by a randomized decision policy of the decision maker under channel Q_1. That is, given a policy $\underline{\gamma}^2 = \{\gamma_t^2(y_{[0,t]}^2), 0 \leq t-1\}$, the decision maker can generate a randomized decision policy $\underline{\gamma}^1 = \{\gamma_t^1(y_{[0,t]}^1), 0 \leq t-1\}$ such that for all $B \in \mathcal{B}(\mathbb{U})$,

$$P(\gamma_t^1(y_{[0,t]}^1) \in B) = \int_{\tilde{y}_t^2} 1_{\{\gamma^2(\tilde{y}_{[0,t]}^2) \in B\}} P(d\tilde{y}_{[0,t]}^2 | y_{[0,t]}^1),$$

where $\tilde{y}_{[0,t]}^2$ has the same probability measure as $y_{[0,t]}^2$, generated recursively for $t \geq 0$. Hence, any policy under Q_2 can be simulated by a decision maker under Q_1 through a randomized policy. Through a dynamic programming argument and by Theorem D.1.1 (Blackwell's Irrelevant Information Theorem), for every randomized policy, there exists a deterministic policy under Q_1 which is at least as good as the randomized policy. Hence, the result follows. □

Remark 4.3.2. Blackwell's informativeness provides a partial order in the space of measurement channels; that is, not every pair of two channels can be compared. We will later see that if the goal is not the minimization of a cost function but that of stochastic stabilization of an open-loop unstable linear system over a communication channel in an appropriate sense, then one can obtain a total order on the space of channels, with the additional flexibility of allowing encoding and decoding operations. ◇

4.4 Topological Characterization of Measurement Channels

In this section, we provide topologies for measurement channels. Such constructions will be used extensively in view of existence of optimal channels as well as optimal quantizers to be considered both in this chapter and later in the book. In the following, we have $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{Y} = \mathbb{R}^m$, and $\mathcal{Q}$ denotes the set of all measurement channels (stochastic kernels) with input space $\mathbb{X}$ and output space $\mathbb{Y}$.

We refer the reader to Appendix B.2 for definitions of convergence of probability measures in the senses of weak convergence, setwise convergence, and total variation.

Definition 4.4.1 (Convergence of Channels [438]). (i) A sequence of channels $\{Q_n\}$ converges to a channel Q *weakly at input* P if $PQ_n \to PQ$ weakly.

(ii) A sequence of channels $\{Q_n\}$ converges to a channel Q *setwise at input* P if $PQ_n \to PQ$ setwise, i.e., if $PQ_n(A) \to PQ(A)$ for all Borel sets $A \subset \mathbb{X} \times \mathbb{Y}$.

(iii) A sequence of channels $\{Q_n\}$ converges to a channel Q in *total variation at input* P if $PQ_n \to PQ$ in total variation, i.e., if $\|PQ_n - PQ\|_{TV} \to 0$. ◇

If we introduce the equivalence relation $Q \equiv Q'$ if and only if $PQ = PQ'$, $Q, Q' \in \mathcal{Q}$, then the convergence notions in Definition 4.4.1 only induce the corresponding topologies on the resulting equivalence classes in $\mathcal{Q}$, instead of $\mathcal{Q}$. Since in most of the development the input distribution P is fixed, there should be no confusion when we talk about the induced topologies (resp. metrics) on $\mathcal{Q}$.

The preceding definition was explicitly dependent on the input distribution P. The next lemma provides a set of sufficient conditions which may be easier to verify, independent of the input distribution. The proof is given in the appendix.

Lemma 4.4.1 ([438]). *(i) If $\{Q_n(\,\cdot\,|x)\}$ converges to $Q(\,\cdot\,|x)$ weakly for P-a.e. x, then $PQ_n \to PQ$ weakly.*

(ii) If $\{Q_n(\,\cdot\,|x)\}$ converges to $Q(\,\cdot\,|x)$ setwise for P-a.e. x, then $PQ_n \to PQ$ setwise.

(iii) If $\{Q_n(\,\cdot\,|x)\}$ converges to $Q(\,\cdot\,|x)$ in total variation for P-a.e. x, then $PQ_n \to PQ$ in total variation. ◇

The conditions in Lemma 4.4.1 are almost universal in the choice of input probability measures; that is, the convergence characterizations will be independent of the input distributions if each of the conditions is replaced with convergence of $\{Q_n(\,\cdot\,|x)\}$ to $Q(\,\cdot\,|x)$ for all $x \in \mathbb{X}$. This is particularly useful when the input distribution is unknown or when it may change. The latter can occur in multistage stochastic control problems.

We should note at this point that total variation is a stringent requirement for convergence. For example, a sequence of discrete probability measures never converges in total variation to a probability measure which admits a density function with respect to the Lebesgue measure. On the other hand, setwise convergence induces a topology on the space of probability measures and channels which is not easy to work with. This is mainly due to the property that the space under this notion of convergence is not metrizable [161, p. 59]. However, the space of probability measures on a complete, separable, metric (Polish) space endowed with the topology of weak convergence is itself a complete, separable, metric space [58]. The Prokhorov metric, for example, can be used to metrize this space. This metric has found many applications in information theory and stochastic control. Furthermore, there are well-known conditions to determine whether a family of probability measures is weakly sequentially compact or not [58]. There are also other advantages of working with weak convergence, as we will see later in the book (see in particular Appendix B.2). Accordingly, one would like to work with weak convergence. However, as we will see, weak convergence is insufficient in a general setup for obtaining continuity. In the following, we provide some examples, taken from [438]:

1. Consider the case where the measurement channel has the form $y_t = Cx_t + v_t$, where $\{v_t\}$ is an i.i.d. noise (measurement disturbance) process. Suppose $v_t \sim f_{\theta_0}$ for some $\theta_0 \in \Theta$, where $\Theta \subset \mathbb{R}^d$ is a parameter set and $\{f_\theta : \theta \in \Theta\}$ is a parametric family of densities such that $f_{\theta_n}(v) \to f_{\theta_0}(v)$ for all $v \in \mathbb{R}^n$ and any sequence of parameters θ_n such that $\theta_n \to \theta_0$. Then by Scheffé's theorem [70] f_{θ_n} converges to f_{θ_0} in the L_1 sense, and consequently, the sequence of corresponding additive channels $Q_n(\,\cdot\,|x)$ converges to the channel $Q(\,\cdot\,|x)$ (corresponding to f_θ) in total variation for all x.
2. Consider again the measurement channel $y_t = Cx_t + v_t$ but assume this time that we only know that v_t has a density (i.e., its measure is absolutely continuous with respect to the Lebesgue measure) f (which is unknown to us). However, suppose that we are provided with a sequence of independent realizations for the noise process. Thus, with these independent observations $v_1, \ldots, v_n$ from the noise process, we can use any of the consistent nonparametric methods, e.g., [115], to obtain an estimate f_n which converges (with probability one) to f in the L_1 sense as $n \to \infty$. The corresponding sequence of estimated channels $Q_n(\,\cdot\,|x)$ converges to the true channel $Q(\,\cdot\,|x)$ in total variation for all x with probability one.
3. Now suppose that the observation channel Q is such that $Q(\,\cdot\,|x)$ admits a conditional density $f(y|x)$ for all $x \in \mathbb{R}^n$. Given observations $(x_1, y_1), \ldots, (x_n, y_n)$ drawn independently from the distribution PQ, there exists a sequence of nonparametric conditional density estimates $f_n(y|x)$ such that as $n \to \infty$

   $$\int \left(\int |f_n(y|x) - f(y|x)|\, dy \right) P(dx) \to 0,$$

   with probability one [183]. This immediately implies that the channels Q_n corresponding to these estimates converge to Q in total variation at input P.
4. Finally, assume again the additive model $y_t = Cx_t + v_t$, where now we do not have any information about the distribution μ of v_t. In this case there are no methods to consistently estimate μ in total variation from independent samples $v_1, \ldots, v_n$ [116]. However, the empirical distribution μ_n of the samples converges weakly to μ with probability one. The corresponding estimated observation channels $Q_n(\,\cdot\,|x)$ converge weakly to the true channel $Q(\,\cdot\,|x)$ for all x with probability one.

4.5 Single Stage: Continuity of the Optimal Cost in Channels

In this section, we study continuity properties under total variation, setwise convergence, and weak convergence, for the single-stage case. Thus, we investigate the continuity of the functional

$$\begin{aligned} J(P,Q) &= \inf_{\underline{\gamma}} E_P^{Q,\underline{\gamma}}\big[c(x_0,u_0)\big] \\ &= \inf_{\gamma\in\mathbf{\Gamma}} \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q(dy|x)P(dx) \end{aligned}$$

in the channel $Q \in \mathcal{Q}$, where $\mathbf{\Gamma}$ is the collection of all Borel-measurable functions mapping $\mathbb{Y}$ into $\mathbb{U}$. Note that by our previous notation, $\underline{\gamma} = \gamma$ is an admissible first-stage control policy. As before, $\mathcal{Q}$ denotes the set of all channels with input space $\mathbb{X}$ and output space $\mathbb{Y}$.

Our results in this section as well as subsequent sections in this chapter will utilize one or more of the assumptions on the cost function c and the (Borel) set $\mathbb{U} \subset \mathbb{R}^k$:

Assumption 4.5.1. *A1. The function $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ is nonnegative, bounded, and continuous on $\mathbb{X}\times\mathbb{U}$.*

A2. The function $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ is nonnegative, measurable, and bounded.

A3. The function $c : \mathbb{X}\times\mathbb{U} \to \mathbb{R}$ is nonnegative, measurable, bounded, and continuous on $\mathbb{U}$ for every $x \in \mathbb{X}$.

A4. $\mathbb{U}$ is a compact set.

A5. $\mathbb{U}$ is a convex set. ◇

Before proceeding further, we look for conditions under which an optimal control policy exists, i.e, when the infimum in $\inf_\gamma E_P^{Q,\gamma}[c(x,u)]$ is a minimum. The following result is proved in the appendix of the chapter.

Theorem 4.5.1 ([438]). *Suppose assumptions A3 and A4 hold. Then, there exists an optimal control policy for any channel Q.* ◇

The following example demonstrates that $J(P,Q)$ may not be sequentially continuous under weak convergence of channels even for continuous loss functions c and compact $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$.

Let $\mathbb{X} = \mathbb{Y} = \mathbb{U} = [0,1]$. Suppose the cost is given as $c(x,u) = (x-u)^2$ and assume that P is a discrete distribution with two atoms:

$$P = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1,$$

where δ_0 is the delta measure at point 0, that is, $\delta_0(A) = 1_{\{0\in A\}}$ for every Borel set A, where 1_E denotes the indicator function of event E. Similarly, δ_1 is the delta measure at point 1. Let $\{Q_n\}$ be a sequence of channels given by

$$Q_n(\,\cdot\,|x) = \begin{cases} \delta_{\frac{1}{n}} & \text{if} \quad x \geq \frac{1}{n}, \\ \delta_0 & \text{if} \quad x < \frac{1}{n}. \end{cases} \tag{4.4}$$

In this case, the optimal control policy, which is unique up to changes in points of measure zero, is

$$\gamma_n(y) = 1_{\{y \geq \frac{1}{n}\}}, \quad n \in \mathbb{N},$$

leading to a cost of 0. We observe that the limit of the sequence $\{Q_n(\,\cdot\,|x)\}$ is given by

$$Q(\,\cdot\,|x) = \delta_0 \quad \text{for all } x \in \mathbb{R}. \tag{4.5}$$

It can now be shown that [438] $Q_n \to Q$ weakly at input P. However, the limit of the sequence of channels cannot distinguish between the inputs, since the channel output always equals a. Thus, even though

$$J(P, Q_n) = 0, \quad \text{for all } n \geq 1,$$

the cost of $Q = \lim_n Q_n$ is

$$J(P, Q) = \frac{1}{4}$$

since, letting $(X, Y) \sim PQ$, we have $\gamma(y) = E[X|Y = y] = 1/2$ for all y.

Upper semi-continuity, however, can be established.

Theorem 4.5.2 ([438]). *Suppose assumptions A1 and A5 hold. If* $\{Q_n\}$ *is a sequence of channels converging weakly at input* P *to a channel* Q*, then*

$$\limsup_{n \to \infty} J(P, Q_n) \leq J(P, Q),$$

that is, $J(P, Q)$ *is upper semicontinuous on* $\mathcal{Q}$ *under weak convergence.* ◇

Proof. Let μ be an arbitrary probability measure on $(\mathbb{X} \times \mathbb{Y}, \mathcal{B}(\mathbb{X} \times \mathbb{Y}))$ and let $\mu_{\mathbb{Y}}$ be its second marginal, i.e., $\mu_{\mathbb{Y}}(A) = \mu(\mathbb{X} \times A)$ for $A \in \mathcal{B}(\mathbb{Y})$. Let $g \in \Gamma$ be arbitrary. By Lusin's theorem [325, Theorem 2.24], there exists a continuous function $f : \mathbb{Y} \to \mathbb{U}$ such that

$$\mu_{\mathbb{Y}}\{y : \; f(y) \neq g(y)\} < \epsilon.$$

Letting $B = \{y : \; f(y) \neq g(y)\}$ we obtain

$$\begin{aligned} \int \big|c(x, g(y)) - c(x, f(y))\big| \mu(dx, dy) &= \int_{\mathbb{X} \times B} \big|c(x, g(y)) - c(x, f(y))\big| \mu(dx, dy) \\ &< \epsilon \cdot c^*, \end{aligned}$$

where $c^* = \sup_{x,u} c(x, u) < \infty$ by assumption A1, so that

$$\int c(x, f(y)) \mu(dx, dy) < \int c(x, g(y)) \mu(dx, dy) + c^* \epsilon. \tag{4.6}$$

Let $\mathcal{C}$ be the set of continuous functions from $\mathbb{Y}$ into $\mathbb{U}$. Define

$$j(\mu,\mathcal{C}) = \inf_{\gamma\in\mathcal{C}} \int c(x,\gamma(y))\mu(dx,dy), \qquad j(\mu,\Gamma) = \inf_{\gamma\in\mathcal{G}} \int c(x,\gamma(y))\mu(dx,dy)$$

and note that $j(\mu,\mathcal{C}) \geq j(\mu,\Gamma)$ since $\mathcal{C} \subset \Gamma$. By (4.6), $j(\mu,\mathcal{C})$ is upper bounded by the right-hand side of (4.6). Since g in (4.6) was arbitrary, we obtain $j(\mu,\mathcal{C}) \leq j(\mu,\Gamma) + c^*\epsilon$, which in turn implies $j(\mu,\mathcal{C}) \leq j(\mu,\Gamma)$ since $\epsilon > 0$ was arbitrary. Hence $j(\mu,\mathcal{C}) = j(\mu,\Gamma)$.

Applying the above first to PQ_n and then to PQ, we obtain

$$\begin{aligned}
\limsup_{n\to\infty} \inf_{\gamma\in\mathcal{G}} \int c(x,\gamma(y))PQ_n(dx,dy) &= \limsup_{n\to\infty} \inf_{f\in\mathcal{C}} \int c(x,f(y))PQ_n(dx,dy) \\
&\leq \inf_{f\in\mathcal{C}} \limsup_{n\to\infty} \int c(x,f(y))PQ_n(dx,dy) \\
&= \inf_{f\in\mathcal{C}} \int c(x,f(y))PQ(dx,dy) \\
&= \inf_{\gamma\in\Gamma} \int c(x,\gamma(y))PQ(dx,dy),
\end{aligned}$$

where the next to last equality holds since PQ_n converges weakly to PQ. □

We state the following for setwise convergence.

Theorem 4.5.3 ([438]). *(i) Under assumption A2, the optimal cost*

$$J(P,Q) := \inf_{\gamma} E_P^{Q,\gamma}[c(x,u)]$$

is sequentially upper semicontinuous on the set of communication channels $\mathcal{Q}$ under setwise convergence.

(ii) $J(P,Q)$ may not be sequentially continuous under setwise convergence of channels even for continuous loss functions and compact $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$.

◇

Proof. (i) Let $\{Q_n\}$ converge setwise to Q at input P. Then

$$\begin{aligned}
\limsup_{n\to\infty} \inf_{\gamma\in\mathbf{\Gamma}} \int c(x,\gamma(y))PQ_n(dx,dy) &\leq \inf_{\gamma\in\mathbf{\Gamma}} \limsup_{n\to\infty} \int c(x,\gamma(y))PQ_n(dx,dy) \\
&= \inf_{\gamma\in\mathbf{\Gamma}} \int c(x,\gamma(y))PQ(dx,dy),
\end{aligned}$$

where the equality holds since c is bounded.

(ii) The following counterexample demonstrates that $J(P, Q)$ may not be sequentially continuous also under setwise convergence of channels even for continuous cost functions and compact $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$.

Let $\mathbb{X} = \mathbb{Y} = \mathbb{U} = [0, 1]$. Assume that X has the two-point distribution

$$P = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1.$$

Let $Q(\,\cdot\,|x) = U([0, 1])$ for all x, so that if $(x, y) \sim PQ$, then y is independent of x and has the uniform distribution on $[0, 1]$. Let $c(x, u) = (x - u)^2$.

By independence, $E[x|y] = E[x] = 1/2$, so

$$\begin{aligned} J(P, Q) &= \min_{\gamma \in \mathbf{\Gamma}} E[(x - \gamma(y))^2] = E[(x - E[x|y])^2] \\ &= \frac{1}{2}\left(1 - \frac{1}{2}\right)^2 + \frac{1}{2}\left(0 - \frac{1}{2}\right)^2 = \frac{1}{4}. \end{aligned}$$

For $n \in \mathbb{N}$ and $k = 1, \dots, n$ consider the intervals

$$L_{nk} = \left[\frac{2k-2}{2n}, \frac{2k-1}{2n}\right), \quad R_{nk} = \left[\frac{2k-1}{2n}, \frac{2k}{2n}\right) \tag{4.7}$$

and define the "square wave" function

$$h_n(t) = \sum_{k=1}^{n} \bigl(1_{\{t \in L_{nk}\}} - 1_{\{t \in R_{nk}\}}\bigr).$$

Since $\int_0^1 h_n(t)\,dt = 0$ and $|h_n(t)| \le 1$, the function $f_n(t) = \bigl(1 + h_n(t)\bigr)1_{\{t \in [0,1]\}}$ is a probability density function. Furthermore, the standard proof of the Riemann–Lebesgue lemma (e.g., [391], Theorem 12.21) can be used almost verbatim to show that

$$\lim_{n \to \infty} \int_0^1 h_n(t) g(t)\,dt = 0 \quad \text{for all } g \in L_1([0, 1], \mathbb{R})$$

and therefore

$$\lim_{n \to \infty} \int_0^1 f_n(t) g(t)\,dt = \int_0^1 g(t)\,dt \quad \text{for all } g \in L_1([0, 1], \mathbb{R}). \tag{4.8}$$

In particular, we obtain that the sequence of probability measures induced by the sequence $\{f_n\}$ converges setwise to $U([0, 1])$.

Now, for every n, introduce a channel as follows:

$$Q_n(\,\cdot\,|x) = \begin{cases} U([0,1]), & x = 0 \\ \sim f_n, & x = 1. \end{cases}$$

Then $Q_n(\cdot|x) \to Q$ setwise for $x = 0$ and $x = 1$, and thus $PQ_n \to PU([0,1])$ setwise. However, letting $(x, y_n) \sim PQ_n$, a simple calculation shows that the optimal policy for PQ_n is

$$\gamma_n(y) = E[x|y_n = y] = \begin{cases} 0, & y \in \bigcup_{k=1}^{n} R_{nk}, \\ \frac{2}{3}, & y \in \bigcup_{k=1}^{n} L_{nk}, \end{cases}$$

and therefore for every $n \in \mathbb{N}$

$$\begin{aligned} J(P, Q_n) &= \min_{\gamma \in \mathbf{\Gamma}} E[(x - \gamma(y_n))^2] \\ &= \frac{1}{2} \int_0^1 (0 - \gamma_n(y))^2 \, dy + \frac{1}{2} \int_0^1 (1 - \gamma_n(y))^2 f_n(y) \, dy \\ &= \frac{1}{6}. \end{aligned}$$

Thus, the optimal cost value is not continuous under setwise convergence. □

We have continuity under the stronger notion of total variation:

Theorem 4.5.4 ([438]). *Under assumption A2, the optimal cost $J(P, Q)$ is continuous on the set of communication channels $\mathcal{Q}$ under the topology of total variation.* ◇

Proof. Assume $Q_n \to Q$ in total variation at input P. Let $\epsilon > 0$ and pick the ϵ-optimal policies γ_n and γ under channels Q_n and Q, respectively. That is, letting $\hat{J}(Q', \gamma') = E_P^{Q', \gamma'}[c(x, u)]$ for any $\gamma' \in \mathbf{\Gamma}$ and $Q' \in \mathcal{Q}$, we have $\hat{J}(Q_n, \gamma_n) < J(P, Q_n) + \epsilon$ and $\hat{J}(Q, \gamma) < J(P, Q) + \epsilon$.

Considering first the case, $J(P, Q_n) < J(P, Q)$, we have

$$\begin{aligned} J(P, Q) - J(P, Q_n) &\leq J(P, Q) - \hat{J}(Q_n, \gamma_n) + \epsilon \\ &\leq \hat{J}(Q, \gamma_n) - \hat{J}(Q_n, \gamma_n) + \epsilon. \end{aligned}$$

By a symmetric argument it follows that

$$|J(P, Q) - J(P, Q_n)| \leq \max\bigl(\hat{J}(Q, \gamma_n) - \hat{J}(Q_n, \gamma_n), \hat{J}(Q_n, \gamma) - \hat{J}(Q, \gamma)\bigr) + \epsilon. \tag{4.9}$$

Now, since c is bounded, it follows from Appendix B.2 that for any $\gamma' \in \mathbf{\Gamma}$,

$$|\hat{J}(Q_n,\gamma')-\hat{J}(Q,\gamma')| = \left| \int c(x,\gamma'(y))PQ_n(dx,dy) - \int c(x,\gamma'(y))PQ(dx,dy) \right|$$
$$\leq \|c\|_\infty \|PQ_n - PQ\|_{TV}.$$

Together with (4.9), this implies that $|J(P,Q_n) - J(P,Q)| \leq \|c\|_\infty \|PQ_n - PQ\|_{TV} + \epsilon$. Since $\epsilon > 0$ was arbitrary, we obtain $|J(P,Q_n) - J(P,Q)| \leq \|c\|_\infty \|PQ_n - PQ\|_{TV}$, and therefore $J(P,Q_n) \to J(P,Q)$ as claimed. □

Remark 4.5.1. We can relax the boundedness condition for c by a *uniform integrability* condition below. Let $Q_n \to Q$ and let for every n, γ_n be optimal for Q_n, γ optimal for Q. If uniformly for $\{Q_n,\gamma_n\}$ and Q,γ, we have that for every $\epsilon > 0$, there exists an $L < \infty$ such that

$$\left| E_{\nu_0}^{Q^n,\gamma^n}[c(x,u)] - E_{\nu_0}^{Q^n,\gamma^n}[c(x,u)1_{(c(x,u)\leq L)}] \right| \leq \epsilon,$$

with c nonnegative and measurable, then continuity holds. See Chap. 11 (Theorem 10.6.4) for further details on such an analysis. ◇

Thus, total variation, although a strong metric, is useful in establishing continuity.

4.6 Single Stage: Existence of Optimal Channels

In this section, again for the single-stage problem, we study characterizations of compactness (or sequential compactness), which will be useful in obtaining existence results and may be useful in obtaining approximation results.

The discussion on weak convergence has already shown us that weak convergence does not induce a strong enough topology under which continuity properties can be obtained. In the following, we will obtain conditions for sequential compactness for the other two convergence notions, that is, for setwise convergence and total variation.

We first discuss setwise convergence. A set of probability measures $\mathcal{M}$ on some measurable space is said to be *setwise sequentially precompact* if every sequence in $\mathcal{M}$ has a subsequence converging setwise to a probability measure (not necessarily in $\mathcal{M}$). For two finite measures ν and μ defined on the same measurable space, we write $\nu \leq \mu$ if $\nu(A) \leq \mu(A)$ for all measurable sets A.

We have the following condition for setwise (pre)compactness (see also pp. 305–306 in [127]).

Lemma 4.6.1 ([66, Theorem 4.7.25]). *Let μ be a finite measure on a measurable space $(\mathbb{T},\mathcal{A})$. Assume a set of probability measures $\Psi \subset \mathcal{P}(\mathbb{T})$ satisfies*

$$P \leq \mu, \quad \text{for all } P \in \Psi.$$

Then Ψ is setwise precompact. ◇

As before, $PQ \in \mathcal{P}(\mathbb{X} \times \mathbb{Y})$ denotes the joint probability measure induced by input P and channel Q, where $\mathbb{X} = \mathbb{R}^n$ and $\mathbb{Y} = \mathbb{R}^m$. A simple consequence of the preceding majorization criterion is the following.

Lemma 4.6.2 ([438]). *Let ν be a finite measure on $\mathcal{B}(\mathbb{X} \times \mathbb{Y})$ and let P be a probability measure on $\mathcal{B}(\mathbb{X})$. Suppose $\mathcal{Q}$ is a set of channels such that*

$$PQ \leq \nu, \quad \text{for all } Q \in \mathcal{Q}.$$

Then $\mathcal{Q}$ is setwise sequentially precompact at input P in the sense that any sequence in $\mathcal{Q}$ has a subsequence $\{Q_n\}$ such that $Q_n \to Q$ setwise at input P for some channel Q. ◇

Proof. By Lemma 4.6.1, the set of joint measures $\mathcal{M} = \{PQ : Q \in \mathcal{Q}\}$ is setwise sequentially precompact, that is, any sequence in $\mathcal{M}$ has a subsequence $\{PQ_n\}$ converging to some $\hat{P}$ setwise. Furthermore, since the first marginal of PQ_n is P for all n, the first marginal of $\hat{P}$ is also P (since $PQ_n(A \times \mathbb{Y}) \to \hat{P}(A \times \mathbb{Y})$ for all $A \in \mathcal{B}(\mathbb{X})$). Now let Q be a regular conditional probability measure satisfying $\hat{P} = PQ$, and the result follows for the subsequential convergence of Q_n. □

For a probability density function p on $\mathbb{R}^N$, we let P_p denote the induced probability measure: $P_p(A) = \int_A p(x)\,dx$, $A \in \mathcal{B}(\mathbb{R}^N)$. The next lemma provides a sufficient condition for precompactness.

Lemma 4.6.3 ([438]). *Let μ be a finite Borel measure on $\mathbb{R}^N$ and let $\mathcal{F}$ be an equicontinuous and uniformly bounded family of probability density functions. Define $\Psi \subset \mathcal{P}(\mathbb{R}^N)$ by*

$$\Psi = \{P_p : P_p \leq \mu,\, p \in \mathcal{F}\}.$$

Then Ψ is precompact under total variation. ◇

Proof. See Sect. 4.12.3. □

The next result is an analogue of Lemma 4.6.2 and has an essentially identical proof.

Lemma 4.6.4 ([438]). *Let $\mathcal{Q}$ be a set of channels such that $\{PQ : Q \in \mathcal{Q}\}$ is a precompact set of probability measures under total variation. Then $\mathcal{Q}$ is precompact under total variation at input P.* ◇

The following theorem, when combined with the preceding results, leads to sufficient conditions for the existence of best and worst channels when the given family of channels $\mathcal{Q}$ is closed under the appropriate convergence notion.

Theorem 4.6.1 ([438]). *(i) There exists a worst channel in $\mathcal{Q}$, that is, a solution to the maximization problem*

$$\sup_{Q \in \mathcal{Q}} J(P, Q) = \sup_{Q \in \mathcal{Q}} \inf_{\gamma} E_P^{Q,\gamma}[c(x, u)],$$

when the set $\mathcal{Q}$ is weakly sequentially compact and assumptions A1, A4, and A5 hold.

(ii) *There exists a worst channel in $\mathcal{Q}$ when the set $\mathcal{Q}$ is setwise sequentially compact and assumption A2 holds.*

(iii) *There exist best and worst channels in $\mathcal{Q}$, that is, solutions to the minimization problem $\inf_{Q\in\mathcal{Q}} J(P,Q)$ and the maximization problem $\sup_{Q\in\mathcal{Q}} J(P,Q)$ when the set $\mathcal{Q}$ is compact under total variation and assumption A2 holds.* ◇

Proof. Under the stated conditions, we have sequential upper semi-continuity or continuity (Theorems 4.5.2, 4.5.3, and 4.5.4) under the corresponding topologies. By sequential compactness, the existence of the cost maximizing (worst) channel follows when $J(P,Q)$ is upper semicontinuous, while the existence of the cost minimizing (best) channel follows when $J(P,Q)$ is continuous in Q. □

Remark 4.6.1. The existence of worst channels is a useful result for the robust control or game-theoretic approach to optimization problems. If the problem is formulated as a game where the uncertainty in the set is regarded as a maximizer and the controller is the minimizer, one could search for a max-min solution, which we have just proven to exist. We note that, in information theory, problems of similar nature have been considered in the context of mutual information games [346]. ◇

4.7 Quantizers as a Class of Channels

In this section, we consider the problem of convergence and optimization of quantizers. Quantization will be an recurrent subject throughout the rest of the book.

We start with the definition of a quantizer.

Definition 4.7.1. An M-cell *vector quantizer*, Q, is a (Borel) measurable mapping from a subset of $\mathbb{X} = \mathbb{R}^n$ to the finite set $\{1,2,\ldots,M\}$, characterized by a measurable partition $\{B_1,B_2,\ldots,B_M\}$ such that $B_i = \{x : Q(x) = i\}$ for $i = 1,\ldots,M$. The B_is are called the cells (or bins) of Q. ◇

Remark 4.7.1.

(i) For later convenience we allow for the possibility that some of the cells of the quantizer are empty.

(ii) Traditionally, in source-coding theory, a quantizer is a mapping $Q : \mathbb{R}^n \to \mathbb{R}^n$ with a finite range. Thus Q is defined by a partition and a reconstruction value in $\mathbb{R}^n$ for each cell in the partition. That is, for given cells $\{B_1,\ldots,B_M\}$ and reconstruction values $\{q^1,\ldots,q^M\} \subset \mathbb{R}^n$, we have $Q(x) = q^i$ if and only if $x \in B_i$. In the definition above, we do not include the reconstruction values. ◇

A quantizer Q with cells $\{B_1, \ldots, B_M\}$ can also be characterized as a stochastic kernel Q from $\mathbb{X}$ to $\{1, \ldots, M\}$ defined by

$$Q(i|x) = 1_{\{x \in B_i\}}, \quad i = 1, \ldots, M,$$

so that $Q(x) = \sum_{i=1}^{M} q^i Q(i|x)$. We denote by $\mathcal{Q}_D(M)$ the space of all M-cell quantizers represented in the channel form. In addition, we let $\mathcal{Q}(M)$ denote the set of (Borel) stochastic kernels from $\mathbb{X}$ to $\{1, \ldots, M\}$, i.e., $Q \in \mathcal{Q}(M)$ if and only if $Q(\,\cdot\,|x)$ is a probability distribution on $\{1, \ldots, M\}$ for all $x \in \mathbb{X}$, and $Q(i|\,\cdot\,)$ is Borel measurable for all $i = 1, \ldots, M$. Note that $\mathcal{Q}_D(M) \subset \mathcal{Q}(M)$ and by our definition $\mathcal{Q}_D(M-1) \subset \mathcal{Q}_D(M)$ for all $M \geq 2$. We note that elements of $\mathcal{Q}(M)$ are sometimes referred to as random quantizers.

Lemma 4.7.1 ([438]). *The set of quantizers $\mathcal{Q}_D(M)$ is setwise sequentially precompact at any input P.* ◇

Proof. The proof follows from Lemma 4.6.2 and the interpretation above viewing a quantizer as a channel. In particular, a majorizing finite measure ν is obtained by defining $\nu = P \times \lambda$, where λ is the counting measure on $\{1, \ldots, M\}$ (note that $\nu(\mathbb{R}^n \times \{1, \ldots, M\}) = M$). Then for any measurable $B \subset \mathbb{R}^n$ and $i = 1, \ldots, M$, we have $\nu(B \times \{i\}) = P(B)\lambda(\{i\}) = P(B)$ and thus

$$PQ(B \times \{i\}) = P(B \cap B_i) \leq P(B) = \nu(B \times \{i\}).$$

Since any measurable $D \subset \mathbb{X} \times \{1, \ldots, M\}$ can be written as the disjoint union of the sets $D_i \times \{i\}$, $i = 1, \ldots, M$, with $D_i = \{x \in \mathcal{X} : (x, i) \in D\}$, the above implies $PQ(D) \leq \nu(D)$. ☐

The following lemma provides a useful result.

Lemma 4.7.2 ([438]). *A sequence $\{Q_n\}$ in $\mathcal{Q}(M)$ converges to a Q in $\mathcal{Q}(M)$ setwise at input P if and only if*

$$\int_A Q_n(i|x)P(dx) \to \int_A Q(i|x)P(dx) \quad \textit{for all } A \in \mathcal{B}(\mathbb{X}) \textit{ and } i = 1, \ldots, M.$$

◇

Proof. The lemma follows by noticing that for any $Q \in \mathcal{Q}(M)$ and measurable $D \subset \mathbb{X} \times \{1, \ldots, M\}$,

$$PQ(D) = \int_D Q(dy|x)P(dx) = \sum_{i=1}^{M} \int_{D_i} Q(i|x)P(dx),$$

where $D_i = \{x \in \mathcal{X} : (x, i) \in D\}$. ☐

The following example shows that the space of quantizers $\mathcal{Q}_D(M)$ is not closed under setwise convergence:

Let $\mathbb{X} = [0,1]$ and P the uniform distribution on $[0,1]$. Recall the definition $L_{nk} = \left[\frac{2k-2}{2n}, \frac{2k-1}{2n}\right)$ in (4.7) and let $B_{n,1} = \bigcup_{k=1}^{n} L_{nk}$ and $B_{n,2} = [0,1] \setminus B_{n,1}$. Define $\{Q_n\}$ as the sequence of 2-cell quantizers given by

$$Q_n(1|x) = 1_{\{x \in B_{n,1}\}}, \quad Q_n(2|x) = 1_{\{x \in B_{n,2}\}}.$$

This then implies that for all $A \in \mathcal{B}([0,1])$,

$$\lim_{n\to\infty} \int_A Q_n(dy|x)P(dx) = \lim_{n\to\infty} \int_0^1 \frac{1}{2} f_n(t)\,dt = \frac{1}{2}P(A),$$

and thus, by Lemma 4.7.2, Q_n converges setwise to Q given by $Q(1|x) = Q(2|x) = \frac{1}{2}$ for all $x \in [0,1]$. However, Q is not a (deterministic) quantizer.

Definition 4.7.2. The class of *finitely randomized quantizers* $\mathcal{Q}_{FR}(M)$ is the convex hull of $\mathcal{Q}_D(M)$, i.e., $Q \in \mathcal{Q}_{FR}(M)$ if and only if there exist $k \in \mathbb{N}$, $Q_1, \ldots, Q_k \in \mathcal{Q}_D(M)$, and $\alpha_1, \ldots, \alpha_k \in [0,1]$ with $\sum_{i=1}^{k} \alpha_i = 1$, such that

$$Q(i|x) = \sum_{j=1}^{k} \alpha_j Q_j(i|x), \quad \text{for all } i = 1, \ldots, M \text{ and } x \in \mathbb{X}.$$

◇

The next result says that $\mathcal{Q}_R(M)$ is the (sequential) closure of the convex hull of $\mathcal{Q}_D(M)$.

Theorem 4.7.1 ([438]). *For any $Q \in \mathcal{Q}(M)$ there exists a sequence $\{\hat{Q}_n\}$ of finitely randomized quantizers in $\mathcal{Q}_{FR}(M)$ which converges to Q setwise at any input P.* ◇

The preceding theorem has important implications in that it tells us that the space of deterministic quantizers is a "basis" for the space of communication channels between $\mathbb{X}$ and $\{1, \ldots, M\}$ in an appropriate sense. For the case when both spaces are finite, we can obtain a result reminiscent of the Birkhoff–von Neumann theorem [381], the proof of which immediately follows from the Krein–Milman theorem [50].

Theorem 4.7.2 ([438]). *Let $\mathbb{X}, \mathbb{Y}$ be finite spaces and let $\mathcal{Q}$ be the space of stochastic kernels (matrices) from $\mathbb{X}$ to $\mathbb{Y}$. Then every $Q \in \mathcal{Q}$ can be expressed as a convex combination of (deterministic) quantizers from $\mathbb{X}$ to $\mathbb{Y}$.* ◇

Remark 4.7.2 (Extreme Point Property of Quantizers). Related to Theorem 4.7.2, it is worth stating a further representation result due to Borkar [69] (see also [76]). Consider the set of probability measures

$$\Theta := \{\zeta \in P(\mathbb{R}^n \times \mathbb{M}) : \zeta = PQ, Q \in \mathcal{Q}\}, \tag{4.10}$$

on $\mathbb{R}^n \times \mathbb{M}$ having fixed input marginal P, equipped with weak topology. This set is the (Borel-measurable) set of the extreme points on the set of probability measures on $\mathbb{R}^n \times \mathbb{M}$ with a fixed input marginal P. Borel measurability of Θ follows from [307] since set of probability measures on $\mathbb{R}^n \times \mathbb{M}$ with a fixed input marginal P is a convex and compact set in a complete separable metric space, and therefore, the set of its extreme points is Borel measurable. Hence, the set of all stochastic kernels from $\mathbb{R}^n$ to $\mathbb{M}$ with fixed input marginal measure P on $\mathbb{R}^n$ is such that any element $\mathcal{K}$ in this space can be expressed in the form

$$\mathcal{K}(A) = \int \xi(dQ) PQ(A), \quad A \in \mathcal{B}(\mathbb{R}^N \times \mathbb{M})$$

for some $\xi \in \mathcal{P}(\Theta)$. ◇

In the following we again consider Euclidean spaces and show that an optimal channel can be replaced with an optimal quantizer without any loss in performance.

Proposition 4.7.1 ([438]). *For any $Q \in \mathcal{Q}(M)$, there exists a $Q' \in \mathcal{Q}_D(M)$ with $J(P, Q') \leq J(P, Q)$. If there exists an optimal channel in $\mathcal{Q}(M)$, then there is a quantizer in $\mathcal{Q}_D(M)$ that is optimal.* ◇

Proof. We need to prove only the first statement. For a policy $\gamma : \{1, \dots, M\} \to \mathbb{U} = \mathbb{X}$ (with finite cost) define for all i,

$$\bar{B}_i = \{x : c(x, \gamma(i)) \leq c(x, \gamma(j)), \quad j = 1, \dots, M\}.$$

Letting $B_1 = \bar{B}_1$ and $B_i = \bar{B}_i \setminus \bigcup_{j=1}^{i-1} B_j$, $i = 2, \dots, M$, we obtain a partition $\{B_1, \dots, B_M\}$ and a corresponding quantizer $Q' \in \mathcal{Q}_D(M)$. It is easy to see that $E_P^{Q',\gamma}[c(x, u)] \leq E_P^{Q,\gamma}[c(x, u)]$ for any $Q \in \mathcal{Q}(M)$. □

The following shows that setwise convergence of quantizers implies convergence under total variation.

Theorem 4.7.3 ([438]). *Let $\{Q_n\}$ be a sequence of quantizers in $\mathcal{Q}_D(M)$ which converges to a quantizer $Q \in \mathcal{Q}_D(M)$ setwise at P. Then, the convergence is also under total variation at P.* ◇

Proof. See Sect. 4.12.4. □

We have seen in the above discussion that, without further restriction, the convergence of quantizers may not lead to desirable continuity properties: We observed that, for example, the space of quantizers is not closed. To alleviate this aspect, in the following, we consider quantizers with convex codecells and an input distribution that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$ [185]. We note that such quantizers are commonly used in practice, and for a large class of cost functions they are known to contain optimal policies among all possible quantizers. For cost functions of the form $c(x, u) = \|x - u\|^2$ for $x, u \in \mathbb{R}^n$, the cells of optimal quantizers (if they exist) will be convex by

Lloyd–Max conditions of optimality [240]; see [185] for further results on convexity of bins for entropy-constrained quantization problems. We note that [1] also considered such cost functions for existence results on optimal quantizers; Graf and Luschgy [167] considered more general norm-based cost functions.

Now, assume $Q \in \mathcal{Q}_D(M)$ with cells $B_1, \ldots, B_M$, each of which is a convex subset of $\mathbb{R}^n$. By the separating hyperplane theorem [243], there exist pairs of complementary closed half-spaces $\{(H_{i,j}, H_{j,i}) : 1 \leq i,j \leq M, i \neq j\}$ such that for all $i = 1, \ldots, M$,

$$B_i \subset \bigcap_{j \neq i} H_{i,j}.$$

Each $\bar{B}_i := \bigcap_{j \neq i} H_{i,j}$ is a closed convex polytope and by the absolute continuity of P one has $P(\bar{B}_i \setminus B_i) = 0$ for all $i = 1, \ldots, M$. One can thus obtain a (P–a.s) representation of Q by the $M(M-1)/2$ hyperplanes $h_{i,j} = H_{i,j} \cap H_{j,i}$.

Let $\mathcal{Q}_c(M)$ denote the collection of M-cell quantizers with convex cells. Since one can represent such a hyperplane h by a vector $(a_1, \ldots, a_m, b) \in \mathbb{R}^{n+1}$ with $\sum_k |a_k|^2 = 1$ such that $h = \{x \in \mathbb{R}^n : \sum_i a_i x_i = b\}$, thus obtaining a parametrization over $\mathbb{R}^{(M-1)(n+1)}$ of all such quantizers in $\mathcal{Q}_c$.

Consider a sequence $\{Q_n\}$ in $\mathcal{Q}_c(M)$. It can be shown (see the proof of Theorem 1 in [185]) that using the above appropriate parametrization of the separating hyperplanes, a subsequence Q_{n_k} can be chosen which converges to a $Q \in \mathcal{Q}_c(M)$ in the sense that $P(B_i^{n_k} \triangle B_i) \to 0$ for all $i = 1, \ldots, M$, where the $B_i^{n_k}$ and the B_i are the cells of Q_{n_k} and Q, respectively. In view of the proof of Theorem 4.7.3, we obtain the following.

Theorem 4.7.4 ([438]). *The set $\mathcal{Q}_c(M)$ is compact under total variation at any input measure P that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$.* ◇

We can now state an existence result for optimal quantization.

Theorem 4.7.5 ([438]). *Let P be absolutely continuous and suppose the goal is to find the best quantizer Q with M cells minimizing $J(P,Q) = \inf_\gamma E_P^{Q,\gamma}[c(x,u)]$ under assumption A2, where Q is restricted to $\mathcal{Q}_c(M)$. Then an optimal quantizer exists.* ◇

Proof. The existence follows from Theorems 4.6.1 and 4.7.4. □

The above result and the topological construction will be useful and used extensively when we search for optimal dynamic quantizers later in the book, in Chap. 12.

Remark 4.7.3. In the quantization literature, finding an optimal quantizer entails finding optimal codecells and corresponding reconstruction points. Our formulation here does not require the existence of optimal reconstruction points (i.e., existence of an optimal policy γ). ◇

Remark 4.7.4 (Further Existence Results on Optimal Quantizers). It is worth noting that for the existence of an optimal quantizer and reconstruction values for $\inf_Q \inf_\gamma E_P^{Q,\gamma}[c(x,u)]$, the condition that the source admits a density can be relaxed, provided that the cost function is lower semicontinuous in u. We refer the reader to [2] for cost functions of the form $c(x,u) = c(|x-u|)$. ◇

4.8 The Multistage Case

We now consider the general stochastic control problem with T stages [as in (4.2)]. It should be noted that the effect of a control policy applied at any given time stage presents itself in two ways, in the cost incurred at the given time stage and the effect on the process distribution (and hence, the estimation error at the controller regarding the true state of the system) at future time stages. This is known as the dual effect of control [43].

The next theorem shows the continuity of the optimal cost in the measurement channel under some regularity conditions. Note that the existence of best and worst channels follows under an appropriate compactness condition as in Theorem 4.6.1(iii). We need the following definition.

Definition 4.8.1. A sequence of channels $\{Q_n\}$ converges to a channel Q *uniformly* in total variation if

$$\lim_{n\to\infty} \sup_{x\in\mathbb{X}} \left\| Q_n(\,\cdot\,|x) - Q(\,\cdot\,|x) \right\|_{TV} = 0.$$

◇

Note that in the special but important case of additive measurement channels, uniform convergence in total variation is equivalent to the weaker condition that $Q_n(\,\cdot\,|x) \to Q(\,\cdot\,|x)$ in total variation for each x. When the additive noise is absolutely continuous with respect to the Lebesgue measure, uniform convergence in total variation is equivalent to requiring that the noise density corresponding to Q_n converges in the L_1 sense to the density corresponding to Q. For example, if the noise density is estimated from n independent observations using any of the L_1 consistent density estimates described in, e.g., [115], then the resulting Q_n will converge (with probability one) uniformly in total variation.

Theorem 4.8.1 ([438]). *Consider the cost function (4.2) with arbitrary $T \in \mathbb{N}$. Suppose assumption A2 holds. Then, the optimization problem is continuous in the observation channel in the sense that if $\{Q_n\}$ is a sequence of channels converging to Q uniformly in total variation, then*

$$\lim_{n\to\infty} J(P, Q_n) = J(P, Q).$$

◇

Proof. Let $\epsilon > 0$ and pick ϵ-optimal policies $\underline{\gamma}^n = \{\gamma_0^n, \gamma_1^n, \ldots, \gamma_{T-1}^n\}$ and $\underline{\gamma} = \{\gamma_0, \gamma_1, \ldots, \gamma_{T-1}\}$ for channels Q_n and Q, respectively. That is, using the notation in (4.2), we have $J(P, Q_n, \underline{\gamma}^n) < J(P, Q_n) + \epsilon$ and $J(P, Q, \underline{\gamma}) < J(P, Q) + \epsilon$. The argument used to obtain (4.9) gives

$$\begin{aligned}&|J(P,Q) - J(P,Q_n)|\\ &\quad \leq \max\Big(J(P,Q,\underline{\gamma}^n) - J(P,Q_n,\underline{\gamma}^n), J(P,Q_n,\underline{\gamma}) - J(P,Q,\underline{\gamma})\Big) + \epsilon.\end{aligned}$$

We will next show that both terms in the maximum above converge to zero. First we consider the term

$$J(P, Q^n, \underline{\gamma}^n) - J(P, Q, \underline{\gamma}^n) = \sum_{t=0}^{T-1} E_P^{Q^n, \underline{\gamma}^n}[c(x_t, u_t)] - E_P^{Q, \underline{\gamma}^n}[c(x_t, u_t)]. \quad (4.11)$$

Under policy $\underline{\gamma}^n = \{\gamma_0^n, \gamma_1^n, \ldots, \gamma_{T-1}^n\}$, we have $u_t = \gamma_t^n(y_{[0,t]}, u_{[0,t-1]})$. We absorb in the notation the dependence of u_t on $\gamma_0^n, \ldots, \gamma_{t-1}^n$ and write $U_t = \gamma_t^n(y_{[0,t]})$.

For $t = 0, \ldots, T-1$ and $k = 0, \ldots, t$ define $\zeta_{k,t}^n : \mathbb{X}^k \times \mathbb{Y}^k \to \mathbb{R}$ by setting

$$\zeta_{t,t}^n(x_{[0,t]}, y_{[0,t]}) := c(x_t, \gamma_t^n(y_{[0,t]}))$$

and defining recursively for $k = t-1, \ldots, 0$

$$\begin{aligned}&\zeta_{k,t}^n(x_{[0,k]}, y_{[0,k]})\\ &:= \int P(dx_{k+1}|x_k, \gamma_k^n(y_{[0,k]}))Q_n(dy_{k+1}|x_{k+1})\zeta_{k+1,t}^n(x_{[0,k+1]}, y_{[0,k+1]}).\end{aligned}$$

Note that $\|\zeta_{t,t}^n\|_\infty \leq \|c\|_\infty$ and thus $\|\zeta_{k,t}^n\|_\infty \leq \|c\|_\infty$ for all $k = t-1, \ldots, 0$.

Fix $0 \leq k \leq t$ and consider a system such that the observation channel is Q at stages $0, \ldots, k-1$ and Q_n at stages $k, k+1, \ldots, t$. Let μ_k^n denote the distribution of the resulting process segment $(x_{[0,k]}, y_{[0,k]})$ under policy $\underline{\gamma}^n$ (by definition $\mu_0^n = PQ_n$). Also under policy $\underline{\gamma}^n$, let ν_k^n denote the distribution of $(x_{[0,k]}, y_{[0,k]})$ if the observation channel is Q for all the stages $0, \ldots, t$. Then we have

$$E_P^{Q^n, \underline{\gamma}^n}[c(x_t, u_t)] = \int \mu_0^n(dx_0, dy_0)\zeta_{0,t}^n(x_0, y_0)$$

and

$$E_P^{Q, \underline{\gamma}^n}[c(x_t, u_t)] = \int \nu_t^n(dx_{[0,t]}, dy_{[0,t]})\zeta_{t,t}^n(x_{[0,t]}, y_{[0,t]}).$$

Note that by construction, for all $k = 1, \ldots, t$,

$$\int \mu_k^n(dx_{[0,k]}, dy_{[0,k]})\zeta_{k,t}^n(x_{[0,k]}, y_{[0,k]})$$
$$= \int \nu_{k-1}^n(dx_{[0,k-1]}, dy_{[0,k-1]})\zeta_{k-1,t}^n(x_{[0,k-1]}, y_{[0,k-1]}).$$

Thus each term in the sum on the right-hand side of (4.11) can be expressed as a telescopic sum, which in turn can be bounded term by term, as follows:

$$\begin{aligned}\left|E_P^{Q^n,\underline{\gamma}^n}[c(x_t,u_t)]-E_P^{Q,\underline{\gamma}^n}[c(x_t,u_t)]\right| &= \left|\sum_{k=0}^{t}\int\mu_k^n(dx_{[0,k]}, dy_{[0,k]})\zeta_{k,t}^n(x_{[0,k]}, y_{[0,k]})\right.\\ &\qquad \left. -\int\nu_k^n(dx_{[0,k]}, dy_{[0,k]})\zeta_{k,t}^n(x_{[0,k]}, y_{[0,k]})\right|\\ &\le \sum_{k=1}^{t}\|\mu_k^n-\nu_k^n\|_{TV}\|\zeta_{k,t}^n\|_\infty\\ &\le \|c\|_\infty\sum_{k=1}^{t}\|\mu_k^n-\nu_k^n\|_{TV}. \end{aligned} \tag{4.12}$$

For any Borel set $B \subset \mathbb{X}^{k+1}\times\mathbb{Y}^{k+1}$, define $B(x_{[0,k]}, y_{[0,k-1]}) = \{y_k \in \mathbb{Y} : (x_{[0,k]}, y_{[0,k]}) \in B\}$, so that

$$\begin{aligned}|\mu_k^n(B)-\nu_k^n(B)| &= \left|\int \nu_{k-1}^n(dx_{[0,k-1]}, dy_{[0,k-1]})\int P(dx_k|x_{k-1},\gamma_{k-1}^n(y_{[0,k-1]}))\right.\\ &\qquad \left.\Big(Q_n(B(x_{[0,k]}, y_{[0,k-1]})|x_k) - Q(B(x_{[0,k]}, y_{[0,k-1]})|x_k)\Big)\right|\\ &\le \sup_{x_k\in\mathbb{X}}\|Q_n(\,\cdot\,|x_k)-Q(\,\cdot\,|x_k)\|_{TV}.\end{aligned}$$

The preceding bound and the uniform convergence of $\{Q_n\}$ imply $\lim_n \|\mu_k^n - \nu_k^n\|_{TV} = 0$ for all k. Combining this with (4.12) and (4.11) gives

$$J(P,Q^n,\underline{\gamma}^n) - J(P,Q,\underline{\gamma}^n) \to 0.$$

Replacing $\underline{\gamma}^n$ with $\underline{\gamma}$ we can use an identical argument to show that $J(P,Q^n,\underline{\gamma}) \to J(P,Q,\underline{\gamma})$. Since $\epsilon > 0$ in (4.11) was arbitrary, the proof is complete. □

We obtained the continuity of the optimal cost on the space of channels equipped with a more stringent notion for convergence in total variation. This result and its proof indicate that further technical complications arise in multistage problems. Likewise, upper semi-continuity under weak convergence and setwise convergence requires more stringent uniformity assumptions.

On the other hand, the concavity property applies directly to the multistage case. That is, $J(P,Q)$ is concave in the space of channels; the proof of this result follows that of Theorem 4.3.1.

Remark 4.8.1. One related approach regarding the multistage case is to consider adaptive observation channels. For example, one may aim to design optimal adaptive quantizers for a control problem. In this case, Markov Decision Process tools can be used for obtaining existence conditions for optimal channels and quantizers. This approach will be adopted in Chap. 10 and structural as well as existence results will be presented for optimal policies. ◇

4.9 Multi-agent Setting

The results for the single-agent setting apply to multi-agent setups as well, provided that the convergence notions are modified accordingly. Consider now a two-agent setup as follows.

$$\begin{aligned} x_{t+1} &= f(x_t, u_t^1, u_t^2, w_t), \\ y_t^i &= g^i(x_t, v_t^i), \quad i = 1, 2, \end{aligned}$$

where the noise variables v_t^1 and v_t^2 are independent. Suppose that g^i induces a channel Q^i for $i = 1, 2$ as described earlier and DM i has only access to y^i. Let $\underline{\gamma} = \{\underline{\gamma}^1, \underline{\gamma}^2\}$ denote the measurable policies of the agents. By the property of conditional independence, the product measure writes as $P \times Q^1 \times Q^2(A \times B^1 \times B^2) = \int_A P(dx)(\int_{B^1} Q^1(dy|x) \int_{B^2} Q^2(dy|x))$ for $A \in \mathcal{B}(\mathbb{X})$, $B^i \in \mathcal{B}(\mathbb{Y}^i), i = 1, 2$. Let us define the following cost functional for a single-stage setup:

$$\begin{aligned} J(P, Q^1, Q^2) &= \inf_{\{\underline{\gamma}^1, \underline{\gamma}^2\}} E_P^{Q^1, Q^2, \underline{\gamma}}\big[c(x, u^1, u^2)\big] \\ &= \inf_{\gamma^1, \gamma^2} \int_{\mathbb{X} \times \mathbb{Y}} c(x, \gamma^1(y^1), \gamma^2(y^2)) Q^1(dy^1|x) Q^2(dy^2|x) P(dx). \end{aligned}$$

First we discuss the comparison of information structures in the sense of Blackwell parallel to Theorem 4.3.2. See Proposition 3.2.3 for a related discussion in view of *garbling*.

It should be evident that if $Q^{2'}$ is a channel which is stochastically degraded with respect to Q^2 and $Q^{1'}$ is stochastically degraded with respect to Q^1, then

$$J(P, Q^1, Q^2) \leq J(P, Q^{1'}, Q^{2'}),$$

since under any fixed policy γ^2, Theorem 4.3.2 applies for DM 1 and likewise for DM 2. As another approach, note that any policy under $(Q^{1'}, Q^{2'})$ can be simulated by a team policy under (Q^1, Q^2) using independent randomization devices and such randomization does not improve the team performance. Furthermore, even if a common randomization device which is independent of the variables x, y^1, y^2 is provided, the result still holds true. We refer the reader to [229] for further analysis of such problems and Theorem 10.5.1 for an example where additional information provided to a team could be redundant (irrelevant) for a team decision problem.

We now discuss continuity. We first introduce the following convergence notion.

Definition 4.9.1. A sequence of channel pairs $\{Q_n^1, Q_n^2\}$ converges to a channel pair Q^1, Q^2 in total variation at P if

$$\left\| P \times Q_n^1 \times Q_n^2 - P \times Q^1 \times Q^2 \right\|_{TV} \to 0.$$

$\diamond$

We then have the following result, whose proof is a direct extension of that of Theorem 4.5.4 by considering the convergence in the product measure under total variation:

Theorem 4.9.1. *If $c(x, u^1, u^2)$ is measurable and bounded, the optimal cost $J(P, Q^1, Q^2)$ is continuous on a set of communication channels $\mathcal{Q}^1 \times \mathcal{Q}^2$ in the sense that if a sequence (Q_n^1, Q_n^2) converges to (Q^1, Q^2) in total variation at P, then*

$$\lim_{n \to \infty} J(P, Q_n^1, Q_n^2) = J(P, Q^1, Q^2).$$

$\diamond$

We note that it is possible to obtain analogous convergence results under various convergence notions for the channel pairs. Under setwise convergence, for example, an extension of upper-semi-continuity result in Theorem 4.5.3 can be obtained. Likewise under Definition 4.8.1 for the channel pairs, Theorem 4.8.1 is applicable in a multi-agent setting.

4.10 Revisiting Nonclassical Information Structures and Lack of Convexity Due to Signaling

We revisit in this section nonclassical information structures discussed in Chap. 3 and bring a different perspective to the difficulties underlying decision problems with such structures, in view of the preceding results in this chapter.

Let us consider dynamic decentralized control systems where multiple controllers with non-shared measurements are present. In such problems, the information structure is generally of nonclassical nature, and we have already seen in Chap. 3 that in such cases there may be an incentive for signaling. Under signaling, the DMs apply their actions to affect the information available at the other decision makers. In this case, the control policies induce stochastic kernels from the exogenous random variable space to the observation space of the signaled DMs. Note that this is very different from the quasi-classical case, when either (1) the signaling DM affects the information of a signaled DM, in which case the signaled decision maker already has access to all the information available to the signaling DM, or (2) it does not affect the information of another DM and the channel output is only dependent on exogenous variables. However, for the nonclassical case, the problem also features an information transmission aspect, and the signaling DM's objective also includes the design of an optimal measurement channel. The discussion in Sect. 4.3 indicates that concavity is unavoidable in such settings.

It is an implication of Theorem 4.3.1 that (as also demonstrated in Chap. 3) stochastic control problems are difficult when signaling is present. In this case, the problem becomes partly a communication problem, and as we have seen, the underlying problem is *non-convex*.

To make this important issue more explicit, let us consider the following example. Consider a two-controller system evolving in $\mathbb{R}^n$:

$$\begin{aligned} x_{t+1} &= Ax_t + B^1 u_t^1 + B^2 u_t^2 + w_t, \\ y_t^1 &= C^1 x_t + v_t^1, \\ y_t^2 &= C^2 x_t + v_t^2, \end{aligned}$$

where w, v^1, v^2 are zero-mean, i.i.d. disturbances and A, B^1, B^2, C^1, C^2 matrices of appropriate dimensions. For $\rho_1, \rho_2 > 0$, let the objective be the minimization of the cost functional

$$J = E\Bigg[\bigg(\sum_{t=0}^{T-1} |x_t|^2 + \rho_1 |u_t^1|^2 + \rho_2 |u_t^2|^2\bigg) + |x_T|^2\Bigg]$$

over control policies of the form:

$$u_t^i = \mu_t^i(y_{[0,t]}^i, u_{[0,t-1]}^i), \quad i = 1, 2; \quad t = 0, 1, \ldots, T-1.$$

As discussed earlier in Chap. 2, a static LQG team problem (i.e., the above with $T = 1$) admits an optimal solution which is linear. The proof for this result has followed the property that the team cost is convex in the joint actions of the DMs and is continuously differentiable, and as a consequence it suffices to find the unique fixed point. This, in turn, leads (under the Gaussian statistics) to linear optimal strategies for the agents.

For a multistage problem (say with $T = 2$), however, the cost is in general no longer *convex* in the action variables of the controllers acting in the first stage $t = 0$, by Corollary 4.3.1 or Theorem 4.3.1. This is because these actions might affect the estimation quality of the other controller in the future stages, if one DM can signal information to the other DM in one stage. We note that this condition is equivalent to $C^1 A^l B^2 \neq 0$ or $C^2 A^l B^1 \neq 0$ with $l + 1$ denoting the delay in signaling with $l = 0$ in the problem considered. In particular, if the controller is allowed to apply a randomized policy (e.g., by possibly using private random information that it has from the past realizations), this induces a conditional probability measure (channel) from the external variables and the initial state of the system to the observation variables at the other decision maker. The optimization problem, as such, is not jointly convex in such policies, and finding a fixed point to the stationarity conditions in the optimal policies does not necessarily lead to the conclusion that such policies are optimal. We will revisit this discussion in Sect. 11.5.

4.11 Conditions for Continuity Under Weak Convergence and Empirical Consistency

We observed in Theorem 4.5.4 that total variation is a metric which provides continuity properties. However, a careful analysis of the proof of Theorem 4.5.4 reveals that we essentially need a uniform convergence property for setwise convergence to also be sufficient for continuity. That is, we wish to have

$$\lim_{n\to\infty} \sup_{\gamma\in\mathcal{F}_s} \left| \int \left(\int Q(dy|x)c(x,\gamma(y)) - \int Q_n(dy|x)c(x,\gamma(y)) \right) P(dx) \right| = 0,$$

for a class of admissible policies $\mathcal{F}_s$ to be able to have continuity under setwise convergence. Thus, one important question of practical interest is the following: What type of stochastic control problems, cost functions, and allowable policies leads to solutions which admit such a uniform convergence principle under setwise convergence? Some partial answers to this can be found in [363].

Likewise, a parallel discussion applies to weak convergence under the assumption that for every Q_n and for Q, corresponding optimal policies γ_n and γ are continuous and are assumed to be from a restricted class of policies $\mathcal{F}_w$. One wants to have

$$\int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma_n(y))Q_n(dy|x)P(dx) \to \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q(dy|x)P(dx).$$

A sufficient condition for this is the following form of uniform weak convergence:

$$\lim_{n\to\infty} \sup_{\gamma\in\mathcal{F}_w} \left| \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q_n(dy|x)P(dx) - \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q(dy|x)P(dx) \right| = 0.$$

Two application areas of the above set of results in networked control would be learning and identification:

When one does not know the system dynamics, such as the observation channel, one typically attempts to learn the channel via test inputs or empirical observations. Let $\{(x_i,y_i),\ i\in\mathbb{N}\}$ be an $\mathbb{X}\times\mathbb{Y}$-valued i.i.d sequence generated according to some distribution μ. Defining for every measurable $B \subset \mathbb{X} \times \mathbb{Y}$ and $n \in \mathbb{N}$, the empirical occupation measures

$$\mu_n(B) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{(x_i, y_i) \in B\}},$$

one has $\mu_n(B) \to \mu(B)$ almost surely (a.s.) by the strong law of large numbers. However, it is generally not true that $\mu_n \to \mu$ setwise a.s. (e.g., μ_n never converges to μ setwise when either x_i or y_i has a nonatomic distribution), in which case μ_n cannot converge to μ in total variation. A number of related examples were considered in Sect. 4.4.

Again by the strong law of large numbers, for any μ-integrable function f on $\mathbb{X} \times \mathbb{Y}$, one has, almost surely,

$$\lim_{n\to\infty} \int f(x, y)\mu_n(dx, dy) = \int f(x, y)\mu(dx, dy).$$

In particular, $\mu_n \to \mu$ weakly with probability one [124].

In the learning theoretic context, the convergence of optimal costs under μ_n to the cost optimal for μ is called the *consistency of empirical risk minimization* (see [373] for an overview). In particular, if the cost function and the allowable control policies $\mathcal{F}_l$ are such that

$$\lim_{n\to\infty} \sup_{\gamma \in \mathcal{F}_l} \left| \int c(x, \gamma(y))\mu_n(dx, dy) - \int c(x, \gamma(y))\mu(dx, dy) \right| = 0,$$

then we arrive at consistency.

A class of measurable functions $\mathcal{E}$ is called a *Glivenko–Cantelli class* [125], if the integrals with respect to the empirical measures converge almost surely to the integrals with respect to the true measure uniformly over $\mathcal{E}$. Thus, if

$$\mathcal{G} = \{\gamma : c(x, \gamma(y)) \in \mathcal{E}\},$$

where $\mathcal{E}$ is a class of Glivenko–Cantelli family of functions, then we could establish consistency. One example of a Glivenko–Cantelli family of real functions on $\mathbb{R}^N$ is the family $\{f : \|f\|_{BL} \leq M\}$ for some $0 < M < \infty$, where $\| \cdot \|_{BL}$ denotes the bounded Lipschitz norm [125].

Thus, if we restrict the class of control policies and cost functions, we can have consistency in learning and robustness to errors in the probabilistic description of a channel.

4.12 Appendix: Proofs

4.12.1 Proof of Lemma 4.4.1

(i) Since $c(x, \cdot)$ is continuous and bounded on $\mathbb{Y}$ for all x, we have

$$\begin{aligned}\lim_{n\to\infty}\int_{\mathbb{X}\times\mathbb{Y}} c(x,y)PQ_n(dx,dy) &= \lim_{n\to\infty}\int_{\mathbb{X}}\left(\int_{\mathbb{Y}} c(x,y)Q_n(dy|x)\right)P(dx)\\ &= \int_{\mathbb{X}}\left(\int_{\mathbb{Y}} c(x,y)Q(dy|x)\right)P(dx)\\ &= \int_{\mathbb{X}\times\mathbb{Y}} c(x,y)PQ(dx,dy),\end{aligned}$$

where first we used Fubini's theorem and then the dominated convergence theorem (see Appendix A) and the fact that $\int_{\mathbb{X}} c(x,y)Q_n(dy|x)$ is bounded and converges to $\int_{\mathbb{X}} c(x,y)Q(dy|x)$ for P-a.e. x.

(ii) Let $A \in \mathcal{B}(\mathbb{X}\times\mathbb{Y})$ and for x, let $A_x = \{y : (x,y) \in A\}$. Similarly to the previous proof,

$$\begin{aligned}\lim_{n\to\infty} PQ_n(A) &= \lim_{n\to\infty}\int_{\mathbb{X}} Q_n(A_x|x)P(dx)\\ &= \int_{\mathbb{X}} Q(A_x|x)P(dx)\\ &= PQ(A)\end{aligned}$$

by the dominated convergence theorem, since $\lim_{n\to\infty} Q_n(A_x|x) = Q(A_x|x)$ for P-a.e. x.

(iii) We have

$$\begin{aligned}&\sup_{A\in\mathcal{B}(\mathbb{X}\times\mathbb{Y})} |PQ_n(A) - PQ(A)|\\ &= \sup_{A\in\mathcal{B}(\mathbb{X}\times\mathbb{Y})}\left|\int_{\mathbb{X}} Q_n(A_x|x)P(dx) - \int_{\mathbb{X}} Q(A_x|x)P(dx)\right|\\ &\le \sup_{A\in\mathcal{B}(\mathbb{X}\times\mathbb{Y})}\int_{\mathbb{X}}\bigl|Q_n(A_x|x) - Q(A_x|x)\bigr|P(dx)\\ &\le \int_{\mathbb{X}}\sup_{B\in\mathcal{B}(\mathbb{Y})}\bigl|Q_n(B|x) - Q(B|x)\bigr|P(dx).\end{aligned}$$

Since $\sup_{B \in \mathcal{B}(\mathbb{Y})} \big|Q_n(B|x) - Q(B|x)\big| \to 0$ for P-a.e. x, an application of the dominated convergence theorem completes the proof. □

4.12.2 Proof of Theorem 4.5.1

We have

$$J(P,Q) = \inf_{\gamma \in \boldsymbol{\Gamma}} \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma(y))Q(dy|y)P(dx).$$

Let $(x,y) \sim PQ$ and let $P(\,\cdot\,|y)$ be the (regular) conditional distribution of x given y. If $(PQ)_{\mathbb{Y}}$ denotes the distribution of y, then

$$\begin{aligned} J(P,Q) &= \inf_{\gamma \in \boldsymbol{\Gamma}} \int_{\mathbb{Y}} \int_{\mathbb{X}} c(x,\gamma(y))P(dx|y)(PQ)_{\mathbb{Y}}(dy) \\ &= \int_{\mathbb{Y}} \left(\inf_{u \in \mathbb{U}} \int_{\mathbb{X}} c(x,u)P(dx|y) \right) (PQ)_{\mathbb{Y}}(dy), \end{aligned}$$

where the validity of the second equality is explained below.

By assumption A3, c is bounded and $c(x,u_n) \to c(x,u)$ if $u_n \to u$ for all x; thus by the dominated convergence theorem

$$\int_{\mathbb{X}} c(x,u_n)P(dx|y) \to \int_{\mathbb{X}} c(x,u)P(dx|y)$$

proving that $g(u,y) = \int_{\mathbb{X}} c(x,u)P(dx|y)$ is continuous in u for each y. Since $\mathbb{U}$ is compact, there exists $\gamma^*(y) \in \mathbb{U}$ such that $g(\gamma^*(y),y) = \inf_{u\in\mathbb{U}} g(u,y)$. A standard argument shows that $\gamma^* : \mathbb{Y} \to \mathbb{U}$ can be taken to be measurable (see, e.g., Appendix D of [194]) and we have

$$J(P,Q) = \int_{\mathbb{X}\times\mathbb{Y}} c(x,\gamma^*(y))Q(dy|x)P(dx).$$

□

4.12.3 Proof of Lemma 4.6.3

By Lemma 4.6.1, Ψ is setwise sequentially precompact and thus any sequence in Ψ has a subsequence $\{P_n\}$ such that $P_n \to P$ setwise for some $P \in \mathcal{P}(\mathbb{R}^N)$. P is clearly absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^N$, and so it admits a density p.

Let p_n be the density of P_n. It suffices to show that

$$\lim_{n\to\infty} \|p_n - p\|_1 = 0 \tag{4.13}$$

since $\|p_n - p\|_{TV} = 2\|p_n - p\|_1 = 2\int |p_n(x) - p(x)|\,dx$.

Pick a sequence of compact sets $K_j \subset \mathbb{R}^N$ such that $K_j \subset K_{j+1}$ for all $j \in \mathbb{N}$ and $\bigcup_j K_j = \mathbb{R}^N$. Since the collection of densities $\{p_n\}$ is uniformly bounded and equicontinuous, it is precompact in the supremum norm on each K_j by the Arzelà–Ascoli theorem [124]. Thus there exist subsequences $\{p_{n_k^j}\}$ such that

$$\lim_{k\to\infty} \sup_{x\in K_j} |p_{n_k^j}(x) - p^j(x)| = 0$$

for some continuous $p^j : K_j \to [0,\infty)$.

Since the K_j are nested, one can choose $\{p_{n_k^{j+1}}\}$ to be a subsequence of $\{p_{n_k^j}\}$ for all $j \in \mathbb{N}$. Then p^{j+1} coincides with p^j on K_j and we can define $\hat{p}$ on $\mathbb{R}^N$ by setting $\hat{p}(x) = p^j(x)$, $x \in K_j$. We can now use Cantor's diagonal method to pick an increasing sequence of integers $\{m_i\}$ which is a subsequence of each $\{n_k^j\}$, and thus

$$\lim_{i\to\infty} p_{m_i}(x) = \hat{p}(x), \quad \text{for all } x \in \mathbb{R}^N. \tag{4.14}$$

Note that by construction the convergence is uniform on each K_j (and $\hat{p}$ is continuous). By uniform convergence $P_{p_{m_i}}(A) \to P_{\hat{p}}(A)$ for all Borel subsets A of K_j. The setwise convergence of P_n to P_p implies $P_{p_{m_i}}(A) \to P_p(A)$ for all Borel sets, so we must have $p = \hat{p}$ almost everywhere. This and (4.14) imply via Scheffé's theorem [59] that

$$\|p_{m_j} - p\|_1 \to 0,$$

which completes the proof. □

4.12.4 Proof of Theorem 4.7.3

Let $B_1^n, \ldots, B_M^n$ be the cells of Q_n. Since $Q_n \to Q$ setwise at input P, we have $PQ_n(B \times \{i\}) \to PQ(B \times \{i\})$ for any $B \in \mathcal{B}(\mathbb{X})$. Since $PQ_n(B \times \{i\}) = \int_B 1_{\{x\in B_i^n\}} P(dx)$, we obtain

$$P(B \cap B_i^n) \to P(B \cap B_i), \quad \text{for all } i = 1, \ldots, M.$$

If $B_1, \ldots, B_M$ are the cells of Q, the above implies $P(B_j \cap B_i^n) \to P(B_j \cap B_i)$ for all $i, j \in \{1, \ldots, M\}$. Since both $\{B_i^n\}$ and $\{B_n\}$ are partitions of $\mathbb{X}$, we obtain

$$P(B_i^n \triangle B_i) \to 0 \quad \text{for all } i = 1, \ldots, M,$$

where $B_i^n \triangle B = (B_i^n \setminus B) \cup (B \setminus B_i^n)$. Then we have

$$\begin{aligned}
&\|PQ_n - PQ\|_{TV} \\
&= \sup_{f:\|f\|_\infty \le 1} \left| \sum_{i=1}^{M} \left(\int_{\mathbb{X}} f(x,i) Q_n(i|x) P(dx) - \int_{\mathbb{X}} f(x,i) Q(i|x) P(dx) \right) \right| \\
&= \sup_{f:\|f\|_\infty \le 1} \left| \sum_{i=1}^{M} \int_{\mathbb{X}} f(x,i) \big(1_{\{x \in B_i^n\}} - 1_{\{x \in B_i\}} \big) P(dx) \right| \\
&\le \sup_{f:\|f\|_\infty \le 1} \sum_{i=1}^{M} \int_{B_i^n \triangle B_i} |f(x,i)| P(dx) \\
&\le \sum_{i=1}^{M} P(B_i^n \triangle B_i) \to 0 \qquad (4.15)
\end{aligned}$$

and convergence in total variation follows. □

4.13 Concluding Remarks

This chapter has looked at the structural and topological properties of some optimization problems in stochastic control in the space of measurement channels and quantizers. Continuity, compactness, and existence results have been established both for measurement channels as well as quantizers (viewed as a suitable subset of measurement channels).

One further main result was that the optimization problem is concave in such channels as well as on the space of information structures. This is a significant observation since the design of information structures is inherently a non-convex problem. One further implication of this result is that in a decentralized control problem, if signaling is present, the original convex problem (which may be convex under a nested, partially nested, or a stochastically nested information structure) loses its convexity.

The restriction to Euclidean state spaces is not essential and many (but not all) of the results in this chapter can be extended to the case where $\mathbb{X}$, $\mathbb{Y}$, and $\mathbb{U}$ are Polish spaces. In particular, all the results in Sects. 4.3 and 4.5 carry through without change, except Theorem 4.5.2. The results of Sect. 4.6 hold for this more general setup (however, in Lemma 4.6.3, we need the additional condition that the space is σ-compact). Likewise, most of the results in Sect. 4.7 on quantization hold more generally (in fact, Theorem 4.7.1 holds for an arbitrary measurable space), but two of the main results, Theorems 4.7.4 and 4.7.5, do require the assumption that $\mathbb{X}$ is a finite-dimensional Euclidean space.

The results here can also be applied to study the topological properties with regard to the space of the input probability measures. For example, if a sequence of priors of a decision maker regarding the state of the world converges in some sense, the implications of this convergence on the optimal costs would follow from a similar machinery as presented in the chapter.

4.14 Bibliographic Notes

On the theory of optimal quantization, [1, 173] study the effects of uncertainties in the input distribution and consider robustness in the quantizer design. Linder [235] and Pollard [309] study the consistency of optimal quantizers based on empirical data for an unknown source. Another study, [366], deals with topological properties and the existence of optimal quantizers in the context of decentralized detection. In [276], some discussions on the topology of information channels are presented. In [408], continuity and other functional properties of minimum mean-square estimation problems under Gaussian channels have been studied.

Charalambous et al. [91,321,343] and Ugrinovskii and Petersen [371] considered similar settings where the uncertain probability measure is considered for the product space in multistage control. They have considered both optimal control and estimation, and the related problem of optimal control design when the channel is unknown. In particular, [343] has studied the existence of optimal continuous estimation policies and worst-case channels under a relative entropy constraint characterizing the uncertainty in the system. In [321], the total variation norm is considered as a measure of uncertainty, and an inf-sup policy has been determined (thus, the setup considered is that of a min-max problem for the generation of optimal control policies). Similarly, there are connections with robust detection, such as those studied by Huber [203] and Poor [311], when the source distribution to be detected belongs to some set.

Blackwell's comparison of information structures or measurement channels has been expanded by Le Cam [85] by introducing an approximation interpretation. The Le Cam distance between two measurement channels Q_1, Q_2 is defined as the infimum over all randomizations at the output of Q_1 of the supremum over all input symbols x of the total variation distance between a randomization at the output of Q_1 and Q_2 at input measure δ_x. Along this spirit, Raginsky [317] relaxed the total variation metric.

Regarding the concavity properties of the optimization problems in measurement channels developed in this chapter and in [438], relevant results in the setting of quantization problems have been studied in György and Linder [184]; see Arrow [15] for a discussion in the context of team theory.

In the chapter, Theorem 4.3.2 follows from [61]. Many of the results presented in this chapter are based on results by Yüksel and Linder [435, 438].

Part II
Stabilization of Networked Control Systems

Chapter 5
Coding for Control and Connections with Information Theory

5.1 Introduction

In this chapter, we study quantizers and encoders. The notion of a quantizer was introduced formally in Sect. 4.7. We will discuss further properties of quantizers, and their performance, and bring a perspective where we view quantizers as decision variables. The chapter is also concerned with the derivation of fundamental bounds in connection with stabilizability of a linear system over a communication channel. The ideas and results presented here will be used throughout the rest of the book.

The chapter introduces the notion of real-time coding and defines the selection of a quantizer function as a decision problem in Sect. 5.2. In Sect. 5.3, a review of basic operational definitions in information theory is presented. Section 5.4 highlights the subtle differences between the performances of optimal coding for a single random variable and the limit performance of a sequence of optimal codes for blocks of random variables as the block length becomes unbounded (a common view adopted in Shannon's formulation of information theory). Performance bounds of quantizers for causal and noncausal coding of unstable processes are studied in Sect. 5.5. Fundamental lower bounds for stabilization are presented in Sect. 5.6.

5.2 Quantization and Real-Time Coding

5.2.1 Real-Time Coding

In real-time applications such as remote control of time-sensitive processes, causality in encoding and decoding is a natural limitation. As discussed earlier, there is a natural causal ordering of events in a controlled process, consisting of measurement, estimation, and actuation. All these events need to take place in real time and not

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_5,

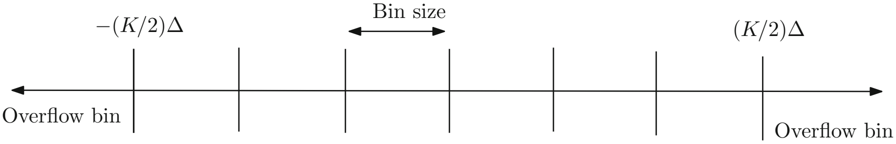

Fig. 5.1 A modified uniform quantizer. There is a single overflow bin

with significant delay. In the following, we provide a characterization of information structures in such systems and obtain fundamental bounds on transmission rates for stabilization of such systems.

Essential in communication problems is the embedding of information into a finite set, possibly with loss of information. This is done through quantization, which is a mapping from a larger alphabet to a smaller alphabet.

Before proceeding further, recall from Definition 4.7.1 that a quantizer Q is a (Borel-measurable) function from a topological space $\mathbb{X}$ to a finite index set $\mathcal{M} := \{1, 2, \ldots, M\}$. We define the bins or cells in such a quantizer as the sets

$$\mathcal{B}_i = \{x \in \mathbb{X} : Q(x) = i\}, \quad i \in \mathcal{M}.$$

Thus, a quantizer partitions its domain set. Occasionally, quantization bins are represented by *reconstruction values*. Traditionally, in source-coding theory, a quantizer is also characterized by a collection of reconstruction values in addition to a set of partitions. According to our model, this corresponds to assigning a sequence of vectors $\{q^i \in \mathbb{A}\}$ (for some set $\mathbb{A}$ which is typically $\mathbb{X}$ itself) such that

$$Q(x) = \sum_{i \in \mathcal{M}} q^i 1_{\{x \in \mathcal{B}_i\}}.$$

Thus, one could regard the above to be a composition of a quantizer and a decoding function $\mathcal{D} : \mathcal{M} \to \mathbb{A}$ defined by

$$\mathcal{D}(k) = \sum_{i \in \mathcal{M}} q^i 1_{\{i = k\}}.$$

When reconstruction values are specified *a priori*, we will explicitly include them in the definition of the quantizer. An example of a quantizer, which will be used extensively later in the book, is the following : A *modified uniform quantizer* $Q_K^{\varDelta} : \mathbb{R} \to \mathbb{R}$ with step size $\varDelta$ and $K + 1$ (with K even) number of bins satisfies the following for $k = 1, 2 \ldots, K$ (see Fig. 5.1):

$$Q_K^{\varDelta}(x) = \begin{cases} (k - \frac{1}{2}(K+1))\varDelta, & \text{if } x \in [(k - 1 - \frac{1}{2}K)\varDelta, (k - \frac{1}{2}K)\varDelta), \\ (\frac{1}{2}(K-1))\varDelta, & \text{if } x = \frac{1}{2}K\varDelta, \\ 0, & \text{if } x \notin [-\frac{1}{2}K\varDelta, \frac{1}{2}K\varDelta], \end{cases} \tag{5.1}$$

where we have $\mathcal{M} = \{1, 2, \dots, K + 1\}$. The quantizer-decoder mapping thus described corresponds to a uniform quantizer with bin size Δ. The interval $[-K/2, K/2]$ is termed the *granular region* of the quantizer, and $\mathbb{R} \setminus [-K/2, K/2]$ is named the *overflow region* of the quantizer (see Fig. 5.1). We will refer to this quantizer as a *modified uniform quantizer*, since the overflow region is assigned a single bin.

Typically, it is assumed that a quantizer is followed by an encoder. Let $\mathbb{M}$ be a finite set, typically taken as $\{0, 1\}$, that is, the binary alphabet. We refer to an encoder $\mathcal{E}$ as a mapping from $\mathcal{M}$ to $\mathbb{M}^*$, where $\mathbb{M}^*$ denotes the set of all finite-length sequences with elements in $\mathbb{M}$. The sequences in $\mathbb{M}^*$ are called *codewords*. If the lengths of the codewords corresponding to elements in $\mathcal{M}$ under an encoder $\mathcal{E}$ are all equal, the encoder is said to be a *fixed-rate* encoder. If the lengths are different, the encoder is said to be *variable rate*. Hence, we define the *fixed-rate* rate of a quantizer or an encoder by the (base-2) logarithm of the number of cells, which is $\lceil \log_2(|\mathcal{M}|) \rceil$ for the quantizer described in (5.1).

Unless explicitly stated, we will always assume that a quantizer is followed by an encoder.

5.2.2 Information Structures for Real-Time Encoders and Controllers: Policies, Actions and Measurability

This subsection considers a typical optimal causal encoding/quantization setup in a networked control system. For simplicity of the setup, we consider only two encoders, for a decentralized system, and use this system to introduce the causality and measurability constraints in quantizer design.

We begin by providing a description of the system model. Consider a partially observed Markov process, defined on a probability space, again, $(\Omega, \mathcal{F}, P)$, and described by the following discrete-time equations for $t \geq 0$:

$$x_{t+1} = f(x_t, u_t, w_t), \tag{5.2}$$

$$y_t^i = g^i(x_t, v_t^i), \tag{5.3}$$

for (Borel)-measurable functions $f, g^i, i = 1, 2$, with $\{x_0, w_t, v_t^i, i = 1, 2\}$ random variables, which are mutually independent across time and space and whose distribution functions are available at the decision makers. We further have $x_t \in \mathbb{X}$, and $y_t^i \in \mathbb{Y}^i$, $u_t \in \mathbb{U}$, where $\mathbb{X}, \mathbb{Y}^i, \mathbb{U}$ are complete, separable, metric spaces (Polish spaces) and, thus, include countable spaces or $\mathbb{R}^n$, $n \in \mathbb{N}$.

Consider a scenario where an encoder, Encoder i, is located at one end of a measurement channel characterized by (5.3), this being for $i = 1, 2$. The encoders transmit their information to a receiver (see Fig. 5.2), over a discrete noiseless channel with finite capacity, and hence, they have to quantize their information.

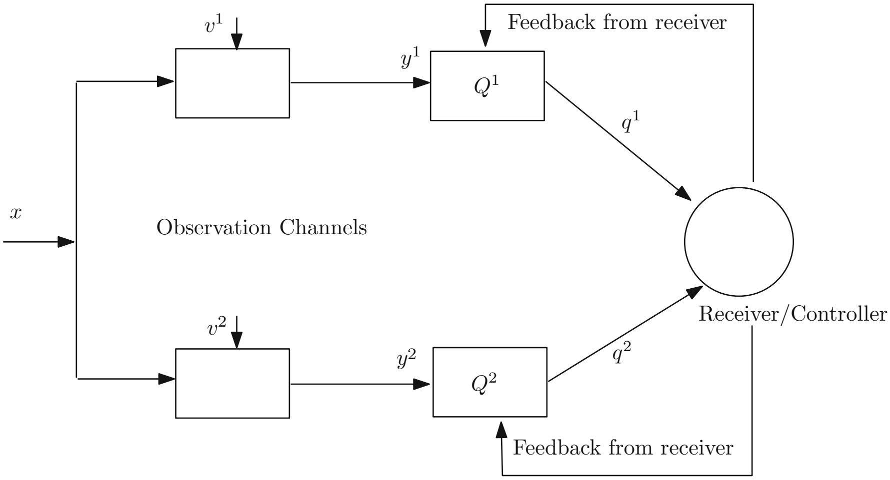

Fig. 5.2 Partially observed source under a decentralized structure

We let $\Pi^{comp,i}$ denote a *composite quantization policy* for Encoder i, defined as a sequence of functions $\{Q_t^{comp,i}, t \geq 0\}$ which are causal such that the quantization output at time t, q_t^i, under $\Pi^{comp,i}$ is generated by a function of its local information, that is, a mapping measurable with respect to the sigma-algebra generated by

$$I_t^i = \{y^i_{[0,t]}, q^i_{[0,t-1]}, z^i_{[0,t-1]}\}, \quad t \geq 1, \qquad I_0^i = \{y_0^i\},$$

to $\mathcal{M}_t^i$, where

$$\mathcal{M}_t^i := \{1, 2, \ldots, |\mathcal{M}_t^i|\},$$

for $0 \leq t \leq T-1$ and $i = 1, 2$. Here z^i denotes some additional side information available, such as feedback from the receiver.

Let $\mathbb{I}_t^i$ denote the space in which the realization I_t^i takes values in, such that $I_t^i \in \mathbb{I}_t^i$, $t \geq 0$. Hence,

$$Q_t^{comp,i} : \mathbb{I}_t^i \to \mathcal{M}_t^i.$$

Alternatively, and equivalently, we can express the policy $\Pi^{comp,i}$ as a composition of a *quantization policy* $\underline{\gamma}^i$ and a *quantizer*. A quantization policy of Encoder i, $\underline{\gamma}^i$, is a sequence of functions $\{\gamma_t^i\}$, such that for each $t \geq 0$, γ_t^i is a mapping from the information space $\mathbb{I}_t^i$ to a space of quantizers $\mathbb{Q}_t^i$ to be clarified further below. A quantizer is subsequently used to generate the quantizer output. That is, for every t and i, $\gamma_t^i(I_t^i) \in \mathbb{Q}_t^i$ and for every realization $I_t^i \in \mathbb{I}_t^i$,

$$Q_t^{comp,i}(I_t^i) = (\gamma_t^i(I_t^i))(I_t^i), \tag{5.4}$$

mapping the information space to $\mathcal{M}_t^i$ in its most general form. Even though there may appear to be duplicated information in (5.4) (since first, a map is used to pick a quantizer, and then the quantizer maps the available information to outputs), it is possible to eliminate any existing informational redundancy: A quantizer action will be generated based on the common information at the encoders and the receiver, and the quantizer will map the relevant private information at the encoder to the quantization output. Such a separation is without any loss in the space of all composite quantization policies. To establish this explicitly, let the information at the receiver at time $t \geq 0$ be $I_t^r = \{q^1_{[0,t]}, q^2_{[0,t]}\}$. Let the common information, under feedback information, at the encoders and the receiver be I_t^c. Thus, we can express any measurable composite quantization policy as

$$Q_t^{comp,i}(I_t^i) = (\gamma_t^i(I_t^c))(I_t^i \setminus I_t^c),$$

mapping the information space to $\mathcal{M}_t^i$.

We note that any composite quantization policy $Q_t^{comp,i}$ can be expressed in the form above; that is, there is no loss in the space of possible such policies, since for any $Q_t^{comp,i}$, one could define

$$\gamma_t^i(I_t^c)(\cdot) := Q_t^{comp,i}(I_t^c, \cdot).$$

Thus, we let DM^i have policy $\underline{\gamma}^i = \{\gamma_t^i\}$ and under this policy generate quantizer actions $\{Q_t^i, t \geq 0\}$, $Q_t^i \in \mathbb{Q}_t^i$ (Q_t^i is the quantizer used at time t). Under action Q_t^i, the encoder generates q_t^i, as the *quantization output* at time t.

The receiver (or the controller), upon receiving the information from the encoders, generates its decision at time t, also causally: An admissible causal receiver policy is a sequence of measurable functions $\underline{\gamma}^0 = \{\gamma_t^0\}$ such that

$$\gamma_t^0 : \prod_{s=0}^{t} \left(\mathcal{M}_s^1 \times \mathcal{M}_s^2\right) \to \mathbb{U}, \qquad t \geq 0,$$

where $\mathbb{U}$ denotes the decision set for the receiver.

In the following, we let the bold letters denote the ensemble of letters in the sense that $\mathbf{Q}_t = \{Q_t^1, Q_t^2\}$ and $\mathbf{q}_{[0,t]} = \{q^1_{[0,t]}, q^2_{[0,t]}\}$. Then, $u_t = \gamma_t^0(\mathbf{q}_{[0,t]})$ for $t \geq 0$.

Case with Noisy Channels

A more general setup that we will consider in the book is the setting which involves noisy channels. We consider first memoryless noisy channels (in the following definitions, we assume feedback is not present; minor adjustments can be made to capture the case with feedback).

Definition 5.2.1. A discrete memoryless channel (DMC) is characterized by a discrete input alphabet $\mathcal{M}$, a discrete output alphabet $\mathcal{M}'$, and a conditional probability mass function $P(q'|q)$, from $\mathcal{M} \times \mathcal{M}'$ to $[0,1]$ which satisfies the following. Let $q_{[0,n]} \in \mathcal{M}^{n+1}$ be a sequence of input symbols, let $q'_{[0,n]} \in \mathcal{M}'^{n+1}$ be a sequence of output symbols, where $q_k \in \mathcal{M}$ and $q'_k \in \mathcal{M}'$ for all k and let P^{n+1}_{DMC} denote the joint mass function on the $n+1$-tuple input and output spaces. It follows that $P^{n+1}_{DMC}(q'_{[0,n]}|q_{[0,n]}) = \prod_{k=0}^{n} P_{DMC}(q'_k|q_k)$, $\forall q_{[0,n]} \in \mathcal{M}^{n+1}$, $q'_{[0,n]} \in \mathcal{M}'^{n+1}$, where q_k, q'_k denote the kth component of the vectors $q_{[0,n]}, q'_{[0,n]}$, respectively. ◇

Definition 5.2.2. A real continuous memoryless channel (CMC) is characterized by a continuous (Borel) input alphabet $\mathcal{M} \subset \mathbb{R}^{s_1}$, a continuous (Borel) output alphabet $\mathcal{M}' \subset \mathbb{R}^{s_2}$ (where $s_1, s_2 \in \mathbb{N}$), and a regular conditional probability measure $p(A|m)$, from $\mathcal{M} \times \mathcal{B}(\mathcal{M}')$ to $\mathbb{R}$, where $\mathcal{B}(\mathcal{M}')$ is the Borel σ-algebra over $\mathcal{M}'$. Let $m_{[0,n]} \in \mathcal{M}^{n+1}$ be a sequence of input symbols, $\{m_0, m_1, \ldots, m_n\}$, and let A_0^n be a sequence of Borel subsets in $\mathcal{B}(\mathcal{M}')$, $\{A_0, A_1, \ldots, A_n\}$, where $m_k \in \mathcal{M}$ and $A_k \in \mathcal{B}(\mathcal{M}')$ for all k. Let p^{n+1}_{CMC} be the probability measure on $\mathcal{B}(\mathcal{M}^{n+1} \times \mathcal{M}'^{n+1})$. A CMC from $\mathcal{M}^{n+1}$ to $\mathcal{M}'^{n+1}$ satisfies the following for all $0 \leq k \leq n$: $p^{n+1}_{CMC}(A_k|m_{[0,k]}, A_{[0,k-1]}) = p(A_k|m_k)$, for almost all $m_{[0,k]} \in \mathcal{M}^{k+1}$, where m_k denotes the kth component of the vector $m_{[0,n]}$. ◇

Channels can also have memory. We state the following for both discrete and continuous-alphabet channels.

Definition 5.2.3. A discrete channel (continuous channel) with memory is characterized by a sequence of discrete (continuous) input alphabets $\mathcal{M}^{n+1}$, discrete (continuous) output alphabets $\mathcal{M}^{n+1}$, and a sequence of regular conditional probability measures $P_n(dq'_{[0,n]}|q_{[0,n]})$, from $\mathcal{M}^{n+1}$ to $\mathcal{M}'^{n+1}$. ◇

Typically, discrete channels admit finite alphabets. Unless stated otherwise, while considering discrete channels, we will assume channels with finite alphabets.

As before, consider a scenario where an encoder, Encoder i, is located at one end of a measurement channel characterized by (5.3), this being so for $i = 1, 2$. The encoders transmit their information to a receiver (see Fig. 5.3), over noisy channels.

In this context, we will also let $\Pi^{comp,i}$ denote a *composite encoding policy* for Encoder i, defined as a sequence of functions $\{Q_t^{comp,i}, t \geq 0\}$ which are causal such that the output at time t, q_t^i, under $\Pi^{comp,i}$ is generated by a function of its local information, that is, a mapping measurable with respect to the sigma-algebra generated by

$$I_t^i = \{y^i_{[0,t]}, q^i_{[0,t-1]}, z^i_{[0,t-1]}\}, \quad t \geq 1, \qquad I_0^i = \{y_0^i\},$$

to $\mathcal{M}_t^i$, where $\mathcal{M}_t^i := \{1, 2, \ldots, |\mathcal{M}_t^i|\}$, for $0 \leq t \leq T-1$ and $i = 1, 2$. Here z^i denotes some additional side information available, such as feedback from the receiver.

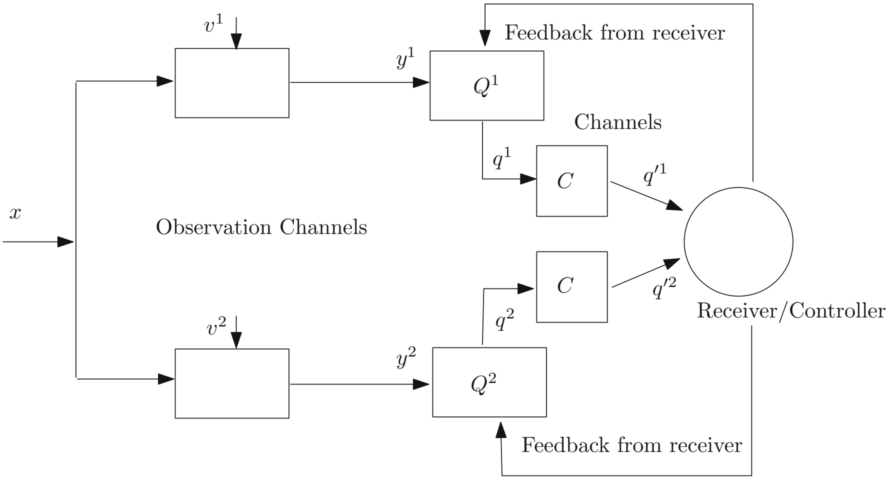

Fig. 5.3 Partially observed source under a decentralized structure with noisy channels

As before, let the information at the receiver at time $t \geq 0$ be $I_t^r = \{q'^1_{[0,t]}, q'^2_{[0,t]}\}$. Let the common information, under feedback information, at the encoders and the receiver be I_t^c. Thus, we can express any measurable composite encoding policy as

$$Q_t^{comp,i}(I_t^i) = (\gamma_t^i(I_t^c))(I_t^i \setminus I_t^c),$$

mapping the information space to $\mathcal{M}_t^i$. Thus, we let DM i have policy $\underline{\gamma}^i$ and under this policy generate quantizer actions $\{Q_t^i, t \geq 0\}$, $Q_t^i \in \mathbb{Q}_t^i$ (Q_t^i is the quantizer used at time t). Under action Q_t^i, the encoder generates q_t^i, as the *quantization output* at time t. The channel maps q_t to q'_t in a stochastic fashion as described in Definitions 5.2.1–5.2.3.

The receiver (or the controller), upon receiving the information from the channel, generates its decision at time t, also causally: An admissible causal receiver policy is a sequence of measurable functions $\underline{\gamma}^0 = \{\gamma_t^0\}$ such that $u_t = \gamma_t^0(\mathbf{q}'_{[0,t]})$ where $u_t \in \mathbb{U}$.

Design Objective

With the above formulation, one typical objective functional for the decision makers would be the following for some $T \in \mathbb{N}$:

$$\inf_{\mathbf{\Pi}^{comp}, \underline{\gamma}^0} E_{\nu_0}^{\mathbf{\Pi}^{comp}, \underline{\gamma}^0} [\sum_{t=0}^{T-1} c(x_t, u_t)],$$

over all policies $\mathbf{\Pi}^{comp}$, $\underline{\gamma}^0$ with initial distribution ν_0 on x_0. Here $c(x_t, u_t)$, is a non-negative function and $u_t = \gamma_t^0(\mathbf{q}'_{[0,t]})$ for $t \geq 0$. As an example, this setup includes, for the case when $\mathbb{X} = \mathbb{R}^n$, the quadratic cost function: $c(x,u) = |x-u|^2$.

Another objective would be to ensure that the dynamical system $\{x_t\}$ is stochastically stable in some appropriate sense.

Both of these considerations will be made precise later in the book.

5.3 Information Theoretic Preliminaries and Performance of Quantizers

In this section, we collect relevant notions and results from information theory.

5.3.1 Information Theoretic Preliminaries

We first present a number of definitions.

Definition 5.3.1. Let x be an $\mathbb{X}$-valued random variable, where $\mathbb{X}$ is a countable set. The *entropy* of x is defined as

$$H(x) = -\sum_{z \in \mathbb{X}} P(z) \log_2(P(z)) \, ,$$

where P is the probability mass function (pmf) of the random variable x. If y is another $\mathbb{X}$-valued random variable, the *conditional entropy of x given $y = y_0$* is defined by

$$H(x \mid y = y_0) = -\sum_{z \in \mathbb{X}} P(z \mid y = y_0) \log_2(P(z \mid y = y_0)) \, ,$$

where $P(\cdot \mid y = y_0)$ is the conditional pmf of x given $y = y_0$. The *conditional entropy of x given y* is defined by the average of such conditional entropies such that

$$H(x \mid y) = \sum_{y_0 \in \mathbb{X}} P(y = y_0) H(x \mid y = y_0).$$

◇

Definition 5.3.2. Let x be an $\mathbb{X}-$valued random variable, where $\mathbb{X} = \mathbb{R}^n$, and the probability measure induced by x is absolutely continuous with respect to the Lebesgue measure. The *differential entropy* of x is defined by

$$h(x) = -\int_{\mathbb{X}} p(x) \log_2(p(x)) dx \, ,$$

where $p(\cdot)$ is the probability density function (pdf) of x. If y is another random variable with density (or mass function) p_y, the *conditional differential entropy of* x *given* $y = y_0$ is defined by

$$h(x \mid y = y_0) = -\int_z p(z \mid y = y_0) \log_2(p(z \mid y = y_0))dz\,,$$

where $p(\cdot \mid y = y_0)$ is the conditional pdf of x given $y = y_0$. The *conditional differential entropy of* x *given* y is defined by the average of such conditional entropies such that

$$h(x \mid y) = \int_{y_0} p_y(y_0) h(x \mid y = y_0) dy_0.$$

◇

An important property of entropy (and differential entropy) is the chain rule: $H(x, y) = H(x) + H(y|x)$, where $H(x, y)$ is the entropy for the vector random variable (x, y). If x and y are independent, $H(y|x) = H(y)$.

A further important result is that conditioning on a random variable cannot increase entropy: $H(x|y) \leq H(x)$; conditioning on a realization, however, may increase the entropy.

Definition 5.3.3. The *mutual information* between a discrete (continuous) random variable x and another discrete (continuous) random variable y, defined on a common probability space is defined as

$$I(x; y) = H(x) - H(x|y)\,,$$

where $H(x)$ is the entropy of x (differential entropy if x is a continuous random variable), and $H(x|y)$ is the conditional entropy of x given y (conditional differential entropy if x is a continuous random variable). ◇

For more general settings, that is, for cases including when the random variables are continuous, discrete or a mixture of the two or for settings where the random variables take values in Polish spaces, mutual information is defined as

$$I(x; y) := \sup_{Q_1, Q_2} I(Q_1(x); Q_2(y)),$$

where Q_1 and Q_2 are quantizers with finitely many bins (see Chap. 5 in [171]). An important relevant result is the following. Let x be a random variable and Q be a quantizer applied to x. Then, $H(Q(x)) = I(x; Q(x)) = h(x) - h(x|Q(x))$.

Definition 5.3.4. The *Kullback-Leibler divergence* or *relative entropy* between two probability measures P_1 and P_2 such that P_1 is absolutely continuous with respect to P_2 is defined as

$$D(P_1||P_2) = \int \log_2(f(x))P_1(dx),$$

where $f(x) = \frac{dP_1}{dP_2}$ is the Radon-Nikodym derivative of P_1 with respect to P_2, that is, for all events A, $P_1(A) = \int_A f(x) P_2(dx)$. $\diamond$

Mutual information between two random variables x (with probability measure P_1) and y (with probability measure P_2) is also given by $D(P_1P_2||P_1 \times P_2)$, that is, the relative entropy between the joint measure and a measure which is the product of the measures of x and y, denoted here by $P_1 \times P_2$.

An important information theoretic property is the data-processing inequality.

Lemma 5.3.1. *Let x, y, z be three random variables which form a Markov chain in the specified order, that is, $x \leftrightarrow y \leftrightarrow z$. Then, $I(x;y) \geq I(x;z)$.* $\diamond$

An immediate consequence of Lemma 5.3.1 is that given two random variables x, y, for any measurable function pairs f, g, the following holds: $I(f(x);g(y)) \leq I(x;y)$.

An important property of differential entropy is the entropy-power inequality, given as follows.

Lemma 5.3.2. *(a) Let x, y be independent random variables and let $z = x + y$. Then,*

$$2^{2h(z)} \geq 2^{2h(x)} + 2^{2h(y)},$$

where equality holds if x, y are Gaussian. b) Let x, y be conditionally independent given u and let $z = x + y$. Then,

$$2^{2h(z|u)} \geq 2^{2h(x|u)} + 2^{2h(y|u)},$$

where $h(\cdot|u)$ denotes the conditional differential entropy. $\diamond$

5.3.2 Fixed or Variable Rates of a Quantizer/Encoder

As discussed earlier, the *fixed-rate* rate of a quantizer or encoder is defined by the (base-2) logarithm of the number of cells, which is $\lceil \log_2(|\mathcal{M}|) \rceil$ for the quantizer described in (5.1). There is also a rate definition in terms of the expected number of bits to be used; this is called the *variable-rate* rate of a quantizer.

Let x be an $\mathbb{X}-$valued random variable and let P denote the probability measure induced by x. Suppose that we wish to represent this variable in terms of binary numbers and an encoder maps $\mathbb{X}$ to a space of binary sequences. The variable-rate rate captures the average number of bits needed to represent the encoder outputs: That is, let $l(x)$ denote the length of a binary string (codeword) which is used to represent the symbol $x \in \mathbb{X}$. Then, the variable-rate rate of this encoder is given by

$$E[l(x)] = \sum_{x \in \mathbb{X}} P(x) \log_2(l(x)).$$

A very important result in source-coding theory is the following.

Proposition 5.3.1 ([103]). *Let x be an $\mathbb{X}-$valued random variable where $\mathbb{X}$ is a finite set. The variable-rate rate of any encoder which leads to a noiseless representation of the random variable satisfies*

$$H(x) \leq E[l(x)].$$

Furthermore, there exists an encoder with variable-rate rate:

$$E[l(x)] \leq H(x) + 1,$$

which leads to a noiseless representation of the source x. ◇

Proposition 5.3.1 thus represents a lower bound on the average rate of quantization outputs when a source is quantized. When a source is quantized, the quantized output can be coded with a variable-rate encoder and the entropy of the quantization output serves as a lower bound on the average length of the codewords. Note that, if an encoder simultaneously encodes a sequence of n i.i.d. random variables $x_{[0,n-1]}$, then, with increasing values of n, entropy becomes an arbitrarily close measure for the minimum average variable-rate information rate needed for noiseless representation since for every n there exists a code with average rate satisfying

$$H(x_0) = \frac{1}{n} H(x_{[0,n-1]}) \leq \frac{1}{n} E[l(x_{[0,n-1]})] \leq H(x_0) + \frac{1}{n}.$$

5.3.3 Rate-distortion Theory

Shannon's rate-distortion function is defined operationally as follows: Given a componentwise $\mathbb{X}$-valued stochastic process $\{x_t, t \in \mathbb{Z}_+\}$, a *rate-distortion* pair (R, D) is achievable if given the source process $\{x_t\}$ and a distortion function $\rho : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$, there exist sequences of

1. Quantizers (encoders): $\mathcal{E}_n : \mathbb{X}^n \to \mathcal{M}(n)$ with $|\mathcal{M}(n)| \leq 2^{Rn}$,
2. Decoders: $\mathcal{D}_n : \mathcal{M}(n) \to \mathbb{X}^n$ such that $\hat{x}_t = \mathcal{D}_n(q_{[0,n-1]})$, $0 \leq t \leq n-1$ with

$$D_n := \frac{1}{n} E[\sum_{t=0}^{n-1} \rho(x_t, \hat{x}_t)] \leq D,$$

with $\lim_{n\to\infty} D_n \leq D$.

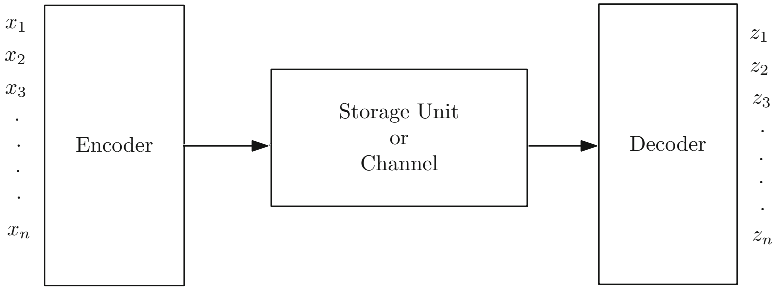

Fig. 5.4 Shannon's setup assumes block coders which are delay insensitive

The quantity

$$\inf\{R : (R, D) \quad \text{is achievable}\}$$

is the (operational) *rate-distortion* function of the source at the distortion level D. The quantity $\inf\{D : (R, D) \quad \text{is achievable}\}$ is the (operational) *distortion-rate function* at rate R.

5.3.4 Channel Coding and Shannon Capacity

Given a channel, a rate R is achievable (transmittable) if there exists a sequence (R, ϵ_n), with $\epsilon_n \to 0$ as $n \to \infty$, such that for every n, there exist

1. A set of messages $\mathcal{M}(n) := \{1, 2, 3 \ldots, M(n)\}$ such that $|\mathcal{M}(n)| = M(n) \geq 2^{Rn}$
2. A channel coder

$$\mathcal{E}_n : \mathcal{M}(n) \to \mathcal{M}^n \tag{5.5}$$

and a decoder: $\mathcal{D}_n : \mathcal{M}'^n \to \mathcal{M}(n)$, with average error probability

$$P_e := \frac{1}{|\mathcal{M}(n)|} \sum_{c \in \mathcal{M}(n)} P(\mathcal{D}_n(q'_{[0,n-1]}) \neq c | c \quad \text{is transmitted}) \leq \epsilon_n$$

Given a channel, the supremum rate R that can be achieved is called the (operational) *Shannon capacity* of the channel.

We emphasize that the above are *operational definitions* for rate distortion and capacity (Fig. 5.4). When the source and the channels belong to further special classes (e.g., those which are ergodic [376, 379]) then the operational definitions stated above also admit what is known as *single-letter mathematical characterizations*.

If the source sequence is independent and identically distributed, the *rate-distortion function*, $R_x(D)$, of a random variable, x, is the minimum (or infimum) value of the mutual information over the class of stochastic kernels from the input to the output, subject to the constraint that distortion is not higher than some level D:

$$R(D) = \inf_{P(d\hat{x}|x): E[\rho(x,\hat{x})] \leq D} I(x;\hat{x}), \tag{5.6}$$

where the infimum is over the space of all stochastic kernels $P(d\hat{x}|x)$ from $\mathbb{X}$ to $\mathbb{X}$. We also note that the inverse kernel $P(dx|\hat{x})$ (from the reconstruction space to the source alphabet space) is known as the *backward test channel* (and thus, this exhibits a duality relationship between the channel-coding problem and the source-coding problem).

A memoryless Gaussian random source constitutes an important example of the above. For a Gaussian source with variance σ_x^2, the rate-distortion function is given by $R(D) = (1/2)\log_2(\sigma_x^2/D)$ for $D \leq \sigma_x^2$.

Likewise, suppose that we have a discrete memoryless channel. The capacity of such a memoryless channel can be expressed as

$$C = \sup_{P(q)} I(q;q'), \tag{5.7}$$

where the supremum is over the set of all admissible channel input distributions $P(q)$.

Two important types of channels are the binary symmetric channel and the binary erasure channel.

A binary symmetric channel is defined by the transition probabilities: $P(q' = 0|q = 1) = P(q' = 1|q = 0) = \epsilon$. The capacity of such a channel is obtained when the input probability is such that $P(q = 0) = P(q = 1) = 1/2$, leading to the capacity value of $1 - H_b(\epsilon)$, where the binary entropy function H_b is defined as

$$H_b(x) = -x\log_2(x) - (1-x)\log_2(1-x),$$

for $x \in [0,1]$.

A binary erasure channel is defined by the transition probabilities $P(q' = 0|q = 0) = P(q' = 1|q = 1) = 1 - \epsilon$ for some $\epsilon \in [0,1]$ and the existence of an erasure symbol e such that $P(q' = e|q = 0) = P(q' = e|q = 1) = \epsilon$. The capacity of such a channel is also obtained when the input probability is such that $P(q = 0) = P(q = 1) = 1/2$, leading to the capacity value of $(1 - \epsilon)$.

Such channels and their generalizations will be revisited in Chap. 8.

Gaussian sources and channels are further important examples. In Chap. 11, these are studied in further detail.

Channels With Feedback

Communication systems may operate with feedback, that is, encoders can use the previous channel outputs to select channel inputs in a causal fashion. In this case, the encoder sequences in (5.5) are to be replaced by

$$\mathcal{E}_n = \left\{ \mathcal{E}_{n,k} : \mathcal{M}(n) \times \mathcal{M}'^{k} \to \mathcal{M}_k \right\},$$

for $0 \leq k \leq n-1$ such that $q_k = \mathcal{E}_{n,k}(c, q'_{[0,k-1]})$. Such feedback does not increase the capacity of memoryless channels, but does in general increase the capacity of channels with memory [103]. Furthermore, feedback may allow for the use of practical coding schemes with significantly less complexity and typically improves the *reliability of channels*, which is a measure of the rate of improvement in error probability as a function of the number of channel uses. The reliability of channels will be considered further in Chap. 8.

5.4 Infinite-Dimensional Coding Versus Finite-Dimensional Coding

The practical or operational applicability of the information theoretic notions (such as capacity and the rate-distortion function) generally requires processing of sequences of random variables with unbounded block length, whereas in real-time control applications (of main concern in this book) excess delay in processing cannot be tolerated. The distinction between the two is quite subtle and we attempt to make this more transparent in the following.

Toward this end, we will emphasize the differences between the settings of Shannon's rate-distortion function and the problem of *distortion-constrained entropy minimization* considered in [184] and [187], among other references.

Consider a real-valued random variable x with probability measure P. Suppose that we wish to quantize this source subject to a constraint on the rate of communication. This constraint can involve either a fixed-rate or a variable-rate measure of information rate.

Let the quantizer have bins $\{\mathcal{B}_i\}$ and let the reconstruction values be $\{\gamma(Q(x))\}$, where γ denotes a decoder function. To economize notation, we will denote the reconstruction value by $\{Q(x)\}$.

Let $J(Q) = E[\rho(x, Q(x))] = \int \rho(x, Q(x))P(dx)$, where $\rho : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ is a distortion function. The distortion under quantizer Q is given by

$$J(Q) = \sum_{i=1}^{|\mathcal{M}|} \int_{\mathcal{B}_i} \rho(x, q^i) P(dx)$$

with $q^i = Q(x)$ for $x \in \mathcal{B}_i$.

We define the distortion-constrained fixed-rate function (or *operational rate-distortion function* [176]) as

$$R_{FQ}(D) = \inf\{\lceil \log_2(|Q(\mathbb{R})|)\rceil : J(Q) \leq D\},$$

where $|Q(\mathbb{R})|$ denotes the cardinality of the quantizer index set.

The entropy under quantizer Q is (which, as discussed earlier, is a measure of the minimum average rate needed for lossless representation of quantizer outputs):

$$H(Q) := -\sum_i P(\mathcal{B}_i) \log_2(P(\mathcal{B}_i)).$$

We define the distortion-constrained entropy function as

$$R_Q(D) = \inf_{Q \in \mathcal{Q}} \{H(Q) : J(Q) \leq D\},$$

where $\mathcal{Q}$ is the space of all quantizers Q, possibly with countably infinite number of bins. For such a setting, the number of quantizer bins in an optimal quantizer can be finite or infinite [186]. It is evident that $R_{FQ}(D) \geq R_Q(D)$, since fixed-rate coding is a special case of variable-rate coding.

Note that in the expression of $R(D)$, see (5.6), the infimization is over the space of stochastic kernels, whereas in the quantization framework, it is over only the space of quantizers. Furthermore, when $\hat{x}$ admits a discrete probability measure and $P(\hat{x}|x) = 1_{\{\hat{x}=Q(x)\}}$, $I(x;\hat{x})$ reduces to $H(\hat{x})$. Thus, the essential difference is the space of optimization. It is a direct consequence of the larger space of optimization in the rate-distortion formulation that $R_Q(D) \geq R(D)$ where the inequality is strict in many of the applications in information theory. Furthermore, even though the rate-distortion function is a convex function of distortion, the distortion-constrained entropy function is not a convex function of distortion (see György and Linder [184, 187], where in the former study [184] an analysis for a uniform source has been explicitly considered). Furthermore, the set of quantizers is not convex. Note also that we have already discussed some further properties of quantizers in Sect. 4.7.

As discussed in Sect. 5.3.3, from a practical or an operational viewpoint, in the rate-distortion theoretic analysis, an increasing sequence of blocks of random variables with unbounded length is considered, and compression is performed based on these sequences. The rate-distortion theoretic performance values are obtained as the limit of the performances attained for these increasing blocks of variables. The rate-distortion function leads to an asymptotically tight lower value, since through a solution to (5.6), a test channel is constructed and the induced *backward test channel* can be used to construct an asymptotically efficient codebook, utilizing the properties of laws of large numbers: Only for quantization of vector-valued processes with ergodic behavior, in the limit of large dimensions, does one observe an equivalence between the operational rate-distortion function, distortion-constrained entropy function, and the rate-distortion function [176].

For a quantization problem, however, a single realization of a random variable is observed, and the output is generated only for the single realization.

We note that, in view of the above, even when an independent sequence is encoded, block encoding may perform strictly better than memoryless quantization. We will, however, see later in Chap. 10 that for causal (zero-delay) coding of independent sources, scalar (memoryless) coding is as good as any causal block encoder of arbitrary length.

Another important difference between finite-length and infinite-length problems is on joint source-channel coding. In the limit of large block lengths, one can design an optimal quantizer/encoder/decoder scheme (for the minimization of some average distortion criterion subject to a rate constraint) by first compressing an ergodic source using a rate-distortion achieving code (up to an arbitrarily small error) and applying a capacity-achieving channel code to map the compression outputs to channel inputs, when the channel is also ergodic. This is known as *source-channel separation*. However, such a setup is not generally optimal for finite block-length settings and the quantization and channel encoding operations need to be analyzed jointly for optimal performance. We will revisit this topic later in Chap. 11.

We end the section by noting that, although typical information theoretic approaches require long blocks, important special cases provide optimal performance even under delay-limited settings: One popular example is the problem of transmitting a Gaussian source over a Gaussian channel, the so-called *Gaussian test channel (GTC) problem*, introduced and discussed in Chap. 3. As stated there, the no-memory property of the GTC is due to the fact that the source and channel pairs are matched in the sense that capacity and rate-distortion achieving pairs are identical, leading to optimality (known as the *matching property*) . We will discuss this topic also further in Chap. 11.

5.5 Noncausal Coding for Stationary and Nonstationary Sources

In the next few chapters we will study information requirements, and coding and control algorithms for stabilization of open-loop unstable linear systems controlled over communication channels. Toward obtaining such results, in the following, we provide a partial review of the contributions in the information theory literature for such sources.

There have been important contributions in the information theory literature on noncausal coding of nonstationary/unstable sources: Consider the following Gaussian autoregressive (AR) process:

$$x_t = -\sum_{k=1}^{m} a_k x_{t-k} + w_t,$$

where $\{w_t\}$ is i.i.d. zero-mean, Gaussian random sequence with variance $E[w_t^2] = \sigma^2$. If the roots of the polynomial $H(z) = 1 + \sum_{k=1}^m a_k z^{-k}$ are all in the interior of the unit circle, then the process is asymptotically stationary and its rate-distortion function (with the distortion being the expected, normalized Euclidean error) is given parametrically by the following [169], obtained by considering the asymptotic distribution of the eigenvalues of the correlation matrix:

$$D_\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min(\theta, \frac{1}{g(w)}) dw,$$

$$R(D_\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \max(1/2(\log \frac{1}{\theta g(w)}), 0) dw,$$

with $g(w) = \frac{1}{\sigma^2} |1 + \sum_{k=1}^m a_k e^{-ikw}|^2$. If at least one root is on or outside the unit circle, the analysis is more involved as the asymptotic eigenvalue distribution contains unbounded components. Gray and Hashimoto (see [169, 192] and [174]) have shown, using the properties of the eigenvalues as well as Jensen's formula for integrations along the unit circle, that $R(D_\theta)$ above should be replaced by

$$R(D_\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \max\left(\frac{1}{2} \log(\frac{1}{\theta g(w)}), 0\right) dw + \sum_{k=1}^m \frac{1}{2} \max(0, \log(|\rho_k|^2), \quad (5.8)$$

where $\{\rho_k\}$ are the roots of the polynomial H.

Thus, an important finding in the above literature is that the logarithms of the unstable poles in such linear systems appear in the rate-distortion formulations, an issue which has also been observed in the networked control literature, which we will discuss further below. It is important to emphasize that the underlying coding schemes are noncausal, that is, the encoder has access to the entire ensemble before the encoding begins, or the coding is a sliding-block/sliding-window scheme with a finite degree of non-causality.

In contrast with information theory, due to the practical motivation of sensitivity to delay, the control theory literature has mainly considered causal/zero-delay coding for unstable (or nonstationary) sources, in the context of networked control systems. For specific settings, Wong and Brockett [406] and Baillieul [37], and for more general contexts, Tatikonda and Mitter [355] (see also [352]) and Nair and Evans [280] have obtained the minimum lower bound needed for stabilization over communication channels under various assumptions on the system noise and channels, sometimes referred to as a *data-rate theorem*. This theorem states that for stabilizability under information constraints, in the mean-square sense, a minimum average rate per time stage needed for stabilizability has to be at least the sum of the logarithms of the unstable poles/eigenvalues in the system, that is

$$R \geq \sum_{k=1}^m \frac{1}{2} \max\left(0, \log(|\rho_k|^2)\right).$$

This result could be compared with (5.8). In particular, together with the information theoretic bound (5.8), these suggest that the rate requirement is not due to causality, but due to the intrinsic evolution of the (differential) entropy in an unstable system. A more general result will be presented in the next section, in Theorem 5.6.1. Further extensions of these results will be provided in later chapters, together with a more comprehensive treatment of the literature in Chap. 8.

More precisely, in the next section, we will provide a derivation of the result above. In Chaps. 7, 8 and 9, we will show, using stochastic stabilization arguments, that the rate bound is tight for a large class of channels and investigate further criteria such as stationarity, ergodicity, and existence of finite moments. In Chap. 8, we will also obtain further converse results for stabilization over a large class of channels with memory.

5.6 Fundamental Bounds on Information Rates for Real-time Stabilization Over Noiseless Channels

As a special case of the model considered earlier in (5.2)–(5.3), we consider here a multidimensional linear system connected over a noiseless channel:

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad y_t = x_t, \tag{5.9}$$

where $x_t \in \mathbb{R}^n$ is the state at time t, u_t is the control input, and $\{w_t\}$ is a sequence of i.i.d. $\mathbb{R}^n$-valued zero-mean second-order random variables. Here A is the system matrix with at least one eigenvalue greater than 1 in magnitude, that is, the system is open-loop unstable. We assume that (A, B) is a stabilizable pair.

Let the quantizer, as described earlier, map its information to a finite set $\mathcal{M}_t$ for $t \geq 0$. We define a quantizer, a quantization policy, and a controller policy accordingly (see Fig. 5.5).

Theorem 5.6.1. *For the problem formulated above, with x_0 a second-order random variable, let $R_{avg}(T) = \frac{1}{T}\sum_{t=0}^{T-1} R_t$ be the per-stage average quantization rate at time $T-1$, where R_t is the quantization rate at time t (thus, encoding is possibly variable rate). For this setup, under any stabilizing control and causal quantization policy, for either*

$$\limsup_{t\to\infty} \frac{1}{t} \log(E[|x_t - E[x_t|q_{[0,t]}]|^2]) \leq 0$$

or

$$\limsup_{t\to\infty} \frac{1}{t} \log(E[|x_t|^2]) \leq 0$$

to hold, the per-stage average quantization rate must satisfy the inequality

$$\liminf_{T\to\infty} R_{avg}(T) \geq \sum_{|\lambda_i|>1} \log_2(|\lambda_i|),$$

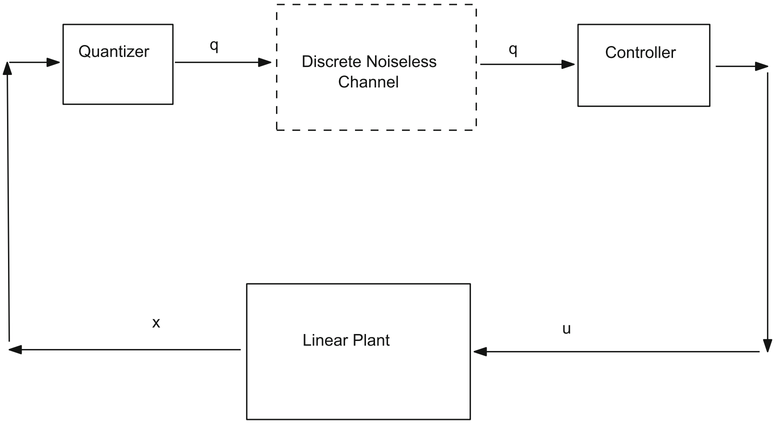

Fig. 5.5 Control over a finite-rate noiseless channel with quantized observations at the controller

where $\{\lambda_i, 1 \leq i \leq n\}$ *are the eigenvalues of A, with each repeated eigenvalue (if any) taken as a distinct element in the sequence.* ◇

Proof. See Sect. 5.7. ☐

Remark 5.6.1. A refinement of the theorem above will be discussed in the context of stabilization over a general class of noisy communication channels in Chap. 8 (see Theorem 8.5.2). ◇

A further general result is the following.

Theorem 5.6.2. *With* x_0 *a second-order random variable, let* $\lim_{T\to\infty} R_{avg}(T) = C$. *For the controlled state process* $\{x_t\}$ *to be asymptotically mean stationary (AMS) (see Definition C.3.5), it must be that* $C \geq \sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$. ◇

The proof of this theorem is almost identical to that of Theorem 8.2.2 (which is presented in Chap. 8) and is therefore omitted.

5.7 Appendix: Proof of Theorem 5.6.1

The proof follows from [422] and [427]. The matrix A can always be *block-diagonalized* with two blocks, where the first block has only stable eigenvalues and the second one unstable eigenvalues. For the stable modes, one need not use the channel, and hence for the remaining discussion and analysis we can assume, without any loss of generality, that A has only unstable eigenvalues.

Let the expected square of the Euclidean norm of the random vector x_t be denoted by D_t and the covariance matrix of the same be denoted by C_t. Thus, $D_t = \text{trace}(C_t)$. We also note that among random vectors with a fixed covariance matrix, the differential entropy is maximized by a jointly Gaussian distribution, which in turn has a finite entropy [103]. Since entropy serves as a lower bound on the average rate (see Proposition 5.3.1), the average per-stage quantization rate satisfies the following for $T \in \mathbb{N}$:

$$
\begin{aligned}
& TR_{avg}(T) \geq H(q_{[0,T-1]}) = \sum_{t=1}^{T-1} H(q_t|q_{[0,t-1]}) + H(q_0) \\
\geq & \sum_{t=1}^{T-1} \Big(H(q_t|q_{[0,t-1]}) - H(q_t|x_t, q_{[0,t-1]}) \Big) + H(q_0) \\
= & \sum_{t=1}^{T-1} I(x_t; q_t|q_{[0,t-1]}) + H(q_0) \\
= & \sum_{t=1}^{T-1} \Big(h(x_t|q_{[0,t-1]}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
= & \sum_{t=1}^{T-1} \Big(h(Ax_{t-1} + w_{t-1} + Bu_{t-1}|q_{[0,t-1]}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
= & \sum_{t=1}^{T-1} \Big(h(Ax_{t-1} + w_{t-1}|q_{[0,t-1]}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
\geq & \sum_{t=1}^{T-1} \Big(h(Ax_{t-1} + w_{t-1}|q_{[0,t-1]}, w_{t-1}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
= & \sum_{t=1}^{T-1} \Big(h(Ax_{t-1}|q_{[0,t-1]}, w_{t-1}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
= & \sum_{t=1}^{T-1} \Big(h(Ax_{t-1}|q_{[0,t-1]}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \qquad (5.10) \\
= & \sum_{t=1}^{T-1} \Big(\log_2(|\det(A)|) + h(x_{t-1}|q_{[0,t-1]}) - h(x_t|q_{[0,t]}) \Big) + H(q_0) \\
= & \Big(\sum_{t=1}^{T-1} \log_2(|\det(A)|) \Big) + h(x_0|q_0) - h(x_{T-1}|q_{[0,T-1]}) + H(q_0) \\
\geq & \Big(\sum_{t=1}^{T-1} \log_2(|\det(A)|) \Big) + h(x_0|q_0) - h(x_{T-1}) + H(q_0) \qquad (5.11)
\end{aligned}
$$

$$\geq \left(\sum_{t=1}^{T-1} \log_2(|\det(A)|) \right) + h(x_0|q_0) - \frac{1}{2} \log \left((2\pi e)^n \det(C_{T-1}) \right) + H(q_0)$$

$$\geq \left(\sum_{t=1}^{T-1} \log_2(|\det(A)|) \right) + h(x_0|q_0) - \frac{1}{2} \log \left((2\pi e)^n (\frac{1}{n} D_{T-1})^n \right) + H(q_0). \tag{5.12}$$

The first inequality follows since discrete entropy is always nonnegative, and the second and third inequalities follow from the fact that conditioning does not increase entropy (see [103]). The fourth inequality follows from the property that the Gaussian measure maximizes the entropy for a given covariance matrix. Equation (5.10) follows from the assumption that $\{w_t\}$ is an i.i.d. process, and (5.11) uses the fact that conditioning does not increase entropy (this last step could be skipped if the goal was to only show the subexponential growth of the estimation error in the sense that $\limsup_{t\to\infty} \frac{1}{t} \log(E[|x_t - E[x_t|q_{[0,t]}]|^2]) \leq 0$). In (5.12) we use $\det(C_{T-1}) \leq (\frac{1}{n}\text{trace}(C_{T-1}))^n$ (which is a consequence of *the inequality of arithmetic and geometric means* applied to eigenvalues). Since $\limsup_{t\to\infty} \frac{1}{t} \log(D_t) \leq 0$ and $h(x_0) < \infty$ we obtain

$$\liminf_{T\to\infty} R_{avg}(T) \geq \sum_{|\lambda_i|>1} \log_2(|\lambda_i|).$$

□

5.8 Concluding Remarks

In this chapter, information structures for encoders and controllers have been introduced, with the setup presented in this chapter to be used throughout the book. In the following chapters of *Part II* of the book, stochastic stabilization under such measurability and information rate constraints will be covered. In *Part III* of the book, optimization of coding and control policies in networked control systems will be studied in detail, under such information constraints.

In the chapter, the connections between real-time coding and information theoretic approaches have been investigated and subtle differences have been highlighted. Further, properties of quantizers and their performance have been studied. The chapter has also been concerned with the derivation of fundamental bounds in connection with stabilizability of a linear system over a communication channel. This problem will be revisited in further generality in Chap. 8.

5.9 Bibliographic Notes

A comprehensive tutorial on quantization and source-coding theory can be found in [176].

Causal coding has been studied in the literature in different contexts and under different assumptions on the classes of sources and encoder types. Some of the relevant results on the topic are the following: When the elements of a sequence $\{x_i\}$, living in a discrete alphabet, are independent and identically distributed, and if the decoder is allowed to perform with delay, then an optimal memoryless coder (possibly randomized) is the optimal causal encoder minimizing the data rate subject to a distortion constraint, which is a result due to Neuhoff and Gilbert [292]. We note that [292] allows for delay at the decoder output to facilitate entropy coding, but not at the encoder/decoder symbol generation. A more restrictive causal coding scheme is what is known as the zero-delay (delayless) coding scheme which does not allow delay in symbol as well as code generation. That is, the zero-delay property is stronger than the causality notion adopted in [292] in that both encoding and decoding are instantaneous.The zero-delay coding scheme was studied in [148], which demonstrated the optimality of memoryless fixed-rate encoders when the source takes values in a finite set, and the encoding is fixed rate. If the source is kth-order Markov, then the optimal fixed-rate, zero-delay coder minimizing any measurable, additive (per-stage) distortion uses only the last k source symbols and the current state at the receiver's memory [396]. The results of [396] were extended in [361] and [250] to systems with noisy feedback, under the assumption of a fixed decoder structure with finite memory. Zero-delay coding for partially observed Markov sources was considered in [425]. The problem of optimal transmission over noisy channels with perfect causal feedback was considered in [385] for the case when the source belongs to a finite and discrete set. In the limit of low distortion (high rate), [237] studied stationary encoding of a stable stationary process and showed that memoryless quantizer followed by a conditional entropy coder is at most 0.25 bits worse than any causal encoder. A relevant problem in causal rate-distortion theory was studied in [352], where under the criterion of *directed mutual information* (see [259]) minimization subject to a distortion constraint and with availability of feedback, optimal causal conditional coding laws were obtained. Further aspects on and differences between real-time coding and standard information theoretic formulations have been described in detail in [359].

Matching of sources and channels as dual problems of capacity of a given channel and the rate-distortion function of a given source has been discussed in [107] and [156]. On a parallel note, the multiterminal source-coding theorems [103], although insightful, are not always applicable in a real-time setting, as the asymptotic partitioning arguments in classical information theory [103] do not apply. In a control context, however, one method to achieve the information theoretic bounds is via binning; see [429, 431] for discussions on binning in a decentralized control context and [312] for a discussion on binning in a general communications context.

In the information theory literature, stochastic stability results were established mostly for stationary sources, which are already in some appropriate sense stable sources. In this literature, the stability of the estimation errors as well as the encoder state processes have been studied. These systems mainly involve causal and noncausal coding (block coding as well as sliding-block coding) of stationary sources [166, 214, 215] and asymptotically mean stationary sources [141]. Real-time settings such as sigma-delta quantization schemes have also been considered in the literature; see, for example, [403], among others. We refer the reader to a review in [174] regarding rate-distortion results for such nonstationary processes and on the methods used in [169] and [192]. Berger [51] obtained the rate-distortion function for Wiener processes and in addition, developed a two-part coding scheme, which was later generalized in [329] and [331] to unstable Markov processes driven by bounded noise. The scheme in [51] exploits the independent increment property of Wiener processes. References [420], [419] obtained ergodicity results for open-loop unstable systems controlled over noisy or noiseless channels.

Uniform scalar quantizers, in addition to their simplicity, are also approximately optimal in the limit of high-rate distortion [163] with respect to the mean-square distortion measure [427]. Gish and Pierce [163] has shown that the uniform quantizer followed by an entropy coder is at most 0.255 bits worse than any optimal quantizer, if one exists, in the context of entropy-constrained causal quantization, a situation which is also applicable in linear control systems. Some other selected relevant papers are [336] and [442].

The fundamental lower bound in Theorem 5.6.1 is a generalization of the results established by Nair and Evans [280] and Tatikonda and Mitter [356], building on [422] and [427]. Part of this chapter follows from [421] and [427].

Chapter 6
Stochastic Stability and Drift Criteria for Markov Chains in Networked Control

6.1 Introduction and Motivation: Why Stochastic Drift Criteria?

One essential aspect of a networked control system is the presence of randomness and uncertainty in measurement information or action transmissions. The evolution of the dynamics is generally event driven and this does not necessarily admit a time-homogeneous dynamics based on a fixed clock or a deterministic sequence. To understand such random event-based updates, it is necessary to develop the requisite mathematical tools to study such processes.

We will develop in this chapter an important mathematical program to study such dynamics; we will refer to it as *random-time state-dependent drift criteria for stabilization*.

We will see later in Chaps. 7–9 that a random-time state-dependent drift-based program is very effective in arriving at stochastic stability results. Furthermore, as we will see in Chap. 12, many decentralized control problems use algorithms which allow for agreement on certain variables of nature at random times, based on which decisions can be generated. Such time characterizations are essential for optimization and stabilization studies of networked and decentralized control systems. Stochastic stabilization and the applicability of ergodic theorems are also important for the development of infinite horizon dynamic optimization algorithms.

This present chapter provides the key mathematical tools needed in the remaining chapters of Part II for stabilization over noiseless as well as noisy channels. We are presenting these results here in this chapter instead of relegating them to an appendix, due to their importance and the prominence (in particular of Theorem 6.2.4) in such applications. The reader could skip or just skim through this chapter if not particularly interested in the mathematical details concerning Markov chains and the martingale arguments which will be applied in the chapters to follow, to establish stochastic stabilization.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_6,

In a nutshell, this chapter presents the drift criteria for stochastic stabilization of Markov chains. The reader is referred to Appendix C for an overview of Markov chains and stochastic stabilization, as well as the comprehensive book by Meyn and Tweedie [271].

6.2 Stochastic Stability and Drift Criteria

Let $\boldsymbol{X} = \{x_t, t \geq 0\}$ denote a Markov chain with state space $\mathbb{X}$. Assume that the state space is a complete, separable, metric space, whose Borel σ-field is denoted $\mathcal{B}(\mathbb{X})$. Let the transition probability be denoted by P, so that for any $x \in \mathbb{X}$, $A \in \mathcal{B}(\mathbb{X})$, the probability of moving from x to A in one step is given by $P(x_{t+1} \in A \mid x_t = x) = P(x, A)$. The n-step transitions are obtained via composition in the usual way, $P(x_{t+n} \in A \mid x_t = x) = P^n(x, A)$, for any $n \geq 1$. The transition law acts on measurable functions $f\colon \mathbb{X} \to \mathbb{R}$ and measures μ on $\mathcal{B}(\mathbb{X})$ via

$$Pf\,(x) := \int_{\mathbb{X}} P(x, dy) f(y), \quad x \in \mathbb{X}, \quad \mu P\,(A) := \int_{\mathbb{X}} \mu(dx) P(x, A), \quad A \in \mathcal{B}(\mathbb{X}).$$

A probability measure π on $\mathcal{B}(\mathbb{X})$ is called invariant if $\pi P = \pi$. That is,

$$\int \pi(dx) P(x, A) = \pi(A), \quad A \in \mathcal{B}(\mathbb{X}).$$

The existence of an invariant probability measure is very important since such a measure represents the asymptotic behavior of the Markov chain. Typically, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k(x, A)$ converges to $\pi(A)$ for some invariant measure π, which may depend on x. If the Markov chain is positive Harris recurrent (see Sect. C.1), there exists a unique invariant probability measure π and furthermore, for every x, $\lim_{n\to\infty} P^n(x, A) = \pi(A)$.

One sufficient, and very general, characterization for the existence of an invariant probability measure is through the use of Lyapunov functions and the drift criteria, which we discuss in the following subsections.

For any initial probability measure ν on $\mathcal{B}(\mathbb{X})$ we can construct a stochastic process with transition law P and satisfying $x_0 \sim \nu$. We let P_ν denote the resulting probability measure on the sample space, with the usual convention $\nu = \delta_x$ when the initial state is $x \in \mathbb{X}$. When $\nu = \pi$, then the resulting process is stationary.

6.2.1 One-stage Foster–Lyapunov Drift Criteria

An increasing family $\{\mathcal{F}_n\}$ of sub-σ-fields of $\mathcal{F}$ is called a filtration (see Sect. C.2). In the following, let $\mathcal{F}_t$ be such a filtration generated by the state sequence, that is, $\mathcal{F}_t = \sigma(x_m, m \leq t), t \geq 0$.

The following results are known as Foster–Lyapunov drift criteria [271].

Theorem 6.2.1 ([271]). *Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain. Suppose moreover that there are functions $V\colon \mathbb{X} \to (0, \infty)$, $\epsilon > 0$, a petite set C, and a constant $b \in \mathbb{R}$, such that the following holds:*

$$E[V(x_{t+1}) \mid \mathcal{F}_t] \leq V(x_t) - \epsilon + b1_{\{x_t \in C\}}. \tag{6.1}$$

Then $\boldsymbol{X}$ is positive Harris recurrent. ◇

Theorem 6.2.2 ([271]). *Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain. Suppose moreover that there are functions $V\colon \mathbb{X} \to (0, \infty)$, $f\colon \mathbb{X} \to [1, \infty)$, a petite set C, and a constant $b \in \mathbb{R}$, such that the following holds:*

$$E[V(x_{t+1}) \mid \mathcal{F}_t] \leq V(x_t) - f(x_t) + b1_{\{x_t \in C\}}. \tag{6.2}$$

Then $\boldsymbol{X}$ is positive Harris recurrent, and moreover $\pi(f) := E_\pi(f(x)) < \infty$, with π being the invariant distribution. ◇

6.2.2 State-dependent Drift Criteria

The following establishes a set of sufficient conditions for positive Harris recurrence under deterministic but state-dependent drift.

Theorem 6.2.3 ([271]). *Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain. Let $V(.)\colon \mathbb{X} \to \mathbb{R}_+$ be a positive-valued functional and $b \in \mathbb{R}_+$. Consider the following inequality for some $\epsilon \geq 0$ and some function $n(\cdot) : \mathbb{X} \to \mathbb{Z}_+$:*

$$E[V(x_{t+n(x_t)}) \mid \mathcal{F}_t] \leq (1-\epsilon)^{(n(x_t))} V(x_t) - n(x_t) + b1_{\{x_t \in C\}},$$

$\forall x \in \mathbb{X}$.

If the inequality holds with $\epsilon = 0$, then a finite invariant measure exists for the Markov chain. If $\epsilon > 0$, then the chain is exponentially ergodic, that is, there exist $r > 1$ and a positive function $M(x)$ such that

$$\lim_{n \to \infty} M(x) r^n ||P^n(x, \cdot) - \pi||_{TV} \to 0.$$

◇

We now partially generalize the above in the following to random-time state-dependent stochastic drift criteria.

6.2.3 Random-time State-dependent Stochastic Drift Criteria

Throughout this subsection, the sequence of stopping times $\{\tau_i : i \in \mathbb{N}_+\}$ is assumed to be nondecreasing, with $\tau_0 = 0$, measurable on the filtration generated by the state process. Additional assumptions are made in the results that follow.

Theorem 6.2.4 ([439]). *Suppose that $\boldsymbol{X}$ is a φ-irreducible and aperiodic Markov chain. Suppose moreover that there are functions $V : \mathbb{X} \to (0,\infty)$, $\delta : \mathbb{X} \to [1,\infty)$, $f : \mathbb{X} \to [1,\infty)$, a small set C, and a constant $b \in \mathbb{R}$, such that the following hold:*

$$\begin{aligned} E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] &\le V(x_{\tau_z}) - \delta(x_{\tau_z}) + b1_{\{x_{\tau_z} \in C\}} \\ E\Big[\sum_{k=\tau_z}^{\tau_{z+1}-1} f(x_k) \mid \mathcal{F}_{\tau_z}\Big] &\le \delta(x_{\tau_z}), \qquad z \ge 0. \end{aligned} \tag{6.3}$$

Then the following hold:

(i) *$\boldsymbol{X}$ is positive Harris recurrent, with unique invariant distribution π.*
(ii) *$\pi(f) := \int f(x)\,\pi(dx) < \infty$.*
(iii) *For any function g that is bounded by f, in the sense that $\sup_x |g(x)|/f(x) < \infty$, we have convergence of moments in the mean, and the Law of Large Numbers holds:*

$$\lim_{t\to\infty} E_x[g(x_t)] = \pi(g),$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} g(x_t) = \pi(g) \qquad a.s.,\ x \in \mathbb{X}.$$

◇

Proof. See Sect. 6.3.1. □

Remark 6.2.1. We note that, for (ii) in Theorem 6.2.4, the condition that $f : \mathbb{X} \to [1,\infty)$, $\delta : \mathbb{X} \to [1,\infty)$, can be relaxed to $f : \mathbb{X} \to [0,\infty)$, $\delta : \mathbb{X} \to [0,\infty)$, provided that one can establish (i), that is, the positive Harris recurrence of the Markov chain $\boldsymbol{X}$ first. ◇

Remark 6.2.2. The assumption of $\tau_0 = 0$ is important for establishing recurrence. If $\tau_0 \neq 0$, additional technical, but mild, conditions would have to be added to verify recurrence. ◇

By taking $f(x) = 1$ for all $x \in \mathbb{X}$, we obtain the following corollary to Theorem 6.2.4.

Corollary 6.2.1. *Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain with filtration $\mathcal{F}_t$. Suppose moreover that there is a function $V : \mathbb{X} \to (0,\infty)$, a petite set C, and a constant $b \in \mathbb{R}$, such that the following hold:*

$$\begin{aligned} &E[V(x_{\tau_{z+1}}) \mid \mathcal{F}_{\tau_z}] \le V(x_{\tau_z}) - 1 + b1_{\{x_{\tau_z} \in C\}}, \\ &\sup_{z\ge 0} E[\tau_{z+1} - \tau_z \mid \mathcal{F}_{\tau_z}] < \infty. \end{aligned} \tag{6.4}$$

Then $\boldsymbol{X}$ is positive Harris recurrent. ◇

To relax the φ-irreducibility assumption, we can impose instead the following continuity assumption: A Markov chain is (weak) Feller if the function Pf is continuous on $\mathbb{X}$, for every continuous and bounded function $f\colon \mathbb{X} \to \mathbb{R}$.

Theorem 6.2.5 ([227,271]). *Suppose that the chain* $\boldsymbol{X}$ *is Feller, living in a locally compact space and that there is a compact set* A *satisfying*

$$\sup_{x \in A} E_x[\tau_A] < \infty,$$

where $\tau_A = \inf(k > 0 : x_k \in A)$*. Then the Markov chain admits an invariant probability measure* π *which is positive on* A. ◇

The following is a consequence of proofs of Theorems 6.2.4 and 6.2.5.

Theorem 6.2.6. *If* (6.4) *holds for a measurable set* C *and a function* $V\colon \mathbb{X} \to (0, \infty)$*, then* C *satisfies*

$$\sup_{x \in C} E[\tau_C] < \infty.$$

◇

In view of the above, we have the following. Note that here the Markov chain need not be irreducible.

Theorem 6.2.7 ([439]). *Suppose that* $\boldsymbol{X}$ *is a Feller Markov chain living in a locally compact space, not necessarily* φ*-irreducible. Then, if* (6.3) *holds with* C *compact, there exists at least one invariant probability measure. Moreover, there exists* $c < \infty$ *such that, under any invariant probability measure* π*,*

$$E_\pi[f(x_t)] = \int_{\mathbb{X}} \pi(dx) f(x) \leq c. \tag{6.5}$$

◇

Finally, we state a result on transience. This will be useful in establishing that bounded range quantizers lead to transience (see Theorem 7.3.1; see also Theorem 8.6.3 which essentially uses a similar argument).

Theorem 6.2.8 (Theorem 8.4.1 in [271]). *Let* $V : \mathbb{X} \to \mathbb{R}_+$*. If there exists a set* A *such that*

$$E[V(x_{t+1}) | x_t = x] \leq V(x), \tag{6.6}$$

for all $x \notin A$ *and* $\exists \bar{x} \notin A$ *such that* $V(\bar{x}) < \inf_{z \in A} V(z)$*, then* $\{x_t\}$ *is not recurrent, in the sense that* $P_{\bar{x}}(\tau_A < \infty) < 1$. ◇

Proof. See Sect. 6.3.2. □

Why Random-time State-dependent Drift Criteria?

Many network protocol and networked control applications share the property that controllers and sensors can access data or act on a system at random times. These times may depend on the availability of communication resources. One example of such a scenario is reported in [419], which will be discussed in detail in Chaps. 7 and 8, for establishing stochastic stability of adaptive quantizers for Markov sources where random stopping times are the instances when the encoder can transmit information to a controller. In this context, there has been a significant amount of research on stochastic stabilization of networked control systems under information constraints. For linear systems driven by unbounded noise, [419] established ergodicity, under fixed-rate constraints, through martingale methods. Yüksel and Başar [432] obtained conditions for the existence of an invariant probability measure for noisy channels, considering deterministic, state-dependent drift criteria, based on the criteria developed in [272]. Further examples in networked control where the theory is applicable are settings in [313, 314, 415] among others which will be discussed further in Chap. 8.

Our motivation here for random time drift comes from such applications in networked control, as well as in information theory with variable length and variable delay decoding, [328, 331], and non-asymptotic information theory [310]. In particular, variable length coding schemes with best error exponents also allow for random delay in decoding [83].

In the chapters to follow, we will define a sequence of increasing stopping times $\{\tau_z, z \in \mathbb{Z}_+\}$, which will denote the random times when important events in the system take place: For control over noiseless channels in Chap. 7 and control over erasure channels in Chap. 8, these will stand for the times when *an informative message* arrives at the controller with no channel error. For control over general noisy channels in Chap. 8, these will be the times when the encoder's uncertainty regarding the state of the system is bounded. Based on these stopping times, we apply the results of this chapter to establish various forms of stochastic stability. In Sects. 7.3.3 and 8.3.1, we will make the connections with stabilization over communication channels more explicit.

6.3 Appendix: Proofs

6.3.1 Proof of Theorem 6.2.4

Proof of Theorem 6.2.4(i)

The proof is similar to the proof of the Comparison Theorem of [271]: Define the sequence $\{M_z : z \geq 0\}$ by $M_0 = V(x_0)$, and for $z \geq 0$,

$$M_{z+1} = V(x_{\tau_{z+1}}) + \sum_{k=0}^{z} (\delta(x_{\tau_k}) - b 1_{\{x_{\tau_k} \in C\}}).$$

Under the assumed drift condition we have

$$E[M_{z+1} \mid \mathcal{F}_{\tau_z}] \le V(x_{\tau_z}) + \sum_{k=0}^{z-1} (\delta(x_{\tau_k}) - b 1_{\{x_{\tau_k} \in C\}}),$$

which implies the supermartingale bound, $E[M_{z+1} \mid \mathcal{F}_{\tau_z}] \le M_z$.

For a measurable subset $C \subset \mathbb{X}$ we denote the first hitting time for the sampled chain,

$$\zeta_C = \min\{z \ge 1 : x_{\tau_z} \in C\}. \tag{6.7}$$

Define $\zeta_C^n = \min(n, \zeta_C)$ for any $n \ge 1$. Then $E[M_{\zeta_C^n}] \le M_0$ for any $n \in \mathbb{Z}$, and

$$E[\sum_{k=0}^{\zeta_C^n - 1} \delta(x_{\tau_k}) | \mathcal{F}_0] \le M_0 + b.$$

Applying the bound $E\Big[\sum_{k=\tau_z}^{\tau_{z+1}-1} f(x_k) \mid \mathcal{F}_{\tau_z}\Big] \le \delta(x_{\tau_z})$ and that $f(x) \ge 1$, the following bound is obtained from the smoothing property of the conditional expectation:

$$\begin{aligned} E[\tau_{\zeta_C^n} \mid \mathcal{F}_0] &= E\Big[\sum_{i=0}^{\zeta_C^n - 1} E[\tau_{i+1} - \tau_i] | \mathcal{F}_0]\Big] \\ &\le E\Big[\sum_{i=0}^{\zeta_C^n - 1} \delta(x_{\tau_i}) | \mathcal{F}_0\Big] \le M_0 + b. \end{aligned}$$

Hence by the monotone convergence theorem,

$$E[\tau_C] \le E[\tau_{\zeta_C}] = \lim_{n\to\infty} E[\tau_{\zeta_C^n} \mid \mathcal{F}_0] \le M_0 + b.$$

Consequently we obtain that $\sup_{x \in C} E[\tau_C] < \infty$, as well as recurrence of the chain, $P_x(\tau_C < \infty) = 1$ for any $x \in \mathbb{X}$. Positive Harris recurrence now follows from [270] Theorem 4.1. □

The following result is key to obtaining moment bounds. The inequality (6.9) is known as *drift condition (V3)* [271]. Define

$$V_f^*(x) := E_x\Big[\sum_{t=0}^{\tau_C - 1} f(x_t)\Big], \quad x \in \mathbb{X}. \tag{6.8}$$

Lemma 6.3.1. *Suppose that* $\boldsymbol{X}$ *satisfies all of the assumptions of Theorem 6.2.4, except that the* ψ*-irreducibility assumption is relaxed. Then, there is a constant* b_f *such that the following bounds hold;*

$$PV_f^* \le V_f^* - f + b_f \mathbb{I}_C, \tag{6.9}$$

$$V_f^*(x) \le V(x) + b_f, \qquad x \in \mathbb{X}. \tag{6.10}$$

◇

Proof. The drift condition (6.9) is given in Theorem 14.0.1 of [271]. The proof of (6.10) is based on supermartingale arguments: Denote $M_0 = V(x_0)$, and for $z \geq 0$,

$$M_{z+1} = V(x_{\tau_{z+1}}) - \sum_{k=0}^{\tau_{z+1}-1} \big(-f(x_k) + b1_{\{x_{\tau_k} \in C\}}\big). \tag{6.11}$$

The supermartingale property for $\{M_z\}$ follows from the assumed drift condition:

$$\begin{aligned} E[M_{z+1} \mid \mathcal{F}_{\tau_z}] = M_z + E\Big[&V(x_{\tau_{z+1}}) - V(x_{\tau_z}) \\ &+ \sum_{k=\tau_z}^{\tau_{z+1}-1} (f(x_k) - b1_{\{x_{\tau_k} \in C\}}) \mid \mathcal{F}_{\tau_z}\Big] \leq M_z. \end{aligned} \tag{6.12}$$

As in the previous proof we bound expectations involving the stopping time ζ_C beginning with its truncation $\zeta_C^n = \min(n, \zeta_C)$.

The supermartingale property gives $E[M_{\zeta_C^n}] \leq M_0$, and once again it follows again by the monotone convergence theorem that V_f^* satisfies the bound (6.10) as claimed. □

Proofs of Theorem 6.2.4(ii) and (iii)

The existence of a finite moment follows from Lemma 6.3.1 and the following generalization of Kac's theorem (see [271, Theorem 10.4.9]):

$$\pi(f) := \int \pi(dx) f(x) = \int_A \pi(dx) E_x\bigg(\sum_{t=0}^{\tau_A - 1} f(x_t)\bigg), \tag{6.13}$$

where A is any set satisfying $\pi(A) > 0$ and $\tau_A = \inf(t \geq 1 : x_t \in A)$. The supermartingale argument above ensures that the expectation under the invariant probability measure is bounded by recognizing C as a recurrent set.

Proof of Theorem 6.2.4(iii) now follows from the ergodic theorem for Markov chains; see [271, Theorem 17.0.1]. □

6.3.2 Proof of Theorem 6.2.8

Let $x = \bar{x}$. Proof follows from observing that

$$
\begin{aligned}
V(x) &\geq \int_y V(y)P(x,dy) \\
&\geq (\inf_{z\in A} V(z))P(x,A) + \int_{y\notin A} V(y)P(x,dy) \\
&\geq (\inf_{z\in A} V(z))P(x,A).
\end{aligned}
$$

It thus follows that

$$P(\tau_A < 2) = P(x,A) \leq \frac{V(x)}{(\inf_{z\in A} V(z))}.$$

Likewise,

$$
\begin{aligned}
V(\bar{x}) &\geq \int_y V(y)P(\bar{x},dy) \\
&\geq (\inf_{z\in A} V(z))P(\bar{x},A) + \int_{y\notin A} (\int_s V(s)P(y,ds))P(\bar{x},dy) \\
&\geq (\inf_{z\in A} V(z))P(\bar{x},A) + \int_{y\notin A} P(\bar{x},dy)((\inf_{s\in A} V(s))P(y,A) + \int_{s\notin A} V(s)P(y,ds)) \\
&\geq (\inf_{z\in A} V(z))P(\bar{x},A) + \int_{y\notin A} P(\bar{x},dy)((\inf_{s\in A} V(s))P(y,A)) \\
&= (\inf_{z\in A} V(z))\Big(P(\bar{x},A) + \int_{y\notin A} P(\bar{x},dy)P(y,A)\Big).
\end{aligned}
\tag{6.14}
$$

Thus, observing that $P(\{\omega : \tau_A(\omega) < 3\}) = \int_A P(\bar{x},dy) + \int_{y\notin A} P(\bar{x},dy)P(y,A)$, we observe that

$$P_{\bar{x}}(\tau_A < 3) \leq \frac{V(\bar{x})}{(\inf_{z\in A} V(z))}.$$

Thus, for any n: $P_{\bar{x}}(\tau_A < n) \leq \frac{V(\bar{x})}{(\inf_{z\in A} V(z))} < 1$. Continuity of probability measures (by defining: $B_n = \{\omega : \tau_A < n\}$ and observing $B_n \subset B_{n+1}$ and that $\lim_n P(\tau_A < n) = P(\cup_n B_n) = P(\tau_A < \infty) < 1$) leads to the result. □

6.4 Concluding Remarks

In this chapter, stochastic random-time drift criteria have been introduced. The random-drift criteria are especially useful in networked control systems, as will be demonstrated later in the book starting with the next chapter. For stabilization, we will observe that a random-time drift based program is very effective in obtaining stochastic stability results. Furthermore, many decentralized control problems use algorithms which allow for agreement at certain variables of nature in random times, based on which decisions can be generated. Such time characterizations are essential for optimization and stabilization studies of networked and decentralized control systems. We will revisit this topic in Chap. 12.

The stochastic stability in finite moments or the existence of a unique invariant probability measure is important also for the applicability of the convex analytic method of Borkar [71] for MDPs (see Appendix D), in view of the applicability of the sample path ergodic theorem.

6.5 Bibliographic Notes

Stochastic stability of Markov chains has a rich and complete theory and forms a foundation for several other general techniques such as dynamic programming and Markov chain Monte Carlo (MCMC) [269].

The *state-dependent* criteria [253, 272] are the basis of the fluid-model (or ODE) approach to stability in stochastic networks and other general models [72, 112, 113, 142, 269].

On the topic of Markov chains, rates of convergence and mixing under random-time drift is an important aspect to be explored for networked control systems. It is apparent that the nature of the drift as well as the distribution of stopping times used for drift will play a role in the rate of convergence. We refer the reader to [100, 121] for rate of convergence analysis when the drift times are deterministic. We also refer the reader to [322] on rates of convergence and geometric ergodicity. Further related random-drift results have been presented in [144].

The methods of random-time drift criteria can also be applied to other models of networked control systems with delay-sensitive information transmission: For such systems, the effects of randomness in delay for transmission of sensor or controller signals (see, e.g., [96, 179, 204, 239, 407]) is an application area where this discussion is relevant. Another related area is *event-triggered feedback control* systems [17, 230, 244, 351], where the event instances constitute the stopping times.

The treatment in this chapter is a condensed discussion from Meyn–Tweedie [271], Hernandez-Lerma and Lasserre [195] and [439]. The results on random-time drift criteria in Sect. 6.2.3 and the proofs are based on Yüksel and Meyn [416, 439]. Theorem 6.2.8 is available in [271] as well as Hairer [189].

Chapter 7
Stochastic Stabilization Over Noiseless Channels

7.1 Introduction

In this chapter, we present conditions for stochastic stabilization of networked linear control systems over noiseless channels with finite capacity. We consider multi-dimensional linear systems within both fully observed and partially observed settings and obtain conditions for stochastic stabilizability. In such systems, there is an information rate constraint on the quantizers. We use the random-time state-dependent drift criteria introduced in Chap. 6 to obtain stabilizing schemes. The rate requirements under such an approach meet the fundamental lower bound given in Theorem 5.6.1.

The specifications considered in this chapter are the existence of an invariant probability measure, as well as the finiteness of certain moments. Section 7.2 establishes the control and the communication models, Sect. 7.3 establishes stochastic stability for a scalar setting, Sect. 7.4 considers multidimensional systems, and Sect. 7.5 investigates partially observed setups.

7.2 Control and Communication Models

We consider stabilization of linear noisy systems described by the dynamics

$$x_{t+1} = Ax_t + Bu_t + w_t, \tag{7.1}$$

where $x_t \in \mathbb{R}^n$, (A, B) is controllable, and the noise process $\{w_t\}$ is a zero-mean i.i.d. sequence of random vectors, with stabilization to be carried out over noiseless channels with finite capacity. We assume also that x_0 is a second-order random variable.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_7,

The conditions we will impose on the zero-mean i.i.d. noise sequence will prove to be important in the goals to be achieved. The noise sequence is required to satisfy one of the following assumptions:

Assumption 7.2.1. *Each* w_t *admits a probability measure* ν *which is absolutely continuous with respect to the Lebesgue measure* λ *on* $\mathbb{R}^n$ *(i.e., the measure admits a density), and for every* $D \in \mathcal{B}(\mathbb{R}^n)$ *with positive Lebesgue measure,* $\nu(D) > 0$*. Furthermore,* $E[|w_t|^2] < \infty$*.* ◇

The second assumption is slightly stronger.

Assumption 7.2.2. *Each* w_t *has a probability measure* ν *which admits a density, and for every* $D \in \mathcal{B}(\mathbb{R}^n)$ *with positive Lebesgue measure,* $\nu(D) > 0$*. Furthermore,* $E[|w_t|^{2+\epsilon}] < \infty$ *for some* $\epsilon > 0$*.* ◇

Finally, we take the noise to be Gaussian.

Assumption 7.2.3. $\{w_t\}$ *is a sequence of Gaussian random variables.* ◇

Under Assumption 7.2.1, we will establish stationarity and ergodicity. Under Assumption 7.2.2 (or the stronger Assumption 7.2.3), we will establish, in addition, the existence of finite moments.

Before discussing the more general case in Sect. 7.4, we first discuss the scalar version described by the following equation:

$$x_{t+1} = ax_t + bu_t + w_t. \tag{7.2}$$

As depicted in Fig. 5.5 and generally described in Sect. 5.2.2, our goal is to stabilize such a system by designing quantization and controller policies, where the quantizers at a given time map the information to a finite set $\mathcal{M}$.

7.3 Stochastic Stability Analysis for a Scalar System

7.3.1 Adaptive Quantizers and a Zooming Scheme

A general class of quantization policies involves *adaptive quantizers*: Following the terminology of Chap. 5, such a quantizer policy picks quantizers whose selection involves memory. Use of such quantizers is motivated by the following result whose proof can be found in the appendix.

Theorem 7.3.1. *Consider (7.2), with* $|a| > 1$*. A fixed quantizer leads to a transient Markov chain in the sense that* $P_x(\tau_S < \infty) < 1$ *where for some* $s > 0$*,* $S = (-\infty, s)$ *is an open set containing the origin,* $x \notin S$ *and* $\tau_S := \inf(t > 0 : x_t \in S)$*.* ◇

Let $\mathbb{S}$ be a set of states for a quantizer state s. Let $F : \mathbb{S} \times \mathbb{R} \to \mathbb{S}$ be a state update function. An adaptive quantizer may have the following state update equations: $s_{t+1} = F(Q_t(x_t), s_t)$. Here, Q_t is the quantizer applied at time t, x_t is the input to the quantizer Q_t, and s_t is the *state* of the quantizer. Such a quantizer is implementable since the updates can be performed at both the encoder and the decoder.

A particular class of adaptive quantizers has been introduced by Goodman and Gersho [166]. One such type has the following form (see (5.1) and Fig. 5.1) with Q_K^{Δ} being a uniform quantizer with $K+1$ bins and bin size Δ and $\bar{Q}$ determining the updates in the bin size of the uniform quantizer as a function of the source and the current bin size:

$$q_t = Q_K^{\Delta_t}(x_t), \qquad \Delta_{t+1} = \Delta_t \hat{Q}(q_t, \Delta_t). \tag{7.3}$$

Here Δ_t characterizes the uniform quantizer, as it is the bin size of the quantizer at time t.

In the analysis, we will consider such adaptive quantizers, whose realizations will be from a class of uniform quantizers, considered earlier in (5.1). Thus, in this setting, we modify the description of a traditional uniform quantizer by assigning the same value when the state is in the overflow region of the quantizer.

The quantizer system is connected over a noiseless channel with a finite capacity to an estimator (controller) as depicted in Fig. 5.5. The controller has only access to the information it has received through the channel. The controller in the model estimates the state and then applies its control action. As such, the problem reduces to a state estimation problem since (7.2) is controllable. Hence, stability of the estimation error is equivalent to stability of the state itself.

7.3.2 *Stochastic Stability Analysis*

In the following, we will consider a discrete noiseless channel with capacity R (hence with $2^R = |\mathcal{M}| \in \mathbb{Z}_+$).

Theorem 7.3.2 ([419]). *Consider an adaptive quantizer applied to the linear control system described by (7.2), under Assumption 7.2.1. If the noiseless channel has capacity,*

$$R = \log_2(\lceil |a| + \epsilon \rceil + 1),$$

for some $\epsilon > 0$, there exists an adaptive quantization policy such that there exists a compact set $S \subset \mathbb{R}^2$ so that with $\tau_S = \min(t > 0 : x_t \in S)$:

$$\sup_{(x,\Delta) \in S} E[\tau_S | x_0 = x] < \infty,$$

which makes S a recurrent set. Furthermore

$$P_{(x,\Delta)}(\tau_S < \infty) = 1$$

for all (x, Δ) pairs visited by the chain. ◇

The proof of the theorem can be found in the appendix to the chapter. For the constructive proof, we consider adaptive quantizers for which the quantizer bins are updated as follows. The quantizer bin sizes will get enlarged until the state process hits the granular region of the quantizer, that is, the joint process hits the set $\{(x_t, \Delta_t) : |h_t| \leq 1\}$, where

$$h_t = x_t/(\frac{K}{2}\Delta_t). \tag{7.4}$$

In this case, the quantizer will be said to be in the *perfect-zoom* phase. Due to the effect of the system noise, occasionally the state will be in the overflow region of the quantizer, leading to an *under-zoom* phase.

Now, with $K = \lceil |a| + \epsilon \rceil$, $R = \log_2(K+1)$, let us introduce $R' = \log_2(K)$. We will consider the following update rules. For $t \geq 0$ and with $\Delta_0 > L$ for some $L \in \mathbb{R}_+$, and $\hat{x}_0 \in \mathbb{R}$, consider

$$u_t = -\frac{a}{b}\hat{x}_t, \quad \hat{x}_t = Q_K^{\Delta_t}(x_t), \quad \Delta_{t+1} = \Delta_t \bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|, \Delta_t). \tag{7.5}$$

If we use $\delta, \epsilon, \alpha > 0$ with $\alpha < 1$ and $L > 0$ such that

$$\begin{aligned}
&\bar{Q}(x,\Delta) = |a| + \delta \quad \text{if} \quad |x| > 1, \\
&\bar{Q}(x,\Delta) = \alpha \quad \text{if} \quad 0 \leq |x| \leq 1, \Delta \geq L, \\
&\bar{Q}(x,\Delta) = 1 \qquad \text{if} \quad 0 \leq |x| \leq 1, \Delta < L,
\end{aligned} \tag{7.6}$$

we will show that a recurrent set exists. We note that the above imply that $\Delta_t \geq L\alpha =: L'$ for all $t \geq 0$.

The stability for such a scheme can be established using random-time stochastic drift conditions. This is because the quantizer helps reduce the uncertainty on the system state only when the state is in the *granular* region of the quantizer. The times when the state is in this region are random stopping times (defined on the filtration generated by the state and quantizer processes).

The following considers the state space for the quantizer bins to be countable, leading to irreducibility and consequently positive Harris recurrence. The proof is given in the appendix.

Theorem 7.3.3 ([419]). *Under the setup of Theorem 7.3.2, for the adaptive quantizer in (7.5), if the quantizer bin sizes are such that their (base-2) logarithms are integer multiples of some scalar s and* $\log_2(\bar{Q}(\cdot,\cdot))$ *take values in integer multiples of s where the integers taken are relatively prime (i.e., they share no common divisors except for* 1*), then the process $\{(x_t, \Delta_t)\}$ is a positive (Harris) recurrent Markov chain (and has a unique invariant distribution).* ◇

The following results are on moment stability, whose proofs can be found in the appendix. Note that more restrictive assumptions (Assumptions 7.2.2 and 7.2.3) are imposed on the tail distribution of the i.i.d. noise process $\{w_t\}$.

Theorem 7.3.4 ([419]). *Under the setups of Theorem 7.3.2, Theorem 7.3.3, and Assumption 7.2.3, it follows that* $\lim_{t\to\infty} E[x_t^2] < \infty$*, and this limit is independent of the initial state of the system. Furthermore,* $\lim_{t\to\infty} E[\Delta_t^2] < \infty$. ◇

This result also holds under Assumption 7.2.2 with essentially the same arguments as in the proof of Theorem 7.3.4. A proof sketch is presented in the appendix.

Theorem 7.3.5 ([209]). *Under the setups of Theorem 7.3.2, Theorem 7.3.3 and Assumption 7.2.2, it follows that* $\lim_{t\to\infty} E[x_t^2] < \infty$*, and this limit is independent of the initial state of the system. Furthermore,* $\lim_{t\to\infty} E[\Delta_t^2] < \infty$. ◇

We also have the following.

Theorem 7.3.6. *Let the assumptions of Theorem 7.3.3 and Assumption 7.2.3 hold. In addition, assume that*

$$|a|^m\Big(\frac{1}{(2^R-1)^m}\Big) < 1,$$

for some $m \in \mathbb{N}$*. Then, with the adaptive quantization policy (7.6) and initial condition* (x_0, Δ_0)*, we have*

$$\lim_{t\to\infty} E[|x_t|^m] = E_\pi[|x_0|^m] < \infty\,.$$

◇

7.3.3 Application of the Theory of Random-time State-dependent Stochastic Drift

We now make explicit the connection with the general theory for random-time stochastic drift introduced in Chap. 6. The Markov chain considered is (x_t, Δ_t) (see Lemma 7.6.1).

Figure 7.1 provides some intuition on the construction of stopping times and the Lyapunov functions, where $h_t = x_t/(2^{R'-1}\Delta_t)$ was introduced in (7.4). The arrows shown in the figure denote the mean one-step increments of (x_t, h_t). That is, the arrow ν with base at (x, h) is defined by

$$\nu = E[(x_{t+1}, h_{t+1}) - (x_t, h_t) \mid (x_t, h_t) = (x, h)].$$

With $F > 0$ fixed and with $F' = F2^{-(R'-1)}$, two sets are used to define the small set in the drift criteria: $C_x = \{x : |x| \le F\}$ for $F > 0$ and $C_\Delta = \{\Delta : \Delta \le F'\}$. Let $C_h = \{h : |h| \le 1\}$, and assume that $F > 0$ is chosen sufficiently large so that $(x_t, \Delta_t) \in C_x \times C_\Delta$ whenever $(x_t, h_t) \in C_x \times C_h$. When x_t is outside C_x and h_t outside C_h (the *under-zoomed phase* of the quantizer), there is a drift of h_t toward

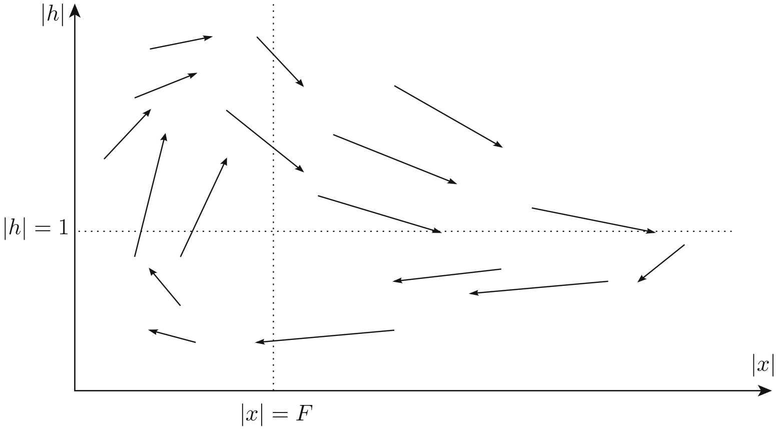

Fig. 7.1 Drift in the Markov Process. When under-zoomed, the error increases on average and the quantizer zooms out; when perfectly zoomed, the error decreases and the quantizer zooms in

C_h. When the process x_t reaches C_h (the *perfectly zoomed phase* of the quantizer), then the process drifts toward C_x.

In the model considered, the controller can receive meaningful information regarding the state of the system when the source lies in the granular region of the quantizer, that is, $x_t \in [-\frac{1}{2}K\varDelta_t, \frac{1}{2}K\varDelta_t]$.

The times at which these events occur form an increasing sequence of *stopping times*. We will apply the drift criteria presented in Sect. 6.2.3 for these random stopping times. In particular, we define the sequence of stopping times as

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k| \leq 1\}, \quad z \in \mathbb{N}. \tag{7.7}$$

These are the times when information reaches the controller regarding the value of the state when the state is in the granular region of the quantizer.

Using this construction of stopping times, we will find appropriate Lyapunov functions (such as $V(x, \varDelta) = \log(\varDelta^2)$ and $V(x, \varDelta) = \varDelta^2$) and drift functions which will lead to stationarity as well as finite moments for the Markov process $\{(x_t, h_t)\}$ through the satisfaction of special instances of Theorem 6.2.4. See the appendix for details.

7.3.4 Simulation

For a simulation study, we consider a linear system with the following dynamics:

$$x_{t+1} = 2.2x_t + u_t + w_t,$$

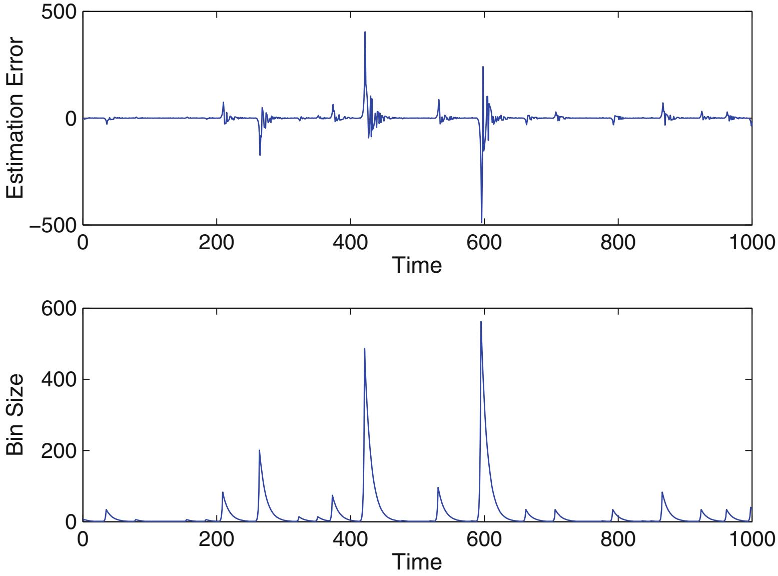

Fig. 7.2 Sample path for a stochastically stable quantizer. The variables picked are as follows: $L' = 1, \epsilon = 0.3, \delta = 0.25, \eta = 0.02$

where $\{w_t\}$ are i.i.d., zero-mean Gaussian with variance 1. We use the zooming quantizer with rate $\log_2(4) = 2$, since 4 is the smallest integer larger than or equal to $\lceil 2.2 \rceil + 1$. We have taken $L' = 1$. Figure 7.2 corroborates the stochastic stability result, by showing the under-zoomed and perfectly zoomed phases, with the peaks in the plots showing the under-zoom phases.

7.4 The Multidimensional Case

The scheme proposed in the previous section is also applicable to the multidimensional setup. Let us revisit (7.1), where $x_t \in \mathbb{R}^n$ is the state at time t, $u_t \in \mathbb{R}^m$ is the control action, and $\{w_t\}$ is a sequence of i.i.d. $\mathbb{R}^n$-valued zero-mean random variables satisfying Assumption 7.2.2. Here A is the system matrix with at least one eigenvalue greater than 1 in magnitude, that is, the system is open-loop unstable. We also assume at this point that the eigenvalues are real.

Without any loss of generality, we assume A to be in Jordan form. Because of this, we allow w_t to have correlated components, that is, the correlation matrix $E[w_t w_t^T]$ is not necessarily diagonal. We may also assume that (A, B) is controllable, and for ease in presentation we assume that B is invertible mainly to make the stopping time analysis easier to pursue (if B is not invertible, the sampled system can be made to have an invertible control matrix, with a periodic scheme of period at most n).

Stabilizability for the diagonalizable case immediately follows from the discussion for scalar systems, since the analysis for the scalar case is applicable to each of the subsystems along the eigenvectors. The possibly correlated noise components will lead to the recurrence analysis discussed earlier. For such a setup, the stopping times can be arranged to be identical for each mode, for the case when the quantizer captures all the state components. Once this is satisfied, the drift conditions will be obtained. The non-diagonalizable Jordan case, however, is more involved. The approach for the scalar system still applies, but it needs to be appropriately generalized.

We now make this discussion more precise. One can consider two approaches for stabilization of a multidimensional system. We first discuss one approach which we will refer to as the *block-coding approach.*

Consider the following system:

$$\begin{bmatrix} x_{t+1}^1 \\ x_{t+1}^2 \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} + B \begin{bmatrix} u_t^1 \\ u_t^2 \end{bmatrix} + \begin{bmatrix} w_t^1 \\ w_t^2 \end{bmatrix}. \tag{7.8}$$

The approach entails quantizing the components in the system according to the adaptive quantization rule provided earlier, that is, we modify the scheme in (7.5) as follows: For $i = 1, 2$, let $R' = R_i' = \log_2(2^{R_i} - 1) = \log_2(K_i)$ (i.e., the same rate is used for quantizing the components with the same eigenvalue). For $t \geq 0$ and with $\Delta_0^1, \Delta_0^2 \in \mathbb{R}$, consider

$$u_t = -B^{-1} A \hat{x}_t,$$

$$\begin{bmatrix} \hat{x}_t^1 \\ \hat{x}_t^2 \end{bmatrix} = \begin{bmatrix} Q_{K_1}^{\Delta_t^1}(x_t^1) \\ Q_{K_2}^{\Delta_t^2}(x_t^2) \end{bmatrix}, \tag{7.9}$$

$$\Delta_{t+1}^1 = \Delta_t^1 \bar{Q}(|h_t^1|, |h_t^2|, \Delta_t^1), \qquad \Delta_{t+1}^2 = \Delta_t^2 \bar{Q}(|h_t^1|, |h_t^2|, \Delta_t^2), \tag{7.10}$$

with, for $i = 1, 2$, $\delta^i, \epsilon^i, \eta^i > 0$, $\eta^i < \epsilon^i$ and $L^i > 0$ such that

$$\bar{Q}(x, y, \Delta) = |\lambda| + \delta^i \quad \text{if} \quad |x| > 1, \quad \text{or} \quad |y| > 1,$$

$$\bar{Q}(x, y, \Delta) = \frac{|\lambda|}{2^{R_i'} - \eta^i} \quad \text{if} \quad 0 \leq |x| \leq 1, |y| \leq 1, \Delta^i > L^i,$$

$$\bar{Q}(x, y, \Delta) = 1 \qquad \text{if} \quad 0 \leq |x| \leq 1, |y| \leq 1, \Delta^i \leq L^i.$$

Note that the above imply that $\Delta_t^i \geq L^i \frac{|\lambda|}{2^{R_i'} - \eta^i} =: L'^i$. We also assume that for some sufficiently large η_Δ, $\Delta_0^1 = \eta_\Delta \Delta_0^2$, which leads to the result that $\Delta_t^1 = \eta_\Delta \Delta_t^2$ for all $t \geq 0$. See Fig. 7.3 for a depiction of the quantizer used at a particular time.

Instead of (7.7), the sequence of stopping times is now defined as follows:

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k^i| \leq 1, i \in \{1, 2, \ldots, n\}\}, \quad z \in \mathbb{Z}_+,$$

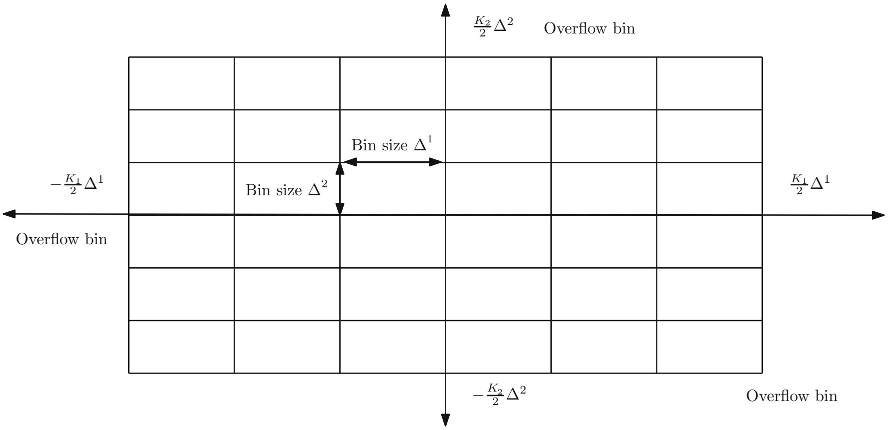

Fig. 7.3 A uniform vector quantizer. There is a single overflow bin

where $h_t^i = \frac{x_t^i}{\Delta_t^i 2^{R_i'-1}}$. Here Δ^i is the bin size of the quantizer in the direction of the eigenvector x^i, with rate R_i'.

With this approach, the drift criterion applies almost identically as it does for the scalar case. The extensions of Theorems 7.3.2, 7.3.3, and 7.3.4 are then immediate.

We thus have the following result. The proof for a single Jordan block with real eigenvalues is in the appendix. The extension to multiple Jordan blocks as well as systems having complex eigenvalues in the real Jordan canonical form follows from similar arguments by coupling the modes having complex conjugate pairs in a single vector quantizer through the establishment of a geometric measure bounding (majorizing) the probability measure of a subsequent stopping time.

Theorem 7.4.1 ([209]). *Consider the multidimensional system (7.1). If the average rate satisfies*

$$R > \sum_{|\lambda_i|>1} \log_2(|\lambda_i|),$$

there exists a stabilizing scheme leading to a Markov chain with a bounded second moment in the sense that $\limsup_{t\to\infty} E[|x_t|_2^2] < \infty$. ◇

Proof. See Sect. 7.6.7. □

Remark 7.4.1. In an alternative scheme, one could consider an approach that we call *sequential stabilization of the scalar components*: We could adopt a lower to upper sequential approach, considering stabilized modes at particular stopping times as noise with a finite moment. Using an inductive argument, one can first start with the lowest mode (in the matrix diagonal) of the system and stabilize that mode so that there is a finite invariant second moment of the state. We can then view this mode as a second-order disturbance for the one upper mode and obtain conditions for stabilizability for this mode, by employing the same approach, since the lower mode is a positive Harris recurrent Markov process. We note that the random process for the upper mode might not have Markov dynamics for its marginal, but the joint system consisting of all modes and quantizer parameters will be Markov. Furthermore, the lower modes, viewed as noise, will no longer be independent from the state process. For example, the effective disturbance affecting the stochastic evolution of a repeated mode is no longer Gaussian. However appropriate moment bounds (such as Jensen's and Hölder's inequality) can be applied to carry out the analysis. Such a sequential stabilization approach will be adopted for noise-free multi-sensor systems in Chap. 9 and for Gaussian channels in Chap. 11. ◇

7.5 The Partially Observed Case

The results in the previous sections have direct counterparts in the partially observed case, where along with (7.1), we have the measurement equation

$$y_t = Cx_t + v_t. \tag{7.11}$$

Here, (A, C) is observable and the noise terms $\{w_t\}, \{v_t\}$ are i.i.d, mutually independent, and Gaussian. The encoder can run an observer, which is given by a Kalman filter. In Appendix D.2 we discuss the reduction to a fully observed model in further detail.

Suppose that $E[w_t w_t^T] = W > 0$ and $E[v_t v_t^T] = V > 0$. We let

$$m_t := E[x_t | y_{[0,t-1]}, u_{[0,t-1]}],$$

$$\Sigma_{t|t-1} := E[(x_t - m_t)(x_t - m_t)^T | y_{[0,t-1]}, u_{[0,t-1]}].$$

Introduce $\tilde{m}_t = E[x_t | y_{[0,t]}, u_{[0,t-1]}]$. In Appendix D.2, we show that

$$\tilde{m}_t = m_t + \Sigma_{t|t-1} C^T (C \Sigma_{t|t-1} C^T + V)^{-1} (y_t - C m_t),$$

or

$$\begin{aligned} \tilde{m}_t &= A\tilde{m}_{t-1} + \Sigma_{t|t-1} C^T (C\Sigma_{t|t-1} C^T + V)^{-1} (CA(x_{t-1} - \tilde{m}_{t-1}) + v_t) \\ \Sigma_{t+1|t} &= A\Sigma_{t|t-1} A^T + W - (A\Sigma_{t|t-1} C^T)(C\Sigma_{t|t-1} C^T + V)^{-1} (C\Sigma_{t|t-1} A^T) \end{aligned} \tag{7.12}$$

with $\Sigma_{0|-1} = E[x_0 x_0']$. The zero-mean variable $x_t - \tilde{m}_t$ is orthogonal to $\mathcal{I}_t = \{y_{[0,t]}, u_{[0,t-1]}\}$, in the sense that the error is independent of the information available at the encoder: given the control actions, the information available at the encoder with regard to the state is Gaussian for all time stages, and consequently independence and orthogonality of $x_t - \tilde{m}_t$ and $\mathcal{I}_t$ are equivalent.

Then, $\tilde{m}_t$ is a fully observed system driven by an independent (but not identical) Gaussian noise process. Asymptotically, however, the independent noise process converges (in total variation) to a stationary distribution since $\Sigma_{t|t-1}$ converges to a constant matrix under the setup considered. The stability analysis for (7.1) can then be applied for (7.12) given the uniform bounds on the noise variances, in view of the convergence argument. Hence, the state process hits a compact set infinitely often, and the state moments are uniformly bounded. As a result, we have the following.

Theorem 7.5.1 ([209]). *Consider the partially observed multidimensional system (7.1)–(7.11). If the average rate satisfies*

$$R > \sum_{|\lambda_i|>1} \log_2(|\lambda_i|),$$

there exists a stabilizing scheme leading to a Markov chain with a bounded second moment in the sense that $\limsup_{t\to\infty} E[|x_t|^2] < \infty$. ◇

Remark 7.5.1. The discussions here and in Sect. 7.4 apply to multi-sensor settings using a quantizer construction similar to that in Sect. 7.4. In this case, the state space is partitioned based on the observable modes at different sensors and coordination among the quantizers for each sensor is established with limited information exchange. In Sect. 9.6.1, this setup is discussed. ◇

7.6 Appendix: Proofs

7.6.1 Proof of Theorem 7.3.1

We provide a proof for the case $a > 1$; a similar line of argument applies for $a < -1$. Suppose that $|u_t| \leq U$ for some finite U. Let a Lyapunov function be picked as $V(x) = \gamma^{-x}$, defined for positive x and with $\gamma > 1$. Define a process $\tilde{x}_{t+1} = \bar{a}\tilde{x}_t + w_t$, where w_t is almost surely the same noise process acting on the original system and $a > \bar{a} > 1$ with $x_0 = \tilde{x}_0$. It follows that $x_t \geq \tilde{x}_t$ for $t \geq 1$ almost surely if the initial condition is greater than $U/(a-\bar{a})$. Let us pick such an initial condition. Observe that for all $\tilde{x} > U/(a-\bar{a})$, $E[V(\tilde{x}_{t+1})|\tilde{x}_t = \tilde{x}] \leq V(\tilde{x})$, since $E[\gamma^{-(\bar{a}\tilde{x}+w_t)}] = \gamma^{-\bar{a}\tilde{x}}E[\gamma^{-w}]$. For all $\tilde{x} \in \{\tilde{x} : \gamma^{(\bar{a}-1)\tilde{x}} > E[\gamma^{-w}]\} =: S^C$ (where $S^C = \mathbb{R} \setminus S$) the Lyapunov condition (6.6) in Theorem 6.2.8 holds. Since the function V is strictly decreasing, $P_{\tilde{x}}(\tau_S < \infty) < 1$ for $\tilde{x} \in S^C$. This implies that $P_x(\tau_S < \infty) < 1$ for the original chain. □

7.6.2 Proof of Theorem 7.3.2

Toward the proof, we will first derive a supporting result.

Lemma 7.6.1. *Let* $\mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$ *denote the Borel* $\sigma-$*field on* $\mathbb{R} \times \mathbb{R}_+$*. Then,*

$$P\Big((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1}), \dots, (x_0, \Delta_0)\Big)$$

$$= P\Big((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1})\Big),$$

$\forall (C \times D) \in \mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$*, i.e.,* (x_t, Δ_t) *is a Markov chain.* ◇

Proof. We observe that $x_{t+1} = ax_t - a\hat{x}_t + w_t$ and $\hat{x}_{t+1} = Q_K^{\Delta_t}(x_{t+1})$. Thus,

$$P\Big((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1}), \dots, (x_0, \Delta_0)\Big)$$

$$= P\Big(x_t \in C | (\Delta_t \in D, x_{t-1}, \Delta_{t-1}), \dots, (x_0, \Delta_0)\Big)$$

$$\times P\Big(\Delta_t \in D | (x_{t-1}, \Delta_{t-1}), \dots, (x_0, \Delta_0)\Big)$$

$$= P\Big(x_t \in C | (x_{t-1}, \Delta_{t-1})\Big) P\Big(\Delta_t \in D | (x_{t-1}, \Delta_{t-1})\Big)$$

$$= P\Big((x_t, \Delta_t) \in (C \times D) | (x_{t-1}, \Delta_{t-1})\Big).$$

The equations above follow from the update equations in the quantizer (7.5). □

Let us define $h_t := \frac{x_t}{\Delta_t 2^{R'-1}}$. Consider the following sets: $C_x = \{x : |x| \leq E\}$, $C_h = \{h : |h| \leq 1\}$, with $E_1 = 2^{R'-1}L'$. Further, let another set be $C'_x = \{x : |x| \leq F\}$, with a sufficiently large F value, to be derived below. We will study the expected number of time stages between visits of $\{(x_t, h_t)\}$ to $C'_x \times C_h$.

We first show that the sequence $\{h_t, t \geq 0\}$ visits C_h infinitely often with probability 1 and the expected length of the excursion is uniformly bounded over all possible values of $(x, h) \in C'_x \times C_h$. Let $V(h_t) = h_t^2$ serve as a Lyapunov function. Define a sequence of stopping times for the perfect-zoom case with (where the initial state is perfectly zoomed)

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k| \leq 1\}, \quad z \in \mathbb{Z}_+.$$

We have that if $|h_t| > 1$ (under-zoomed), $E[h_{t+1}^2|h_t, x_t] \leq \frac{(a^2+\frac{E[w_1^2]}{|x_t|^2})}{(a+\delta)^2}(h_t)^2$. Since when $|h_t| > 1$, we have that $|x_t| > 2^{R'-1}L\alpha$, it follows that

$$E[h_{t+1}^2|h_t, x_t] \leq (\frac{a^2 + \frac{E[w_1^2]}{E_1^2}}{(a+\delta)^2})(h_t)^2.$$

If $|h_t| \leq 1$, then

$$E[h_{t+1}^2] \leq \frac{a^2\frac{(\Delta_t)^2}{4} + E[w_1^2]}{(\Delta_t 2^{R'-1})^2}(\frac{1}{\alpha})^2 \leq \frac{a^2\frac{L'^2}{4} + E[w_1^2]}{(L'2^{R'-1})^2}(\frac{1}{\alpha})^2 =: K_1,$$

where $L' = L\frac{|a|}{|a|+\epsilon-\eta}$ (this is a lower bound on Δ_t). Hence, it follows that

$$E[h_{t+1}^2 - h_t^2|h_t, x_t] \leq -\rho h_t^2 + K_1 1_{\{|h_t|\leq 1\}}, \tag{7.13}$$

where 1_U is the indicator function for event U with $\rho = 1 - \frac{(a^2+\frac{E[w_1^2]}{E^2})}{(a+\delta)^2}$. Since for $A, B > 0$, $A^2 + B^2 \leq (A+B)^2$ it follows that letting L' such that $\sqrt{\frac{E[w_t^2]}{L'2^{R'-1}}} < \delta$ will ensure $\rho > 0$. Now, let us define $K_1' := K_1 + 1$, $M_0 := V(h_0)$, and for $t \geq 1$,

$$M_t := V(h_t) - \sum_{i=0}^{t-1}(-\rho + K_1' 1_{\{h_i \in C_h\}}).$$

Define a stopping time: $\tau^N = \min(N, \min\{i > 0 : V(h_i) + \rho t \geq N\}, \min\{i > 0 : V(h_i) \leq 1\})$. Since, $E[M_{t+1}|(x_s, h_s), s \leq t] \leq M_t, \quad \forall t \geq 0$, it follows that $\{M_t\}$ is a supermartingale sequence. The stopping time τ^N is bounded and the supermartingale sequence is integrable for $t \leq \tau^N$. Hence, we have, by the Martingale Optional Sampling Theorem (see Appendix C.2): $E[M_{(\tau^N)}] \leq E[M_0]$. Thus, we obtain $E[\sum_{i=0}^{\tau^N-1} \rho] \leq V(h_0) + K_1' E[\sum_{i=0}^{\tau^N-1} 1_{\{h_i \in C_h\}}]$, leading to $\rho E[\tau^N - 1 + 1] \leq V(h_0) + K_1'$, and by the monotone convergence theorem,

$$\rho \lim_{N\to\infty} E[\tau^N] = \rho E[\tau] \leq V(h_0) + K_1' \leq 1 + K_1',$$

$$E[\tau_{z+1} - \tau_z] \leq (1 + K_1')/\rho, \tag{7.14}$$

uniformly for $h_{\tau_z} \in C_h$. The above also indicates that if the initial condition is not zoomed, in finite time the process will reach a perfectly zoomed state almost surely. By the strong Markov property (x_{τ_z}, h_{τ_z}) is also a Markov chain as $\{\tau_z < n\} \in \mathcal{F}_n$, the filtration generated by the quantizer state and the quantizer output at time n. The probability that $\tau_{z+1} \neq \tau_z + 1$ is upper bounded by

$$\begin{aligned}
&P\bigg(\{|a|\Delta_{\tau_z}/2 + w_{\tau_z} \geq (2^{R'-1})\Delta_{\tau_z}\frac{1}{\alpha}\} \\
&\qquad\qquad \bigcup\{-|a|\Delta_{\tau_z}/2 + w_{\tau_z} \leq -(2^{R'-1})\Delta_{\tau_z}\frac{1}{\alpha}\}\bigg) \\
&= P\bigg((w_{\tau_z})^2 > (\frac{\Delta_{\tau_z}}{2}(\frac{2^{R'}}{\alpha} - |a|))^2\bigg) \\
&=: P_e(\Delta_{\tau_z}).
\end{aligned} \tag{7.15}$$

If $\tau_{z+1} \neq \tau_z + 1$, this means that the error is increasing on the average and the system is once again under-zoomed at time $t = \tau_z + 1$: $x_{\tau_z+1} = ax_{\tau_z} + w_{\tau_z}$ with $\Delta_{\tau_z+1} = \frac{|a|}{|a|+\epsilon-\eta}\Delta_{\tau_z}$ (when $\Delta_{\tau_z} \geq L$). With some positive probability, the quantizer will still be in the perfect zoom phase: $\tau_{z+1} = \tau_z + 1$. In case perfect-zoom is lost, there is a uniform bound on when the zoom is expected to be recovered. It follows that, conditioned on increment in the error, until the next stopping time, the process will increase exponentially, and hence

$$x_{\tau_{z+1}} = a^{\tau_{z+1}-\tau_z}(x_{\tau_z} + \sum_{t=0}^{\tau_{z+1}-\tau_z-1} a^{-t-1}w_{t+\tau_z}).$$

We now show that there exist $\psi > 0$, $|G| < \infty$ such that

$$E[\log(\Delta^2_{\tau_{z+1}})|\Delta_{\tau_z}, h_{\tau_z}] \leq \log(\Delta^2_{\tau_z}) - \psi + G1_{\{|\Delta_{\tau_z}| \leq F\}}. \tag{7.16}$$

Now, it follows that

$$\begin{aligned}
&E[\log(\Delta^2_{\tau_{z+1}})|\Delta_{\tau_z}, h_{\tau_z}] \\
&\leq (1 - P_e(\Delta_{\tau_z}))\bigg(2\log(\alpha) + \log(\Delta^2_{\tau_z})\bigg) \\
&\quad + P_e(\Delta_{\tau_z})E\bigg[2(\tau_{z+1} - \tau_z)\log(|a| + \delta) + 2\log(\Delta_{\tau_z})\bigg].
\end{aligned}$$

We now proceed to further upper bound $E[\log(\Delta^2_{\tau_{z+1}})|\Delta_{\tau_z}, h_{\tau_z}]$. Toward this end, we have

$$P\bigg(d^2_{\tau_z} > (\frac{\Delta_{\tau_z}}{2}(\frac{2^{R'}}{\alpha} - |a|))^2\bigg) \leq \frac{E[d^2_{\tau_z}]}{(\frac{\Delta_{\tau_z}}{2}(\frac{2^{R'}}{\alpha} - |a|))^2} \leq \frac{E[d^2_{\tau_z}]}{(K_2\Delta_{\tau_z})^2},$$

where we have used Markov's inequality, with $K_2 = (\frac{1}{2}(\frac{2^{R'}}{\alpha} - |a|))$. It then follows that $P_e(\Delta_{\tau_z}) \leq \frac{E[d_{\tau_z}^2]}{(K_2 \Delta_{\tau_z})^2}$. Given the uniform bound in (7.14), for (7.16) to hold, it suffices that the following equation is satisfied for large enough Δ_{τ_z} values, for some $\psi > 0$:

$$P_e(\Delta_{\tau_z})\left\{(2(1+K_1')/\rho)\log(|a|+\delta)\right\} + 2\log(\alpha) \leq -\psi < 0. \quad (7.17)$$

Thus, we have

$$E[\log(\Delta_{\tau_{z+1}}^2)|\Delta_{\tau_z}, h_{\tau_z}] \leq \log(\Delta_{\tau_z}^2) - \psi + G1_{\{|\Delta_{\tau_z}| \leq F'\}},$$

with

$$F' = \frac{\sqrt{E[w_1^2]}\sqrt{(2(1+K_1')/\rho)\log(|a|+\delta)}}{K_2\sqrt{-\psi - 2\log(\alpha)}},$$

and $G = 2\log(F') + 2((1+K_1')/\rho)\log(|a|+\delta) + \psi$. Hence, we have obtained another drift condition for the sampled Markov chain.

Together with (7.14), this leads to, by Theorem 6.2.6 (which follows from Theorem 6.2.4 and Corollary 6.2.1, except the irreducibility condition), the following result: The newly constructed process Δ_{τ_z} hits the set $\{\Delta_t : |\Delta_t| \leq F'\}$ infinitely often with finite expected return time. This is equivalent to x_{τ_z} hitting the set $C_x' = \{x : |x| \leq F := 2^{R'-1}F'\}$ with finite expected return time.

In the above, we assumed that the initial state is perfectly zoomed. If the initial condition is not in a perfect-zoom phase, with probability 1, in finite time the state process will move to this phase by (7.13) and the subsequent discussion. □

7.6.3 *Proof of Theorem 7.3.3*

Before proceeding further, let us first recall the following.

Lemma 7.6.2 (Bézout's Lemma). *[12] Let $\tilde{A} \in \mathbb{N}, \tilde{B} \in \mathbb{N}$. Let $\mathbb{I}$ be the set of all integers that can be obtained by summing positive integer multiples of elements in $\{-\tilde{A}, \tilde{B}\}$. If $\tilde{A}, \tilde{B}$ are relatively prime, then $\mathbb{I} = \mathbb{Z}$, that is, $\mathbb{I}$ is the set of all integers. If r is the greatest common divisor of $\{-\tilde{A}, \tilde{B}\}$, then I contains integer multiples of r.* ◇

We now show that the set of admissible quantizers forms a communication class under the hypothesis of the theorem: Since we have $\Delta_{t+1} = \bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|, \Delta_t)\Delta_t$, it follows that

$$\log_2(\Delta_{t+1})/s = \log_2(\bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|), \Delta_t)/s + \log_2(\Delta_t)/s$$

is also an integer. Furthermore, since the source process x_t is Lebesgue-irreducible (as the system noise admits a probability density function with positive mass on every nonempty open set) and there is a uniform lower bound L' on bin sizes, the error process takes values in any of the admissible quantizer bins with nonzero probability. Let the values taken by $\log_2(\bar{Q}(|\frac{x_t}{\Delta_t 2^{R'-1}}|, \Delta_t))/s$ be $\{-\tilde{A}, 0, \tilde{B}\}$, with $\tilde{A}, \tilde{B}$ relatively prime. Thus, for all $l, k \in \mathbb{Z}_+, l, k \geq \frac{\log_2(L')}{s}$, there exist $N_A, N_B \in \mathbb{Z}_+$ such that $l - k = -N_A\tilde{A} + N_B\tilde{B}$.

Consider two integers $k, l \geq \frac{\log_2(L')}{s}$. In particular, if at time 0, the quantizer is perfectly zoomed and $\Delta_0 = 2^{sk}$, then there exists a sequence of events consisting of $N_{\tilde{B}}$ zoom-out events and $N_{\tilde{A}}$ zoom-in events taking place with nonzero values.

Consider first the case where $k > \frac{\log_2(L')}{s} + N_A\tilde{A}$. We show that the probability of N_A occurrences of perfect zoom and N_B occurrences of under-zoom phases is bounded away from zero. This set of occurrences includes the event that in the first N_A time stages perfect zoom occurs and later, successively, N_B times the under-zoom phase occurs. The probability of the first sequence is lower bounded by

$$\left(P(w_t \in [-(\alpha 2^{R'} - |a|)\Delta_t/2, (\alpha 2^{R'} - |a|)\Delta_t/2])\right)^{N_A}.$$

The probability of the second sequence is lower bounded by the product of an under-zoom event, lower bounded by

$$P(x_t \in (\alpha 2^{R'-1}\Delta_t/2, \alpha 2^{R'-1}\Delta_t/2 + S])$$

with $\Delta_t = 2^{sk}$, for some $S > 0$, which is lower bounded by

$$P\left(w_t \in ((\alpha 2^{R'} + |a|)\Delta_t/2 + \Delta_t/2, (\alpha 2^{R'} - |a|)\Delta_t/2 + \Delta_t/2 + S)\right)^{N_A}$$

and the product of a sequence of under-zoom events, each lower bounded by

$$P\left(x_t \in (\alpha 2^{R'-1}\Delta_t/2, \alpha 2^{R'-1}\Delta_t/2 + S] \;\middle|\; x_{t-1} \in (\alpha 2^{R'-1}\Delta_{t-1}/2, \alpha 2^{R'-1}\Delta_{t-1}/2 + S]\right). \quad (7.18)$$

Hence, the probability of the sequence of under-zoom events is lower bounded by (with $\Delta_{N_A} = 2^{sk}$):

$$\begin{aligned}
&\prod_{n=1}^{N_B} P\Bigg(w_{N_A+n} \in \Bigg(((|a| + \delta)^n \Delta_{N_A} + |a|(((|a| + \delta)^{n-1}\Delta_{N_A} 2^{R'-1} + S), \\
&\qquad\qquad ((|a|+\delta)^n \Delta_{N_A} + |a|(((|a| + \delta)^{n-1}\Delta_{N_A} 2^{R'-1} + S) + S\Bigg)\Bigg) \\
&\times P(w_1 \in ((\alpha 2^{R'} + |a|)\Delta_{N_A}/2 + \Delta_{N_A}/2]), (\alpha 2^{R'} - |a|)\Delta_{N_A}/2 + \Delta_{N_A}/2) + S) \\
&> 0. \qquad (7.19)
\end{aligned}$$

A similar analysis can be performed when $k < \frac{\log_2(L')}{s} + N_A A$, by considering a sequence of events where periodically zoom-out and zoom-in events occur (this is useful for obtaining a nonzero lower bound on the events of perfect-zoom phases, once the state is under-zoomed) and finally, zoom-out events occur leading to a non-zero probability for the events. As such, for any two integers k, l and for some $p > 0$, $P(\log_2(\Delta_{t+p}) = ls | \log_2(\Delta_t) = ks) > 0$.

Now, for some probability measure $\mathcal{K}$ on positive integers, $E \subset \mathbb{R}$, and Δ an admissible bin size,

$$\sum_{n \in \mathbb{N}_+} \mathcal{K}(n) P\Big((x_n, \Delta_n) \in (E \times \{\Delta\}) \Big| x_0, \Delta_0\Big) \geq K_{\Delta_0, \Delta} \psi(E, \Delta).$$

Here $K_{\Delta_0, \Delta}$, denoting a lower bound on the probability of visiting Δ from Δ_0 in some finite time, is nonzero by (7.19) and ψ is a function to be further discussed below.

Let $t > 0$ be the time stage for which $\Delta_t = \Delta$ and thus by the construction in (7.19) and with $|h_{t-1}| \leq 1$: $|ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2 = \frac{\Delta_t}{|a|/(2^{R'} - \eta)}$. Thus, it follows that, for $A_1, B_1 \in \mathbb{R}$, $A_1 < B_1$,

$$\begin{aligned} & P\Big(x_t \in [A_1, B_1] \Big| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1}\Big) \\ & = P\Big(w_{t-1} \in [A_1 - (ax_{t-1} + bu_{t-1}), B_1 - (ax_{t-1} + bu_{t-1})] \\ & \qquad \Big| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1}\Big) \\ & \geq \min\Big(P(w_{t-1} \in [A_1 - \frac{\Delta}{2}(|a|/\alpha), B_1 - \frac{\Delta}{2}(|a|/\alpha)], \\ & \qquad P(w_{t-1} \in [A_1 + \frac{\Delta}{2}(|a|/\alpha), B_1 + \frac{\Delta}{2}(|a|/\alpha)]\Big) > 0. \end{aligned} \tag{7.20}$$

Define the finite set $C'_\Delta := \{\Delta : L' \leq |\Delta| \leq F', \frac{\log_2(\Delta)}{s} \in \mathbb{N}\}$. We now show that the recurrent set $C_x \times C'_\Delta$ is *petite* and hence small. We note that under aperiodicity and irreducibility (see Appendix C), every petite set is small and we will establish aperiodicity toward the end of the proof.

Now, the chain satisfies the recurrence property that $P_{(x,\Delta)}(\tau_{C_x \times C'_\Delta} < \infty) = 1$ for any admissible (x, Δ). This follows from (7.13) and the drift conditions in (7.16). Once a state which is perfectly zoomed, that is, satisfying $|x_t| \leq 2^{R'-1}\Delta_t$, is visited, the stopping time analysis can be used to verify that from any initial condition the recurrent set is visited in finite time with probability 1. In view of (7.19), we have that the chain is irreducible.

To establish petiteness, we establish the following uniform countable additivity condition (see Sect. C.1): Now, the set $C_x \times C'_\Delta$ satisfies (C.2), since for any given bin size Δ' in the countable space constructed above, we have that

$$\begin{aligned}
&\lim_{n\to\infty} \sup_{(x,\Delta)\in C_x \times C'_\Delta} P((x_{t+1}, \Delta_{t+1}) \in (B_n \times \Delta')|x_t = x, \Delta_t = \Delta) \\
&= \lim_{n\to\infty} \sup_{(x,\Delta)\in C_x \times C'_\Delta} P((ax + bu_t + w_t, \Delta_{t+1}) \in (B_n \times \Delta')|x_t = x, \Delta_t = \Delta) \\
&= \lim_{n\to\infty} \sup_{(x,\Delta)\in C_x \times C'_\Delta} P\Bigg((w_t, \Delta_{t+1}) \in \Bigg((B_n - (ax + bu_t)) \times \Delta'\Bigg) \\
&\qquad\qquad\qquad\qquad\qquad\qquad \Bigg| x_t = x, \Delta_t = \Delta\Bigg) \\
&= 0. \qquad\qquad (7.21)
\end{aligned}$$

This follows from the fact that the Gaussian random variable w_1 satisfies

$$\lim_{n\to\infty} \sup_{A_n} P(w_1 \in A_n) = 0,$$

for any sequence $A_n \downarrow \emptyset$, since a Gaussian measure admits a uniformly bounded density function.

If the integers $\tilde{A}, \tilde{B}$ are relatively prime, then by Lemma 7.6.2, the communication class will form the set of integers except those leading to $\Delta \leq L'$. We finally show that the Markov chain is aperiodic. This follows from the fact that the smallest admissible state for the quantizer, $\Delta^* = L'$, can be visited in subsequent time stages with non-zero probability, since

$$\min_{|x|\leq \Delta^*/2} P(w_t \in [-2^{R'-1}\Delta^* - x, 2^{R'-1}\Delta^* - x]) > 0.$$

Now, we can connect these results with Theorem 7.3.2, Theorem 6.2.4, and Corollary 6.2.1 and establish the positive Harris recurrence property for the chain and the existence of a unique invariant probability measure. □

7.6.4 Proof of Theorem 7.3.4

Toward the proof, we state a number of supporting results:

Lemma 7.6.3. *Let $z = 0, \tau_z = 0$ and $\tau_1 = \tau_{z+1} - \tau_z$. It follows that, for $k \in \mathbb{Z}_+$,*

$$P(\tau_1 > k|x_0, \Delta_0) \leq Ce^{-((\zeta^{k-1}N - 1/2)\Delta_0)^2/(2\sigma'^2)},$$

uniformly for $|h_0| \leq 1$ *with*

$$\sigma'^2 = \frac{E[w_1^2]}{1-|a|^{-2}}, \quad \zeta = \frac{|a|+\delta}{|a|}, \quad N = (2^{R'-1}\frac{\alpha}{|a|}),$$

$$C = \sigma' \frac{2}{\sqrt{2\pi}(2N-1)\Delta_0/2}.$$

◇

Proof. Let P denote the probability conditioned on x_0, Δ_0. It follows that for $k \geq 1$,

$$\begin{aligned}
&P(\tau_1 > k) \\
&= P\Big(\bigcap_{t=1}^{k}\{x_t \notin [-(|a|+\delta)^{t-1}2^{R'-1}\alpha\Delta_0, 2^{R'-1}(|a|+\delta)^{t-1}\alpha\Delta_0]\}\Big) \\
&\leq P\Big(x_k \notin [-(|a|+\delta)^{k-1}2^{R'-1}\alpha\Delta_0, 2^{R'-1}(|a|+\delta)^{k-1}\alpha\Delta_0]\Big) \qquad (7.22) \\
&= P\Big(a^k(x_0 + (b/a)u_0 + \sum_{i=0}^{k-1} a^{-i-1}w_i) \\
&\qquad\qquad \notin [-(|a|+\delta)^{k-1}2^{R'-1}\alpha\Delta_0, 2^{R'-1}(a+\delta)^{k-1}\alpha\Delta_0]\Big) \\
&= P\Big((x_0 + (b/a)u_0 + \sum_{i=0}^{k-1} a^{-i-1}w_i) \\
&\qquad\qquad \notin [-(\frac{|a|+\delta}{|a|})^{k-1}2^{R'-1}\frac{\alpha}{|a|}\Delta_0, 2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-1}\frac{\alpha}{|a|}\Delta_0]\Big) \\
&\leq 2P\Big(\sum_{i=0}^{k-1} a^{-i-1}w_i \geq (2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-1}\frac{\alpha}{|a|} - 1/2)\Delta_0\Big) \qquad (7.23) \\
&\leq Ce^{-((\zeta^{k-1}N-1/2)\Delta_0)^2/(2\sigma'^2)}, \qquad (7.24)
\end{aligned}$$

uniformly over $|h_0| \leq 1$ with σ'^2, ζ, N and C as defined in the statement of the lemma. Here, (7.22) follows from the chain property of a probability measure, (7.23) follows from the fact that $|ax_0 + bu_0| \leq |a|\Delta_0/2$, since the state is perfectly zoomed at time 0. Inequality (7.24) follows by bounding the complementary error function: For $u > 0$: $\int_u^\infty \mu(dx) \leq \int_u \frac{x}{u}\mu(dx)$, $\mu(.)$ being the Gaussian measure. □

Now, since a decaying exponential decays faster than a polynomial and the bound above is decreasing in Δ_0, there exists an $M < \infty$ such that, with $r > (|a|+\delta)^2$, and for all $\Delta_0 \geq L'$,

$$P(\tau_1 = k) < Ce^{-((\zeta^{k-2}N-1/2)\Delta_0)^2/(2\sigma'^2)} \leq Mr^{-k}, \quad \forall k \in \mathbb{Z}_+. \qquad (7.25)$$

Lemma 7.6.4. *If for some* $\gamma > 0$, $b < \infty$, *the following holds:*

$$\gamma E[\sum_{k=0}^{\tau_1 - 1} \Delta_k^2 | x_0, \Delta_0] \leq \Delta_0^2 - E[\Delta_{\tau_1}^2 | x_0, \Delta_0] + b 1_{\{(\Delta_0, h_0) \in (C'_x \times C_h)\}},$$

then $\lim_{t \to \infty} E[\Delta_t^2 | x_0, \Delta_0] < \infty$. ◇

The proof of this result follows from Theorem 6.2.4. Let us now note that (with a simple bounding argument in the last inequality)

$$\begin{aligned}
E[\sum_{t=0}^{\tau_1 - 1} \Delta_t^2 | x_0, \Delta_0] &= \sum_{l=1}^{\infty} P(\tau_1 = l) \sum_{k=0}^{l-1} E[\Delta_k^2 | \tau_1 = l, x_0, \Delta_0] \\
&\leq \Delta_0^2 \sum_{l=1}^{\infty} P(\tau_1 = l) \sum_{k=0}^{l-1} (|a| + \delta)^{2k} \\
&\leq \Delta_0^2 \left(\sum_{l=1}^{\infty} M(r^{-l}) \frac{(|a| + \delta)^{2l} - 1}{(|a| + \delta)^2 - 1} \right) \\
&\leq M \Delta_0^2 \left(\frac{1}{1 - r^{-1}(|a| + \delta)^2} - \frac{1}{1 - r^{-1}} \right) \frac{1}{(|a| + \delta)^2 - 1}.
\end{aligned}$$

Toward obtaining a bound on $E[\Delta_{\tau_1}^2 | x_0, \Delta_0]$,

$$\begin{aligned}
&E[\Delta_{\tau_1}^2 | x_0, \Delta_0] \\
&= P(\tau_1 = 1) E[\Delta_{\tau_1}^2 | \tau_1 = 1, x_0, \Delta_0] + P(\tau_1 > 1) E[\Delta_{\tau_1}^2 | \tau_1 > 1, x_0, \Delta_0] \\
&\leq P(\tau_1 = 1) \Delta_0^2 (\frac{|a|}{|a| + \epsilon - \eta})^2 + P(\tau_1 > 1) \sum_{k=2}^{\infty} P(\tau_1 = k) E[\Delta_k^2 | \tau_1 = k, x_0, \Delta_0] \\
&\leq P(\tau_1 = 1) \Delta_0^2 (\frac{|a|}{|a| + \epsilon - \eta})^2 + P(\tau_1 > 1) \sum_{k=2}^{\infty} M r^{-k} (|a| + \delta)^{2k} \Delta_0^2 \\
&\leq P(\tau_1 = 1) \Delta_0^2 (\frac{|a|}{|a| + \epsilon - \eta})^2 \\
&\qquad + P(\tau_1 > 1) M ((|a| + \delta)^2 r^{-1})^2 \frac{\Delta_0^2}{1 - r^{-1}(|a| + \delta)^2}.
\end{aligned}$$

Thus, we require, for some $\gamma > 0$, and sufficiently large Δ_0

$$\begin{aligned}
&\left(P(\tau_1 = 1) \Delta_0^2 \alpha^2 + P(\tau_1 > 1) M ((|a| + \delta)^2 r^{-1})^2 \frac{\Delta_0^2}{1 - r^{-1}(|a| + \delta)^2} \right) \\
&< -\gamma M \Delta_0^2 \left((\frac{1}{1 - r^{-1}(|a| + \delta)^2} - \frac{1}{1 - r^{-1}}) \frac{1}{(|a| + \delta)^2 - 1} \right) + \Delta_0^2.
\end{aligned}$$

Since $\lim_{\Delta_0 \to \infty} P(\tau_1 > 1) = 0$, by (7.17): $P_e(\Delta_{\tau_z}) \leq E[d^2_{\tau_z}]/(K_2 \Delta_{\tau_z})^2$, for some sufficiently small γ, say,

$$\gamma < \frac{(1-\alpha^2)}{\left(M(\frac{1}{1-r^{-1}(|a|+\delta)^2} - \frac{1}{1-r^{-1}}) \frac{1}{(|a|+\delta)^2-1} \right)},$$

the desired stability result follows for $\{\Delta_t\}$, that is, $\lim_{t \to \infty} E[\Delta_t^2 | x_0, \Delta_0] < \infty$. We now have the final result:

Lemma 7.6.5. *If* $R > \log_2(\lceil |a| + \epsilon \rceil + 1)$, *then* $\lim_{t \to \infty} E[x_t^2 | x_0, \Delta_0] < \infty$. ◇

Proof. First we observe that for some $\kappa > 0$,

$$\kappa E[\sum_{t=0}^{\tau_1 - 1} x_t^2 | x_0, \Delta_0] \leq (2^{2(R'-1)}) \Delta_0^2.$$

To see this, note that

$$E\left[\sum_{t=0}^{\tau_1-1} |x_t|^2 \,\middle|\, x_0, \Delta_0 \right] = E\left[\sum_{t=0}^{\infty} 1_{\{t<\tau_1\}} |x_t|^2 \,\middle|\, x_0, \Delta_0 \right]$$

$$\leq \sum_{t=0}^{\infty} \left(E[(1_{\{t<\tau_1\}})^{1+\chi} | x_0, \Delta_0] \right)^{\frac{1}{1+\chi}} \left(E[|x_t|^{2(\frac{1+\chi}{\chi})} | x_0, \Delta_0] \right)^{\frac{\chi}{1+\chi}}, \tag{7.26}$$

for some $\chi > 0$, by Hölder's inequality. Now, $E[|x_t|^{2(\frac{1+\chi}{\chi})} | x_0, \Delta_0] = E[|a^t(x_0 + (\sum_{i=0}^{t-1} a^{-i-1} w_i))|^{2(\frac{1+\chi}{\chi})}] \leq B_2(\Delta_0^2 2^{2(R'-1)})^{\frac{1+\chi}{\chi}} |a|^{2t(\frac{1+\chi}{\chi})}$ for some $B_2 < \infty$. Here, we use the fact the random variable $(h_0 + \frac{\sum_{i=0}^{\infty} a^{-i-1} w_i}{2^{R'-1} \Delta_0})$ has a Gaussian distribution, with its expected fixed moments uniform on $\Delta_0 \geq L'$. Thus,

$$E[\sum_{t=0}^{\tau_1-1} |x_t|^2 | x_0, \Delta_0] \leq (\Delta_0^2 2^{2(R'-1)}) \sum_{t=0}^{\infty} \left(P(\tau_1 \geq t+1 | x_0, \Delta_0) \right)^{\frac{1}{1+\chi}} \left(B_2^{\frac{\chi}{1+\chi}} |a|^{2t} \right)$$
$$< \zeta_{B_2} (2^{R'-1} \Delta_0)^2,$$

for $\zeta_{B_2} < \infty$.

Hence, for some $\epsilon > 0$, by taking

$$\delta(x, \Delta) = \epsilon \Delta^2, \quad f(x, \Delta) = \frac{\epsilon}{\zeta_{L_2} 2^{2(R'-1)}} x^2,$$

C a compact set and $V_2(x, \Delta) = \Delta^2$, Theorem 6.2.4 applies and $\lim_{t \to \infty} E[x_t^2] < \infty$. □

Thus, with $R > \log_2(\lceil |a| + \epsilon \rceil + 1)$, stability with a finite second moment is achieved. Finally, the limit is independent of the initial distribution since the chain is irreducible, by Theorem 7.3.3. This completes the proof of the theorem. □

7.6.5 Proof of Theorem 7.3.5

Proof follows from essentially the same steps as in the proof of Theorem 7.3.4. With τ_z as defined, we can obtain the following for $k > 1$:

$$\begin{aligned}
&P(\tau_1 > k) \\
&= P\Big(\bigcap_{t=1}^{k} \{x_t \notin [-(|a|+\delta)^{t-1} 2^{R'-1} \alpha \Delta_0, 2^{R'-1}(|a|+\delta)^{t-1}\alpha\Delta_0]\} \Big) \\
&\leq P\Big(x_k \notin [-(|a|+\delta)^{k-1} 2^{R'-1} \alpha \Delta_0, 2^{R'-1}(|a|+\delta)^{k-1}\alpha\Delta_0] \Big) \qquad (7.27) \\
&\leq 2P\Big(\sum_{i=0}^{k-1} a^{-i-1} w_i \geq (2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-1}\frac{\alpha}{|a|} - 1/2)\Delta_0 \Big) \qquad (7.28) \\
&\leq 2 \frac{E[(\sum_{i=0}^{k-1} a^{-i-1} w_i)^{2+\epsilon}]}{\Big((2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-1}\frac{\alpha}{|a|} - 1/2)\Delta_0 \Big)^{2+\epsilon}} \qquad (7.29) \\
&\leq M r^{-(2+\epsilon)k}, \qquad (7.30)
\end{aligned}$$

for M and $r \in (\rho\frac{|a|+\delta}{|a|}, \frac{|a|+\delta}{|a|})$ with $\rho < 1$ and which can be made arbitrarily close to 1. In (7.29), we apply Markov's inequality. (7.30) follows due to the geometric expression in the denominator in (7.29) and a polynomial bound in k for $E[(\sum_{i=0}^{k-1} a^{-i-1} w_i)^{2+\epsilon}]$. Now, with $r^{-1}(|a|+\delta)^2 < 1$, by taking, if necessary, δ sufficiently large, we can obtain the following:

$$\begin{aligned}
&E\Big[\sum_{t=0}^{\tau_1 - 1} |x_t|^2 \Big| x_0, \Delta_0 \Big] = E\Big[\sum_{t=0}^{\infty} 1_{\{t<\tau_1\}} |x_t|^2 \Big| x_0, \Delta_0 \Big] \\
&\leq \sum_{t=0}^{\infty} \Big(E[(1_{\{t<\tau_1\}})^{1+\chi} | x_0, \Delta_0] \Big)^{\frac{1}{1+\chi}} \Big(E[|x_t|^{2(\frac{1+\chi}{\chi})} | x_0, \Delta_0] \Big)^{\frac{\chi}{1+\chi}}, \qquad (7.31)
\end{aligned}$$

for some $\chi > 0$, by Hölder's inequality. Now,

$$E[|x_t|^{2(\frac{1+\chi}{\chi})} | x_0, \Delta_0] \leq B_2 (\Delta_0^2 2^{2(R'-1)})^{\frac{1+\chi}{\chi}} |a|^{2t(\frac{1+\chi}{\chi})},$$

for some $B_2 < \infty$. Thus,

$$E[\sum_{t=0}^{\tau_1-1}|x_t|^2|x_0,\Delta_0]\leq(\Delta_0^2 2^{2(R'-1)})\sum_{t=0}^{\infty}\Big(P(\tau_1 \geq t+1|x_0,\Delta_0)\Big)^{\frac{1}{1+\chi}}\Big(B_2^{\frac{\chi}{1+\chi}}|a|^{2t}\Big)$$
$$< \zeta_{B_2}(2^{R'-1}\Delta_0)^2$$

for some finite ζ_{B_m}. Hence, as before, with some $\epsilon > 0$ and

$$\delta(x,\Delta) = \epsilon\Delta^2, \quad f(x,\Delta) = \frac{\epsilon}{\zeta_{B_2}2^{2(R'-1)}}|x|^2,$$

C a compact set, and $V_2(x,\Delta)=\Delta^2$, Theorem 6.2.4 applies and $\lim_{t\to\infty} E[|x_t|^2]<\infty$. □

7.6.6 Proof of Theorem 7.3.6

The proof follows closely that of Theorem 7.3.4. First observe that

$$E\left[\sum_{t=0}^{\tau_1-1}|x_t|^m \,\middle|\, x_0,\Delta_0\right] = E\left[\sum_{t=0}^{\infty}1_{\{t<\tau_1\}}|x_t|^m \,\middle|\, x_0,\Delta_0\right]$$
$$\leq \sum_{t=0}^{\infty}\Big(E[(1_{\{t<\tau_1\}})^{1+\chi}|x_0,\Delta_0]\Big)^{\frac{1}{1+\chi}}\Big(E[|x_t|^{m(\frac{1+\chi}{\chi})}|x_0,\Delta_0]\Big)^{\frac{\chi}{1+\chi}}, \quad (7.32)$$

for some $\chi > 0$, by Hölder's inequality. As before, $E[|x_t|^{m(\frac{1+\chi}{\chi})}|x_0,\Delta_0] \leq B_m(\Delta_0^m 2^{m(R'-1)})^{\frac{1+\chi}{\chi}}|a|^{mt(\frac{1+\chi}{\chi})}$ for some $B_m < \infty$.

Thus,

$$E[\sum_{t=0}^{\tau_1-1}|x_t|^m|x_0,\Delta_0]\leq(\Delta_0^m 2^{m(R'-1)})\sum_{t=0}^{\infty}\Big(P(\tau_1\geq t+1|x_0,\Delta_0)\Big)^{\frac{1}{1+\chi}}\Big(B_m^{\frac{\chi}{1+\chi}}|a|^{mt}\Big)$$
$$<\zeta_{B_m}(2^{R'-1}\Delta_0)^m$$

for some finite ζ_{B_m}. With $\epsilon > 0$, with

$$\delta(x,\Delta) = \epsilon\Delta^m, \quad f(x,\Delta) = \frac{\epsilon}{\zeta_{B_m}2^{2(R'-1)}}|x|^m,$$

C a compact set, and $V_m(x,\Delta) = \Delta^m$, Theorem 6.2.4 applies and $\lim_{t\to\infty} E[|x_t|^m]<\infty$. □

7.6.7 Proof of Theorem 7.4.1

We assume, without any loss of generality, that all the eigenvalues are unstable. Let P denote P_{x_0,Δ_0}. We first note that, by an application of the union bound,

$$\begin{aligned}
&P(\tau_1 > k) \leq P(\cup_{k=1}^{n}(|h_k^i| > 1)|\text{zoom until} \quad k)) \\
&\leq \sum_{k=1}^{n} P(|h_k^i| > 1|\text{zoom until} \quad k).
\end{aligned}$$

Consider a two-dimensional Jordan block example considered in (7.8). We have already obtained a bound on $P(|h_k^2| > 1)$ in (7.22), as this mode evolves free from upper modes. In the following, since $|x_0^1 - \hat{x}_0^1| < \Delta_0^1/2$, we will, by an abuse of notation, let $|x_0^1| \leq \Delta_0^1/2$ and, likewise, $|x_0^2| \leq \Delta_0^2/2$. Let $\delta^1 = \delta^2 = \delta$. For $P(|h_k^1| > 1)$, we have the following:

$$\begin{aligned}
&P(|h_k^1| > 1) \\
&= P\Big(x_k^1 \notin [-(|\lambda| + \delta)^{k-1} 2^{R_1'-1} \alpha \Delta_0^1, 2^{R_1'-1}(|\lambda| + \delta)^{k-1} \alpha \Delta_0^1]\Big) \\
&= P\Big(\lambda^k (x_0^1 + \sum_{i=0}^{k-1} \lambda^{-i-1}(w_i^1 + x_i^2)) \\
&\qquad\qquad \notin [-(|\lambda| + \delta)^{k-1} 2^{R_1'-1} \alpha \Delta_0^1, 2^{R_1'-1}(a + \delta)^{k-1} \alpha \Delta_0^1]\Big) \\
&= P\Big(\lambda^k (x_0^1 + \sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1} \lambda^{-i-1}(\lambda^i x_0^2 + \sum_{j=0}^{i-1} \lambda^{i-j-1} w_j^2)) \\
&\qquad\qquad \notin [-(|\lambda| + \delta)^{k-1} 2^{R_1'-1} \alpha \Delta_0^1, 2^{R_1'-1}(\lambda + \delta)^{k-1} \alpha \Delta_0^1]\Big) \\
&= P\Big(|\lambda^k (x_0^1 + \sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1} \lambda^{-i-1}(\lambda^i x_0^2 + \sum_{j=0}^{i-1} \lambda^{i-j-1} w_j^2))| \\
&\qquad\qquad > (|\lambda| + \delta)^{k-1} 2^{R_1'-1} \alpha \Delta_0^1\Big) \\
&= P\Big(\Big|(x_0^1 + \sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1} \lambda^{-i-1}(\lambda^i x_0^2 + \sum_{j=0}^{i-1} \lambda^{i-j-1} w_j^2))\Big| \\
&\qquad\qquad > (\frac{|\lambda| + \delta}{|\lambda|})^{k-1} \frac{\alpha}{|\lambda|} 2^{R_1'-1} \Delta_0^1\Big)
\end{aligned}$$

$$
\begin{aligned}
&\leq P\Big(|\sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1}\sum_{j=0}^{i-1} \lambda^{-j-2} w_j^2))| \\
&> (\frac{|\lambda| + \delta}{|\lambda|})^{k-1} \frac{\alpha}{|\lambda|} 2^{R_1'-1} \Delta_0^1 - \Delta_0^1/2 - k\lambda^{-1}\Delta_0^2 \Big) \\
&\leq P\Big(|(\sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1}\sum_{j=0}^{i-1} \lambda^{-j-2} w_j^2)| \\
&\qquad > \Delta_0^1/2\Big((\frac{|\lambda| + \delta}{|\lambda|})^{k-1} \frac{\alpha}{|\lambda|} 2^{R_1'} - (1 + k|\lambda|^{-1}\eta_\Delta) \Big) \Big). \qquad (7.33)
\end{aligned}
$$

Under Assumption 7.2.2, the bound in (7.33) can be used to obtain

$$
\begin{aligned}
&P\Big(|(\sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1}\sum_{j=0}^{i-1} \lambda^{-j-2} w_j^2)|^{2+\epsilon} \\
&\qquad > \Big(\Delta_0^1/2(\frac{|\lambda| + \delta}{|\lambda|})^{k-1} \frac{\alpha}{|\lambda|} 2^{R_1'} - (1 + k|\lambda|^{-1}\eta_\Delta) \Big)^{2+\epsilon} \Big). \qquad (7.34)
\end{aligned}
$$

Now, it can be shown that $|(\sum_{i=0}^{k-1} \lambda^{-i-1} w_i^1 + \sum_{i=0}^{k-1}\sum_{j=0}^{i-1} \lambda^{-j-2} w_j^2)|^{2+\epsilon} \leq C_B k^{2(2+\epsilon)}$ for some $C_B < \infty$. By an application of Markov's inequality and with the condition $R_1' > \log_2(\frac{|a|}{\alpha})$, this leads to an expression similar to (7.30). Thus, the drift criteria and the analysis in Sect. 7.6.4 follows, and the finite moment conditions are satisfied.

A parallel argument applies for dimensions higher than 2, leading to the desired conclusion. For a complete proof, see [209]. □

7.7 Concluding Remarks

In this chapter, we considered constructions of quantizers leading to stochastic stability for an open-loop unstable system driven by noise. In particular, we showed that such quantizers are rate efficient, in addition to being simple and fixed rate.

7.8 Bibliographic Notes

There has been a significant amount of research on quantizer design for networked control systems. We will discuss and review relevant contributions in detail in the next chapter.

Zooming-type adaptive quantizers have been considered by Goodman and Gersho [166] and later in a control context by Brockett and Liberzon [81] (see also [234] among many others) for remote stabilization of open-loop unstable, noise-free systems with arbitrary initial conditions. There is a large body of literature on quantizer design in the communications and information theory communities. One key reference on adaptive quantization is the work by Goodman and Gersho [166], where an adaptive quantizer was introduced and its stochastic stability was investigated when the source fed to the quantizer is a second-order i.i.d. sequence. Zooming-type quantizers of Brockett and Liberzon form a special class of the adaptive quantization scheme of Goodman and Gersho. Kieffer and Dunham [215] have studied the stochastic stability of a number of coding schemes when the source fed to the quantizer is also stochastically stable, but not necessarily i.i.d. For the setting of this chapter, however, the schemes in [166] and [215] are not directly applicable, as the process considered here is open-loop unstable (as well as Markovian).

Nair and Evans [280] considered stability under the assumption that the quantizer is variable rate and showed that for a setup with system noise which has unbounded support for its probability measure, on the average it suffices to use an average rate given in Theorem 5.6.1 per channel use to achieve a form of stability, when the channel is noiseless. They used asymptotic quantization theory to obtain a time-varying scheme, where the quantizer is used at certain intervals at a very high rate, and at other time stages, the quantizer is not used. The authors show that there exists a policy such that $\limsup_{t\to\infty} E[|x_t|^2] < \infty$. Reference [419] presented an approach which uses fixed rate, meeting the lower bounds presented in Theorem 5.6.1, and which leads to positive Harris recurrence of the state with finite moments, that is, $\lim_{t\to\infty} E[x_t^2] < \infty$. Higher-dimensional settings have been considered in [209].

Part of this chapter is based on [419, 421] and [209].

Chapter 8
Stochastic Stabilization Over Noisy Channels

8.1 Introduction

In this chapter, we discuss several generalizations of the results of Chap. 7 and consider stochastic stabilization of linear systems over various types of noisy channels: these are erasure channels, general discrete channels (memoryless as well as with memory) with noiseless feedback, and a class of discrete and continuous channels without any feedback (Gaussian channels will be considered in detail in Chap. 11). We will present strong forms of stability and ergodicity, which will come at the expense of further complexities in the design and analysis when compared with noiseless channels considered earlier.

As in Chap. 7, this chapter identifies conditions on the channels for which there exist coding and control policies such that a controlled linear system with state x_t is stochastically stable in one or more of the following senses:

- The state $\{x_t\}$ and the coding and control parameters lead to a positive Harris recurrent Markov chain.
- $\{x_t\}$ is asymptotically mean stationary (AMS) and satisfies Birkhoff's sample path ergodic theorem (see Definition C.3.5 and Appendix C for an overview of ergodic theory).
- $\lim_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} |x_t|^2$ exists and is finite almost surely (this will be referred to as *quadratic stability*).

We will obtain a tight converse theorem on stabilizability in the AMS sense over a large class of noisy channels. We will see that the Shannon capacity provides the appropriate measure on whether a system can be stabilized or not under arbitrary admissible coding and control policies.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_8,

We then establish constructive proofs for various channels. Our stabilization analysis starts with the erasure channel, which is perhaps the simplest nontrivial noisy channel. This channel is also important in many engineering applications, since it is used as a common model, such as in control over communication networks arising in many industrial systems. Our treatment of the erasure channel allows us to obtain stronger results regarding stochastic stability of linear systems when compared with general noisy channels considered later in the chapter. We will see that the Shannon capacity of the erasure channel is an almost sufficient condition for the positive Harris recurrence of the state and the quantizer process. For quadratic (finite second moment) stabilization, however, the conditions on the channel are more stringent. Similar conclusions will be arrived at in the context of more general discrete memoryless channels (DMCs).

Every imperfect information transmission problem features a nonclassical information pattern. However, an information structure in a setup for control over a noisy channel with noiseless, instantaneous feedback leads to settings where the information at the receiver is nested in that at the encoder. For such a system, we saw in Chap. 7 that one can obtain stochastic stability results by establishing stopping times measurable with respect to the information at the receiver, as well as the controller. In this chapter, for the cases with noiseless feedback, we will define appropriate stopping times and extend the proof program in Chap. 7 to such a context.

The presence of a noisy channel with noisy feedback brings up further challenges since the agents (encoders, decoders, and controllers) do not have nested information. In this case, the *dual role* of control is present, as the control policy might affect the estimation error of the controller with respect to the state of the system. In the chapter, we study problems in these settings also and establish stochastic stability results. We consider stabilization of open-loop unstable linear time-invariant (LTI) stochastic systems when communications between the plant and the controller (forward communication), and between the controller and the plant (reverse communication) are conducted over noisy channels which are either discrete memoryless or continuous memoryless. A new coding scheme allowing a version of the random-time state-dependent drift considered in Chap. 6 to be applicable will be presented, and differences between finite-alphabet and continuous-alphabet channels will be highlighted.

The treatment in this chapter starts with stabilization over channels with noiseless feedback in Sect. 8.2, where we obtain converse theorems on stabilizability over a noisy channel. We then consider stabilization over erasure channels in Sect. 8.3, and later more general DMCs in Sect. 8.4, and establish ergodicity and quadratic stability results using the theory developed in Chap. 6. Channels with memory and multidimensional systems are investigated in Sect. 8.5. Section 8.6 investigates the case with noisy reverse channels without any feedback. An appendix, constituting Sect. 8.7, includes proofs of some of the main results.

8.2 Stabilization Over Noisy Channels with Noiseless Feedback and a Converse Theorem

8.2.1 Control and Communication Model

We first consider a scalar LTI discrete-time system, with the multidimensional case relegated to Sect. 8.5. Here, the scalar system is described by

$$x_{t+1} = ax_t + bu_t + w_t, \quad t \in \mathbb{Z}_+, \tag{8.1}$$

where x_t is the state at time t, u_t is the control input, the initial state x_0 is a zero-mean second-order random variable, and $\{w_t\}$ is a sequence of zero-mean i.i.d. Gaussian random variables, also independent of x_0. We assume that the system is open-loop unstable and controllable, that is, $|a| \geq 1$ and $b \neq 0$.

This system is connected over a noisy channel to a controller, as shown in Fig. 8.1. The controller has access to the information it has received through the channel. A source coder maps the source symbols, state values, to corresponding channel inputs. The quantizer outputs are transmitted through a channel, after passing through a channel encoder. We recall here Definitions 5.2.1 and 5.2.3.

The receiver has access to noisy versions of the quantizer/coder outputs for each time, which we denote by $q'_t \in \mathcal{M}'$. The quantizer and the channel encoder policies are causal so that the channel input at time t, q_t, is generated using the information vector I^e_t available at the encoder:

$$I^e_t = \{I^e_{t-1}, x_t, q_{t-1}, q'_{t-1}\}, \quad t \in \mathbb{N}, \quad I^e_0 = \{x_0\}.$$

That is, the encoder has access to noiseless feedback from the channel output. The control policy at time t, also causal, is measurable with respect to the sigma-algebra generated by I^c_t:

$$I^c_t = \{I^c_{t-1}, q'_t\}, \quad t \in \mathbb{N}, \quad I^c_0 = \emptyset,$$

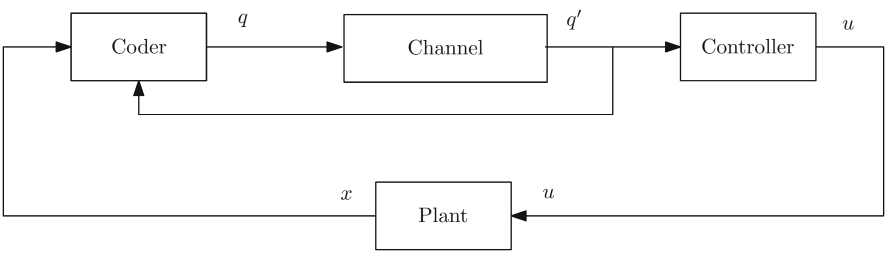

Fig. 8.1 Control of a system over a noisy channel

and is a mapping to $\mathbb{R}$. We assume that the source model, channel model, the probabilistic description of the noise variables, and the initial state distribution ν_0 are available to both the quantizer/channel encoder and the controller.

We will call such coding and control policies *admissible*.

We start our analysis with a converse theorem. This will be discussed first in the context of DMCs.

8.2.2 Converse Theorem on Stochastic Stability Over a Discrete Memoryless Channel

We have the following converse theorem, which generalizes Theorem 5.6.1. This result is further generalized to a class of channels with memory in Theorem 8.5.2.

Theorem 8.2.1 ([422]). *Suppose that a linear plant given as in (8.1) controlled over a DMC, under some admissible coding and controller policy, satisfies the condition*

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0, \tag{8.2}$$

where h denotes the entropy function. Then, the channel capacity C must satisfy

$$C \geq \log_2(|a|).$$

◇

Proof. This is a special case of Theorem 8.5.2 which deals with the multidimensional case. □

Remark 8.2.1. The condition (8.2) is a weak one. For example, a stochastic process whose second moment grows subexponentially in time, $\liminf_{T\to\infty} \frac{\log(E[x_T^2])}{T} \leq 0$, satisfies this condition. ◇

We now present a supporting result due to Matveev.

Proposition 8.2.1 ([260]). *Suppose that a linear plant given as in (8.1) is controlled over a DMC. If*

$$C < \log_2(|a|),$$

then

$$\limsup_{T\to\infty} P(|x_T| \leq b(T)) \leq \frac{C}{\log_2(|a|)},$$

for all $b(T) > 0$ such that $\lim_{T\to\infty} \frac{1}{T}\log_2(b(T)) = 0$. ◇

Proof. See the proof of Proposition 8.5.1 in Sect. 8.7.2, which considers the more general setting of channels with memory. □

With this lemma at hand, we can prove the following.

Theorem 8.2.2 ([424]). *Suppose that a linear plant given as in (8.1) is controlled over a DMC. If, under some causal encoding and controller policy, the state process is AMS, then the channel capacity C must satisfy*

$$C \geq \log_2(|a|).$$

◇

Proof. See the proof of Theorem 8.5.3 in Sect. 8.7.3. □

In the following sections, we will observe that the condition $C \geq \log_2(|a|)$ in Theorems 8.2.1 and 8.2.2 is almost sufficient as well for stability in the AMS sense. Toward this goal, we first discuss the erasure channel with feedback and then consider more general DMCs, followed by a class of channels with memory. We will also investigate quadratic stability.

8.3 Stochastic Stabilization Over Erasure Channels with Feedback

For the linear system (8.1), we first consider a particular and important class of DMCs, the erasure channel, in the following.

The details of the setup considered are specified as follows: The channel source consists of state values from $\mathbb{R}$. The source output is, as before, quantized. We use the same set of quantizers as in the previous chapter, defined by (5.1), repeated below:

$$Q_K^{\Delta}(x) = \begin{cases} (k - \frac{1}{2}(K+1))\Delta, & \text{if } x \in [(k-1-\frac{1}{2}K)\Delta, (k-\frac{1}{2}K)\Delta), \\ (\frac{1}{2}(K-1))\Delta, & \text{if } x = \frac{1}{2}K\Delta, \\ 0, & \text{if } x \notin [-\frac{1}{2}K\Delta, \frac{1}{2}K\Delta]. \end{cases} \tag{8.3}$$

The quantizer outputs are transmitted through a memoryless erasure channel, after being subjected to a bijective mapping, which is performed by the channel encoder. The channel encoder maps the quantizer output symbols to corresponding channel inputs $q \in \mathcal{M} := \{1, 2 \ldots, K+1\}$. A channel encoder at time t, denoted by $\mathcal{E}_t$, maps the quantizer outputs to $\mathcal{M}$ such that $\mathcal{E}_t(Q_t(x_t)) = q_t \in \mathcal{M}$.

The controller/decoder has access to noisy versions of the encoder outputs for each time, which we denote by $\{q'\} \in \mathcal{M} \cup \{e\}$, with e denoting the erasure symbol, generated according to a probability distribution for every fixed $q \in \mathcal{M}$. The channel transition probabilities are given by

$$P(q' = i | q = i) = p, \quad P(q' = e | q = i) = 1 - p, \quad i \in \mathcal{M}.$$

At each time $t \in \mathbb{Z}_+$, the controller/decoder applies a mapping $\mathcal{D}_t : \mathcal{M} \cup \{e\} \to \mathbb{R}$, given by

$$\mathcal{D}_t(q'_t) = \mathcal{E}_t^{-1}(q'_t) \times 1_{\{q'_t \neq e\}} + 0 \times 1_{\{q'_t = e\}}.$$

Let $\{\Upsilon_t\}$ denote a binary sequence of i.i.d. random variables, representing the erasure process in the channel, where the event $\Upsilon_t = 1$ indicates that the signal is transmitted with no error through the channel at time t. Let $p = E[\Upsilon_t]$ denote the probability of success in transmission.

The following key assumptions are imposed throughout this section: Given $K \geq 2$ introduced in the definition of the quantizer, define the *rate variable*s

$$R := \log_2(K+1) \quad R' = \log_2(K), \tag{8.4}$$

We fix positive scalars δ, α satisfying

$$|a| 2^{-R'} < \alpha < 1, \tag{8.5}$$

and

$$\alpha(|a| + \delta)^{p^{-1}-1} < 1. \tag{8.6}$$

Similar to (7.5) and (7.6), we consider the following update rules. For $t \in \mathbb{Z}_+$ and with $\Delta_0 \in \mathbb{R}$ selected arbitrarily, consider:

$$\begin{aligned} u_t &= -\frac{a}{b}\hat{x}_t, \\ \hat{x}_t &= \mathcal{D}_t(q'_t) = \Upsilon_t Q_K^{\Delta_t}(x_t), \\ \Delta_{t+1} &= \Delta_t \bar{Q}(\Delta_t, |\frac{x_t}{\Delta_t 2^{R'-1}}|, \Upsilon_t). \end{aligned} \tag{8.7}$$

Here, $\bar{Q} : \mathbb{R} \times \mathbb{R} \times \{0,1\} \to \mathbb{R}$ is defined below, where $L > 0$ is a constant:

$$\begin{aligned} \bar{Q}(\Delta, h, p) &= |a| + \delta \quad \text{if} \quad |h| > 1, \quad \text{or} \quad p = 0, \\ \bar{Q}(\Delta, h, p) &= \alpha \quad \text{if} \quad 0 \leq |h| \leq 1, p = 1, \Delta > L, \\ \bar{Q}(\Delta, h, p) &= 1 \quad \text{if} \quad 0 \leq |h| \leq 1, p = 1, \Delta \leq L \end{aligned}$$

The update equations above imply that

$$\Delta_t \geq L\alpha =: L'. \tag{8.8}$$

Without any loss of generality, we assume that $L' \geq 1$.

We note that given the channel output $q'_t \neq e$, the controller can simultaneously deduce the realization of Υ_t and the event $\{|h_t| > 1\}$, where $h_t := \frac{x_t}{\Delta_t 2^{R'-1}}$. This is due to the fact that if the channel output is not the erasure symbol, the controller knows that the signal is received with no error. If $q'_t = e$, however, then the controller applies 0 as its control input and enlarges the bin size of the quantizer.

Stochastic Stability and Positive Harris Recurrence

Lemma 8.3.1. *Under (8.7), the process* $(x_t, \varDelta_t)$ *is a Markov chain.* ◇

Proof. The system's state evolution can be expressed as

$$x_{t+1} = ax_t - a\hat{x}_t + w_t,$$

where $\hat{x}_t = \varUpsilon_t Q_K^{\varDelta_t}(x_t)$. It follows that the process $(x_t, \varDelta_t)$ evolves as a nonlinear state space model:

$$\begin{aligned} x_{t+1} &= a(x_t - \varUpsilon_t Q_K^{\varDelta_t}(x_t)) + w_t, \\ \varDelta_{t+1} &= \varDelta_t \bar{Q}(\varDelta_t, |\frac{x_t}{2^{R'-1}\varDelta_t}|, \varUpsilon_t), \end{aligned} \tag{8.9}$$

in which $(w_t, \varUpsilon_t)$ is i.i.d.. Thus, the pair $(x_t, \varDelta_t)$ forms a Markov chain. □

The following result is established in the Appendix to the chapter, based on the stochastic stability results of Chap. 6.

Proposition 8.3.1 ([439]). *If* (8.4) *holds, then there exists a compact set* $A \times B \subset \mathbb{R}^2$ *satisfying the finite-mean return property*

$$\sup_{(x,\varDelta)\in A\times B} E_{x,\varDelta}[\tau_{A\times B}] < \infty,$$

and the recurrence condition $P_{(x,\varDelta)}(\tau_{A\times B} < \infty) = 1$ *for any admissible* $(x, \varDelta)$. ◇

A result on the existence and uniqueness of an invariant probability measure is the following. It basically establishes irreducibility and aperiodicity, which leads to positive Harris recurrence, by Proposition 8.3.1.

Theorem 8.3.1 ([439]). *For an adaptive quantizer satisfying* (8.4)*, suppose that the quantizer bin sizes are such that their base-2 logarithms are integer multiples of some scalar s, and* $\log_2(\bar{Q}(\,\cdot\,))$ *takes values in integer multiples of s. Then the process* $(x_t, \varDelta_t)$ *forms a positive Harris recurrent Markov chain. If the integers taken are relatively prime (i.e., they share no common divisors except* 1*), then the invariant probability measure is independent of the value of the integer multiplying s.* ◇

We note that the (Shannon) capacity of such an erasure channel is given by $\log_2(K+1)p$ [103]. From (8.4) to (8.6), the following is obtained.

Theorem 8.3.2. *If* $\log_2(K)p > \log_2(|a|)$*, then* α, δ *exist such that the conditions of Theorem 8.3.1 are satisfied.* ◇

Remark 8.3.1. Thus, the Shannon capacity of the erasure channel is an almost sufficient condition for the positive Harris recurrence of the state and the quantizer process. We will see that under the weaker notions of ergodicity or asymptotic mean stationarity, this result applies to a large class of memoryless channels and a class of channels with memory and feedback (see Theorem 8.4.3): There is a direct relationship between asymptotic mean stationarity and the Shannon capacity of the channel used in the system. As we shall see, however, for quadratic or finite moment stability, Shannon capacity is typically not sufficient. ◇

Quadratic Stability

Under slightly stronger conditions, we obtain a finite second moment:

Theorem 8.3.3 ([439]). *Suppose that the assumptions of Theorem 8.3.1 hold, and in addition the following bound holds:*

$$a^2\left(1 - p + \frac{p}{(2^R - 1)^2}\right) < 1. \tag{8.10}$$

Then, for each initial condition (x_0, Δ_0)*,*

$$\lim_{t\to\infty} E[x_t^2] = E_\pi[x_0^2] < \infty\,.$$

◇

By the positive Harris recurrence property, the above holds also in a sample-path sense almost surely, leading to quadratic stability.

The above theorem also has a converse in the sense that the rate condition in (8.10) is tight (up to the transmission of an additional symbol).

Theorem 8.3.4 ([273]). *A necessary condition for mean square stability is*

$$a^2\left(1 - p + \frac{p}{2^{2R}}\right) < 1.$$

◇

Proof. By the property that the Gaussian measure maximizes the entropy over the set of all random variables with a fixed variance, we have

$$E[x_t^2 | q'_{[0,t-1]}] \geq \frac{1}{2\pi e} 2^{2h(a^t x_0 | q'_{[0,t-1]})}. \tag{8.11}$$

Conditioned on the event L_k of k successful transmissions at $t \geq k$ transmissions, there are kR bits at the controller by time t. Since $kR \geq I(x_0; q'_{[0,t-1]}, L_k) = h(x_0) - h(x_0 | q'_{[0,t-1]}, L_k)$, it follows that $2^{2h(a^t x_0 | q'_{[0,t-1]}, L_k)} \geq a^{2t} 2^{2h(x_0)} / 2^{2kR}$. Conditional entropy is an average of these realized conditional entropies (see Definition 5.3.2). A necessary condition for the boundedness of the second moment is then $\liminf_{t\to\infty} E[\frac{a^{2t}}{2^{2R_t}}] < 1$, where R_t is the total number of bits successfully received by time t. By considering combinatorially the events of erasures up until time t, we obtain

$$a^{2t}\sum_{k=0}^{t}\binom{t}{k}p^k(1-p)^{t-k}\frac{1}{2^{2Rk}} = (a^2(\frac{p}{2^{2R}} + 1 - p))^{2t} < 1,$$

as a necessary condition. □

Thus, the sufficiency condition in Theorem 8.3.3 almost meets this bound except for the additional symbol transmitted for the under-zoom events. We note that the average rates can be made arbitrarily close to the rate required in (8.3.4) by sampling the control system with larger periods. Such a relaxation of the sampling period, however, would lead to a process which is not Markov, but an AMS process which is quadratically stable.

We now consider the mth moment case. In this case, the proof is almost identical to that of Theorem 8.3.3 and is thus omitted.

Theorem 8.3.5 ([439]). *Consider the scalar system in (8.1). Let $m \in \mathbb{N}$, suppose that the assumptions of Theorem 8.3.1 hold, and in addition we have the inequality,*

$$|a|^m\left(1 - p + \frac{p}{(2^R-1)^m}\right) < 1.$$

Then, with the adaptive quantization policy considered and given the initial condition (x_0, Δ_0),

$$\lim_{t\to\infty} E[|x_t|^m] = E_\pi[|x_0|^m] < \infty\,.$$

◇

8.3.1 Connections with Random-time Drift Criteria

As it was done in Sect. 7.3.3, we point out the connections between the results above and the random-time drift criteria.

The process (x_t, Δ_t) will once again form a Markov chain. Now, in the model considered, the controller can receive meaningful information regarding the state of the system when two events occur concurrently: the channel carries information with no error, and the source lies in the granular region of the quantizer, that is, $x_t \in [-\frac{1}{2}K\Delta_t, \frac{1}{2}K\Delta_t)$ and $\Upsilon_t = 1$. The times at which both of these events occur form an increasing sequence of random stopping times, defined as

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k| \le 1, \Upsilon_k = 1\}, \quad z \in \mathbb{N}.$$

In the proofs of the stability theorems, we will apply Theorem 6.2.4 for these stopping times. These are the times when information reaches the controller regarding the value of the state when the state is in the granular region of the quantizer. As in Fig. 7.1, we introduce $C_x = \{x : |x| \le F\}$, $C_h = \{h : |h| \le 1\}$, for some sufficiently large F value, which will serve as a small set as well as a recurrent set.

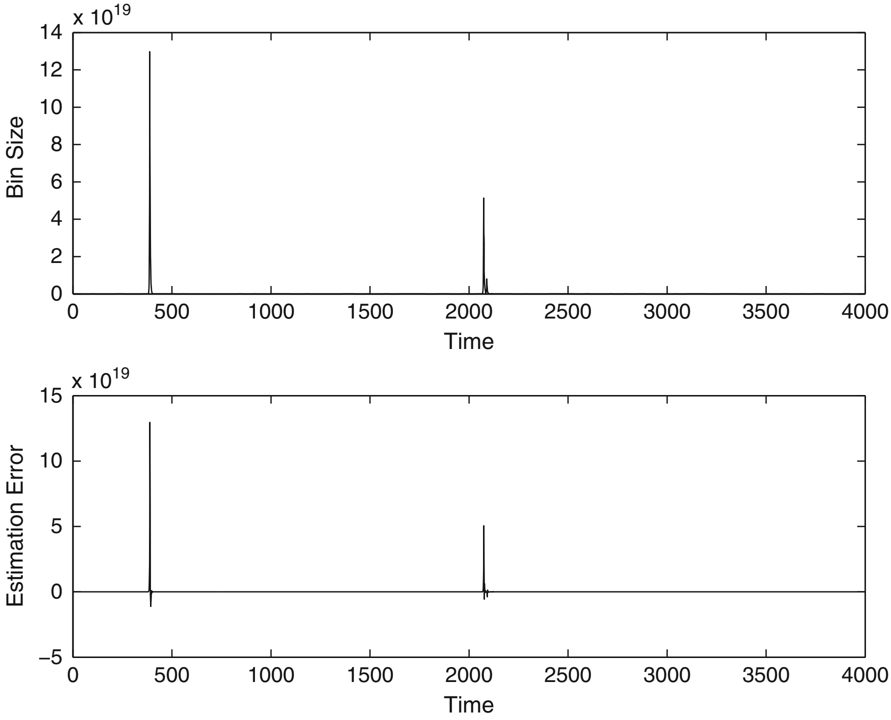

Fig. 8.2 Sample path for a stochastically stable system with a 5-bin quantizer

In view of Theorem 6.2.4, first without an irreducibility assumption, we will establish recurrence of the set $C_x \times C_h$ by defining a Lyapunov function of the form $V(x_t, \Delta_t) = \frac{1}{2}\log_2(\Delta^2) + B_0$ for some $B_0 > 0$. Later we will establish the irreducibility of the Markov chain by imposing a countability condition. The details are given in the appendix.

8.3.2 Simulation

Consider the linear system (8.1) with $a = 2.5$, $b = 1$, and $\{w_t\}$ a sequence of i.i.d., zero-mean Gaussian variables with $E[w_t^2] = 1$. The erasure channel has error probability $1 - p = 0.1$. For stability with a finite second moment, we employ a quantizer with rate

$$\log_2(\lceil \sqrt{\frac{p}{\frac{1}{a^2} - (1-p)}} \rceil + 1) = \log_2(5)$$

bits or a uniform quantizer with 5 bins. Here, we have taken $L' = 1$. Figures 8.2 and 8.3 illustrate the stochastic stability results presented in Theorems 8.3.1 and 8.3.3. The plots show the under-zoomed and perfectly zoomed phases, with the peaks in the plots showing the under-zoom phases. For the plot with 5 levels, the system is positive Harris recurrent, since the update equations are such that $\alpha = 0.629, \delta = 0.025$, and $\log_2(\bar{Q}(\cdot)) \in \{-0.6744, 0, 1.363\}$. These values satisfy the irreducibility condition since $-0.6744 = -(1/2)1.363$, and the communication conditions are satisfied. Furthermore,

$$\alpha(|a| + \delta)^{\frac{1}{p}-1} < 0.698 < 1.$$

One notable aspect is that increasing the bit rate by only two bits in Fig. 8.3 leads to a much more desirable sample path. Moreover, by increasing the rate, the severity of rare events is reduced. In view of these observations, the sensitivity of performance to the bit rate emerges as a relevant problem.

8.4 Stochastic Stabilization Over DMCs with Feedback

In this section, we consider more general DMCs. The construction will be more tedious, but the essence of the analysis for the erasure channel applies to this setting as well.

We first note that the condition $C \geq \log_2(|a|)$ in Theorem 8.2.1 is almost sufficient for strong forms of stability.

Theorem 8.4.1 ([422]). *For the existence of a compact coordinate recurrent set (see Definition C.3.4), the following is sufficient: The channel capacity C satisfies: $C > \log_2(|a|)$.* ◇

Proof. See Sect. 8.7.7. □

We now consider the following update algorithm which is used in the proof of Theorem 8.4.1. Let n be a given block length. Consider a class of uniform quantizers, defined by two parameters, with bin size $\Delta > 0$, and an even number $K(n) \geq 2$ (see Fig. 5.1). Define the uniform quantizer as follows: For $k = 1, 2 \ldots, K(n)$,

$$Q^{\Delta}_{K(n)}(x) = \begin{cases} (k - \frac{1}{2}(K(n)+1))\Delta, & \text{if } x \in [(k-1-\frac{1}{2}K(n))\Delta, (k-\frac{1}{2}K(n))\Delta), \\ (\frac{1}{2}(K(n)-1))\Delta, & \text{if } x = \frac{1}{2}K(n)\Delta, \\ \mathcal{Z}, & \text{if } x \notin [-\frac{1}{2}K(n)\Delta, \frac{1}{2}K(n)\Delta]. \end{cases}$$

where $\mathcal{Z}$ is the overflow symbol in the quantizer. Let $\{x : Q^{\Delta}_{K(n)}(x) \neq \mathcal{Z}\}$ be the *granular region* of the quantizer.

At every sampling instant $t = kn, k = 0, 1, 2, \ldots$, the source coder $\mathcal{E}^s_t$ quantizes output symbols in $\mathbb{R} \cup \{\mathcal{Z}\}$ to a set $\mathcal{M}(n) = \{1, 2, \ldots, K(n)+1\}$. A channel encoder $\mathcal{E}^c_t$ maps the elements in $\mathcal{M}(n)$ to corresponding channel inputs

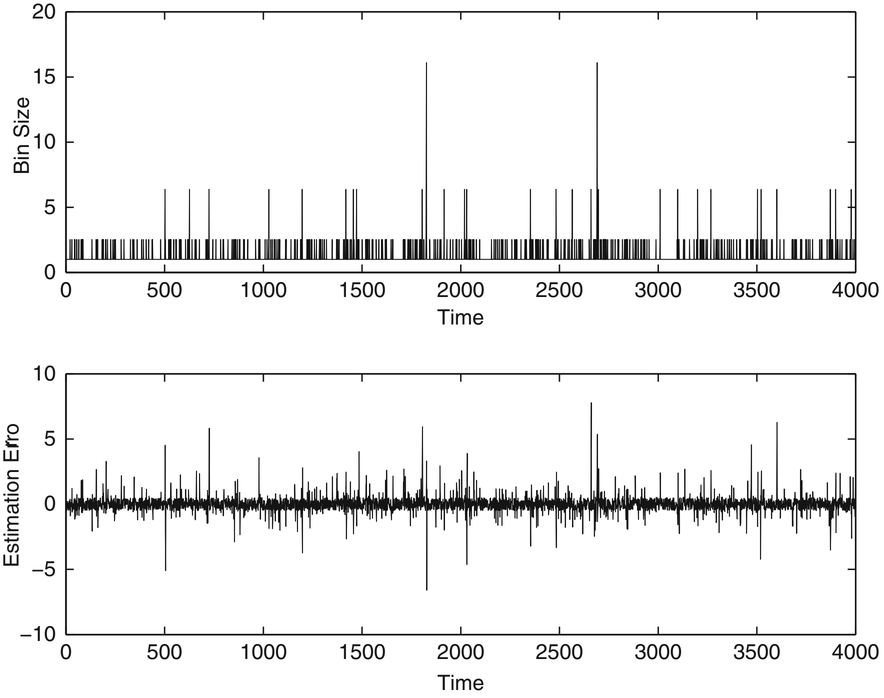

Fig. 8.3 Sample path with a 17-bin quantizer, a more desirable path

$q_{[kn,(k+1)n-1]} \in \mathcal{M}^n$. For each time $t = kn - 1, k = 1, 2, 3, \ldots$, the channel decoder applies a mapping $\mathcal{D}_{tn} : \mathcal{M}'^n \to \mathcal{M}(n)$, such that

$$c'_{(k+1)n-1} = \mathcal{D}_{kn}(q'_{[kn,(k+1)n-1]}).$$

Finally, the controller runs an estimator:

$$\hat{x}_{kn} = (\mathcal{E}^s_{kn})^{-1}(c'_{(k+1)n-1}) \times 1_{\{c'_{(k+1)n-1} \neq \mathcal{Z}\}} + 0 \times 1_{\{c'_{(k+1)n-1} = \mathcal{Z}\}}.$$

Hence, when the decoder output is the overflow symbol, the estimation output is 0.

As in the previous two chapters, at time kn the bin size Δ_{kn} is taken to be a function of the previous state $\Delta_{(k-1)n}$ and the past n channel outputs. Further, the encoder has access to the previous channel outputs, thus making such a quantizer implementable at both the encoder and the decoder.

With $K(n) > \lceil |a|^n \rceil$, $R = \log_2(K(n) + 1)$, let us introduce $R'(n) = \log_2(K(n))$ and let

$$R'(n) > n \log_2(\frac{|a|}{\alpha}),$$

for some $\alpha, 0 < \alpha < 1$ and $\delta > 0$. When clear from the context, we will drop the index n in $R'(n)$. We will consider the following update rules in the controller actions and the quantizers. For $t \geq 0$ and with $\Delta_0 > L$ for some $L \in \mathbb{R}_+$, and $\hat{x}_0 \in \mathbb{R}$, consider, for $t = kn$, $k \in \mathbb{N}$,

$$\begin{aligned} u_t &= -1_{\{t=(k+1)n-1\}} \frac{a^n}{b} \hat{x}_{kn}, \\ \Delta_{(k+1)n} &= \Delta_{kn} \bar{Q}(\Delta_{kn}, c'_{(k+1)n-1}), \end{aligned} \tag{8.12}$$

where c' denotes the decoder output variable. If we use $\delta > 0$ and $L > 0$ such that

$$\begin{aligned} \bar{Q}(\Delta, c') = (|a| + \delta)^n \quad & \text{if} \quad c' = \mathcal{Z}, \\ \bar{Q}(\Delta, c') = \alpha^n \quad & \text{if} \quad c' \neq \mathcal{Z}, \Delta \geq L, \\ \bar{Q}(\Delta, c') = 1 \quad & \text{if} \quad c' \neq \mathcal{Z}, \Delta < L, \end{aligned} \tag{8.13}$$

we will show in Sect. 8.7.7 that a recurrent set exists. Note that the above implies that $\Delta_t \geq L\alpha^n =: L'$ for all $t \geq 0$.

Thus, we have three main events: When the decoder output is the overflow symbol, the quantizer is zoomed out (with a coefficient of $(|a| + \delta)^n$). When the decoder output is not the overflow symbol $\mathcal{Z}$, the quantizer is zoomed in (with a coefficient of α^n) if the current bin size is greater than or equal to L, and otherwise the bin size does not change.

In the following, we make the quantizer bin size set countable and as a result establish the irreducibility of the sampled process (x_{tn}, Δ_{tn}).

Theorem 8.4.2. *For an adaptive quantizer satisfying the conditions of Theorem 8.4.1, suppose that the quantizer bin sizes are such that their logarithms are integer multiples of some scalar s, and $\log_2(\bar{Q}(\,\cdot\,))$ takes values in integer multiples of s. Suppose the integers taken are relatively prime (that is they share no common divisors except for 1). Then the sampled process (x_{tn}, Δ_{tn}) forms a positive Harris recurrent Markov chain at sampling times on the set of admissible quantizer bins and state values.* ◇

Proof. See Sect. 8.7.8. □

Theorem 8.4.3. *Under the conditions of Theorems 8.4.1 and 8.4.2, the process $\{x_t, \Delta_t\}$ is n-stationary, n-ergodic, and hence AMS. That is, if the channel capacity C satisfies: $C > \log_2(|a|)$, there exists a coding and control policy such that the process $\{x_t, \Delta_t\}$ is n-stationary, n-ergodic, and AMS.* ◇

Proof. The proof follows from the observation that a positive Harris recurrent Markov chain is recurrent and stationary. It uses the property that if a sampled process is a positive Harris recurrent Markov chain and if the intersampling time is fixed, with a time-homogenous update in the intersampling times, then the process is mixing, $n-$ergodic, and $n-$stationary. Details are given in Sect. 8.7.9. □

Quadratic Stability and Finite Second Moment

We now discuss quadratic and finite moment stability (see Definition C.3.7). For a given coding scheme with block length n, a message set $\mathcal{M}(n) = \{1, 2, \ldots, K(n)+1\}$, and a decoding function $\gamma : \mathcal{M}'^n \to \{1, 2, \ldots, K(n)+1\}$, we have three types of errors:

- Type I-A: Error from a granular symbol to another granular symbol. A bound for such an error is

$$P^e_{g|g}(n) := \max_{c \in \mathcal{M}(n) \setminus \mathcal{Z}} P(\gamma(q'_{[0,n-1]}) \neq c, \gamma(q'_{[0,n-1]}) \neq \mathcal{Z} | c \quad \text{is transmitted})$$

- Type I-B: Error from a granular symbol to $\mathcal{Z}$

$$P^e_{g|g}(n) := \max_{c \in \mathcal{M}(n) \setminus \mathcal{Z}} P(\gamma(q'_{[0,n-1]}) = \mathcal{Z} | c \quad \text{is transmitted})$$

- Type II: Error from $\mathcal{Z}$ to a granular symbol

$$P^e_{g|\mathcal{Z}}(n) := P(\gamma(q'_{[0,n-1]}) \neq \mathcal{Z} | \mathcal{Z} \quad \text{is transmitted})$$

Type II error will be shown to be crucial in the analysis of the error exponent. Types I-A and I-B will play an important role in establishing the drift properties. The following theorem captures the main results for establishing quadratic stability.

Theorem 8.4.4 ([422]). *A sufficient condition for quadratic stability (for the joint (x_t, Δ_t) process) over a DMC is that:*

$$\lim_{n \to \infty} (\frac{1}{n} \log(P^e_{\mathcal{Z}|g}(n)) + 2\log(|a| + \delta) < 0,$$

$$\lim_{n \to \infty} (\kappa \frac{1}{n} \log(P^e_{g|\mathcal{Z}}(n)) + 2\log(|a| + \delta) < 0,$$

$$\lim_{n \to \infty} (\kappa \frac{1}{n} \log(P^e_{g|g}(n)) + 2\log(|a| + \delta) + 2\kappa \log(\alpha) < 0,$$

$$R'(n) > n \log_2(|a|/\alpha)$$

and

$$\kappa < \frac{1}{\log_{\frac{|a|+\delta}{|a|}}(\frac{|a|+\delta}{\alpha})}.$$

◇

Proof. See Sect. 8.7.10. □

Let

$$\bar{P}_e(n) := \max_{c \in \mathcal{M}(n)} P(\gamma(q'_{[0,n-1]}) \neq c | c \quad \text{is transmitted}).$$

When the block length is clear from context, we drop the explicit dependence on n. We have the following corollary to Theorem 8.4.4.

Corollary 8.4.1. *A sufficient condition for quadratic stability (for the joint* (x_t, Δ_t) *process) over a DMC is:*

$$\lim_{n\to\infty} (\kappa \frac{1}{n} \log(\bar{P}_e(n)) + 2\log(|a| + \delta) < 0,$$

with rate $R'(n) > n\log_2(\frac{|a|}{\alpha})$. ◇

Remark 8.4.1. For a DMC with block length n, Shannon's random coding [150] leads to

$$P_e(n) \leq e^{-nE(R)+o(n)},$$

uniformly for all codewords $c \in \{1, 2, \ldots, \mathcal{M}(n)\}$ with c' being the decoder output (thus, the random exponent also applies uniformly over the set). Here $\frac{o(n)}{n} \to 0$ as $n \to \infty$ and $E(R) > 0$ for $0 < R < C$. Thus, under the above conditions, the exponent under random coding should satisfy $E(R) > \frac{2\log_2(|a|+\delta)}{\kappa}$. ◇

Remark 8.4.2 (Converse Results for Quadratic Stability). For quadratic stability over erasure channels, we observed in Theorem 8.3.4 that a converse theorem exists, and the proposed scheme achieves the converse result. For general DMCs, however, a tight converse result on quadratic stabilizability is not yet available. One primary reason is that the error exponents of fixed length block codes with noiseless feedback for general DMCs are not currently known. We note here that the channel reliability or error exponent of DMCs is typically improved with feedback, unlike the capacity of DMCs. Some partial results have been reported in [118] (in particular, the sphere packing upper bound is tight for a class of symmetric channels for rates above a critical rate even with feedback), see Chap. 10 of [107] among others. In Remark 8.6.2, a discussion on converse results for quadratic stability, which also applies to the setting in this section, is given. Related references addressing partial results include [267, 268] which consider lower bounds on estimation error moments for transmission of a single variable over a noisy channel (in the context of this chapter, this single variable may correspond to the initial state x_0). A further related notion for quadratic stability is the notion of *any-time capacity* introduced by Sahai and Mitter (see [328, 331]) which is discussed further in Sect. 8.9. ◇

Zero-Error Transmission for $\mathcal{Z}$

An important practical setup would be the case when $\mathcal{Z}$ is transmitted with no error and is not confused with messages from the granular region. The following captures this.

Assumption 8.4.1. $P^e_{\mathcal{Z}|g}(n) = P^e_{g|\mathcal{Z}}(n) = 0$ *for* $n \geq n_0$ *for some* $n_0 \in \mathbb{N}$. ◇

Theorem 8.4.5. *Under Assumption 8.4.1, a sufficient condition for quadratic stability is*

$$\lim_{n\to\infty} (\bar{P}_e(n))(|a| + \delta)^{2n} < 1,$$

with rate $R'(n) > n\log_2(\frac{|a|}{\alpha})$ *and* $\kappa > 1/2$. ◇

Proof. See Sect. 8.7.11. □

It is worth emphasizing that the reliability of sending the symbol $\mathcal{Z}$ for the under-zoom phase allows a relaxation in the overall channel reliability requirements.

8.5 Channels with Memory and Multidimensional Sources

Definition 8.5.1. Channels are said to be of **Class A** type, if:

(i) they satisfy the following Markov chain condition:

$$q'_t \leftrightarrow q_t, q_{[0,t-1]}, q'_{[0,t-1]} \leftrightarrow \{x_0, w_t, t \geq 0\},$$

for all $t \geq 0$, and

(ii) their capacity with feedback is given by

$$C = \lim_{T\to\infty} \max_{\{P(q_t|q_{[0,t-1]}, q'_{[0,t-1]}), \quad 0\leq t\leq T-1\}} \frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}),$$

where the directed mutual information is defined by

$$I(q_{[0,T-1]} \to q'_{[0,T-1]}) = \sum_{t=1}^{T-1} I(q_{[0,t]}; q'_t | q'_{[0,t-1]}) + I(q_0; q'_0).$$

◇

DMCs naturally belong to this class. For DMCs, feedback does not increase the capacity [103]. Such a class also includes finite state stationary Markov channels which are indecomposable [306], and non-Markov channels which satisfy certain symmetry properties [106]. Further examples can be found in [111, 357].

Theorem 8.5.1 ([422]). *Suppose that a linear plant given by (8.1) is controlled over a **Class A** type noisy channel with feedback. If the channel capacity (with feedback) is less than* $\log_2(|a|)$*, then*

(i) The following condition:

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0$$

cannot be satisfied under any policy

(ii) The state process cannot be AMS under any policy.

◇

Proof. This is a special case of Theorems 8.5.2 and 8.5.3, and hence its proof is postponed until later. □

Remark 8.5.1. The result above is *negative*, but one can also obtain a positive one: If the channel capacity is greater than $\log_2(|a|)$ and there is a positive error exponent (uniform over all transmitted messages, as in Theorem 14 of [306]), then there exists a coding scheme leading to an AMS state process provided that the channel restarts itself with the transmission of every new block (either independently or as a Markov process). ◇

Remark 8.5.2. If the channel is not information stable, then information spectrum methods lead to pessimistic realizations of capacity (known as the $\liminf$ *in probability* of the normalized information density, see [357, 379]). We do not consider such channels in this book, although the approach here is generalizable to some cases when the channel state is Markov and the worst-case initial input state is considered as in [306]. ◇

Higher-Order Plants

The result for the scalar problem has a natural counterpart in the multidimensional setting. Consider the linear system described by

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{8.14}$$

where $x_t \in \mathbb{R}^N$ is the state at time t, $u_t \in \mathbb{R}^m$ is the control input, and $\{w_t\}$ is a sequence of zero-mean i.i.d. $\mathbb{R}^d$-valued Gaussian random vectors. Here A is the square system matrix with at least one eigenvalue greater than or equal to 1 in magnitude, that is, the system is open-loop unstable. Furthermore, (A, B) and (A, G) are controllable pairs.

In the following we assume that all eigenvalues $\{\lambda_i, 1 \leq i \leq N\}$ of A are unstable, that is, have magnitudes greater than or equal to 1. There is no loss here since if some eigenvalues are stable, by a similarity transformation, the unstable modes can be decoupled from the stable ones and one can instead consider a lower-dimensional system; stable modes are already recurrent.

Consider a multidimensional linear system as in (8.14) with all eigenvalues unstable, that is, $|\lambda_i| \geq 1$ for $i = 1, \ldots, N$. We have the following results:

Theorem 8.5.2. *For such a system controlled over a **Class A** type noisy channel with feedback, if the channel capacity (with feedback) satisfies*

$$C < \sum_i \log_2(|\lambda_i|),$$

there does not exist a stabilizing coding and control scheme with the property $\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0$. ◇

Proof. See Sect. 8.7.1. □

Proposition 8.5.1. *For such a system controlled over a **Class A** type noisy channel with feedback, if*

$$C < \log_2(|A|),$$

then,

$$\limsup_{T\to\infty} P(|x_T| \leq b(T)) \leq \frac{C}{\log_2(|A|)} > 0,$$

for all $b(T) > 0$ *such that* $\lim_{T\to\infty} \frac{1}{T} \log_2(b(T)) = 0$. ◇

Proof. See Sect. 8.7.2 □

With this lemma, we state the following.

Theorem 8.5.3. *Consider such a system controlled over a **Class A** type noisy channel with feedback. If there exists some encoding and controller policy so that the state process is AMS, then the channel capacity (with feedback) C must satisfy*

$$C \geq \log_2(|A|).$$

◇

Proof. See Sect. 8.7.3 □

For sufficiency, we will assume that A is a diagonalizable matrix (a sufficient condition for which is that its eigenvalues are distinct real).

Theorem 8.5.4 ([422]). *Consider a multidimensional system with a diagonalizable matrix* A. *If the Shannon capacity of the DMC used in the controlled system satisfies*

$$C > \sum_{|\lambda_i|>1} \log_2(|\lambda_i|),$$

there exists a stabilizing scheme in the AMS sense. ◇

Proof. See Sect. 8.7.12. □

Regarding channels with memory, the discussions in Remark 8.5.1 also apply for this setting.

Remark 8.5.3. Theorem 8.5.4 can be extended to the case where the matrix A is not diagonalizable, in the same spirit as in Theorem 7.4.1, by constructing stopping times in view of the coupling between modes sharing a common eigenvalue. ◇

8.6 Stabilization with Noisy Forward and Feedback/Reverse Channels

In the spirit of the coverage so far in the chapter, we consider in this section stabilization of open-loop unstable LTI stochastic systems when communications between the plant and the controller (forward communication), and the controller and the plant (reverse communication) are conducted over noisy channels which are either discrete memoryless or continuous memoryless. A new coding scheme allowing a version of the state-dependent drift considered in Chap. 6 to be applicable is presented. To facilitate the analysis, we will consider continuous-time systems.

8.6.1 *Formulation*

We consider here the class of LTI continuous-time scalar systems described by

$$dx'_{t'} = (\mu x'_{t'} + b' u'_{t'})dt' + dB_{t'}, \quad t' \geq 0, \tag{8.15}$$

where $x'_{t'}$ is the state; $B_{t'}$, $t' \geq 0$, is the standard Brownian motion process; $u'_{t'}$ is the (applied) control, which is assumed to be piecewise constant over intervals of length T_s (which is a constant sampling period); the initial state x_0 is Gaussian; and $\mu > 0$, thus making the open-loop system unstable. After sampling, with period T_s, we have the discrete-time system

$$x_{t+1} = a x_t + b u'_t + w_t\,, \quad t = 0, 1, \ldots\,, \tag{8.16}$$

where t is the discrete-time variable, defined through the relationship $t' = tT_s$; $x_t = x'_{tT_s}$ is the state x' at the sampling times; $\{w_t\}$ is a sequence of zero-mean i.i.d. Gaussian random variables; $a = e^{\mu T_s}$; $b = b'(e^{\mu T_s} - 1)/\mu$; and $E[w_t^2] = (e^{2\mu T_s} - 1)/2\mu$.

We refer to the channel carrying the signal from the plant to the controller as the *forward channel* and the one carrying the signal from the controller to the plant as the *reverse (feedback) channel* (see Fig. 8.4). The plant is controlled over the reverse channel and the controller receives information over the forward channel, both of which are noisy.

We consider here both discrete and continuous (alphabet) channels, which are memoryless.

When the forward channel is a DMC, we will let M_f denote the set of sensor symbols, with $|M_f|$ denoting its cardinality, and N_f be the number of channel uses. Then, the coding rate for the forward channel is defined as $R_f = \log_2(|M_f|)/N_f$. Likewise, for the reverse channel, for the DMC case, the coding rate is given by $R_r = \log_2(|M_r|)/N_r$ with obvious corresponding meanings for M_r and N_r. The plant output (state) is quantized (by the source coder) and turned into a bit

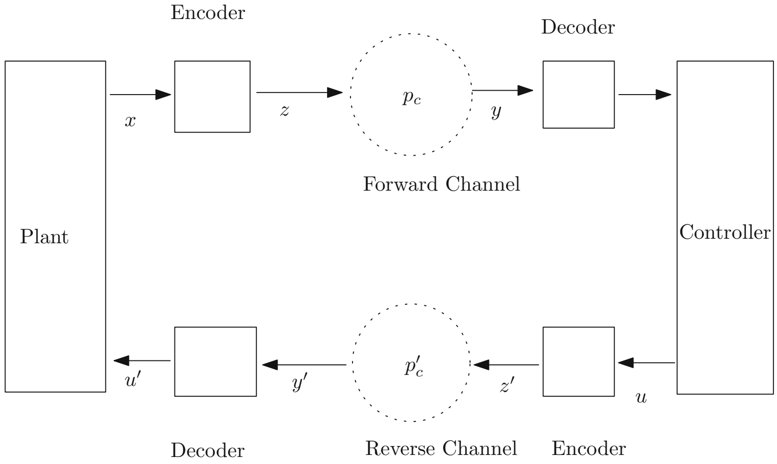

Fig. 8.4 Control over noisy forward and reverse channels

stream for each quantization symbol (by the channel encoder) before being inputed to the forward channel (see Fig 8.4). Likewise, the controller output is quantized and turned into a bit stream before being inputed to the reverse channel.

We note that the analysis to be carried out for the scalar system is applicable to multivariable systems through a block coding approach or a sequential stabilization approach as was discussed in Chap. 7 and Remark 7.4.1.

Referring to Fig. 8.4, in this setup both the sensor and the controller act as both transmitters and receivers because of the closed-loop structure. For the DMC case, we model the forward source-channel encoder as a stochastic kernel $p_s(z_t|x_t), x_t \in \mathbb{R}$, $z_t \in \mathcal{Z}$ (with $\mathcal{Z}$ being the channel alphabet), between the source output and the channel input; hence $p_s(z_t|x_t)$ is a collection of (conditional) probability mass functions parametrized by $x_t \in \mathbb{R}$. The forward channel is a memoryless stochastic kernel, $p_c(y_t|z_t)$, between the channel input and the channel output, where $y_t \in \mathcal{Y}$, the output channel alphabet. The channel output is acted on by the controller in a *memoryless* fashion, so that we have another well-defined stochastic kernel, $p(u_t|y_t)$, which is the probability for control at time t to be $u_t \in \mathcal{U}$ given that the output of the forward channel at time t is $y_t \in \mathcal{Y}$.

The reverse channel also has a source-channel encoder, $p'_s(z'_t|u_t), z'_t \in \mathcal{Z}'$, channel mapping $p'_c(y'_t|z'_t), y'_t \in \mathcal{Y}'$, and a channel decoder $p'_d(u'_t|y'_t), u'_t \in \mathcal{U}'$. Appropriate adjustments are made to the interpretations of these different stochastic kernels in the case of continuous memoryless channels (CMCs).

For the DMC case, a quantizer is used to obtain a countable representation of the input source. Here, the quantizer bins, $\mathcal{B}_i$, are taken to be non-overlapping semi-open intervals, $\mathcal{B}_i = [\delta_i, \delta_{i+1})$ for $i > 0$, with $\delta_i < \delta_{i+1}, i = 0, 1, 2, \ldots$, such that δ_0 is at the origin, where $\{\delta_i\}$ are termed "bin edges." We consider "symmetric quantizers," in the sense that if $(\delta_i, \delta_{i+1}]$ is a quantization bin, where $0 < \delta_i < \delta_{i+1}$, then $\mathcal{B}_{-i} := (-\delta_{i+1}, -\delta_i]$ is also a quantization bin. We define the encodable state set $S_x \subset \mathbb{R}$ as the set of elements which are represented by some codeword; $S_x := \bigcup_i \mathcal{B}_i$. Such a definition applies to the encodable control set, S_c, as well. Suppose that the state is within the encodable state set and is in the ith bin of the quantizer. The source-coding output at the plant sensor will represent this state as q_i and send the ith index over the channel. After a joint mapping of the channel and the channel decoder, the controller will receive the transmitted index i as index j with probability $p(j|i)$. The controller will apply its control over index j, computing Q'_j. Thus, the controller-decoder and the controller-encoder can be regarded as a single (composite) mapping. The controller transmits the control signal through the reverse channel to the plant which would interpret this value as Q'_l with probability $p'(l|j)$, by a mapping through the reverse channel. Given that the state is in the ith bin, the plant will receive the control Q'_l with probability $\sum_j p'(l|j)p(j|i)$. Thus, the applied control will be $u'_t = Q'_l$ with probability $\sum_j p'(l|j)p(j|i)$, the probability of the state being in the ith bin being $p(i) = \text{Prob}(x \in \mathcal{B}_i)$. For CMCs, however, we do not use a quantizer; we denote the joint channel encoder and the channel as a stochastic kernel, $p(A|x)$, for $x \in \mathbb{R}$, and $A \in \mathcal{B}(\mathbb{R})$. The control is a deterministic function of the channel decoder, mapping $\mathbb{R}$ into $\mathbb{R}$. The control signals are sent back to the plant, via a reverse channel encoder and a reverse channel. Upon the arrival of the reverse channel output, the plant decoder generates the decoded control signals, $u' \in \mathbb{R}$.

8.6.2 Necessary Conditions for Stabilization

Conditions on Capacities

The following theorem shows that there is a relationship between the capacities of the forward and reverse channels and the existence of an invariant probability measure under memoryless policies.

Theorem 8.6.1 ([432]). *Consider the system described by (8.16), and let C_f and C_r denote, respectively, the forward and reverse channel capacities. Then, for the existence of an invariant probability measure with a finite second moment, we need* $\min(C_f, C_r) \geq \log_2(|a|)$, *when memoryless policies are considered.* ◇

Proof. See Sect. 8.7.13. □

Remark 8.6.1. The above result captures the requirements on channel capacities under memoryless policies, for the existence of an invariant probability measure for the controlled Markov process. This should be contrasted with the analysis earlier in the context of Theorem 8.2.1 for asymptotic mean stationarity (AMS). For the AMS property, we saw that, under any policy, the capacity requirement is both necessary and (almost) sufficient; however, the setting did not necessarily allow for a Markov process. Recall also that Theorem 8.3.2 stated that for stochastic stabilization over an erasure channel under policies with feedback, the Shannon capacity is almost sufficient not only for the AMS property but also for positive Harris recurrence. Such a strong result was also established in Chap. 7 for noiseless channels. Such a positive result will also be established for Gaussian channels with noiseless feedback later in this chapter as well as in Chap. 11. ◇

We note that Theorem 8.5.2 applies also for memoryless channels with noisy or noiseless feedback since such channels also belong to **Class A** (see Definition 8.5.1), the reason being that feedback does not increase the capacity of such channels [103]. Hence, for the system described by (8.16), the forward channel's capacity should be at least $\log_2(|a|)$ for stabilization in the sense of Theorem 8.5.2 under any coding or control policy. We can make this result stronger, by regarding the controller in Fig. 8.4 as an intermediate encoder, and y'_t as the channel output from the effective channel consisting of the encoder, controller, and the channels. With such an interpretation, we obtain the following.

Theorem 8.6.2. *Consider the system described by (8.16), and let C_f and C_r denote, respectively, the forward and reverse channel capacities. Then, for*

$$\liminf_{t\to\infty} \frac{1}{t} h(x_t) \leq 0,$$

under any causal coding and control policy, we need $\min(C_f, C_r) \geq \log_2(|a|)$. ◇

Proof. See Sect. 8.7.14. □

Structural Conditions

We now present a result on the structure of the encoder, which has significant practical implications that will be elaborated on later. This result could be seen as an extension of Theorem 7.3.1, to the present context. This result also highlights an important difference between continuous and discrete (finite-alphabet) channels.

Theorem 8.6.3 ([432]). *For a discrete-time linear system as in (8.16), with $|a| > 1$, with channel transitions forming an irreducible Markov chain, if either the encodable control set or the encodable state set is bounded, the Markov chain is transient.* ◇

Proof. See Sect. 8.7.15. □

The restriction alluded to in Theorem 8.6.3 above leads to significant complexity on encoding for control over a discrete noisy channel, since there needs to be a matching between the entire state space (which requires at least a countably infinite number of codewords) and a finite-symbol channel. Such a complexity does not arise, however, in a CMC, as we will see later.

This now motivates us to introduce the following.

Definition 8.6.1. An open-loop unstable system is *escape-free* if all the state symbols are encoded in such a way that given $x \in X$, there exists in the quantizer a reconstruction level, q, such that $x < q$. ◇

Let $\tau_{[-M,M]} := \inf\{t > 0 : |x_t| \in [-M, M]\}$ for $M > 0$. By Theorem 8.6.3, for the chain to be positive Harris recurrent, i.e., to have $P(\tau_{[-M,M]} < \infty | x_0 = x) = 1$, $\forall x \in X$, an unstable system has to be escape-free. Such a condition is not required for a stable system, since such a system is always recurrent. A system controlled over a continuous channel can always be made escape-free, and if the capacity is sufficiently large then the system can be stabilized. We will see that using a dynamic structure, *escape-freeness* can be assured by considering a side channel which can be either continuous or a discrete one with finite capacity that can transmit variable length codes through variable sampling.

8.6.3 Stabilization Over Discrete Channels and State-dependent Sampling

Further Discussion on Error Exponents and Channel Reliability Requirements

Before proceeding with the analysis of DMCs, we first review a few relevant results on reliability of channels. Earlier, the random coding exponent was considered in Remark 8.4.1. Here, we discuss further bounds on error exponents primarily in view of pairwise errors between symbols.

Let $\mathcal{C} = \{c_0, c_1, \ldots, c_{M-1}\}$ be a codebook of cardinality M, where each codeword is of length N. Let $p(y|c_m)$ be the conditional probability of y being received given that input to the channel is c_m. Suppose that the decoding rule is such that the mth codeword is the output if $p(y|c_m) > p(y|c_n)$, for all $n \neq m$ (in case of an equality, we can declare an error to obtain an upper bound on the error event). This rule corresponds to the maximum likelihood (ML) decoding [65]. The set of output symbols which would lead to a decoding of c_m is thus given by $R_m = \{y : p(y|c_m) > p(y|c_n), n \neq m\}$. It follows that, if $y \notin R_m$, for at least one $n \neq m$, say n', such that $p(y|c_{n'}) \geq p(y|c_m)$, we have $\sum_{n \neq m} \frac{p(y|c_n)}{p(y|c_m)} \geq 1$. Let $s > 0$. It follows that,

$$\begin{aligned}
\sum_{y \in R_n} p(y|c_m) &\leq \sum_{y \in R_n} p(y|c_m) \left(\frac{p(y|c_n)}{p(y|c_m)} \right)^{\frac{s}{s+1}} \\
&= \sum_{y \in R_n} (p(y|c_m))^{\frac{1}{s+1}} (p(y|c_n))^{\frac{s}{s+1}} \\
&= \exp\left(\log\left[\sum_{y \in R_n} (p(y|c_m))^{\frac{1}{s+1}} (p(y|c_n))^{\frac{s}{s+1}} \right] \right) \\
&=: \exp\big(- d(c_m, c_n; s) \big),
\end{aligned} \tag{8.17}$$

which defines the quantity $d(c_m, c_n; s)$ whose minimum overall codeword pairs and $s > 0$ is called the *minimum distance of a code*, as introduced in [337]. For $s = 1$, we have the Bhattacharyya distance between two symbols c_m and c_n, which we denote (by some abuse of notation) by $d(c_m, c_n)$; note that $d(c_m, c_n; 1) = d(c_m, c_n)$. The minimum of this quantity over all $c_m, c_n \in \mathcal{C}$, $m \neq n$, is called the *minimum Bhattacharyya distance of a code* $\mathcal{C}$ and is denoted by $d(\mathcal{C})$. We let $E_L(N, R) := \frac{1}{N} d(\mathcal{C})$ denote the minimum Bhattacharyya distance for a codebook $\mathcal{C}$ with length N and rate R and assume that the limit $\lim_{N\to\infty} E_L(N, R)$ exist for all rates. The probability of error between two different codewords (i.e., $p(n|m), m \neq n; c_m, c_n \in \mathcal{C}$) can be upper bounded using $E_L(N, R)$: $p(n|m) \leq e^{-NE_L(N,R)}$. We note that, for low coding rates, the minimum Bhattacharyya distance is closely related to the Gilbert bound; for details see [298, 380].

For rates which are not low, however, a more useful bound is the random coding exponent [152], considered earlier in Remark 8.4.1. One difference is that in the random coding bound, the exponent is strictly positive for rates below capacity, while this may not be so for the Gilbert bound. We also note that the random coding bound can be used to obtain a uniform bound on the errors for all transmitted messages as well, as observed in Remark 8.4.1.

An upper bound on the error exponent is given by the sphere packing error exponent. This exponent is related to the maximum over the minimum distance described above. Hence a lower bound can be obtained on the average probability of error, $p_e := \sum_m p(c_m)p(y \notin R_m|c_m)$, where $c_m \in \mathcal{C}$, which is

$$p_e \geq e^{-N\{E_{sp}(R-o_1(N))+o_2(N)\}},$$

known as the *sphere packing bound* [151, 337], with $o_1(N), o_2(N) \to 0$ with increasing N.

Asymptotic Stability in the Absence of System Noise

We first have the following result, a proof of which can be found in the appendix.

Theorem 8.6.4. *Let* $C \subset \mathbb{Z}$, $L < \infty$, *and* δ_i, $\forall i \in \mathbb{Z}$, *be the bin edges of a symmetric quantizer. Let* $1_{\{i\in C\}}$ *be the indicator function for* $i \in C$. *For a discrete channel, if the following drift condition holds for some sufficiently small* $\epsilon > 0$, *and for all bins:*

$$\max\left(|(\sum_l \sum_j p(j|i)p'(l|j)(a\delta_i + bQ'_l))|, |(\sum_l \sum_j p(j|i)p'(l|j)(a\delta_{i+1} + bQ'_l))|\right) < \delta_i - \epsilon + L1_{\{i\in C\}}, \tag{8.18}$$

then C *is a recurrent set in the sense that* $\sup_{x\in\cup_{i\in C}\mathcal{B}_i} E[\min(t > 0 : x_t \in \cup_{i\in C}\mathcal{B}_i)|x_0 = x] < \infty$. ◇

Proof. See Sect. 8.7.16. □

We now study the case when the channels are discrete and noiseless. If the channels are noiseless, the above leads (with $\epsilon = 0, L = 0$) to a logarithmic quantizer [132] .

Corollary 8.6.1. *Consider a symmetric quantizer at the sensor. Let both the forward and reverse channels be noiseless. To lead to a drift toward the origin, quantizer bin edges (on the positive real line) have to satisfy* $\delta_{i+1} \leq (1 + 2/|a|)\delta_i$.
◇

Before studying the stability conditions, however, we first note the following relationship between reliability and delay.

Let us fix the forward and reverse channel rates, $R_f = \log_2(|M_f|)/N_f$ and $R_r = \log_2(|M_r|)/N_r$. We penalize the codelengths in both channels by a possibly linear term in the sampling period. It then takes longer to send more bits, that is, reliability competes with delay. The following theorem says that if the controller waits long enough, stability can be achieved. To separate out the difficulty that comes about due to the escape-freeness requirement, we consider here first a one-stage problem; the more general system and control setup will be considered subsequently.

Theorem 8.6.5 ([432]). *Consider the scalar continuous-time system (8.15) but without the driving Brownian motion process. Let the probability distribution of the initial state* x_0 *have a bounded support set. Let the sampling period be a function of block lengths:* $T_s = \alpha N_f + \beta N_r$*;* α, β *be possibly depending on the codelengths, and the number of symbols in the state and control be* $K = |\mathcal{X}'| = |\mathcal{U}| = |\mathcal{U}'|$*. Let the rates* $R_f = \log_2(K)/N_f$ *and* $R_r = \log_2(K)/N_r$ *be kept constant as* N_f *and* N_r *grow. If the system and channel parameters satisfy the following three conditions:*

$$\lim_{N_f \to \infty} (R_f + 2\mu\alpha - E_L^f(N_f, R_f)) + (2\mu\beta R_f/R_r) < 0,$$

$$\lim_{N_f \to \infty} (R_r + 2\mu\beta - E_L^r(N_r, R_r)) + (2\mu\alpha R_r/R_f) < 0,$$

$$\frac{\alpha}{R_f} + \frac{\beta}{R_r} < \frac{1}{\mu}, \tag{8.19}$$

then $\lim_{T_s \to \infty} E[x_{T_s}^2] = 0$. ◇

Proof. See Sect. 8.7.17. □

Let $a = e^{\mu T_s}$ as before. We have the following observations regarding the result of Theorem 8.6.5 above. Positivity of the random coding error exponent (which is the case when rate R is strictly less than the Shannon capacity C of the memoryless channel) does not directly lead to stability, and the exponent actually has to be larger than a specific positive quantity. The condition (8.19) is the quantization rate requirement: $K > |a|$, also considered in Chap. 7.

Remark 8.6.2. We note that one could obtain a converse bound by obtaining a lower bound on the pairwise error probabilities provided that a uniform lower bound on the distortion for every transmitted message in the event of an error is satisfied with the coding policy. This provides a lower bound on the second moment conditioned on any transmitted symbol, which can then readily be used in the setting here leading to a converse result. The sphere packing exponent, $E_{sp}(R)$, maximizing the minimum distance, can be used for such an analysis. Note the earlier Remark 8.4.2, which had related discussions. ◇

Remark 8.6.3. In the remainder of the section we will consider the bound obtained via the Bhattacharyya distance. Note that one could generate parallel results using the random coding exponent, which is tight at high rates (and equals the sphere packing exponent bound for rates greater than a critical rate). ◇

Asymptotic Stability in the Presence of System Noise and Variable Length Coding Through State-dependent Sampling

We now consider the sampled system (8.16) driven by i.i.d. noise, which is a more realistic scenario, where a finite sampling period is given, and the amount of data to be sent over a sampling period is finite. In this case, asymptotic analysis of Theorem 8.6.5 becomes inapplicable. We already know from Theorem 8.6.3 that for stability the encodable set has to be unbounded; however, the number of bits that can be transmitted per unit time over the channel is finite. This dichotomy can be resolved using a coding scheme based on binning, where the coset of the code is transmitted and the particular bin is transmitted using an additional side channel.

Suppose that $K = 2^{N_f R_f}$ symbols can be transmitted during each unit time stage. Partition the entire state space into bins, group K adjacent elements into one larger bin, indexed by I, and represent them by a single channel codebook. We refer to this ensemble of bins as a *Codebin*. Hence, a total of $2^{N_f R_f}$ codewords are used to represent the entire state space.

Thus, we have $\text{Codebin(I)} = \{x : \delta_{IN_f R_f} \leq x < \delta_{(I+1)N_f R_f}\}$. We denote the bin indices by δ_{nI+i}, which means that the edge belongs to Codebin I and is represented by the ith channel codeword. We say the source code is in mode I, if the state is in Codebin I. The reconstruction value of each bin is assumed to be its midpoint, so that $Q_i = (\delta_i + \delta_{i+1})/2$.

We assume that the controller and the sensor can transmit the index information of the Codebin, over variable periods by using either explicit variable length codes or a *timing channel* (see Fig. 8.5). Timing channel is noiseless and carries the binary signal of starting the encoding, effectively carrying the index information. In such a scheme, Codebins are generated according to the number of sampling periods required to send the side channel information. Thus, the effective sampling period will vary. Here, the number of Codebins for a given period will grow exponentially with the sampling period. However, in this case the system will no longer be first-order Markovian.

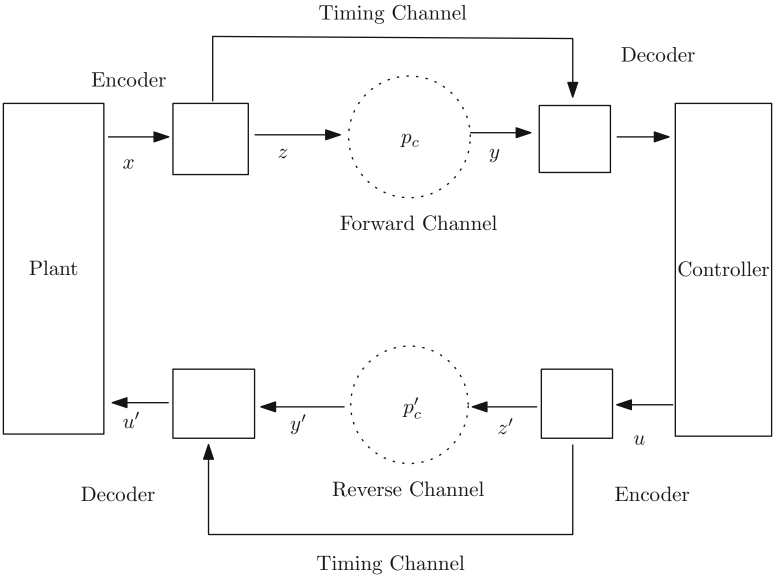

Fig. 8.5 Variable encoding via a time-channel as side channel

Given state x, let $n(x)$ be the number of sampling periods it takes for the transmission of the side channel symbol (or timing instant). Note that $n(x)$ is a causally measurable stopping time, which is in fact deterministic. Consider the sampling process given by the dynamics: $T_k = T_{k-1} + n(x_{T_{k-1}})$. The sampled process, x_{T_k}, is also Markovian. With this observation, we can obtain a state-dependent drift condition (see Theorem 6.2.3), to study stability. Here, we provide only conditions for the existence of an invariant measure; the approach for stronger conditions is identical and the extension is merely technical. If the transmission of the state and the control signals takes k times as long as it does in the fixed length case, the effective sampling period in the variable length encoding scheme is $kT_s, k \in \mathbb{Z}_+$. Thus the system will be open loop during kT_s seconds. These considerations lead to the following.

Theorem 8.6.6 ([432]). *Consider the scalar continuous-time system described by (8.15), which is to be remotely controlled over discrete channels. Let the forward and the reverse side channels be noiseless and the side channel symbols be transmitted over variable durations (as described above). Further, let* $U_k(\gamma) := \gamma_k^{2(kN_f R_f)}$*, and*

$$\Upsilon(k) := \Big(e^{-kN_f E_L^f(N_f,R_f) - kN_r E_L^r(N_r,R_r)} 2^{kN_f R_f} + e^{-kN_f E_L^f(N_f,R_f)} + e^{-kN_r E_L^r(N_r,R_r)} \Big) 2^{kN_f R_f}. \tag{8.20}$$

Suppose $\lim_{k\to\infty} \frac{\Upsilon(k)U_k(\gamma)}{e^{-2\mu k T_s}} < 1$. *If for a sequence* $\{\gamma_k \geq 1, k \in \mathbb{Z}_+\}$, $\exists k_0 > 0$ *such that* $\forall k > k_0$, *the following holds:*

$$\gamma_k < 1 + 2e^{-k\mu(\alpha N_f + \beta N_r)} \sqrt{1 - \frac{\Upsilon(k)U_k(\gamma)}{e^{-2\mu k T_s}}},$$

and

$$N_f R_f > \log_2(|a|),$$

then there exists a coding scheme leading to positive Harris recurrence for the sampled Markov chain. The source coder is a symmetric logarithmic quantizer with sequentially decreasing expansion ratios γ_k *used for symbols transmitted in the kth time stage.* ◇

Proof. See Sect. 8.7.18 in the appendix. □

We note that since the system noise is Gaussian, the Markov chain is Lebesgue-irreducible. Furthermore, if there is a compact set which is recurrent, then all open subsets of this compact set are also recurrent. Now, even though the transition kernel is not continuous for all $x \in \mathbb{R}$, for a small set $C \in \mathcal{B}(\mathbb{R})$ in one of the bins close to the origin, $P(x_{t+1} \in A | x_t = x)$, $A \in \mathcal{B}(\mathbb{R})$, will be uniformly continuous in $x \in C$ and bounded from below by some nontrivial measure. As such, such a set can serve as a petite set.

Remark 8.6.4. We note that, in the above construction, the original chain is no longer Markov. The sampled chain is Markov, however, and furthermore, the original system is such that there exists a compact set which satisfies the finite-mean return property $\sup_{x\in S} E[\tau_S] < \infty$. ◇

Remark 8.6.5. In the above, we required the time-channel to transmit noise-free the integer $k \in \mathbb{Z}_+$ (for kT_s as the open-loop duration during transmission) in k time stages. A prefix-free, uniquely decodable code can be used for the time-channel, such as, with 1 denoting the stopping bit, $1, 01, 001, 0001, \ldots, 0\ldots01, \ldots$. Hence, it suffices to send one bit of information to stop decoding. One needs an arbitrarily small but nonzero zero-error capacity, since for every zero-error capacity of ϵ, there exists a sampling period T_s such that $1/T_s < \epsilon$ and reliable transmission of one bit per time stage is possible. One can adjust the petite set as a function of T_s. We also note that the conditions in Theorem 8.6.6 simplify when one or both of the channels are noiseless, since the probability of error is zero (and the exponent is infinite). ◇

If the goal is the existence of a finite moment, more stringent criteria on channels will be needed. Nonetheless, the same techniques can be applied using the drift criteria. We also note that the coding construction can be arbitrary, as long as the drift conditions are satisfied.

8.6.4 Stabilization Over Continuous-Alphabet Channels

We now consider CMCs and obtain achievable rates for control over continuous alphabet channels. For continuous alphabet channels, there is no restriction on the values transmitted over the channel; there may be constraints over the input distributions, but arbitrary values can be fed into the channel. Thus, the system can always be designed to be escape-free. Here, we will study two special channels: Gaussian and continuous erasure channels.

Stabilizing Rates Over Gaussian Channels

For Gaussian channels, we associate power constraints with the encoder outputs, P_f and P_r, for the forward and the reverse channels, respectively. The objective is to develop coding schemes that will lead to an invariant density with a finite second moment. Here, one does not face the difficulty of explicitly using a finite codelength, for $\mathcal{X}'$ is the entire real line, and rare events are transmitted with higher magnitude signals, whose contribution to the expected power is limited. This was the main difficulty we observed in DMCs in the design of variable length codes in a control context. There, rare events had to be represented in longer codewords to prevent the Markov chain from becoming transient.

In the following theorem, we restrict the encoders, the sensor, and the controller to be scaling their inputs. We further restrict the controller and the decoders to be linear in their arguments, and obtain the decoder and the controller that minimize the invariant second moment of the state, leading to a stabilizing configuration.

Theorem 8.6.7 ([432]). *Suppose that the sensor encoder and the controller encoder have average power constraints P_f and P_r, respectively. Further suppose that the encoders and the decoders, and also the controller itself, are restricted to be linear and memoryless. Then, the optimal such policy at the input of the plant which minimizes the steady-state variance is*

$$u'_t = -\frac{a}{b}\sqrt{P_f P_r}\left(\frac{\sqrt{E[x_t^2]}}{\sqrt{P_f + \sigma_w^{f^2}}\,(P_r + \sigma_w^{r^2})}\right)y'_t,$$

where $y'_t = z'_t + w^r_t$, and $\sigma_w^{f^2}$ and $\sigma_w^{r^2}$ are the channel noise variances for the forward and the reverse channels, respectively. If the forward and the reverse channel capacities satisfy the condition

$$2^{-2C_f} + 2^{-2C_r} - 2^{-2C_f - 2C_r} < 1/a^2,$$

then the steady-state variance is finite. ◇

Proof. See Sect. 8.7.19 in the appendix. □

Note that the lower bound on the capacities is $\log_2(|a|)$. This leads us to the following Corollary to Theorem 8.6.7:

Corollary 8.6.2. *As C_f (respectively, C_r) $\to \infty$, the condition on C_r (respectively, C_f) becomes $C_r > \log_2(|a|)$ (respectively, $C_f > \log_2(|a|)$). Hence, if either the forward channel or the reverse channel is noiseless, then memoryless and linear coders are as good as any other coder for the optimum rate (or power) for the existence of an invariant probability measure for the state process.* ◇

Hence, linear policies are almost optimal as they meet the lower bound, when one of the channels becomes very reliable. It will be shown in Chap. 11 that when noiseless feedback is available, innovation coders are optimal. The corollary above shows that, if the goal is to have stability, memoryless schemes might as well be used without much loss in performance.

We will consider Gaussian channels in further generality in Chap. 11.

Continuous Erasure Channels

Consider forward and reverse erasure channels which lose packets with probabilities e_f and e_r, respectively. Consider also the case where the packets can be sent without a need of quantization, i.e., the erasure channel codebook set is the real line (thus the information theoretic capacity is infinite). In this case we have the following result.

Proposition 8.6.1. *Consider the unstable plant in (8.16), along with forward and reverse erasure channels. If the forward and the reverse channel packet loss probabilities satisfy the inequality $e_f + e_r - e_f e_r < 1/a^2$, then there exist memoryless policies such that $\lim_{t\to\infty} E[x_t^2] < \infty$.* ◇

Proof. If there is erasure in the forward channel as well as in the reverse channel, a control of zero value can be applied. For a nonzero control to have any effect, both channels have to transmit successfully. Hence, with probability $(1 - e_f)(1 - e_r)$, we have

$$x_{t+1} = ax_t - b(a/b)x_t + w_t.$$

Then the evolution of the second moment will be

$$E[x_{t+1}^2] = [1 - (1 - e_f)(1 - e_r)]a^2 E[x_t^2] + E[w_t^2],$$

which is stable if the condition in the statement of the proposition holds. □

8.7 Appendix: Proofs

8.7.1 Proof of Theorem 8.5.2

For channels of the type **Class A** (which includes the DMCs as a special case), the capacity is given by

$$C = \lim_{T\to\infty} \max_{\{P(q_i|q_{[0,t-1]},q'_{[0,t-1]})\}} \frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}),$$

where

$$I(q_{[0,T-1]} \to q'_{[0,T-1]}) = \sum_{t=1}^{T-1} I(q_{[0,t]}; q'_t | q'_{[0,t-1]}),$$

Let us define $R_T = \max_{\{P(q_t|q_{[0,t-1]},q'_{[0,t-1]}),0\le t\le T-1\}} \frac{1}{T}\sum_{t=0}^{T-1} I(q'_t; q_{[0,t]}| q'_{[0,t-1]})$. Observe that for $t > 0$:

$$\begin{aligned} I(q'_t; q_{[0,t]}|q'_{[0,t-1]}) &= H(q'_t|q'_{[0,t-1]}) - H(q'_t|q_{[0,t]}, q'_{[0,t-1]}) \\ &= H(q'_t|q'_{[0,t-1]}) - H(q'_t|q_{[0,t]}, x_t, q'_{[0,t-1]}) \qquad (8.21) \\ &\ge H(q'_t|q'_{[0,t-1]}) - H(q'_t|x_t, q'_{[0,t-1]}) \\ &= I(x_t; q'_t|q'_{[0,t-1]}). \qquad (8.22) \end{aligned}$$

Here, (8.21) follows from the assumption that the channel is of **Class A** type (and the control actions are determined by channel outputs). It follows that since for two sequences $\{a_n\}, \{b_n\}$ with $a_n \ge b_n$, we have $\limsup_n a_n \ge \limsup_n b_n$, and R_T is assumed to have a limit

$$\begin{aligned} &\lim_{T\to\infty} R_T \\ &\ge \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{t=1}^{T-1} I(x_t; q'_t|q'_{[0,t-1]})) + I(x_0; q'_0)\right) \\ &= \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{t=1}^{T-1} \left(h(x_t|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]})\right) + I(x_0; q'_0)\right) \\ &= \limsup_{T\to\infty} \frac{1}{T}\left(\sum_{t=1}^{T-1} \left(h(Ax_{t-1} + Gw_{t-1} + Bu_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]})\right) \right. \\ &\qquad\qquad \left. + I(x_0; q'_0)\right) \end{aligned}$$

$$= \limsup_{T\to\infty} \frac{1}{T}\Bigg(\sum_{t=1}^{T-1} \Big(h(Ax_{t-1} + Gw_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \Big) + I(x_0;q'_0) \Bigg) \tag{8.23}$$

$$\geq \limsup_{T\to\infty} \frac{1}{T}\Bigg(\sum_{t=1}^{T-1} \Big(h(Ax_{t-1} + Gw_{t-1}|q'_{[0,t-1]}, w_{t-1}) - h(x_t|q'_{[0,t]}) \Big) + I(x_0;q'_0) \Bigg) \tag{8.24}$$

$$= \limsup_{T\to\infty} \frac{1}{T}\Bigg(\sum_{t=1}^{T-1} \Big(h(Ax_{t-1}|q'_{[0,t-1]}, w_{t-1}) - h(x_t|q'_{[0,t]}) \Big) + I(x_0;q'_0) \Bigg)$$

$$= \limsup_{T\to\infty} \frac{1}{T}\Bigg(\sum_{t=1}^{T-1} \Big(h(Ax_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \Big) + I(x_0;q'_0) \Bigg) \tag{8.25}$$

$$= \limsup_{T\to\infty} \frac{1}{T}\Bigg(\sum_{t=1}^{T-1} \Big(\log_2(|A|) + h(x_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \Big) + I(x_0;q'_0) \Bigg)$$

$$= \limsup_{T\to\infty} \frac{1}{T}\Bigg(\Big(\sum_{t=1}^{T-1} \log_2(|A|) \Big) + h(x_0|q'_0) - h(x_{T-1}|q'_{[0,T-1]}) + I(x_0;q'_0) \Bigg)$$

$$= \log_2(|A|) - \liminf_{T\to\infty} \left(\frac{1}{T} h(x_{T-1}|q'_{[0,T-1]}) \right) \tag{8.26}$$

$$\geq \log_2(|A|) - \liminf_{T\to\infty} \left(\frac{1}{T} h(x_{T-1}) \right) \geq \log_2(|A|). \tag{8.27}$$

Equality (8.23) follows from the fact that the control action is a function of the past channel outputs, (8.24) follows from the fact that conditioning does not increase entropy, and (8.25) follows from the observation that $\{w_t\}$ is an independent process. Equation (8.27) follows from conditioning. The other (intermediate) lines follow from the properties of mutual information. By the hypothesis, $\liminf_{t\to\infty} \frac{1}{t} h(x_t) \leq 0$, it must be that $\lim_{T\to\infty} R_T \geq \log_2(|A|)$. Thus, the capacity also needs to satisfy this bound. □

8.7.2 Proof of Proposition 8.5.1

The proof follows from that of Theorem 8.5.2, with the following differences (where we build on [260]). Define the event

$$\mathcal{S}_T = \{\omega : |x_0| \leq K, |\sum_{k=0}^{T-1} A^{-k-1} G w_k - \zeta_k| \leq 1\},$$

where in this context $|\cdot|$ denotes the l_2 norm, ζ_k is a deterministic sequence, and K is sufficiently large, such that the event $\mathcal{S}_T$ has positive probability. Define $\bar{x}_t = x_t - \sum_{k=0}^{t-1} A^{t-k-1} B u_t$ to be the control-free state such that $\bar{x}_0 = x_0$ and $\bar{x}_{t+1} = A\bar{x}_t + G w_t$. By Definition 8.5.1, we note that the capacity expression satisfies

$$\begin{aligned} C &= \lim_{T\to\infty} \max_{\{P(q_t|q_{[0,t-1]},q'_{[0,t-1]}), \quad 0\leq t\leq T-1\}} \frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}) \\ &= \lim_{T\to\infty} \max_{\{P(q_t|q_{[0,t-1]},q'_{[0,t-1]}), \quad 0\leq t\leq T-1\}} \frac{1}{T} I(q_{[0,T-1]} \to q'_{[0,T-1]}|\mathcal{S}_T), \end{aligned}$$

where the conditional directed information is given by

$$I(q_{[0,T-1]} \to q'_{[0,T-1]}|\mathcal{S}_T) = \sum_{t=1}^{T-1} I(q_{[0,t]}; q'_t|q'_{[0,t-1]}, \mathcal{S}_T) + I(q_0; q'_0|\mathcal{S}_T).$$

This result is a consequence of the equivalence of the expressions under a capacity achieving channel input policy and the Markov chain condition in Definition 8.5.1. Now, instead of (8.21), the following can be established:

$$\begin{aligned} I(q'_t; q_{[0,t]}|q'_{[0,t-1]}, \mathcal{S}_T) &= H(q'_t|q'_{[0,t-1]}, \mathcal{S}_T) - H(q'_t|q_{[0,t]}, q'_{[0,t-1]}, \mathcal{S}_T) \\ &= H(q'_t|q'_{[0,t-1]}, \mathcal{S}_T) - H(q'_t|q_{[0,t]}, \bar{x}_T, q'_{[0,t-1]}, \mathcal{S}_T) \qquad (8.28) \\ &\geq H(q'_t|q'_{[0,t-1]}, \mathcal{S}_T) - H(q'_t|\bar{x}_T, q'_{[0,t-1]}, \mathcal{S}_T) \\ &= I(\bar{x}_T; q'_t|q'_{[0,t-1]}, \mathcal{S}_T). \end{aligned}$$

Here, (8.28) holds by conditioning and by the fact that the system variables do not affect the channel as a consequence of Definition 8.5.1. We also note that here the events are conditioned on the event realization $\mathcal{S}_T$. Thus, $R_T \geq \sum_{t=0}^{T-1} I(\bar{x}_T; q'_t|q'_{[0,t-1]}, \mathcal{S}_T) = I(\bar{x}_T; q'_{[0,T-1]}|\mathcal{S}_T)$. Let us write

$$I(\bar{x}_T; q'_{[0,T-1]}|\mathcal{S}_T) = h_{\mathcal{S}_T}(\bar{x}_T) - h_{\mathcal{S}_T}(\bar{x}_T|q'_{[0,T-1]}),$$

where the notation $h_{\mathcal{S}_T}(\cdot)$ denotes the conditioning on the event realization $\mathcal{S}_T$. The following then holds:

$$h_{\mathcal{S}_T}(\bar{x}_T|q'_{[0,T-1]}) \geq h_{\mathcal{S}_T}(\bar{x}_T) - T R_T. \qquad (8.29)$$

Furthermore,

$$\begin{aligned} h_{\mathcal{S}_T}(\bar{x}_T) &= h_{\mathcal{S}_T}(A^T(x_0 + \sum_{k=0}^{T-1} A^{-k-1} G w_k)) \\ &= \log_2(|A|)T + h_{\mathcal{S}_T}((x_0 + \sum_{k=0}^{T-1} A^{-k-1} G w_k)). \end{aligned} \tag{8.30}$$

Observe that since u_t is a function of $q'_{[0,t]}$ for all $t \geq 0$,

$$h_{\mathcal{S}_T}(\bar{x}_T | q'_{[0,T-1]}) = h_{\mathcal{S}_T}(x_T | q'_{[0,T-1]}) \leq h_{\mathcal{S}_T}(x_T, \mathcal{Y} | q'_{[0,T-1]}),$$

where $\mathcal{Y}$ is a binary random variable which is 1 if $|x_T| \leq b(T)$ and 0 otherwise. Let

$$P(\mathcal{Y} = 1) = P(|x_T| \leq b(T)) =: p_T.$$

Then,

$$\begin{aligned} h_{\mathcal{S}_T}(x_T, \mathcal{Y} | q'_{[0,T-1]}) &= h_{\mathcal{S}_T}(x_T | q'_{[0,T-1]}, \mathcal{Y}) + h_{\mathcal{S}_T}(\mathcal{Y} | q'_{[0,T-1]}) \\ &\leq h_{\mathcal{S}_T}(x_T | q'_{[0,T-1]}, \mathcal{Y}) + 1, \end{aligned}$$

since $\mathcal{Y}$ is binary. We have that

$$\begin{aligned} h_{\mathcal{S}_T}(x_T | q'_{[0,T-1]}, \mathcal{Y}) &\leq 1 + p_T \frac{n}{2} \log_2(2\pi e b^2(T)) \\ &\quad + (1 - p_{T-1}) h_{\mathcal{S}_T}\left(x_T \middle| q'_{[0,T-1]}, |x_T| \geq b(T) \right) \end{aligned}$$

and

$$\begin{aligned} & h_{\mathcal{S}_T}\left(x_T \middle| q'_{[0,T-1]}, |x_T| > b(T) \right) \\ &= h_{\mathcal{S}_T}\left(A^T(x_0 + \sum_{k=0}^{T-1} A^{-k-1}(w_k + u_k)) \middle| q'_{[0,T-1]}, |x_T| > b(T) \right) \\ &\leq \log_2(|A|)T + \frac{n}{2} \log_2(2\pi e (K+1)^2). \end{aligned} \tag{8.31}$$

Here (8.31) follows from the fact that the Gaussian measure maximizes the differential entropy among random variables with a fixed covariance matrix, which in this case is further bounded by the effects of the event $\mathcal{S}_T$. We have then, by (8.29)–(8.31) and by adjusting T and $\mathcal{S}_T$ accordingly,

$$\begin{aligned} \liminf_{T \to \infty} \frac{1}{T} \Bigg(& 1 + (1 - p_T)(\log_2(|A|)T + \frac{n}{2} \log_2(2\pi e (K+1)^2)) \\ & + p_T \frac{n}{2} \log_2(2\pi e b^2(T)) \Bigg) \end{aligned}$$

$$\geq \liminf_{T\to\infty}\left(\log_2(|A|)T + h_{\mathcal{S}_T}((x_0 + \sum_{k=0}^{T-1} A^{-k-1}Gw_k)) - TR_T\right)$$
$$= \log_2(|A|) - C,$$

where the last step follows since $|h_{\mathcal{S}_T}((x_0 + \sum_{k=0}^{T-1} A^{-k-1}Gw_k))| < \infty$ (uniformly over T). It then follows that

$$\limsup_{T\to\infty} P(|x_T| \leq b(T)) \leq \frac{C}{\log_2(|A|)},$$

for all $b(T)$ such that $\log_2(b(T))/T = 0$. This completes the proof. □

8.7.3 Proof of Proposition 8.5.3

If the process is AMS (see Sect. C.3.1), then there exists a stationary measure $\bar{P}$ such that

$$\lim_{T\to\infty} \sum_{k=1}^{T} P(T^{-k}D) = \bar{P}(D), \tag{8.32}$$

for all (cylinder) events D. Let for $b_B \in \mathbb{R}_+$, $B \in \mathcal{B}(\mathbb{R}^N)$ be given by $B = \{x : |x| \leq b_B\}$ and $X_n(z) = z_n$ be the coordinate function (see Sect. C.3.1) where $z = \{z_0, z_1, z_2, \cdots\}$.

If, by Proposition 8.2.1,

$$\limsup_{T\to\infty} P(|x_T| \leq b_B) \leq \frac{C}{\log_2(|A|))} < 1, \tag{8.33}$$

holds for all $b_B \in \mathbb{R}_+$, then $\bar{P}_n(B) < \frac{C}{\log_2(|A|)}$ for all compact B, where $\bar{P}_n$ is the marginal probability on the nth coordinate defined as

$$\bar{P}_n(B) = \bar{P}(x : |X_n(x)| \leq b_B).$$

But then $\bar{P}_n$, as an individual probability measure, must be tight [58]; therefore, for every $\delta > 0$ there exists $b_B < \infty$ such that $\bar{P}_n(B) \geq 1-\delta$. But, by (8.32), this would imply that $\limsup_{t\to\infty} P(T^{-t}B) = \limsup_{t\to\infty} P(|x_t| \in B) \geq 1 - \delta$, leading to a contradiction with (8.33) for $\delta < 1 - \frac{C}{\log_2(|A|)}$. Hence, the AMS property cannot be achieved. □

8.7.4 Proof of Proposition 8.3.1

As in (8.12), we introduce a sequence of stopping times as

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k > \tau_z : |h_k| \leq 1, p_k = 1\}, \quad z \in \mathbb{N}.$$

By the strong Markov property and the nature of the stopping times, (x_{τ_z}, h_{τ_z}) is also a Markov chain. In the following, we show that there exist $b_0 > 0$, $b_1 < \infty$ such that

$$E[\log(\varDelta^2_{\tau_{z+1}})|\varDelta_{\tau_z}, h_{\tau_z}] \leq \log(\varDelta^2_{\tau_z}) - b_0 + b_1 1_{\{|\varDelta_{\tau_z}| \leq F\}}. \tag{8.34}$$

We first bound the probability $P(\tau_{z+1} - \tau_z \geq k|\varDelta_{\tau_z}, h_{\tau_z})$ from above and below in the following two lemmas:

Lemma 8.7.1. *The discrete probability measure* $P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z})$ *has the upper bound*

$$P(\tau_{z+1} - \tau_z \geq k|x_{\tau_z}, \varDelta_{\tau_z}) \leq (1-p)^{k-1} + G_k(\varDelta_{\tau_z}),$$

where $G_k(\varDelta_{\tau_z}) \to 0$ *as* $\varDelta_{\tau_z} \to \infty$ *uniformly in* x_{τ_z}. ◇

Proof. For $k \in \mathbb{N}$, let

$$\varTheta_k := P(\tau_{z+1} - \tau_z \geq k \mid x_{\tau_z}, \varDelta_{\tau_z}) = P_{x_{\tau_z}, \varDelta_{\tau_z}}(\tau_{z+1} - \tau_z \geq k). \tag{8.35}$$

Without any loss of generality, let $z = 0, \tau_0 = 0$, so that $\varTheta_k = P_{x_0, \varDelta_0}(\tau_1 \geq k)$. The probability $\varTheta_k$ for $k \geq 2$ is bounded as follows:

$$\begin{aligned}
\varTheta_k &= P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-1} (\varUpsilon_s = 0) \cup (|h_s| > 1)\right) \\
&\leq P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-1} (\varUpsilon_s = 0) \cup (|x_s| \geq 2^{R'-1}(|a| + \delta)^{s-1}\alpha\varDelta_0)\right) \\
&= P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-1} (\varUpsilon_s{=}0) \cup (|a^s(x_0{+}\sum_{i=0}^{s-1} a^{-i-1}w_i)| \geq (|a|{+}\delta)^{s-1}2^{R'-1}\alpha\varDelta_0)\right) \\
&\leq P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-2} (\varUpsilon_s = 0) \cup (|h_s| > 1) \,\middle|\, \varUpsilon_{k-1} = 0\right)(1-p) \\
&\quad + P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-2} (\varUpsilon_s = 0) \cup (|h_s| > 1)\right. \\
&\qquad\qquad \left.\middle| \{|a^{k-1}(x_0 + \sum_{i=0}^{k-2} a^{-i-1}w_i)| \geq (|a| + \delta)^{k-2}2^{R'-1}\alpha\varDelta_0\}\right) \\
&\quad \times P_{x_0,\varDelta_0}\left(|a^{k-1}(x_0 + \sum_{i=0}^{k-2} a^{-i-1}w_i)| \geq (|a| + \delta)^{k-2}2^{R'-1}\alpha\varDelta_0\right)
\end{aligned} \tag{8.36}$$

$$
\begin{aligned}
&\leq P_{x_0,\varDelta_0}\Big(\bigcap_{s=1}^{k-2}(\varUpsilon_s=0)\cup(|h_s|>1)\,\Big|\,\varUpsilon_{k-1}=0\Big)(1-p)\\
&\qquad+P_{x_0,\varDelta_0}\Big(|a^{k-1}(x_0+\sum_{i=0}^{k-2}a^{-i-1}w_i)|\geq(|a|+\delta)^{k-2}2^{R'-1}\alpha\varDelta_0\Big)\\
&=P_{x_0,\varDelta_0}(\tau_1\geq k-1)(1-p)\\
&\qquad+P_{x_0,\varDelta_0}\Big(|a^{k-1}(x_0+\sum_{i=0}^{k-2}a^{-i-1}w_i)|\geq(|a|+\delta)^{k-2}2^{R'-1}\alpha\varDelta_0\Big). \quad (8.37)
\end{aligned}
$$

In the above derivation, (8.36) follows from the following: For any three events M, C, D in a common probability space,

$$
P\big(M\cap(C\cup D)\big)=P\big((M\cap C)\cup(M\cap D)\big)\leq P\big(M\cap C\big)+P\big(M\cap D\big)
$$

Now, observe that for $k\geq 2$,

$$
\begin{aligned}
&P_{x_0,\varDelta_0}\Big(|(x_0+\sum_{i=0}^{k-2}a^{-i-1}w_i)|\geq(\frac{|a|+\delta}{|a|})^{k-2}2^{R'-1}\frac{\alpha}{|a|}\varDelta_0\Big)\\
&\leq 2P_{x_0,\varDelta_0}\Big(\sum_{i=0}^{k-2}a^{-i-1}w_i>(2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-2}\frac{\alpha}{|a|}-\frac{1}{2})\varDelta_0\Big)\\
&\leq C\exp\Big(-\frac{((\xi^{k-2}N-1/2)\varDelta_0)^2}{2\sigma'^2}\Big), \quad (8.38)
\end{aligned}
$$

where (8.38) follows from (8.5), for this condition ensures that the term

$$
(2^{R'-1}(\frac{|a|+\delta}{|a|})^{k-2}\frac{\alpha}{|a|}-\frac{1}{2})
$$

is positive for $k\geq 2$ and bounding the complementary error function by the following: $\int_q^\infty \mu(dx)\leq q^{-1}\int_q^\infty x\mu(dx)$, for $q>0$. In the above derivation, the constants are as follows:

$$
\sigma'^2=\frac{E[w_1^2]}{1-|a|^{-2}},\quad \xi=\frac{|a|+\delta}{|a|},\quad N=\frac{2^{R'-1}}{(|a|)/\alpha},\quad C=2\sigma'\frac{1}{\sqrt{2\pi}(2N-1)\varDelta_0/2}.
$$

Let us define:

$$
\varXi_k:=\frac{((\xi^{k-2}N-1/2)\varDelta_0)^2}{2\sigma'^2},\qquad \tilde{\varXi}_k:=\frac{((\xi^{k}N-1/2)\varDelta_0)^2}{2\sigma'^2}.
$$

We can bound the probability Θ_k defined in (8.35). Since a decaying exponential decays faster than any decaying polynomial, for each $m \in \mathbb{N}_+$, there exists an $M < \infty$ such that for all $k \in \mathbb{N}$,

$$Ce^{-\Xi_k} \leq M\tilde{\Xi}_k^{-m}. \tag{8.39}$$

Thus, we have that

$$P_{x_0,\Delta_0}\left((x_0+\sum_{i=0}^{k-2} a^{-i-1}w_i)\geq(\frac{|a|+\delta}{|a|})^{k-2}2^{R'-1}\frac{\alpha}{|a|}\Delta_0\right)\leq M\tilde{\Xi}_k^{-m}. \tag{8.40}$$

Now $\Theta_1 = 1$ by definition, and for $k > 1$,

$$\Theta_k \leq \Theta_{k-1}(1-p) + Ce^{-\Xi_k}. \tag{8.41}$$

We obtain,

$$\Theta_k \leq (1-p)^{k-1} + G_k(\Delta_{\tau_0}), \tag{8.42}$$

where

$$G_k(\Delta_{\tau_0}) := \sum_{s=1}^{k-1} M(1-p)^{k-s-1}\tilde{\Xi}_s^{-m}. \tag{8.43}$$

It now follows that

$$\begin{aligned}
G_k(\Delta_{\tau_0}) &= \sum_{s=1}^{k-1} M(1-p)^{k-s-1}\tilde{\Xi}_s^{-m} \\
&= \Delta_0^{-2m}(1-p)^{k-1}\sum_{s=1}^{k-1} M(1-p)^{-s}\left((\xi^s N - 1/2)^2/(2\sigma'^2)\right)^{-m} \\
&\leq \Gamma_m \Delta_0^{-2m}(1-p)^{k-1}\frac{\left((1-p)\xi^{2m}\right)^{-k} - 1}{\left((1-p)\xi^{2m}\right)^{-1} - 1},
\end{aligned} \tag{8.44}$$

with $\Gamma_m = M(N - \frac{1}{2\xi})^{-2m}(2\sigma'^2)^m < \infty$. Now if m is taken such that

$$(1-p)\xi^{2m} > 1, \tag{8.45}$$

then $\lim_{\Delta_0 \to \infty} G_k(\Delta_0) = 0$, and for all $k \in \mathbb{N}$,

$$\Theta_k \leq (1-p)^{k-1}\left(1 + \Gamma_m \Delta_{\tau_0}^{-2m}\frac{1}{1-((1-p)\xi^{2m})^{-1}}\right) \tag{8.46}$$

□

Lemma 8.7.2. *The discrete probability measure* $P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z})$ *has the lower bound*

$$P(\tau_{z+1} - \tau_z \geq k | x_{\tau_z}, \varDelta_{\tau_z}) \geq (1-p)^{k-1},$$

for all realizations of $x_{\tau_z}, \varDelta_{\tau_z}$. ◇

Proof. This follows since

$$P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-1} (\varUpsilon_s = 0) \cup (|h_s| > 1)\right) \geq P_{x_0,\varDelta_0}\left(\bigcap_{s=1}^{k-1} (\varUpsilon_s = 0)\right). \qquad \square$$

As a consequence of Lemmas 8.7.1 and 8.7.2, the probability below tends to $(1-p)^{k-1}p$ as $\varDelta_{\tau_z} \to \infty$:

$$\begin{aligned} &P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z}) \\ &= P(\tau_{z+1} - \tau_z \geq k \mid x_{\tau_z}, \varDelta_{\tau_z}) - P(\tau_{z+1} - \tau_z \geq k+1 | x_{\tau_z}, \varDelta_{\tau_z}). \end{aligned} \tag{8.47}$$

We can now invoke Theorem 6.2.6. With the candidate Lyapunov function $V_0(x_t, \varDelta_t) = \log(\varDelta_t^2) + B_0$, for $\varDelta_{\tau_z} > L$, we have that

$$\begin{aligned} E[V_0(x_{\tau_{z+1}}, \varDelta_{\tau_{z+1}}) \mid x_{\tau_z}, \varDelta_{\tau_z}] = B_0 + P(\tau_{z+1} - \tau_z = 1)\left(2\log(\alpha) + \log(\varDelta_{\tau_z}^2)\right) \\ + \sum_{k=2}^{\infty} \log(\varDelta_{\tau_z+k}^2) P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z}). \end{aligned}$$

Thus, the drift satisfies:

$$\begin{aligned} &E[V_0(x_{\tau_{z+1}}, \varDelta_{\tau_{z+1}}) \mid x_{\tau_z}, \varDelta_{\tau_z}] - V_0(x_{\tau_z}, \varDelta_{\tau_z}) \\ &= \sum_{k=1}^{\infty} 2\log((|a| + \delta)^{(k-1)}\alpha) P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z}) \\ &= 2\log(\alpha) + 2\sum_{k=1}^{\infty} (k-1)\log(|a| + \delta) P(\tau_{z+1} - \tau_z = k \mid x_{\tau_z}, \varDelta_{\tau_z}). \end{aligned} \tag{8.48}$$

By (8.43), the summability of $\sum_{k=1}^{\infty} G_k(\varDelta_{\tau_z})$, and the dominated convergence theorem, we have

$$
\begin{aligned}
\lim_{\Delta_{\tau_z}\to\infty} & \sum_{k=1}^{\infty}(k-1)((1-p)^{k-1}+G_k(\Delta_{\tau_z})-(1-p)^k) \\
&= \sum_{k=1}^{\infty}\lim_{\Delta_{\tau_z}\to\infty}(k-1)((1-p)^{k-1}+G_k(\Delta_{\tau_z})-(1-p)^k) \\
&= \sum_{k=1}^{\infty}p(1-p)^{k-1}(k-1)=p^{-1}-1.
\end{aligned}
$$

Provided that (8.6) holds, it follows that for some $b_0>0$

$$
\begin{aligned}
\lim_{\Delta_{\tau_z}\to\infty} & \big\{E[V_0(x_{\tau_{z+1}},\Delta_{\tau_{z+1}}) \mid x_{\tau_z},\Delta_{\tau_z}]-V_0(x_{\tau_z},\Delta_{\tau_z})\big\} \\
&= 2\lim_{\Delta_{\tau_z}\to\infty}\big\{\log(\alpha)+\sum_{k=1}^{\infty}(k-1)\log(|a|+\delta)P(\tau_{z+1}-\tau_z=k \mid x_{\tau_z},\Delta_{\tau_z})\big\} \\
&\leq -b_0\,.
\end{aligned}
\tag{8.49}
$$

For Δ_{τ_z} in a compact set and lower bounded by L' defined by (8.8), $E[\log(\Delta^2_{\tau_{z+1}}) \mid x_{\tau_z},\Delta_{\tau_z}]$ is uniformly bounded. This follows from the representation of the drift given in (8.48). Finally, since,

$$
G_k(\Delta_{\tau_0})\leq(1-p)^{k-1}\Gamma_m\Delta_{\tau_0}^{-2m}\frac{1}{1-((1-p)\xi^{2m})^{-1}},
$$

it follows that $\sum_{k=1}^{\infty}G_k(\Delta_{\tau_0})k<\infty$ and as a result

$$
\sup_{x_{\tau_z},\Delta_{\tau_z}} E_{x_{\tau_z},\Delta_{\tau_z}}[\tau_{z+1}-\tau_z]<\infty. \tag{8.50}
$$

Consequently, under the bound (8.6), there exist $b_0>0$, $b_1<\infty, F'>0$ such that

$$
E[V_0(x_{\tau_{z+1}},\Delta_{\tau_{z+1}})|x_{\tau_z},\Delta_{\tau_z}]\leq V_0(x_{\tau_z},\Delta_{\tau_z})-b_0+b_1 1_{\{|\Delta_{\tau_z}|\leq F'\}}. \tag{8.51}
$$

With (8.50) and (8.51), Corollary 6.2.1 leads to positive Harris recurrence.

Together with Lemmas 6.2.6 and 8.7.1, (8.50) and (8.51) imply the existence of a set $C_x\times C_\Delta$ satisfying the property that

$$
\sup_{(x,\Delta)\in C_x\times C_\Delta} E_{x,\Delta}[\tau_{C_x\times C_\Delta}]<\infty.
$$

To verify the property $P_{(x,\Delta)}(\tau_{C_x\times C_\Delta}<\infty)=1$ for any admissible (x,Δ), the argument follows, as before, from the construction of

$$
\Theta_k(\Delta,x):=P(\tau_1\geq k \mid x,\Delta),
$$

where

$$\tau_1 = \inf(k > 0 : |x_k| \leq 2^{R'-1}\Delta_k, x_0 = x, \Delta_0 = \Delta)$$

and observing that $\Theta_k(\Delta, x)$ is majorized by a geometric measure with similar steps. Once a state which is perfectly zoomed, that is, which satisfies $|x_t| \leq 2^{R'-1}\Delta_t$, is visited, the stopping time analysis can be used to verify that from any initial condition the recurrent set is visited in finite time with probability 1. □

8.7.5 Proof of Theorem 8.3.1

Let the values taken by $\log_2(\bar{Q}(\cdot))/s$ be $\{-\tilde{A}, 0, \tilde{B}\}$. Let

$$\mathbb{L}_{z_0,\tilde{A},\tilde{B}} := \{n \in \mathbb{N}, n \geq \log_2(L')/s : \exists N_{\tilde{A}} \in \mathbb{N}, N_{\tilde{B}} \in \mathbb{N}, n = -N_{\tilde{A}}\tilde{A} + N_{\tilde{B}}\tilde{B} + z_0\}.$$

By (8.7),

$$\log_2(\Delta_{t+1})/s = \log_2(\bar{Q}(\Delta_t, |h_t|, \Upsilon_t))/s + \log_2(\Delta_t)/s$$

is also an integer. We will establish that $\mathbb{L}_{z_0,\tilde{A},\tilde{B}}$ forms a communication class, where $z_0 = \log_2(\Delta_0)/s$ is the initial condition of the parameter for the quantizer. Furthermore, since the source process x_t is “Lebesgue-irreducible” (for the system noise admits a probability density function that is positive everywhere) and there is a uniform lower bound L' on bin sizes, the error process takes values in any of the admissible quantizer bins with nonzero probability. In view of these, we now establish that the Markov chain is irreducible.

Given $l, k \in \mathbb{L}_{z_0,\tilde{A},\tilde{B}}$, there exist $N_{\tilde{A}}, N_{\tilde{B}} \in \mathbb{N}$ such that $l - k = -N_{\tilde{A}}\tilde{A} + N_{\tilde{B}}\tilde{B}$. In particular, if at time 0, the quantizer is perfectly zoomed and $\Delta_0 = 2^{sk}$, then there exists a sequence of events consisting of $N_{\tilde{B}}$ erasure events (simultaneously satisfying $|h_t| \leq 1$) and consequently $N_{\tilde{A}}$ zoom-in events taking place with probability at least:

$$\left((1-p)P(w_t \in [-(2^{R'} - a)L'/2, (2^{R'} - a)L'/2])\right)^{N_{\tilde{B}}}$$
$$\left(pP(w_t \in [-(\alpha 2^{R'} - a)L'/2, (\alpha 2^{R'} - a)L'/2])\right)^{N_{\tilde{A}}} > 0, \quad (8.52)$$

so that $P(\Delta_{N_{\tilde{A}}+N_{\tilde{B}}} = 2^{sl} | \Delta_0 = 2^{ks}, x_0) > 0$, uniformly in x_0. In the following we will consider this sequence of events.

Now, for some distribution $\mathcal{K}$ on positive integers, $E \subset \mathbb{R}$ and Δ an admissible bin size,

$$\sum_{n \in \mathbb{N}_+} \mathcal{K}(n) P\Big((x_n, \Delta_n) \in (E \times \{\Delta\}) \,\Big|\, x_0, \Delta_0\Big) \geq K_{\Delta_0,\Delta}\psi(E, \Delta).$$

Here $K_{\Delta_0,\Delta}$, denoting a lower bound on the probability of visiting Δ from Δ_0 in some finite time, is nonzero by (8.52) and ψ is a positive function as the following argument shows: Let $t > 0$ be the time stage for which $\Delta_t = \Delta$ and thus by the construction in (8.52), with $|h_{t-1}| \leq 1$: $|ax_{t-1}+bu_{t-1}| \leq |a|\Delta_{t-1}/2 = (|a|/\alpha)\frac{\Delta}{2}$. Thus, it follows that for $A_1, B_1 \in \mathbb{R}$, $A_1 < B_1$,

$$\begin{aligned}
&P\biggl(x_t \in [A_1, B_1] \biggm| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1}\biggr) \\
&= P\biggl(ax_{t-1} + bu_{t-1} + w_{t-1} \in [A_1, B_1] \biggm| |ax_{t-1} + bu_{t-1}| \leq |a|\Delta_{t-1}/2, \Delta_{t-1}\biggr) \\
&\geq \min_{|z|\leq \frac{\Delta}{2}(|a|/\alpha)} \biggl(P(w_{t-1}{\in}[A_1{-}z, B_1{-}z]\biggr){>}0. \qquad (8.53)
\end{aligned}$$

Now, define the finite set $C'_\Delta := \{\Delta : L' \leq |\Delta| \leq F', \frac{\log_2(\Delta)}{s} \in \mathbb{N}\}$. The chain satisfies the recurrence property that $P_{(x,\Delta)}(\tau_{C_x \times C'_\Delta} < \infty) = 1$ for any admissible (x, Δ). This follows, as before, from the construction of

$$\Theta_k(\Delta, x) := P(\tau_1 \geq k \mid x, \Delta),$$

where

$$\tau_1 = \inf(k > 0 : |x_k| \leq 2^{R'-1}\Delta_k, x_0 = x, \Delta_0 = \Delta)$$

and observing that $\Theta_k(\Delta, x)$ is majorized by a geometric measure. Once a state which is perfectly zoomed, that is which satisfies $|x_t| \leq 2^{R'-1}\Delta_t$, is visited, the stopping time analysis can be used to verify that from any initial condition the recurrent set is visited in finite time with probability 1. In view of (8.52), we have that the chain is irreducible.

We can show that the set $C_x \times C'_\Delta$ is small by first showing that this set is *petite* and the property that under aperiodicity and irreducibility, every petite set is small. As before in Sect. 7.6.3, to establish the petiteness property, we can establish the uniform countable additivity condition (C.2) through (7.21). Therefore, every compact set $\{x, \Delta : |x| \leq F_0, |h| \leq 1, \Delta \in S_\Delta\}$ for some finite F_0 and finite S_Δ is petite in view of the discussion above.

If the integers $\tilde{A}, \tilde{B}$ are relatively prime, then by Lemma 7.6.2, the communication class will include the bin sizes whose logarithms are integer multiples of a constant except those leading to $\Delta < L'$.

We finally can show that the Markov chain is aperiodic. This follows from the fact that the smallest admissible state for the quantizer, $\Delta^* = L'$, can be visited in subsequent time stages with nonzero probability, since

$$\biggl(\min_{|x|\leq \Delta^*/2} P(w_t \in [-2^{R'-1}\Delta^* - x, 2^{R'-1}\Delta^* - x])\biggr)p > 0.$$

□

8.7.6 Proof of Theorem 8.3.3

First, let us note that by (8.42) and (8.46), for every $\kappa > 0$, we can find $\varDelta_0$ such that

$$\lim_{t\to\infty} \frac{P(\tau_1 \geq t|x_0, \varDelta_0)}{(1-p+\kappa)^{t-1}} = 0.$$

We can pick $\kappa > 0$ such that $(1-p+\kappa)|a+\delta|^2 < 1$. Such κ, δ exist since, by hypothesis, $(1-p)|a|^2 < 1$.

We now observe that for all x_0 such that $|h_0| \leq 1$:

$$\begin{aligned}
&\lim_{\varDelta_0\to\infty} \frac{E[V_2(x_{\tau_1}, \varDelta_{\tau_1}) \mid x_0, \varDelta_0]}{V_2(x_0, \varDelta_0)} = \lim_{\varDelta_0\to\infty} \frac{E[\varDelta_{\tau_1}^2 \mid x_0, \varDelta_0]}{\varDelta_0^2} \\
&= \lim_{\varDelta_0\to\infty} \frac{1}{\varDelta_0^2} \sum_{k=1}^{\infty} P(\tau_1 = k) E[\varDelta_k^2 | \tau_1 = k, x_0, \varDelta_0] \\
&= \lim_{\varDelta_0\to\infty} \alpha^2 \sum_{k=1}^{\infty} P(\tau_1 = k)(|a|+\delta)^{2(k-1)} \\
&= p\alpha^2 \frac{1}{1-(1-p)(|a|+\delta)^2},
\end{aligned} \tag{8.54}$$

where the last equality follows from Lemma 8.7.1 and the dominated convergence theorem. Now, if (8.10) holds, we can find α such that $R' > \log_2(|a|/\alpha)$, and

$$\frac{p\alpha^2}{1-(1-p)(|a|+\delta)^2} < 1, \tag{8.55}$$

and simultaneously (8.6) is satisfied. We note that (8.55) implies (8.6) since by Jensen's inequality, $\log(p\alpha^2 + (1-p)(|a|+\delta)^2) > p\log(\alpha^2) + (1-p)\log((|a|+\delta)^2)$.

To establish the required drift equation, we first establish the following bound for all $z \geq 0$:

$$\kappa E[\sum_{m=\tau_z}^{\tau_{z+1}-1} x_m^2 \mid x_0, \varDelta_0] \leq \varDelta_{\tau_z}^2 2^{2(R'-1)}, \tag{8.56}$$

for some $\kappa > 0$.

Let, for ease of notation, $\tau_z = 0$. Observe that by Hölder's inequality

$$\begin{aligned}
&E[\sum_{t=0}^{\tau_1-1} x_t^2 | x_0, \varDelta_0] = E[\sum_{t=0}^{\infty} 1_{\{t<\tau_1\}} x_t^2 | x_0, \varDelta_0] \\
&\leq \sum_{t=0}^{\infty} \left(E[(1_{\{t<\tau_1\}})^{1+\chi} | x_0, \varDelta_0] \right)^{\frac{1}{1+\chi}} \left(E[x_t^{2(\frac{1+\chi}{\chi})} | x_0, \varDelta_0] \right)^{\frac{\chi}{1+\chi}},
\end{aligned} \tag{8.57}$$

for some $\chi > 0$. Now

$$
\begin{aligned}
&E[x_t^{2(\frac{1+\chi}{\chi})}|x_0, \Delta_0] = E[a^{2t(\frac{1+\chi}{\chi})}(x_0 + \sum_{i=0}^{t-1} a^{-i-1} w_i)^{2(\frac{1+\chi}{\chi})}|x_0, \Delta_0] \\
&\le |a|^{2t(\frac{1+\chi}{\chi})} E[(x_0 + \sum_{i=0}^{\infty} a^{-i-1} w_i)^{2\frac{1+\chi}{\chi}}|x_0, \Delta_0] \\
&= |a|^{2t(\frac{1+\chi}{\chi})} (2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} E[(\frac{x_0 + \sum_{i=0}^{\infty} a^{-i-1} w_i}{2^{R'-1}\Delta_0})^{2\frac{1+\chi}{\chi}}|x_0, \Delta_0] \\
&= |a|^{2t(\frac{1+\chi}{\chi})} (2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} E[(h_0 + \frac{\sum_{i=0}^{\infty} a^{-i-1} w_i}{2^{R'-1}\Delta_0})^{2\frac{1+\chi}{\chi}}|x_0, \Delta_0] \\
&< L_2 (2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} |a|^{2t(\frac{1+\chi}{\chi})},
\end{aligned}
\tag{8.58}
$$

for some $L_2 < \infty$, where the last inequality follows since for every fixed $|h_0| \le 1$, the random variable $(h_0 + \frac{\sum_{i=0}^{\infty} a^{-i-1} w_i}{2^{R'-1}\Delta_0})$ has a Gaussian distribution with finite moments, uniform on $\Delta_0 \ge L'$.

Thus,

$$
\begin{aligned}
&E[\sum_{t=0}^{\tau_1 - 1} x_t^2 | x_0, \Delta_0] \\
&\le \sum_{t=0}^{\infty} \Big(E[(1_{\{t<\tau_1\}})^{1+\chi}|x_0, \Delta_0] \Big)^{\frac{1}{1+\chi}} \Big(L_2 (2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} |a|^{2t(\frac{1+\chi}{\chi})} \Big)^{\frac{\chi}{1+\chi}} \\
&= (2^{R'-1}\Delta_0)^2 \sum_{t=0}^{\infty} \Big(P(\tau_1 \ge t+1 | x_0, \Delta_0) \Big)^{\frac{1}{1+\chi}} \Big(L_2 |a|^{2t(\frac{1+\chi}{\chi})} \Big)^{\frac{\chi}{1+\chi}} \\
&< \zeta_{L_2} (2^{R'-1}\Delta_0)^2
\end{aligned}
$$

for some $\zeta_{L_2} < \infty$. The last inequality is due to the fact there exists $\kappa > 0$ such that

$$\lim_{t\to\infty} \frac{P(\tau_1 \ge t | x_0, \Delta_0)}{(1 - p + \kappa)^{t-1}} = 0,$$

and we can pick $\chi > 0$ with $(1 - p + \kappa)|a|^{2(1+\chi)} < 1$. Such χ and κ exist by the hypothesis that $(1-p)|a|^2 < 1$. Hence, with $0 < \epsilon < 1 - \frac{p\alpha^2}{1-(1-p)(|a|+\delta)^2}$,

$$\delta(x, \Delta) = \epsilon \Delta^2, \quad f(x, \Delta) = \frac{\epsilon}{\zeta_{L_2} 2^{2(R'-1)}} x^2,$$

C a compact set and $V_2(x, \Delta) = \Delta^2$, Theorem 6.2.4 applies and $\lim_{t\to\infty} E[x_t^2] < \infty$. □

8.7.7 *Proof of Theorem 8.4.1*

Auxiliary Results and the Stopping Time Analysis

This section presents an important supporting result on stopping time distributions, which is key in the use of Theorem 6.2.4 for the stochastic stability results. We begin with the following.

Lemma 8.7.3. *Let* $\mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$ *denote the Borel* σ*-field on* $\mathbb{R} \times \mathbb{R}_+$. *Then,*

$$P\Big((x_{kn}, \Delta_{kn}) \in (C \times D) | (x_{(k-1)n}, \Delta_{(k-1)n}), \ldots, (x_0, \Delta_0)\Big)$$
$$= P\Big((x_{kn}, \Delta_{kn}) \in (C \times D) | (x_{(k-1)n}, \Delta_{(k-1)n})\Big),$$

$\forall (C \times D) \in \mathcal{B}(\mathbb{R} \times \mathbb{R}_+)$, *i.e.,* (x_{tn}, Δ_{tn}) *is a Markov chain.* ◇

The above follows from the observations that the channel is memoryless, the encoding only depends on the most recent samples of the state and the quantizer, and the control policies use the channel outputs received in the last block, which stochastically only depend on the parameters in the previous block.

Let us define $h_t := \frac{x_t}{\Delta_t 2^{R'-1}}$. We will say that the quantizer is perfectly zoomed when $|h_t| \leq 1$ and under-zoomed otherwise.

Define a sequence of stopping times for the perfect-zoom case with (where the initial state is perfectly zoomed at τ_0)

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{kn > \tau_z : |h_{kn}| \leq 1\}, \quad z, k \in \mathbb{Z}_+. \tag{8.59}$$

As discussed in Sect. 8.4, there will be three types of errors (of Types I-A, I-B, II).

Lemma 8.7.4. *The discrete probability distribution* $P(\tau_{z+1} - \tau_z | x_{\tau_z}, \Delta_{\tau_z})$ *is asymptotically, in the limit of large* Δ_{τ_z}*, dominated (majorized) by a geometrically distributed measure. That is, for* $k \geq \lceil 1/\kappa \rceil + 1$,

$$\begin{aligned} &P(\tau_{z+1} - \tau_z \geq kn | x_{\tau_z}, \Delta_{\tau_z}) \\ &\leq \Xi(\Delta_{\tau_z}) \Big((1 - P^e_{g|g} - P^e_{\mathcal{Z}|g})(eP_e^{(\kappa)})^{k-2} \\ &\qquad + P^e_{g|g}(eP_e^{(\kappa - \frac{1-\kappa}{k-2})})^{k-2} + (P^e_{\mathcal{Z}|g})(eP_e^{(\kappa + \frac{\kappa}{k-2})})^{k-2} \Big). \end{aligned} \tag{8.60}$$

where $\Xi(\Delta_{\tau_z}) < \infty$ *and* $\Xi(\Delta_{\tau_z}) \to 1$ *as* $\Delta_{\tau_z} \to \infty$ *for every fixed* n*, uniformly in* $|h_0| \leq 1$ *and*

$$\kappa < \frac{1}{\log_{\frac{|a|+\delta}{|a|}}(\frac{|a|+\delta}{\alpha})}. \tag{8.61}$$

◇

Proof. Let for $k \in \mathbb{N}$,

$$\Theta_k := P(\tau_{z+1} - \tau_z \geq kn | x_{\tau_z}, \Delta_{\tau_z}). \tag{8.62}$$

Without any loss of generality, let $z = 0$ and $\tau_0 = 0$, so that $\Theta_k = P(\tau_1 \geq kn | x_0, \Delta_0)$.

Before proceeding with the proof, we highlight the technical difficulties that will arise when the quantizer is in the under-zoom phase. As elaborated on earlier, the errors at time 0 are crucial for obtaining the error bounds: At time 0, at most with probability $P^e_{g|g}(n)$, an error will take place so that the quantizer will be zoomed in, yet an incorrect control signal will be applied. With probability at most $P^e_{\mathcal{Z}|g}(n)$, an error will take place so that no control action is applied and the quantizer is zoomed out. At consecutive time stages, until the next stopping time, the quantizer should ideally zoom out but an error takes place with probability $P^e_{g|\mathcal{Z}}(n)$ and forces the quantizer to be zoomed in and a control action to be applied. Our analysis below will address all of these issues.

In the following we will assume that the probabilities are conditioned on particular x_0, Δ_0 values, to ease the notation.

We first consider the case where there is an intra-granular, Type I-A, error at time 0, which takes place at most with probability $P^e_{g|g}$ (this happens to be the worst-case error for the stopping time distribution). Now,

$$\begin{aligned}
&P(\tau_1 \geq kn | \text{Type I-A error at time } 0) \\
&= P\Big(\bigcap_{m=1}^{k-1} (|h_{mn}| > 1) | \text{Type I-A error at time } 0\Big) \\
&= P\Big(\bigcap_{m=1}^{k-1} (|x_{mn}| \geq 2^{R'-1}(|a| + \delta)^{(m-s_m-1)n} \alpha^{(1+s_m)n} \Delta_0)\Big) \\
&= P\Big(\bigcap_{m=1}^{k-1} (|a^{mn}(x_0 + \sum_{i=0}^{mn-1} a^{-i-1}(w_i + u_i))| \\
&\qquad\qquad \geq 2^{R'-1}(|a| + \delta)^{(m-s_m-1)n} \alpha^{(1+s_m)n} \Delta_0)\Big) \\
&= P\Big(\bigcap_{m=1}^{k-1} (|(x_0 + \sum_{i=0}^{mn-1} a^{-i-1}(w_i + u_i))| \geq \\
&\qquad\qquad \frac{2^{R'-1}\alpha^n}{|a^n|} (\frac{|a| + \delta}{|a|})^{(m-1)n} (\frac{\alpha}{|a| + \delta})^{(s_m)n} \Delta_0)\Big).
\end{aligned} \tag{8.63}$$

In the above, s_m is the number of errors in the transmissions that have taken place until (but not including) time m, except for the one at time 0. An error at time 0 would possibly lead to a further enlargement of the bin size with nonzero probability, whereas no error at time 0 leads to a strict decrease in the bin size.

The study for the number of errors is crucial for analyzing the stopping time distributions. In the following, we will condition on the number of erroneous transmissions for k successive block codings for the under-zoomed phase. Suppose that for $k > 1$, there are s_k total erroneous transmissions in the time stages $\{n, 2n, \ldots, (k-1)n\}$ when the state is in fact under-zoomed, but the controller interprets the received transmission as a successful one. Thus, we take $s_1 = 0$.

Let $\zeta_1, \zeta_2, \ldots, \zeta_{s_{k-1}}$ be the time stages when errors take place:

$$\zeta_{t+1} : \min(\min(m > \zeta_t : c'_{nm} \neq c_{nm}), k-1), \quad \zeta_0 = 0,$$

such that $\zeta_{s_{k-1}+1} = k-1$ or $\zeta_{s_{k-1}} = k-1$ and define $\eta_t = \zeta_{t+1} - \zeta_t$.

In the time interval $[\zeta_t n+1, \zeta_{t+1}n-1]$ the system is open-loop, that is, there is no control signal applied, as there is no misinterpretation by the controller. However, there will be a nonzero control signal at times $\{\zeta_k n, k \geq 0\}$. These are, however, upper bounded by the ranges of the quantizers at the corresponding time stages. That is, when an erroneous interpretation by the controller arises, the control applied $-(a^n/b)u_{(\zeta_z+1)n-1}$ lives in the set: $\{a^n(-2^{R'-1} + k - (1/2))\Delta_{\zeta_z}, 1 \leq k \leq 2^{R'}\}$.

From (8.63), we have that

$$\begin{aligned}
&P\Big(\bigcap_{m=1}^{k-1} (|h_{mn}| > 1) | \text{Type I-A error at time } 0\Big) \\
&\leq P\Big(\bigcap_{m=1}^{k-1} \Big(|a^{mn}(x_0 + \sum_{i=0}^{mn-1} a^{-i-1}(w_i + u_i))| \\
&\qquad\qquad \geq 2^{R'-1}(|a|+\delta)^{(m-s_m-1)n}\alpha^{(1+s_m)n}\Delta_0\Big)\Big) \\
&\leq P\Big(\bigcup_{p=0}^{k-2} \Big(\{s_{k-1} = p\} \\
&\qquad \bigcap \Big\{ \bigcap_{m=1}^{p} (|a^{\zeta_m n}(x_0 + \sum_{i=0}^{\zeta_m n-1} a^{-i-1}w_i + \sum_{i=0}^{m-1} a^{(-\zeta_i-1)n}u_{(\zeta_i+1)n-1})| \\
&\qquad\qquad\qquad \geq 2^{R'-1}(|a|+\delta)^{(\zeta_m-s_m-1)n}\alpha^{(1+s_m)n}\Delta_0)\Big\}\Big)\Big) \\
&= \sum_{p=0}^{k-2} \binom{k-2}{p} (P^e_{g|\mathcal{Z}})^p (1-(P^e_{g|\mathcal{Z}}))^{k-1-p} 1_{\{s_{k-1}=p\}} \\
&\qquad \times P\Big(\bigcap_{m=1}^{p} (|a^{\zeta_m n}(x_0 + \sum_{i=0}^{\zeta_m n-1} a^{-i-1}w_i + \sum_{i=0}^{m-1} a^{(-\zeta_i-1)n}u_{(\zeta_i+1)n-1})| \\
&\qquad\qquad \geq 2^{R'-1}(|a|+\delta)^{(\zeta_m-s_m-1)n}\alpha^{(1+s_m)n}\Delta_0 | s_{k-1} = p\Big).
\end{aligned} \tag{8.64}$$

Since the control signal $u_{(\zeta_i+1)n-1}$ lives in $\{a^n(-2^{R'-1}+k-(1/2))\Delta_{(\zeta_i)n}, 1 \le k \le 2^{R'}\}$, conditioned on having s_{k-1} errors in the transmissions, the bound writes as

$$
\begin{aligned}
&P\Big\{ \bigcap_{m=1}^{p} \Big(|a^{\zeta_m n}(x_0 + \sum_{i=0}^{\zeta_m n-1} a^{-i-1} w_i + \sum_{i=0}^{m-1} a^{(-\zeta_i-1)n} u_{(\zeta_i+1)n-1})| \\
&\qquad\qquad \ge 2^{R'-1}(|a|+\delta)^{(\zeta_m-s_m-1)n}\alpha^{(1+s_m)n}\Delta_0 \mid s_{k-1}=p \Big) \Big\} \\
&\le \min_{0 \le m \le s_{k-1}} \Big\{ \\
&P\Big(|\sum_{i=0}^{\zeta_m n-1} a^{-i-1} w_i| \ge 2^{R'-1}(\frac{|a|+\delta}{|a|})^{(\zeta_m-s_m-1)n}(\frac{\alpha}{|a|})^{(1+s_m)n}\Delta_0 \\
&\quad - \sum_{i=1}^{m-1} |a|^n(\frac{|a|+\delta}{|a|})^{(\zeta_i-s_i-1)n}(\frac{\alpha}{|a|})^{(s_m+1)n}(2^{R'-1}-1/2)\Delta_0 \Big| s_{k-1}=p \Big) \Big\} \\
&\quad -|x_0| - (2^{R'-1}-1/2)\Delta_0 \qquad\qquad (8.65) \\
&\le \min_{0 \le m \le s_{k-1}} \Big\{ \\
&P\Big(|\bar{d}| \ge 2^{R'-1}(\frac{|a|+\delta}{|a|})^{(\zeta_m-s_m-1)n}(\frac{\alpha}{|a|})^{(1+s_m)n}\Delta_0 \\
&\quad -|x_0| - (2^{R'-1}-1/2)\Delta_0 \\
&\quad - \sum_{i=1}^{m-1} (\frac{|a|+\delta}{|a|})^{(\zeta_m-s_m-1)n}(\frac{\alpha}{|a|})^{(1+s_m)n}(2^{R'-1}-1/2)\Delta_0 \Big| s_{k-1}=p \Big) \Big\},
\end{aligned}
\tag{8.66}
$$

where $\bar{d} = \sum_{i=0}^{\infty} a^{-i-1} w_i$ is a zero-mean Gaussian random variable with variance $\frac{E[d^2]a^{-2}}{1-a^{-2}}$. In the above (8.65) corresponds to the worst-case where, even when the quantizer is zoomed the controller incorrectly picks the worst case control signal and the chain rule for total probability: For two events A, B: $P(A,B) \le \min(P(A), P(B))$. Now, let us consider $m = k-1$. In this case,

$$
\begin{aligned}
P\Big(|\bar{d}| \ge 2^{R'-1}(\frac{|a|+\delta}{|a|})^{(\zeta_m-s_m-1)n}(\frac{\alpha}{|a|})^{(1+s_m)n} \\
-|x_0| - (2^{R'-1}-1/2)\Delta_0 \\
- \sum_{i=1}^{m-1} (\frac{|a|+\delta}{|a|})^{(\zeta_m-s_m-1)n}(\frac{\alpha}{|a|})^{(1+s_m)n}(2^{R'-1}-1/2)\Delta_0 \Big)
\end{aligned}
$$

$$\begin{aligned}\leq P\Bigg(|\bar{d}| \geq 2^{R'-1}(\frac{|a|+\delta}{|a|})^{(\zeta_m - s_m)n}(\frac{\alpha}{|a|})^{(s_m)n}(\frac{\alpha}{|a|+\delta})^{n} \\ \times\Bigg(1 - \sum_{i=1}^{m-1}(\frac{|a|+\delta}{|a|})^{(\zeta_i-\zeta_m)n}(\frac{\alpha}{|a|})^{(s_i-s_m)n}(1-2^{-R'})\Bigg)\varDelta_0 \\ -2^{R'}\varDelta_0\Bigg| s_{k-1} = p\Bigg),\end{aligned}$$

where in the inequality we observe that $|x_0| \leq 2^{R'-1}\varDelta_0$, since the state is zoomed at this time.

In bounding the stopping time distribution, we will consider the condition that

$$(k-1) - (s_{k-1}+1)(\log_{\frac{(|a|+\delta)}{|a|}}(\frac{|a|+\delta}{\alpha})) > \alpha_A(s_{k-1}+1), \tag{8.67}$$

for some arbitrarily small but positive α_A, to be able to establish that

$$\Bigg(1 - \sum_{i=1}^{m-1}(\frac{|a|+\delta}{|a|})^{(\zeta_i-\zeta_m)n}(\frac{\alpha}{|a|})^{(s_i-s_m)n}(1-2^{-R'})\Bigg) > 0 \tag{8.68}$$

and that $(\frac{|a|+\delta}{|a|})^{(\zeta_{k-1}-s_{k-1})n}(\frac{\alpha}{|a|+\delta})^{(s_{k-1})n} > 2$ for sufficiently large n. Now, there exists an m such that $(\frac{|a|+\delta}{|a|})^{\zeta_m n}(\frac{\alpha}{|a+\delta|})^{s_m n} \geq (\frac{|a|+\delta}{|a|})^{(\zeta_{k-1}-s_{k-1})n}(\frac{\alpha}{|a|+\delta})^{(s_{k-1})n}$ and for this m,

$$\begin{aligned}\Bigg(1 - \sum_{i=1}^{m-1}(\frac{|a|+\delta}{|a|})^{n(\zeta_i-\zeta_m-(\log_{\frac{(|a|+\delta)}{|a|}}(\frac{|a|+\delta}{\alpha}))(s_i-s_m))}\Bigg) \\ \geq (1 - \sum_{i=1}^{m-1}(\frac{|a|+\delta}{|a|})^{-\alpha_A i n}) > 0.\end{aligned} \tag{8.69}$$

This follows from considering a conservative configuration among an increasing subsequence of times $\{\zeta_{i_1}, \dots, \zeta_{i_n}\}$, $i_n \leq k-1$, such that for all elements i_j of this sequence:

$$(\frac{|a|+\delta}{|a|})^{\zeta_{i_j} n}(\frac{\alpha}{|a+\delta|})^{s_{i_j} n} \geq (\frac{|a|+\delta}{|a|})^{(\zeta_{k-1}-s_{k-1})n}(\frac{\alpha}{|a|+\delta})^{(s_{k-1})n}$$

and for $i \leq m$, with $\zeta_m \in \{\zeta_{i_1}, \dots, \zeta_{i_n}\}$, $(\zeta_m - \zeta_i - (s_m - s_i)(\log_{\frac{(|a|+\delta)}{|a|}}(\frac{|a|+\delta}{\alpha})) \geq \alpha_A(s_m - s_i)$. Such an ordered sequence provides a conservative configuration which yet satisfies (8.69), by considering if needed, m to be an element in the sequence with a lower index value for which this is satisfied. This has to hold at least for one time ζ_m, since $k-1$ satisfies (8.67). Such a construction ensures that (8.68) is uniformly bounded from below for every k since $\sum_{i=1}^{\infty}(\frac{|a|+\delta}{|a|})^{-\alpha_A in} < 1$ for n large enough.

Hence, by (8.67), for some constant $B_b > 0$, the following holds:

$$P\Bigg(|\bar{d}| \geq B_b\Delta_0\Bigg((\frac{|a|+\delta}{a})^{\zeta_m n}(\frac{\alpha}{(|a|+\delta)})^{(s_m+1)n}\Bigg)$$

$$\leq 2\frac{\sigma'}{B_b\Delta_0\Bigg((\frac{|a|+\delta}{a})^{\zeta_m n}(\frac{\alpha}{|a|+\delta})^{(s_m+1)}\Bigg)}e^{-\Bigg(B_b\Delta_0(\frac{|a|+\delta}{a})^{\zeta_m n}(\frac{\alpha}{(|a|+\delta)})^{(s_m+1)n}\Bigg)^2/2\sigma'^2}. \tag{8.70}$$

The above follows from bounding the complementary error function by the following: $\int_q^\infty \mu(dx) \leq \int_q^\infty \frac{x}{q}\mu(dx)$, for $q > 0$ when μ is a zero-mean Gaussian measure. In the above derivation $\sigma'^2 = E[w_1^2]|a|^{-2}/(1-|a|^{-2})$. The above can be further upper bounded by, for any $r > 0$:

$$M_r(\Delta_0)r^{-\Bigg((\frac{|a|+\delta}{a})^{\zeta_m n}(\frac{\alpha}{|a|+\delta})^{(s_m+1)n}\Bigg)}$$

$$+\Bigg(1 - 1_{\{\zeta_m-(s_m+1)(\log_{\frac{(|a|+\delta)}{|a|}}(\frac{|a|+\delta}{\alpha}))>\alpha_A(s_m+1)\}}\Bigg) \tag{8.71}$$

with $M_r(\Delta_0) \to 0$ as $\Delta_0 \to \infty$ exponentially: $M_r(\Delta_0)\Delta_0^p \to 0$, for any $p \in \mathbb{N}_+$, due to the exponential dependence of (8.70) in Δ_0. Thus, combined with (8.67), conditioned on having s_{k-1} errors and a Type I-A error at time 0, we have the following bound on (8.65):

$$M_r(\Delta_0)r^{-\Bigg((\frac{|a|+\delta}{a})^{(k-1-s_{k-1})n}(\frac{\alpha}{|a|})^{(s_{k-1}+1)n}\Bigg)} + 1_{\{\zeta_{k-1}\leq\frac{(s_{k-1}+1)}{\kappa}\}} \tag{8.72}$$

with $\kappa = \frac{1}{\log_{\frac{(|a|+\delta)}{|a|}}(\frac{|a|+\delta}{\alpha})+\alpha_A}$. We observe that the number of errors needs to satisfy the following relation for the above bound in (8.71) to be less than 1: $k-1 > (1+s_{k-1})/\kappa$. Finally, the probability that the number of incorrect transmissions exceeds $\kappa(k-1)-1$ is exponentially low, as we observe below.

Let, as before, $P_e(n) = P^e_{g|\mathcal{Z}}(n)$. We consider below Chernoff-Sanov's theorem [59]: The sum of Bernoulli error events leads to a binomial distribution. Let for $1 > \zeta > 0$, $D(\zeta, P_e) = \zeta\log(\zeta/P_e) + (1-\zeta)\log(\frac{1-\zeta}{1-P_e})$. Then, the following upper bound holds [103], for every $k > 3$:

$$
\begin{aligned}
&P\Big(\sum_{t=1}^{k-2} 1_{\{\text{Type II Error}\}} \geq \kappa(k-1)-1\Big)\\
&= P\Big(\sum_{t=1}^{k-2} 1_{\{\text{Type II Error}\}} \geq (k-2)(\kappa - \frac{1-\kappa}{k-2})\Big)\\
&\leq e^{-(k-2)D((\kappa-\frac{1-\kappa}{k-2}),P_e)}. \qquad (8.73)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&P\Big(\sum_{t=1}^{k-2} 1_{\{\text{Type II Error}\}} \geq (\kappa - \frac{1-\kappa}{k-2})(k-2)\Big)\\
&\quad \leq (e^{H((\kappa-\frac{1-\kappa}{k-2}))} P_e^{(\kappa-\frac{1-\kappa}{k-2})})^{k-2} \leq (eP_e^{(\kappa-\frac{1-\kappa}{k-2})})^{k-2}, \qquad (8.74)
\end{aligned}
$$

with $H(z) = -z\log(z) - (1-z)\log(1-z) \leq 1$. We could bound the following summation:

$$
\begin{aligned}
&\sum_{s_{k-1}=0}^{\lfloor\kappa(k-1)\rfloor-1} \binom{k-2}{s_{k-1}} M_r(\varDelta_0) r^{-\left((k-1)-(s_{k-1}+1)/\kappa\right)n} (P_e)^{s_{k-1}}(1-P_e)^{k-1-s_{k-1}}\\
&\leq M_r(\varDelta_0)(1-P_e)^{k-1}\left(\sum_{s_{k-1}=0}^{\lfloor(\kappa-\frac{1-\kappa}{k-2})(k-2)\rfloor} \binom{k-2}{s_{k-1}}\right)(\frac{P_e}{1-P_e})^{\kappa(k-1)-1} \qquad (8.75)\\
&\leq M_r(\varDelta_0)(2P_e^{(\kappa-\frac{1-\kappa}{k-2})})^{(k-2)}, \qquad (8.76)
\end{aligned}
$$

where (8.75) holds since r can be taken to be $r > (\frac{1-P_e}{P_e})^{\kappa}$ by taking $\varDelta_0$ to be large enough and in the summations s_{k-1} taken to be $\kappa(k-1)-1$.

Thus, from (8.64) we have computed a bound on the stopping time distributions through (8.74) and (8.76). Following similar steps for the Type I-B error and no error cases at time 0, we obtain the bounds on the stopping time distributions as follows:

- Conditioned an error in the granular region (Type I-A) at time 0, the condition for the number of errors is that $k-1 > (1+s_{k-1})/\kappa$, and by adding (8.74) and (8.76), the stopping time is dominated by:

$$
\begin{aligned}
P(\tau_1 \geq kn) &\leq M_r(\varDelta_0)(2P_e^{(\kappa-\frac{1-\kappa}{k-2})})^{(k-2)} + (eP_e^{(\kappa-\frac{1-\kappa}{k-2})})^{k-2}\\
&\leq \varXi(\varDelta_0)(eP_e^{(\kappa-\frac{1-\kappa}{k-2})})^{k-2} \qquad (8.77)
\end{aligned}
$$

for $\varXi(\varDelta) = M_r(\varDelta) + 1$ which goes to 1 as $\varDelta \to \infty$.

- Conditioned on the error that $\mathcal{Z}$ is the decoding output at time 0, in the above, the condition for the number of errors is that $k - 1 > s_{k-1}/\kappa$, and we may replace the exponent term $(\kappa - \frac{1-\kappa}{k-2})$ with $(\kappa + \frac{\kappa}{k-2})$ and the stopping time is dominated by

$$P(\tau_1 \geq kn) \leq \Xi(\Delta_0)(eP_e^{(\kappa + \frac{\kappa}{k-2})})^{(k-2)} \tag{8.78}$$

for $\Xi(\Delta) = M_r(\Delta) + 1$ which goes to 1 as $\Delta \to \infty$.
- Conditioned on no error at time 0 and the rate condition $R' > \log_2(|a|/\alpha)$, the condition for the number of errors is that $k - 1 > 1 + s_{k-1}/\kappa$, and we may replace the exponent term $(\kappa - \frac{1-\kappa}{k-2})$ with κ.

 The reason for this is that $|x_0 - \hat{x}_0| \leq \Delta_0/2$ and the control term applied at time n reduces the error. As a result (8.65) writes as

$$P\Bigg(\cap_{m=1}^{p} (|a^{\zeta_m n}(x_0 + \sum_{i=0}^{\zeta_m n - 1} a^{-i-1} w_i + \sum_{i=0}^{m-1} a^{(-\zeta_i - 1)n} u_{(\zeta_i+1)n-1})|$$

$$\geq 2^{R'-1}(|a| + \delta)^{(\zeta_m - s_m - 1)n} \alpha^{(1+s_m)n} | s_{k-1} = p\Bigg)$$

$$\leq \min_{0 \leq m \leq s_{k-1}} \Bigg\{$$

$$P\Bigg\{ |\bar{d}| \geq \Bigg(2^{R'-1}(\frac{\alpha}{|a|})^n \Bigg((\frac{|a|+\delta}{|a|})^{(\zeta_m - s_m - 1)n} (\frac{\alpha}{|a|})^{s_m n} - 2^{-R'}(\frac{\alpha}{|a|})^{-n} \Bigg) \Delta_0$$

$$- \sum_{i=1}^{m-1} (2^{R'-1}(\frac{\alpha}{|a|})^n (\frac{|a| + \delta}{|a|})^{(\zeta_i - s_i - 1)n} (\frac{\alpha}{|a|})^{s_i n} (1 - 2^{-R'})\Delta_0) \Bigg) \Bigg\} \Bigg\}.$$

Since $2^{R'-1}(\frac{\alpha}{|a|})^n > 1$, the effect of the additional 1 in the exponent for $(\frac{\alpha}{|a|})^{s_m + 1}$ can be excluded, unlike the case with $P^g_{e|e}$ above in (8.67).

 As a result, the stopping time is dominated by

$$P(\tau_1 \geq kn) \leq \Xi(\Delta_0)(eP_e^{\kappa})^{k-2}, \tag{8.79}$$

for $\Xi(\Delta) = M_r(\Delta) + 1$ which goes to 1 as $\Delta \to \infty$.

This completes the proof of the lemma. □

Proof of Theorem 8.4.1

Once we have the Markov chain by Lemma 8.7.3, and the bound on the distribution of the sequence of stopping times defined in (8.59) by Lemma 8.7.4, we can invoke

Theorems 6.2.4 or 6.2.6 with Lyapunov functions $V(x, \varDelta) = \log_2(\varDelta^2)$, $f(x, \varDelta)$ taken as a constant, and C a compact set.

As mentioned in Remark 8.4.1, for a DMC with block length n, Shannon's random coding method satisfies:

$$P_e(n) := \max_{c \in \{1,2,\ldots,M(n)\}} P(c' \neq c | c \quad \text{is transmitted}) \leq e^{-nE(R)+o(n)},$$

with c' being the decoder output. Here $\frac{o(n)}{n} \to 0$ as $n \to \infty$ and $E(R) > 0$ for $0 < R < C$. Thus, by Lemma 8.7.4, we have that

$$E[\tau_1 | x_0, \varDelta_0] = \sum_{k=1}^{\infty} P(\tau_1 \geq k) \leq K'_{\varDelta_0}(n) < \infty, \tag{8.80}$$

for some finite number $K'_{\varDelta_0}(n)$. The finiteness of this expression follows from the observation that for $k-2 > \frac{1-\kappa}{\kappa}$, the exponent in $e^{-n(\kappa - \frac{1-\kappa}{k-2})(E(R) - \frac{o(n)}{n})}$ becomes negative. Furthermore, $K'_{\varDelta_0}(n)$ is monotone decreasing in $\varDelta_0$ since $M_r(\varDelta)$ is decreasing in $\varDelta$.

We now apply the random-time drift result in Theorem 6.2.4 and Corollary 6.2.1 below. First, observe that the probability that $\tau_{z+1} \neq \tau_z + n$ is upper bounded by the probability below:

$$\begin{aligned}
&P^e_{g|g} + (1 - P^e_{g|g} - P^e_{\mathcal{Z}|g})2P\Big(\bar{d} \geq (2^{R'}(\frac{\alpha}{|a|})^n - 1)\varDelta_0/2\Big) \\
&\quad +2P^e_{\mathcal{Z}|g}P\Big(\bar{d} > (2^{R'-1}(|a| + \delta)^n)\varDelta_0 - |a^n x_0|\Big) \\
&\leq P^e_{g|g} + (1 - P^e_{g|g} - P^e_{\mathcal{Z}|g})2P\Big(\bar{d} \geq (2^{R'}(\frac{\alpha}{|a|})^n - 1)\varDelta_0/2\Big) \\
&\quad +P^e_{\mathcal{Z}|g}2P\Big(\bar{d} > 2^{R'-1}((|a| + \delta)^n - |a|^n)\varDelta_0\Big) \\
&=: \varUpsilon(\varDelta_{\tau_0}).
\end{aligned} \tag{8.81}$$

Note that, provided that $R'(n) > n\log_2(|a|/\alpha)$, $\lim_{\varDelta_0 \to \infty} \varUpsilon(\varDelta_{\tau_0}) = P^e_{g|g}$.

We now pick the Lyapunov function $V(x, \varDelta) = \log_2(\varDelta^2)$ and $f(x, \varDelta)$ a constant to obtain the following:

$$\begin{aligned}
&E[\log(\varDelta^2_{\tau_{z+1}}) | x_{\tau_z}, \varDelta_{\tau_z}] \\
&= E[\log(\varDelta^2_{\tau_{z+1}})(1_{\{\text{Type I-A error at } \tau_z\}} + 1_{\{\text{Type I-B error at } \tau_z\}} \\
&\qquad\qquad +1_{\{\text{no error at } \tau_z\}}) | x_{\tau_z}, \varDelta_{\tau_z}]
\end{aligned}$$

$$
\begin{aligned}
&\leq (1 - P^e_{\mathcal{Z}|g} - P^e_{g|g})\Bigg(n\log_2(\alpha) \\
&\qquad\qquad +nE[\log_2((|a+\delta)^{2(\tau_{z+1}-1)})1_{\{\tau_{z+1}>\tau_z+n\}})|\text{no error}]\Bigg) \\
&+P^e_{\mathcal{Z}|g}\Bigg(n\log_2(|a|+\delta) \\
&\qquad\qquad +nE[\log_2((|a+\delta)^{2(\tau_{z+1}-1)})1_{\{\tau_{z+1}>\tau_z+n\}})|\text{Type I-B error}]\Bigg) \\
&+P^e_{g|g}\Bigg(n\log_2(\alpha) + nE[\log_2((|a+\delta)^{2(\tau_{z+1}-1)})1_{\{\tau_{z+1}>\tau_z+n\}}|\text{Type I-A error}]\Bigg) \\
&\quad + \log_2(\Delta^2_{\tau_z})
\end{aligned}
$$

$$
\begin{aligned}
&= \log(\Delta^2_{\tau_z}) + n\Bigg((1-P^e_{\mathcal{Z}|g})\log_2(\alpha) + P^e_{\mathcal{Z}|g}\log_2(|a|+\delta)\Bigg) \\
&\qquad +nE\Bigg[\log_2\Bigg((|a+\delta)^{2(\tau_{z+1}-1)}\Bigg)1_{\{\tau_{z+1}>\tau_z+n\}})\Bigg] \\
&\leq \log(\Delta^2_{\tau_z}) + n\Bigg((1-P^e_{\mathcal{Z}|g})\log_2(\alpha) + P^e_{\mathcal{Z}|g}\log_2(|a|+\delta)\Bigg) \\
&+n\Bigg(P(\tau_{z+1}>\tau_z+n)\Bigg)^{\frac{\chi}{1+\chi}} \\
&\qquad\qquad \times\Bigg(\sum_{k=2}^{\infty} P(\tau_{z+1}=\tau_z+kn)((k-1)\log_2(|a|+\delta))^{1+\chi}\Bigg)^{\frac{1}{1+\chi}} \qquad (8.82) \\
&\leq \log(\Delta^2_{\tau_z}) + n\Bigg((1-P^e_{\mathcal{Z}|g})\log_2(\alpha) + P^e_{\mathcal{Z}|g}\log_2(|a|+\delta)\Bigg) \\
&\quad + (\Upsilon(\Delta_{\tau_z}))^{\frac{\chi}{1+\chi}}n\Bigg(\sum_{k=2}^{\infty} P(\tau_{z+1}=\tau_z+kn)((k-1)\log_2(|a|+\delta))^{1+\chi}\Bigg)^{\frac{1}{1+\chi}},
\end{aligned}
\tag{8.83}
$$

where $\chi > 0$ is an arbitrarily small positive number. In (8.82) we use the facts that (i) zooming out for all time stages after $\tau_z + n$ provides a worst-case sequence, and (ii) by Hölder's inequality for a random variable X and an event $\mathbb{A}$ the following holds:

$$
E[X1_{\mathbb{A}}] \leq (E[|X|^{1+\chi}])^{\frac{1}{1+\chi}}(E[1_{\mathbb{A}}^{\frac{1+\chi}{\chi}}])^{\frac{\chi}{1+\chi}} = (E[|X|^{1+\chi}])^{\frac{1}{1+\chi}}(P(\mathbb{A}))^{\frac{\chi}{1+\chi}}.
$$

Now, the last term in (8.83) will converge to zero with n large enough and as $\Delta_{\tau_z} \to \infty$ for some $\chi > 0$, since by Lemma 8.7.4 $P(\tau_{z+1} = \tau_z + kn)$ is bounded by a geometric measure and the expectation of $((\tau_{z+1} - \tau_z - 1)\log_2(|a| + \delta))^{1+\chi}$ is finite (and decreasing in Δ_0). The second term in (8.83) is negative with $P^e_{\mathcal{Z}|g}$ sufficiently small.

As a result, for some sufficiently large F, the inequality

$$E[\log(\Delta^2_{\tau_{z+1}})|\Delta_{\tau_z}, h_{\tau_z}] \le \log(\Delta^2_{\tau_z}) - b_0 + b_1 1_{\{|\Delta_{\tau_z}| \le F\}} \tag{8.84}$$

holds for some positive b_0 and finite b_1. Here, b_1 is finite since $K'(n)$ is finite. With the uniform boundedness of (8.80) over the sequence of stopping times, this implies by Theorem 6.2.6 that $\{(x, \Delta) : |\Delta_{\tau_z}| \le F, |\frac{x}{2^{R'-1}\Delta}| \le 1\}$ is a recurrent set.

□

8.7.8 Proof of Theorem 8.4.2

In this section, we establish irreducibility and the existence of a small set, to be able to invoke Theorem 6.2.4. The following follows the same approach as in Sect. 7.6.3. In this setting, (x_{tn}, Δ_{tn}) form the Markov chain, as was observed in Lemma 8.7.3.

Let the values taken by $\log_2(\bar{Q}(\Delta_{tn}, c'_{(t+1)n-1}))/s$ be $\{-\tilde{A}, 0, \tilde{B}\}$. Here $\tilde{A}, B$ are relatively prime. Let $\mathbb{L}_{z_0,\tilde{A},\tilde{B}}$ be defined as

$$\{n \in \mathbb{N}, n \ge \log_2(L')/s : \exists N_A, N_B, n = -N_A\tilde{A} + N_B\tilde{B} + z_0\},$$

where $z_0 = \log_2(\Delta_0)/s$ is the initial value of the parameter for the quantizer. We note that since $\tilde{A}, B$ are relatively prime, then as before, by Bézout's Lemma, the communication class will include the bin sizes whose logarithms are integer multiples of a constant except those leading to $\Delta < L'$: Since we have $\Delta_{(t+1)n} = \bar{Q}(\Delta_{tn}, c'_{(t+1)n-1})\Delta_{tn}$, it follows that

$$\log_2(\Delta_{(t+1)n})/s = \log_2(\bar{Q}(\Delta_{tn}, c'_{(t+1)n-1}))/s + \log_2(\Delta_{tn})/s$$

is also an integer. Furthermore, since the source process $\{x_{tn}\}$ is Lebesgue-irreducible, and there is a uniform lower bound L' on bin sizes, the error process takes values in any of the admissible quantizer bins with nonzero probability. Consider two integers $k, l \ge \frac{\log_2(L')}{s}$. For all $l, k \in \mathbb{L}_{z_0,\tilde{A},\tilde{B}}$, there exist $N_A, N_B \in \mathbb{N}$ such that $l - k = -N_A\tilde{A} + N_B\tilde{B}$. We can show that the probability of N_A occurrences of perfect zoom, and N_B occurrences of under-zoom phases is bounded away from zero. This set of occurrences includes the event that in the first N_A time stages perfect zoom occurs and later, successively, N_B times under-zoom phase occurs. Considering worst possible control realizations and errors, the probability of this event is lower bounded by

$$\left(P\left(\tilde{d} \in [-2^{R'(n)-1}L' - |a|^n L', 2^{R'(n)-1}L' - |a|^n L']\right)(P^e(\mathcal{Z}|i))\right)^{N_B}$$

$$\left(P\left(\tilde{d}\in[-(\alpha^n 2^{R'} - a^n)L'/2, (\alpha^n 2^{R'} - a^n)L'/2])(1-P_e)\right)^{N_A} > 0, \qquad (8.85)$$

where $\tilde{d} = \sum_{i=0}^{n-1} a^{n-i-1} w_i$ is a Gaussian random variable. The above follows from considering the sequence of zoom-ins and zoom-outs and the behavior of $a^n(x_{tn} - \hat{x}_{tn}) + \tilde{d}$. In the above discussion, $P^e(\mathcal{Z}|i)$ is the conditional error on the zoom symbol given the transmission of granular bin i, with the lowest error probability (if the lowest such an error probability is zero, an alternative sequence of events can be provided through events concerning the noise variables leading to zooming). Thus, for any two such integers k, l and for some $r > 0$, $P(\log_2(\Delta_{(t+r)n}) = ls \mid \log_2(\Delta_{tn}) = ks) > 0$.

We can establish the irreducibility and aperiodicity of the sampled Markov chain, by following the approach in Sect. 7.6.3. The rest of the argument for petiteness follows that of Sect. 7.6.3 and (7.21) for the sampled chain, in view of the uniform countable additivity condition. Hence, we can establish that the set $C_x \times C'_\Delta = \{(x, \Delta) : L' \leq \Delta \leq F, |h| \leq 1\}$ is petite.

Aperiodicity of the sampled chain follows from the fact that the smallest admissible state for the quantizer, L', can be visited in subsequent time stages, since

$$P(\tilde{d} \in [-2^{R'(n)-1}L'/|a|^n - L', -2^{R'(n)-1}L'/|a|^n + L']) > 0. \qquad \square$$

8.7.9 Proof of Theorem 8.4.3

By Kolmogorov's extension theorem, it suffices to check that the property holds for finite-dimensional cylinder sets, since these sets generate the σ-algebra on which the stochastic process measure is defined. Suppose first that the sampled Markov chain is stationary. Consider two elements:

$$\begin{aligned}
&P(x_{t+1+n} \in A_1, x_{t+2+n} \in A_2) \\
&= \int_{x_{\lfloor \frac{t+1+n}{n} \rfloor n}} P(dx_{\lfloor \frac{t+1+n}{n} \rfloor n}, x_{t+1+n} \in A_1, x_{t+2+n} \in A_2) \\
&= \int_{x_{\lfloor \frac{t+1+n}{n} \rfloor n}} P(x_{t+1+n} \in A_1, x_{t+2+n} \in A_2 | x_{\lfloor \frac{t+1+n}{n} \rfloor n}) \\
&\qquad \times P(dx_{\lfloor \frac{t+1+n}{n} \rfloor n}) \\
&= \int_{x_{\lfloor \frac{t+1}{n} \rfloor n}} P(x_{t+1} \in A_1, x_{t+2} \in A_2 | x_{\lfloor \frac{t+1}{n} \rfloor n}) P(dx_{\lfloor \frac{t+1}{n} \rfloor n}).
\end{aligned}$$

The above holds since the marginals $P(dx_{\lfloor \frac{t+1}{n} \rfloor n})$ and $P(dx_{\lfloor \frac{t+1+n}{n} \rfloor n})$ are equal because the sampled Markov chain is positive Harris recurrent and is assumed to be stationary, and the dynamics for inter-block times are time homogeneous Markov. The above is applicable for any finite-dimensional set, thus for any element in the sigma-field generated by the finite-dimensional sets, on which the stochastic process is defined. Now, let for some event A, $T^{-n}A = A$, where T denotes the shift operation (see Sect. C.3.1). Then

$$P(A) = \lim_{k \to \infty} P(A \cap T^{-kn}A) = \lim_{k \to \infty} P(A)P(T^{-kn}A|A). \tag{8.86}$$

Note that a positive Harris recurrent Markov chain admits a unique invariant distribution and for every $x_0 \in \mathbb{R}$

$$\lim_{k \to \infty} P(x_{kn} \in A|x_0) = \pi(A),$$

where $\pi(\cdot)$ is the unique invariant probability measure. Since such a Markov chain forgets its initial condition, by (8.86), for $A = T^{-n}A$: $P(A) = P(A)P(A)$, thus, $P(A) \in \{0, 1\}$, and the process is n-ergodic. □

8.7.10 Proof of Theorem 8.4.4

We begin with the following result, which is a consequence of Theorem 6.2.4:

Lemma 8.7.5. *Under the conditions of Theorem 8.4.2, we have that, if for some* $\gamma > 0$, $b < \infty$, *the following holds:*

$$\gamma E[\sum_{k=0}^{(\tau_1/n)-1} \Delta_{kn}^2 | x_0, \Delta_0] \leq \Delta_0^2 - E[\Delta_{\tau_1}^2 | x_0, \Delta_0] + b 1_{\{(\Delta_0, h_0) \in (C'_x \times C_h)\}},$$

then $\lim_{k \to \infty} E[\Delta_{kn}^2] < \infty$. ◇

Now, under the hypotheses of Theorem 8.4.2 and observing that Type I-B and I-A errors are worse than the no error case at time 0 for the stopping time tail distributions, we write

$$E[\sum_{t=0}^{(\tau_1/n)-1} \Delta_{tn}^2 \mid x_0, \Delta_0]$$

$$\leq \Delta_0^2 P_{g|g}^e \left(\sum_{l=2}^{\infty} P(\tau_1 = ln | \text{Type I-A error}) \sum_{k=1}^{(l-1)} (|a| + \delta)^{2(k-1)n} \alpha^{2n} \right)$$

$$+\Delta_0^2 P_{\mathcal{Z}|g}^e \left(\sum_{l=2}^{\infty} P(\tau_1 = ln | \text{Type I-B error}) \sum_{k=1}^{l-1} (|a| + \delta)^{2kn} \right)$$

$$
\begin{aligned}
&+\Delta_0^2(1 - P^e_{g|g} - P^e_{\mathcal{Z}|g}) \\
&\times\Bigg(\sum_{l=2}^{\infty} P(\tau_1 = ln|\text{no error at time } 0)\sum_{k=1}^{l-1}(|a| + \delta)^{2(k-1)n}\alpha^{2n}\Bigg) \\
&\qquad +\Delta_0^2 P(\tau_1 = n) \\
\leq{}& \Delta_0^2 P^e_{g|g}\frac{(|a|+\delta)^{(2/\kappa)n}}{(|a|+\delta)^{2n}-1}\Xi(\Delta_0)\Bigg(\sum_{l=2}^{\infty}(e^{(l-2)})P_e^{(\kappa)(l-1-\frac{1}{\kappa})}(|a|+\delta)^{2n(l-1-\frac{1}{\kappa})}\alpha^{2n}\Bigg) \\
&+\Delta_0^2 P^e_{\mathcal{Z}|g}\frac{(|a|+\delta)^{2n}}{1-(|a|+\delta)^{-2n}}\Xi(\Delta_0)\Bigg(\sum_{l=2}^{\infty}(eP_e^{\kappa})^{l-2}(|a|+\delta)^{2(l-2)n}\Bigg) \\
&+\Delta_0^2(1-P^e_{g|g}-P^e_{\mathcal{Z}|g})(|a|+\delta)^{2n}\Xi(\Delta_0)\Bigg(\sum_{l=2}^{\infty}(eP_e^{\kappa})^{l-2}\frac{(|a|+\delta)^{2(l-2)n}}{(|a|+\delta)^{2n}-1}\alpha^{2n}\Bigg) \\
&\qquad +\Delta_0^2 P(\tau_1 = n) \\
\leq{}& \zeta_1\Delta_0^2 \qquad (8.87)
\end{aligned}
$$

for some finite ζ_1. To derive the above, we use the property that $H(\kappa) \leq 1$ as well as the relations (8.77)–(8.79). We now establish that $\lim_{\Delta_0\to\infty} E[\Delta^2_{\tau_1} \mid x_0, \Delta_0]/\Delta_0^2 < 1$. This is a crucial step in the application of Theorem 6.2.4.

Following similar steps as in (8.87), the following upper bound on $(E[\Delta^2_{\tau_1} | x_0, \Delta_0]/\Delta_0^2)$ is obtained:

$$
\begin{aligned}
&(1 - P^e_{g|g} - P^e_{\mathcal{Z}|g}) \\
&\quad\times\Bigg(\alpha^{2n} + \frac{1}{\Delta_0^2}E[\Delta^2_{\tau_1}1_{\{\tau_1>n\}}|\text{no error at time } 0]\Bigg) \\
&+(P^e_{g|g})\Bigg(\alpha^{2n}(1 + (|a| + \delta)^{2n} + \ldots(|a| + \delta)^{2(\lfloor\frac{1}{\kappa}\rfloor)n}) \\
&\quad + \sum_{k=\lceil\frac{1}{\kappa}\rceil+1}^{\infty} e^{k-2}P_e^{(\kappa(k-1-\frac{1}{\kappa}))}(|a| + \delta)^{2(k-1-\frac{1}{\kappa})n}(\alpha)^{2n} \\
&\qquad\quad \times(|a| + \delta)^{(2/\kappa)n}\Xi(\Delta_0)\Bigg) \\
&+P^e_{\mathcal{Z}|g}\Bigg((|a| + \delta)^{2n} + \Xi(\Delta_0)\frac{P_e^{\kappa}(|a| + \delta)^{2n}}{1 - P_e^{\kappa}(|a| + \delta)^{2n}}\Bigg). \qquad (8.88)
\end{aligned}
$$

Now note that

$$
\lim_{\Delta_0\to 0} P(\tau_1 > n|\text{no error at time } 0, x_0, \Delta_0) = 0,
$$

uniformly in x_0 with $|h_0| \leq 1$ and given the rate condition $R'(n) > n \log_2(|a|/\alpha)$ by (8.81). Therefore, the first term in (8.88) converges to 0 as $\Delta_0 \to \infty$, since $\lim_{n\to\infty} \kappa \frac{1}{n} \log(P_e) + 2\log_2(|a| + \delta) < 0$ and we have the following upper bound

$$\left((|a| + \delta)^{2n} \sum_{k=2}^{\infty} (eP_e^{\kappa})^{k-2} (|a| + \delta)^{2n(k-2)} (\alpha)^{2n}\right) < \infty,$$

for sufficiently large n.

For the second term in (8.88), the convergence of the first term above is ensured with $\lim_{n\to\infty} P^e_{g|g}(|a| + \delta)^{(2/\kappa)n} \alpha^{2n} \to 0$ and $P_e(|a| + \delta)^{(2/\kappa)n} \to 0$ as $n \to \infty$. By combining the second and the third terms, the desired result is obtained.

To show that $\lim_{m\to\infty} E[x^2_{mn}] < \infty$, we first show that for some $\kappa > 0$,

$$\kappa E[\sum_{m=0}^{(\tau_1/n)-1} x^2_{mn} \mid x_0, \Delta_0] \leq \Delta_0^2 2^{2(R'-1)}. \tag{8.89}$$

Now,

$$\begin{aligned}
&E[\sum_{m=0}^{(\tau_1/n)-1} x^2_{mn} \mid x_0, \Delta_0] \\
&= E\left[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn} \left((x_0 + \sum_{i=0}^{tn-1} a^{-i-1} w_i) + (\sum_{i=0}^{t-1} a^{-i-1} u_i)\right)^2 \middle| x_0, \Delta_0\right] \\
&\leq 2E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn} (x_0 + \sum_{i=0}^{tn-1} a^{-i-1} w_i)^2 | x_0, \Delta_0] \\
&\quad +2E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn} (\sum_{i=0}^{tn-1} a^{-i-1} u_i)^2 | x_0, \Delta_0],
\end{aligned} \tag{8.90}$$

which follows from the observation that for two real-valued random variables y, z, $E[(y + z)^2] \leq 2E[y^2] + 2E[z^2]$.

Let us first consider the component: $(a^t(x_0 + \sum_{i=0}^{tn-1} a^{-i-1} w_i))^2$.

$$\begin{aligned}
&E[\sum_{t=0}^{(\tau_1/n)-1} (a^{tn}(x_0 + \sum_{i=0}^{tn-1} a^{-i-1} w_i))^2 | x_0, \Delta_0] \\
&= E[\sum_{t=0}^{\infty} 1_{\{t<\tau_1/n\}} (a^{tn}(x_0 + \sum_{i=0}^{tn-1} a^{-i-1} w_i))^2 | x_0, \Delta_0]
\end{aligned}$$

$$\begin{aligned}&\leq \sum_{t=0}^{\infty} \left(E[(1_{\{t<\tau_1/n\}})^{1+\chi}|x_0,h_0] \right)^{\frac{1}{1+\chi}} \\ &\qquad \left(E[(a^{tn}(x_0 + \sum_{i=0}^{tn-1} a^{-i-1}w_i))^{2(\frac{1+\chi}{\chi})}|x_0,\Delta_0] \right)^{\frac{\chi}{1+\chi}},\end{aligned} \tag{8.91}$$

for some $\chi > 0$, by Hölder's inequality.

Moreover, for some $B_2 < \infty$,

$$\begin{aligned}&E[a^{2tn(\frac{1+\chi}{\chi})}(x_0 + \sum_{i=0}^{tn-1} a^{-i-1}w_i)^{2(\frac{1+\chi}{\chi})}|x_0,\Delta_0] \\ &\leq |a|^{2tn(\frac{1+\chi}{\chi})} E[(x_0 + \sum_{i=0}^{\infty} a^{-i-1}w_i)^{2\frac{1+\chi}{\chi}}|x_0,\Delta_0] \\ &= |a|^{2tn(\frac{1+\chi}{\chi})}(2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} E[(\frac{x_0 + \sum_{i=0}^{\infty} a^{-i-1}w_i}{2^{R'-1}\Delta_0})^{2\frac{1+\chi}{\chi}}|x_0,\Delta_0] \\ &= |a|^{2tn(\frac{1+\chi}{\chi})}(2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}} E[(h_0 + \frac{\sum_{i=0}^{\infty} a^{-i-1}w_i}{2^{R'-1}\Delta_0})^{2\frac{1+\chi}{\chi}}|x_0,\Delta_0] \\ &< B_2(2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}}|a|^{2tn(\frac{1+\chi}{\chi})},\end{aligned} \tag{8.92}$$

where the last inequality follows since for every fixed $|h_0| \leq 1$, the random variable $h_0 + (\sum_{i=0}^{\infty} a^{-i-1}w_i)/(2^{R'-1}\Delta_0)$ has a Gaussian distribution with its individual moments uniformly bounded on $\Delta_0 \geq L'$.

Thus,

$$\begin{aligned}&E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn}(x_0 + \sum_{i=0}^{tn-1} a^{-i-1}w_i)^2|x_0,\Delta_0] \\ &\leq \sum_{t=0}^{\infty} \left(E[(1_{\{t<\tau_1/n\}})^{1+\chi}|x_0,\Delta_0] \right)^{\frac{1}{1+\chi}} \left(B_2(2^{R'-1}\Delta_0)^{2\frac{1+\chi}{\chi}}|a|^{2tn(\frac{1+\chi}{\chi})} \right)^{\frac{\chi}{1+\chi}} \\ &= \sum_{t=0}^{\infty} \left(\Xi(\Delta_0)(eP_e^{(\kappa-\frac{1-\kappa}{t-1})})^{t-1} \right)^{\frac{1}{1+\chi}} \left(B_2^{\frac{\chi}{1+\chi}}\Delta_0^2|a|^{2tn} \right) \\ &= \sum_{t=0}^{\infty} \left(\Xi(\Delta_0)(eP_e^{(\kappa-\frac{1-\kappa}{t-1})})^{t-1}|a|^{2tn(1+\chi)} \right)^{\frac{1}{1+\chi}} \left(B_2^{\frac{\chi}{1+\chi}}\Delta_0^2 \right) \\ &< \zeta_{B_2}\Delta_0^2,\end{aligned} \tag{8.93}$$

for some finite ζ_{B_2}. In the discussion above we have used the fact that we can pick $\chi > 0$ such that $(P_e)^{\kappa}|a|^{2n(1+\chi)} < 1$. Such a χ exists by the hypothesis that $\lim_{n\to\infty} P_e^{\kappa}(|a|+\delta)^{2n} = 0$.

We now consider the second term in (8.90). Since u_i is the quantizer output which is bounded in magnitude in proportion with Δ_i, the second term writes as:

$$\begin{aligned}
&E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn}(\sum_{i=0}^{tn-1} a^{-i-1}u_i)^2|x_0,\Delta_0] \\
&\leq E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn}(\sum_{i=0}^{t-1} a^{-in}2^{(R'-1)}(|a|+\delta)^{in}\Delta_0)^2|x_0,\Delta_0] \\
&\leq \left(\frac{1}{(1-(\frac{|a|+\delta}{|a|})^n)}\right)^2 E[\sum_{t=0}^{(\tau_1/n)-1} a^{2tn}\left(2^{(R'-1)}(\frac{|a|+\delta}{|a|})^{tn}\Delta_0\right)^2|x_0,\Delta_0] \\
&\leq \Delta_0^2\tilde{\zeta}_B
\end{aligned} \tag{8.94}$$

for some finite $\tilde{\zeta}_B$, by the bound on the stopping time and arguments presented earlier.

Now, with (8.88), (8.90), and (8.93–8.94), we can apply Theorem 6.2.4: With some $\epsilon > 0$ [whose existence is justified by (8.88)],

$$\delta(x,\Delta) = \epsilon\Delta^2, \quad f(x,\Delta) = \frac{\epsilon}{2\zeta_{B_2}+2\tilde{\zeta}_B}x^2,$$

C a compact set and $V(x,\Delta)=\Delta^2$, Theorem 6.2.4 applies and $\lim_{t\to\infty} E[x_{tn}^2]<\infty$.

Thus, with average rate strictly larger than $\log_2(|a|)$, stability with a finite second moment is achieved. Finally, the limit is independent of the initial distribution since the sampled chain is irreducible, by Theorem 8.4.2. Now, if the sampled process has a finite second moment, the average second moment for the state process satisfies

$$\lim_{N\to\infty}\frac{1}{N}E[\sum_{k=0}^{N-1} x_k^2] = \frac{1}{n}E_{\pi}[\sum_{k=0}^{n-1} x_k^2|x_0,\Delta_0]$$

and is also finite, where E_π denotes the expectation under the invariant probability measure for x_0, Δ_0. By the ergodic theorem for Markov chains (see Theorem 6.2.4), the above holds almost surely, and as a result

$$\lim_{N\to\infty}\frac{1}{N}\left(\sum_{k=0}^{N-1} x_k^2\right) = \frac{1}{n}E_{\pi}[\sum_{k=0}^{n-1} x_k^2|x_0,\Delta_0] < \infty \quad a.s.$$

□

8.7.11 Proof of Theorem 8.4.5

Proof follows from the observation that the number of errors in channel transmission when the state is under-zoomed, s, is zero. No errors take place in the phase when the quantizer is being zoomed out.

Following (8.88), the only term which survives is

$$\alpha^{2n} + P^{e}_{g|g}\Bigg(\alpha^{2n}\Bigg(1 + (|a| + \delta)^{2n} + \ldots (|a| + \delta)^{2(\lfloor \frac{1}{\kappa} \rfloor)n}\Bigg)$$

which is to be less than 1. We can pick $\kappa > 1/2$ for this case. Now,

$$\lim_{\Delta \to \infty} P(\tau \geq kn|x_0, \Delta_0) \leq P(\bar{d} > \frac{(|a| + \delta)^{(k-1)n}\alpha^n}{|a|^{kn}}\Delta_0 2^{R'-1}) = 0$$

for $k > \frac{1}{\kappa}$. Hence, $\lim_{n \to \infty} P^{e}_{g|g}(|a| + \delta)^{2n} \to 0$ is sufficient, since $2 > \frac{1}{\kappa}$. The proof is complete once we recognize $\bar{P}_e$ as $P^{e}_{g|g}$. □

8.7.12 Proof of Theorem 8.5.4

We provide a sketch of the proof since the analysis follows from the scalar case, except for the construction of an adaptive vector quantizer and the associated stopping time distribution.

Consider the following system:

$$\begin{bmatrix} x^1_{t+1} \\ x^2_{t+1} \\ \vdots \\ x^N_{t+1} \end{bmatrix} = \Lambda \begin{bmatrix} x^1_t \\ x^2_t \\ \vdots \\ x^N_t \end{bmatrix} + \tilde{B}u_t + \tilde{G}w_t, \tag{8.95}$$

where $\Lambda = \mathrm{diag}(\lambda^i)$ is a diagonal matrix, obtained via a similarity transformation: $\Lambda = U^{-1}AU$ and $\tilde{B} = U^{-1}B, \tilde{G} = U^{-1}G$, where U consists of the eigenvectors of the matrix A. We can assume without any loss of generality that $\tilde{B}$ is invertible since otherwise, by the controllability assumption, we can sample the system with a period of at most N to obtain an invertible control matrix.

The approach now is to quantize the components in the system according to the adaptive quantization rule provided earlier, except for a joint mapping for the overflow region. We modify the scheme in (8.12) as follows: Let for $i = 1, 2, \ldots, n$, $R'_i(n) = \log_2(2^{R_i(n)} - 1) = \log_2(K_i(n))$. The vector quantizer quantizes uniformly the marginal variables and we define the overflow region as the quantizer outside the granular region: $\prod_{i=1}^{N}[-2^{R_i'(n)-1}\Delta^i, 2^{R_i'(n)-1}\Delta^i]$ and for $i = 1, 2, \ldots, N$

$$Q_{K_i}^{\Delta_t^i}(x) = \mathcal{Z} \quad \text{if} \quad x \notin \prod_{k=1}^{N}[-2^{R_k'(n)-1}\Delta^k, 2^{R_k'(n)-1}\Delta^k],$$

and for $x \in \prod_{i=1}^{N}[-2^{R_i'(n)-1}\Delta^i, 2^{R_i'(n)-1}\Delta^i]$, the quantizer quantizes the marginal according to (8.12). Hence, here Δ^i is the bin size of the quantizer in the direction of the eigenvector x^i, with rate $R_i'(n)$. For $1 \le i \le N$,

$$\begin{aligned} u_t &= -1_{\{t=(k+1)n-1\}}\tilde{B}^{-1}\Lambda^n \hat{x}_{kn}, \\ \hat{x}_t^i &= Q_{K_i}^{\Delta_t^i}(x_t^i), \\ \Delta_{t+1}^i &= \Delta_t^i \bar{Q}^i(\Delta_t^i, c'_{(t+1)n-1}), \end{aligned} \tag{8.96}$$

with $\delta^i > 0$, $\alpha^i < 1$, and $L^i > 0$ such that

$$\begin{aligned} \bar{Q}^i(\Delta^i, c') &= (|\lambda^i| + \delta)^n && \text{if} \quad c' = \mathcal{Z}, \\ \bar{Q}^i(\Delta^i, c') &= (\alpha^i)^n && \text{if} \quad c' \neq \mathcal{Z}, \Delta \ge L^i, \\ \bar{Q}^i(\Delta^i, c') &= 1 && \text{if} \quad c' \neq \mathcal{Z}, \Delta < L^i, \end{aligned}$$

and $R_i'(n) > n\log_2(|\lambda^i|/\alpha^i)$.

Instead of (8.59), the sequence of stopping times is defined as follows. With $\tau_0 = 0$, define

$$\tau_{z+1} = \inf\{kn > \tau_z : |h_{kn}^i| \le 1, i = 1, 2\ldots, N\}, \quad k, z \in \mathbb{Z}_+,$$

where $h_t^i = \frac{x_t^i}{\Delta_t^i 2^{R_i'-1}}$. Now, we observe that for the N-dimensional system:

$$\begin{aligned} & P(\tau_1 > kn | x_0, \Delta_0) \\ &= P\Big(\bigcap_{t=1}^{k}\{\bigcup_{i=1}^{N}(|h_t^i| > 1)\} \mid x_0, \Delta_0\Big) \\ &\le P\Big(\bigcup_{i=1}^{N}(|h_{kn}^i| > 1)|\text{zoom until} \quad k) \mid x_0, \Delta_0\Big) && (8.97) \\ &\le \sum_{i=1}^{N} P(|h_{kn}^i| > 1|\text{zoom until} \quad k, x_0, \Delta_0), && (8.98) \end{aligned}$$

where we have applied the chain rule of probability in (8.97) and the union bound in (8.98). However, for each of the dimensions, $P(|h_{kn}^i| > 1|\text{zoom until} \quad kn, x_0, \Delta_0)$

is dominated by an exponential measure. Furthermore, $P(\tau_1 > n|x_0, \Delta_0)$ still converges to 0 provided the rate condition $R'_i(n) > \log_2(|\lambda^i|/\alpha^i)$ is satisfied for every i, since $P(\tau_1 > n|x_0, \Delta_0) \le \sum_{i=1}^{N} P(|h_n^i| > 1|x_0, \Delta_0)$. Therefore, analogous results to (8.83) and (8.84) are applicable. Once one imposes a countability condition for the bin size spaces as in Theorem 8.4.2, the desired ergodicity properties are established. □

8.7.13 Proof of Theorem 8.6.1

Let P_t be the probability measure at time t, generated as a result of the Markovian recursions $P_t(D) = \int_X P(x, D)P_{t-1}(dx), \quad \forall D \in \mathcal{B}(\mathbb{R}), t \ge 1$. Suppose that we have an invariant probability measure P for the Markovian process, with a finite second moment. Then, the entropy of the invariant density is also finite (which is bounded by the entropy of the Gaussian density with the same second moment). Furthermore, by the irreducibility of the Markovian process, there is a unique invariant probability measure.

Since the system is driven by a Gaussian noise process, it follows that P_t admits a density, which we will refer to as p_t and which can be expressed as a convolution of the Gaussian measure with another probability measure. That is,

$$p_t(x) = \int_{\mathbb{R}} \nu(x - z)\tilde{p}_{t-1}(z)dz,$$

where $\tilde{p}_{t-1}$ is a piecewise continuous density function and ν is the Gaussian density function. The density function $\tilde{p}_{t-1}$ is piecewise continuous due to the effect of quantization and channel errors. It is thus a simple exercise to show that $p_t(x)$ is uniformly continuous and is such that for any open set $B \subset \mathbb{R}$, $P_t(x \in B) > 0$. This implies that P_t is absolutely continuous with respect to P (which in turn is absolutely continuous with respect to the Lebesgue measure). Let p be the density corresponding to P.

This consequently implies that the (Kullback-Leibler) divergence (see Definition 5.3.4) $D(p_t||p) < \infty$, for some $t > 1$.

This, together with the uniqueness of the invariant measure, lead to the conclusion that, by Theorem 4 of Harremoës and Holst [191] (see also [48]),

$$\lim_{t\to\infty} D(p_t||p) = 0.$$

Furthermore, by the f-norm ergodic theorem ([271], Chap. 14),

$$\int p_t(x)\log_2(p(x))dx \to \int p(x)\log_2(p(x))dx.$$

Therefore, $\int p_t(x)\log_2(p_t(x))dx$ converges to $\int p(x)\log_2(p(x))dx$. Hence, $\lim_{t\to\infty} h(x_t)$ exists and is finite.

Since $x_{t+1} = ax_t + bu'_t + w_t$, conditioning does not increase entropy, and $\{w_t\}$ is i.i.d.,

$$\begin{aligned} h(x_{t+1}) &\geq h(x_{t+1}|u'_t) = h(a(x_t - (b/a)u'_t) + w_t|u'_t) \\ &= h(ax_t + w_t|u'_t) \geq h(ax_t + w_t|u'_t, w_t) \\ &= h(ax_t|u'_t) = \log_2(|a|) + h(x_t|u'_t), \end{aligned} \tag{8.99}$$

which implies $h(x_{t+1}) - h(x_t|u'_t) \geq \log_2(|a|)$. Since $I(x_t; u'_t) = h(x_t) - h(x_t|u'_t)$, we have $I(x_t; u'_t) \geq h(x_t) + \log_2(|a|) - h(x_{t+1})$, and since $\lim_{t\to\infty} h(x_{t+1}) - h(x_t) = 0$, we have

$$\liminf_{t\to\infty} I(x_t; u'_t) \geq \log_2(|a|).$$

We now study the asymptotic behavior of the mutual information $I(x_t; u'_t) = h(x_t) + h(u'_t) - h(x_t, u'_t)$. There exists a limit for $h(x_t)$ from the discussion above. There exists an invariant limit density for u'_t with a finite second moment, since the control is a result of a stationary stochastic kernel. Thus, $h(x_t, u'_t)$ is the joint entropy process of a stable stationary pair and thus has a limit by the arguments above. Hence, $I(x_t; u'_t)$ has a limit, and $\lim_{t\to\infty} I(x_t; u'_t) \geq \log_2(|a|)$. Therefore, for some arbitrary stationary distribution, the mutual information is lower bounded by $\log_2(|a|)$. From the data processing inequality (see, e.g., [171], Chapter 9) and the fact that the policies included are memoryless, we have $I(x_t; u'_t) \leq I(z_t; y_t)$ and $I(x_t; u'_t) \leq I(z'_t; y'_t)$.

Since capacities of memoryless channels are achieved by the maximizing source distributions, $C_f = \sup_{p(z_t)} I(z_t; y_t)$, $C_r = \sup_{p(z'_t)} I(z'_t; y'_t)$, and the capacities of the individual channels upper bound the joint capacity, we have $\min(C_f, C_r) \geq \log_2(|a|)$. □

8.7.14 Proof of Theorem 8.6.2

We regard the controller in Fig. 8.4 as an intermediate encoder, and y'_t as the channel output from the effective channel consisting of the encoder, controller, and the channels. First note that the directed information satisfies

$$I(x_{[0,t-1]} \to y'_{[0,t-1]}) = \sum_{k=1}^{T-1} I(x_{[0,k]}; y'_k|y'_{[0,k-1]}) + I(x_0; y'_0),$$

and since u'_k is determined by $y'_{[0,k]}$, and the control actions affect the evolution of x_t in an additive fashion, we have that $I(x_{[0,k]}; y'_k|y'_{[0,k-1]}) = I(\bar{x}_{[0,k]}; y'_k|y'_{[0,k-1]})$, where $\bar{x}_t = a\bar{x}_t + w_t$ is the control-free process driven by the same noise realizations as $\{x_t\}$ and initial condition x_0. Thus,

$$I(x_{[0,t-1]} \to y'_{[0,t-1]}) = I(\bar{x}_{[0,t-1]} \to y'_{[0,t-1]}).$$

Then,

$$\frac{1}{t} I(\bar{x}_{[0,t-1]} \to y'_{[0,t-1]}) \leq \frac{1}{t} I(\bar{x}_{[0,t-1]}; y'_{[0,t-1]}),$$

and by the *data processing inequality* (see Lemma 5.3.1), it follows that

$$\frac{1}{t} I(\bar{x}_{[0,t-1]}; y'_{[0,t-1]}) \leq \min(C_f, C_r). \tag{8.100}$$

This relation holds for every $t \in \mathbb{N}$. The proof then follows from that of Theorem 8.5.2, by applying (8.21)–(8.27). □

8.7.15 *Proof of Theorem 8.6.3*

Since the chain is irreducible, all sets with positive Lebesgue measure are visited in finite time with probability 1. Due to this observation, the estimation error of the controller regarding the state, $e_t := x_t - E[x_t | I^c_{t-1}]$ (where I^c_{t-1} is the information at the controller at time $t-1$), will be visiting a set $T_k := [R, 2^k R]$, with $R > 0$, in finite time with probability one from any given initial condition. Now suppose that the initial state is in a set T_k.

Further, without any loss of generality, let the bounded encodable control set be given by $S_e = \{u : |u| < M/b'\}$. We study the exit time of the process x_t from T_k. We have $|b'u'_t| < M, \forall t \geq 0$. Define a process $dv_t = \gamma v_t dt + dB_t$, with $v_0 = x_0 \in T_k$, where $x_0 > R > M/(\mu - \gamma)$ and $\mu > \gamma > 0$ and B_t having the same sample path as the disturbance in the original system. Let $\tau^N := \min(\inf\{t : x_t \leq R\}, N)$ and $\tau' := \min(\inf\{t : v_t \leq R\}, N)$. Further, define $\tau'^N_k := \min(\inf\{t : v_t \notin T_k\}, N)$ and $e_t := x_t - v_t$. Since $(\mu x_t + b'u'_t) > \gamma v_t$, for $0 \leq t \leq \min({\tau^N}_k, \tau'^N_k)$, e_t is almost surely positive, for $0 \leq t \leq \min(\tau^N, \tau'^N)$. Since $x_t > v_t$ almost surely, the exit time satisfies ${\tau'}^N_k \leq \tau^N_k$ almost surely. Now, let $f(x) = e^{-2\gamma x}$. Note that $f(x)$ is continuously differentiable and bounded over the set of interest. Hence, we can apply Dynkin's formula [295], from which it follows that

$$E_{x_0}[f(v_{{\tau'}^N_k})] = f(x_0) + E_{x_0}\Big[\int_0^{{\tau'}^N_k} Af(v_s) ds\Big],$$

where A is the generator function [295], given by $Af(x) = \gamma f_x + (1/2) f_{xx}$, where f_x denotes the partial derivative of f with respect to x, and f_{xx} is its second partial derivative. Since $A(e^{-2\gamma x}) = 0$ and $E[{\tau'}^N_k] < \infty$, we have the expectation $E_{x_0}[f(v_{{\tau'}_k^N})] = f(x_0)$. With $N \to \infty$, ${\tau'}_k^N \to \tau'_k$, it follows that $E_{x_0}[f(v_{{\tau'}_k})] = f(x_0)$. Since the process is driven by a Brownian process, the process exits a compact set in finite time almost surely. Hence, $p_R e^{-2\gamma R} + (1 - p_R)e^{-\gamma 2^{k+1} R} = e^{-2\gamma x_0}$, where p_R denotes the exit probability at R. Therefore, we have that p_R is bounded in k, and $\gamma > 0$. Hence, it follows that $\lim_{k \to \infty} p_R(k) = e^{-2\gamma x_0}/e^{-2\gamma R} < 1$. Thus, the probability of the exit time for the newly defined process satisfies $P(\tau' < \infty) < 1$, and as a result, $P(\tau < \infty) < 1$. Thus, the Markov chain is transient. Due to the open-loop instability of the dynamics, the return time of the process to a compact set around the origin has probability less than one. The proof for the encodable state set follows from similar arguments. □

8.7.16 Proof of Theorem 8.6.4

The expectation $\sum_l \sum_j p(j|i)p'(l|j)[ax + bQ'_l]$ is linear in x; therefore, the maximum value of $|\sum_l \sum_j p(j|i)p'(l|j)[ax + bQ'_l]|$ for $x \in [\delta_i, \delta_{i+1}]$ is achieved at one of the end points of each quantization bin (as $|f(x)|$ is a convex function, if $f(x)$ is linear). Thus, by ensuring the drift condition for the bin edges, a uniform decrease in the Lyapunov value for all x will be attained.

Let $V(x) = |x|$. The result of the theorem follows from the construction of a supermartingale sequence as follows: Let $M_0 = V(x_0)$ and for $n \geq 1$, $M_n = V(x_n) + \sum_{k=0}^{n-1}(\epsilon - L 1_{\{x_k \in \cup_{i \in C} \mathcal{B}_i\}})$. The sequence $\{M_n\}$ forms a supermartingale: For any finite n, $E[M_{n+1}|\sigma(x_1, x_2, \ldots, x_n)] \leq M_n$. Let $\tau = \min(k > 0 : x_k \in \cup_{i \in C} \mathcal{B}_i)$. Let for $n \in \mathbb{Z}$, $\tau^n = \min(\tau, \min(t > 0 : t + V(x_t) \geq n))$ be a stopping time. Then, $E[M_{\tau^n}|x_0 = x] \leq M_0$, and hence $(\sup_{x \in \cup_{i \in C} \mathcal{B}_i} |x|)E[\tau^n|x_0 = x] \leq \frac{M_0 + L}{\epsilon}$, and by the monotone convergence theorem (see Sect. B.1), it follows that $\sup_{x \in \cup_{i \in C} \mathcal{B}_i} E[\tau|x_0 = x] < \infty$. □

8.7.17 Proof of Theorem 8.6.5

Suppose that the continuous-time system is sampled with a period T_s, to lead to $x_{t+1} = ax_t + bu'_t$, where $a = e^{\mu T_s}$ and b is the sampled data control coefficient, as introduced earlier.

In the following, $K(N_f, R_f)$ is the number of codewords, where we suppress the dependence on N_f and R_f.

Let $Q'_i = E[x|x \in [\delta_i, \delta_{i+1})]$, which is the center of moment of the corresponding bin (centroid). Let $\pi_0(x)$ be the probability measure on the initial state. Then, the distortion is given by

$$D = \sum_{l=1}^{K} \sum_{j=1}^{K} \sum_{i=1}^{K} p(j|i)p'(l|j) \int_{\delta_i}^{\delta_{i+1}} a^2(x - (b/a)Q'_l)^2 \pi_0(dx),$$

which can be upper bounded as a function of the codelengths in the forward and reverse channels:

$$
\begin{aligned}
D \ &\leq e^{(2\mu\alpha - E_L^f(N_f,R_f))N_f} e^{(2\mu\beta - E_L^r(N_r,R_r))N_r} \\
&\quad \sum_{l=1,l\neq j}^{K} \sum_{j=1,j\neq i}^{K} \sum_{i=1}^{K} \int_{\delta_i}^{\delta_{i+1}} (x-(b/a)Q'_l)^2 \pi_0(dx) \\
&\quad + e^{(2\mu\alpha)N_f} e^{(2\mu\beta - E_L^r(N_r,R_r))N_r} \sum_{i=1}^{K} \sum_{l=1,l\neq i}^{K} \\
&\quad \int_{\delta_i}^{\delta_{i+1}} (x-(b/a)Q'_l)^2 \pi_0(dx) \\
&\quad + e^{(2\mu\alpha - E_L^f(N_f,R_f))N_f} e^{(2\mu\beta N_r)} \sum_{j=1}^{K} \sum_{i=1,i\neq j}^{K} \\
&\quad \int_{\delta_i}^{\delta_{i+1}} (x-(b/a)Q'_j)^2 \pi_0(dx) \\
&\quad + (e^{2\mu})^{\alpha N_f + \beta N_r} \sum_{i=1}^{K} \int_{\delta_i}^{\delta_{i+1}} (x-(b/a)Q'_i)^2 \pi_0(dx). \qquad (8.101)
\end{aligned}
$$

The conditions $(R_f + 2\mu\alpha - E_L^f(N_f,R_f)) < 0$ and $(R_r + 2\mu\beta - E_L^r(N_r,R_r)) < 0$ guarantee the convergence of the first term above. Note that the last term in (8.101) is just the quantization error, and using asymptotic quantization theory [427], the distortion is inversely proportional with the square of the number of symbols under a Lloyd-Max quantizer, and this needs to compensate the growth in T_s, which is satisfied by the third condition in (8.19). We note here that the third condition in (8.19) can also be expressed as $\lim_{N_f\to\infty} \frac{K}{e^{\mu(\alpha N_f + \beta N_r)}} > 1$.

What remains to be done is the analysis of the cross terms (second and the third terms in the summation). For the term

$$
e^{(2\mu\alpha)N_f + (2\mu\beta - E_L^r(N_r,R_r))N_r} \sum_{i=1}^{K} \sum_{l=1,l\neq i}^{K} \int_{\delta_i}^{\delta_{i+1}} (x-(b/a)Q'_l)^2 \pi_0(dx),
$$

using the centroid property of $Q'_i, \forall i$, we obtain

$$
\begin{aligned}
&e^{(2\mu\alpha)N_f + (2\mu\beta - E_L^r(N_r,R_r))N_r} \sum_{i=1}^{K} \sum_{l=1,l\neq j}^{K} \\
&\int_{\delta_i}^{\delta_{i+1}} \Big((x-(b/a)Q'_i)^2 + ((b/a)Q_i - (b/a)Q'_l)^2 \Big) \pi_0(dx),
\end{aligned}
$$

for asymptotic boundedness of which it suffices to have $(2\mu\alpha)N_f + (R_r + 2\mu\beta - E^r_L(N_r, R_r))N_r < 0$. Likewise, for the other cross term, we need $(R_f + 2\mu\alpha - E^f_L(N_f, R_f))N_f + (2\mu\beta N_r) < 0$. □

8.7.18 Proof of Theorem 8.6.6

The forward channels in the Codebins are encoded here in k time stages. Due to the increase in the effective sampling period, the sampled Brownian motion noise has a larger variance. Let C_k be the set of Codebins, which are represented by codes that are transmitted over k transmissions. For the finite-mean return property for a compact set, as well as for the existence of an invariant distribution for the sampled chain, by Theorem 6.2.3, it suffices that for some $\epsilon > 0$ and $k_0 > 0$:

$$E[x^2_{t+k(x_t)T_s}|x_t] - x^2_t < -k(x_t), \quad \forall x_t \in \text{Codebin}(I), I \in C_k, k \geq k_0. \tag{8.102}$$

We have that there will be $2^{kN_fR_f}$ codewords which are transmitted in k sampling periods. Here, γ_k is the logarithmic quantizer ratio for this set of codewords. Let us consider the following condition for some $\epsilon > 0$ and $k_0 > 0$:

$$E[x^2_{t+kT_s}|x_t] < (1-\epsilon)x^2_t, \quad \forall x_t \in \text{Codebin}(I), I \in C_k, k \geq k_0. \tag{8.103}$$

We now write $E[x^2_{t+kT_s}] = e^{2kT_s\mu}E[(x_t + (b/a)u'_t)^2] + D_k(B)$, where $D_k(B) := (e^{2\mu kT_s} - 1)/2\mu$ is the variance of the standard Brownian process integrated over a period of kT_s. Thus we need

$$e^{2\mu kT_s}E[(x_t + (b/a)u'_t)^2] < [(1-\epsilon)x^2_t - D_k(B)], \qquad x_t \in S^C.$$

We can bound the distortion term following the results of Theorem 8.6.5. The distortion is equal to

$$E[(x_t + (b/a)u'_t)^2] = \sum_{l=1}^{K}\sum_{j=1}^{K} p(j|i)p'(l|j)e^{2kT_s\mu}(x_t - (b/a)Q'_l)^2.$$

The probability of error consists of the summation of the probability of errors in both channels or in one of the channels. We can bound the probability of error by the following:

$$\Upsilon(k) = \Big(e^{-kN_fE^f_L(N_f,R_f) - kN_rE^r_L(N_r,R_r)}2^{kN_fR_f} + e^{-kN_fE^f_L(N_f,R_f)} + e^{-kN_rE^r_L(N_r,R_r)}\Big)2^{kN_fR_f}.$$

For any i and l, the worst-case distortion is upper bounded by $\gamma_k^{2(kN_fR_f)}x_t^2$, and the quantization (source-coding) error is upper bounded by $(1/4)x_t^2(e^{2k\mu})^{T_s}(\gamma_k-1)^2$. Hence, if we have

$$e^{2\mu kT_s}\Big(\Upsilon(k)x_t^2+(\gamma_k-1)^2x_t^2/4\Big)<(1-\epsilon)x_t^2-D_k(B), \tag{8.104}$$

for $x_t\in C_k, k\geq k_0$, (8.103) will be satisfied. It now follows that (8.103) implies (8.102), since $\lim_{|x|\to\infty}k(x)/x=0$. This follows since γ_k can be as large as

$$(1+2\sqrt{e^{-2\mu kT_s}-\Upsilon(k)U_k(\gamma)})=(1+2e^{-\mu kT_s}\sqrt{1-\frac{\Upsilon(k)U_k(\gamma)}{e^{-2\mu kT_s}}}),$$

$\lim_{k\to\infty}\frac{\Upsilon(k)U_k(\gamma)}{e^{-2\mu kT_s}}<1$, and that, as a result,

$$\lim_{k\to\infty}\frac{(1+2\sqrt{e^{-2\mu kT_s}-\Upsilon(k)U_k(\gamma)})^{2^{kN_fR_f}}}{k}=\infty,$$

when $N_fR_f>\log_2(|a|)$. This implies that $\lim_{|x|\to\infty}\frac{k(x)}{x}=0$ (and also that the entire state space is encoded). □

8.7.19 *Proof of Theorem 8.6.7*

We have

$$\begin{aligned}
z_t &= \sqrt{P_f}x_t/||x_t||, \quad y_t=z_t+w_t, \quad x'_t=\alpha_ty_t, u_t=\rho_tx'_t,\\
z'_t &= \sqrt{P_r}u_t/||u_t||, \quad y'_t=z'_t+w'_t, \quad u'_t=\beta_ty'_t.
\end{aligned}$$

Let $||x_t||^2=E[x_t^2]$ and $||u_t||^2=E[u_t^2]$. The control applied, u'_t, can be written as follows:

$$u'_t=\beta_t\sqrt{P_r}\Bigg(\frac{\sqrt{P_f}x_t}{||x_t||\sqrt{P_f+{\sigma^f}^2_w)}}+\frac{w_t}{\sqrt{P_f+{\sigma^f}^2_w}}\Bigg)+\beta_tw'_t.$$

Using this in the system equation $x_{t+1}=ax_t+bu'_t+w_t$, we obtain

$$\begin{aligned}
E[x_{t+1}^2]=E\Bigg[\Bigg((a+\frac{b\beta_t\sqrt{P_rP_f}}{||x_t||\sqrt{P_f+{\sigma^f}^2_w}})x_t\Bigg)^2\Bigg]&\\
+b^2\beta_t^2{\sigma^f}^2_wP_r/(P_f+{\sigma^f}^2_w)+b^2\beta_t^2{\sigma^r}^2_w+E[w_t^2].&
\end{aligned} \tag{8.105}$$

Define $\rho_t := b\beta_t$. The minimization of $E[x_{t+1}^2]$ with respect to ρ_t yields

$$\rho_t = -\frac{a||x_t||\sqrt{P_f P_r}}{(P_r + {\sigma^r}_w^2)\sqrt{P_f + {\sigma^f}_w^2}}. \tag{8.106}$$

From this, the optimal expression for β_t follows. If we plug this optimal β_t value in the expression for $\|x_{t+1}\|^2 = E[x_{t+1}^2]$, we obtain

$$\begin{aligned}||x_{t+1}||^2 &= ||x_t||^2 a^2 \left(1 - \frac{P_r P_f}{(P_f + {\sigma^f}_w^2)(P_r + {\sigma^r}_w^2)}\right)^2 \\ &+ ||x_t||^2 a^2 \left(\frac{{\sigma^f}_w^2 P_r^2 P_f}{(P_f + {\sigma^f}_w^2)^2 (P_r + {\sigma^r}_w^2)^2}\right) \\ &+ ||x_t||^2 a^2 \frac{P_r P_f}{(P_f + {\sigma^f}_w^2)(P_r + {\sigma^r}_w^2)^2} {\sigma^r}_w^2 + E[w_t^2].\end{aligned}$$

Upon recognizing the capacity expression in the following:

$$P_f/(P_f + {\sigma^f}_w^2) = 1 - 2^{-2C_f}, \qquad P_r/(P_r + {\sigma^r}_w^2) = 1 - 2^{-2C_r},$$

we arrive at

$$\begin{aligned}E[x_{t+1}^2] = a^2 \Bigg\{ &(1 - 2^{-2C_f})(1 - 2^{-2C_r})(2^{-2C_f} - 2^{-2C_f - 2C_r} + 2^{-2C_r}) \\ &+ \left(1 - (1 - 2^{-2C_f})(1 - 2^{-2C_r})\right)^2 \Bigg\} E[x_t^2] + E[w_t^2].\end{aligned}$$

Rearranging the first term, it follows that the condition

$$1 - (1 - 2^{-2C_f})(1 - 2^{-2C_r}) < 1/a^2$$

implies stability. □

8.8 Concluding Remarks

This chapter considered stochastic stabilization over noisy channels of linear systems driven by unbounded noise and established tight conditions for asymptotic mean stationarity and ergodicity, as well as sufficient conditions for stability in the sense of having a finite average second moment for the state process. One message was that Shannon capacity and the stochastic stabilizability (in the sense of the AMS property or ergodicity) are inherently related. This relationship leads to stronger results in the context of discrete noiseless channels (which were studied in Chap. 7) and erasure channels which were covered in this chapter: For these channels, we established not only asymptotic mean stationarity and n-ergodicity but also stationarity and ergodicity. However, the full generality of this relationship became apparent later in this chapter when we investigated more general noisy channels. We have also presented conditions for quadratic and finite second moment stability.

We note that the assumption that the system noise is Gaussian can be relaxed. For the second moment stability, a sufficiently light tail which would provide a geometric bound on the stopping times as in (8.77) through (8.70) will be sufficient. For the AMS property, this is not needed. For a noiseless DMC, Chap. 7 established that a finite second moment for the system noise is sufficient for the existence of an invariant probability measure. To establish irreducibility, we require, however, that the noise admits a density which is positive everywhere.

We observed in the development that three types of errors are critical. These bring up the importance of unequal error coding schemes with feedback. Recent results in the literature [68] have focused on fixed length without feedback and variable length with feedback, and further research in this arena could be useful for networked control problems.

In the absence of noiseless feedback, for both DMCs and CMCs, this chapter has presented necessary conditions on the channel capacities for the existence of a controller that will lead to $\limsup_{T\to\infty} E[x_T^2] < \infty$. We have obtained capacity bounds and achievable forward and reverse rate regions leading to a positive recurrent Markov chain. We have shown that if the underlying closed-loop system is described by an irreducible Markov chain (such as in the case when the system noise has unbounded support for its probability distribution with an everywhere positive density function), then the entire state space and the control space have to be encoded. We observed that control over discrete channels requires an intricate design. We showed the inadequacy of fixed length encoding schemes and introduced the notion of escape-freeness, which required the design to use variable length encoding. The design stabilizes the system via variable rate sampling and uses properties of the sampled Markov chain. We saw that continuous alphabet channels are simpler to analyze since they can always be designed to be escape-free, without resorting to variable length sampling. However, since continuous channels are not as widely used as discrete channels, the analysis provided for DMCs is particularly important in applications.

The value of information channels in optimization and control problems (beyond stabilization) is an important problem in view of applications in networked control systems. Further research from the information theory community for non-asymptotic or finite delay coding results will provide useful applications and insight for such problems. In this context, suppose we have a channel where agreement on a binary event in finite time is possible between the encoder and the decoder. Binary events may include synchronization of encoding times and agreement on zooming times. If the following assumption holds, then such agreements are possible in finite expected time: The channel is such that there exist input letters x_1, x_2, x_3, x_4 so that $D(P(\cdot|x_1)||P(\cdot|x_2)) = \infty$ and $D(P(\cdot|x_3)||P(\cdot|x_4)) = \infty$. Here, x_1 can equal to x_4 and x_2 can equal to x_3, which is a property satisfied by, for example, the erasure channel. Note that the above condition is weaker than having a nonzero zero-error capacity, but stronger than what Burnashev's [83, 284, 411] method requires, since there are more hypotheses to be tested. In such a setting, one could use variable length encoding schemes. Such a design will allow the encoder and the decoder to have transmission in three phases: zooming, transmission, and error confirmation. Using random-time, state-dependent stochastic drift, we may find alternative schemes for stochastic stabilization.

8.9 Bibliographic Notes

In Sect. 5.5, a brief overview of the literature was presented. In the following, we provide some further discussion. Nair and Evans [280], Tatikonda and Mitter [355], and Wong and Brockett [406] obtained the minimum lower bound needed for stabilization over noisy channels under a set of assumptions on the system noise and channels, also known as the *data-rate theorem*. This theorem states that for stabilizability under information constraints, in the mean-square sense, a minimum rate needed for stabilizability has to be at least the sum of the logarithms of the unstable poles/eigenvalues in the system, that is,

$$\sum_{k=1}^{m} \frac{1}{2} \max\left(0, \log(|\lambda_k|^2)\right). \tag{8.107}$$

Martins and Dahleh [256] established that when a channelis present in a controlled linear system, under stationarity assumptions, the rate requirement in (8.107) is necessary for having finite second moments for the state variable. A related argument was made in [432] under the assumption of invariance conditions for the controlled state process under memoryless policies and finite second moments.

The problem of control (as well as estimation) over noisy channels with or without feedback has been considered in a large number of publications: [41, 98, 104, 258, 261, 266, 273, 355] among others. Most of the constructive results involve Gaussian channels or erasure channels (some modeled as infinite capacity erasure channels as in [205, 335]). We now discuss some of these in the following.

For coding and information transmission for unstable linear systems, there is an important difference between continuous alphabet and finite-alphabet (discrete) channels as discussed in [432]: When the space is continuous alphabet, we do not necessarily need to consider adaptation in the encoders. On the other hand, when the channel is finite alphabet, and the system is driven by unbounded noise, a bounded range quantizer (a quantizer with bounded granular region) leads to almost sure instability, a topic which will be discussed further in Chap. 10 and which was already discussed in Chap. 7 in the context of discrete-time systems. This was perhaps first recognized in view of the unboundedness of second moments in Proposition 5.1 in [280], and the transience of such a controlled state process was established in Theorem 4.2 in [432]. Hokayem et al. [110] and Ramponi et al. [319] studied conditions for stabilization when the control actions are uniformly bounded and the controlled multidimensional system is marginally stable and is driven by noise with unbounded support.

Nair and Evans [280] considered a class of quantizer policies for such unstable linear systems driven by noise, with unbounded support set for its probability measure, and controlled over noiseless channels; they obtained necessary and sufficient conditions for the boundedness of the following expression:

$$\limsup_{t\to\infty} E[|x_t|^2] < \infty.$$

A stronger result was obtained in [419], by establishing the existence of a limit

$$\lim_{t\to\infty} E[|x_t|^2] < \infty,$$

and obtaining a scheme which made the state process and the encoder process stochastically stable in the sense that the joint process is a positive Harris recurrent Markov chain and the sample path ergodic theorem is applicable.

It should be stressed that the notion of stochastic stability is very important in characterizing the conditions on the channel. Matveev and Savkin, in [261, 266], considered stabilization in the following almost sure sense, when the system noise is bounded:

$$\limsup_{t\to\infty} |x_t| < \infty \quad a.s.,$$

and observed that one needs the zero-error capacity (with feedback) to be greater than a particular lower bound. A similar observation was made in [331]. When the system is driven by noise which admits a probability measure with unbounded support, the stability requirement above is impossible for an infinite horizon problem, even when the system is open-loop stable, since for any bound, there exists almost surely a realization of a noise variable which will be larger.

Sahai [328] and Sahai and Mitter [331] considered systems driven by bounded noise and considered a number of stability criteria: almost sure stability, moment stability ($\limsup_{t\to\infty} E[|x_t|^p] < \infty$) as well as *stability in probability* in the following sense: For every $p > 0$, there exists a ζ such that $P(|x_t| > \zeta) < p$ for all $t \in \mathbb{N}$. The authors, in these papers, also introduced a characterization for reliability for controlling unstable processes, namely, *any-time capacity*, as the characterization of channels for which the following criterion can be satisfied:

$$\limsup_{t\to\infty} E[|x_t|^p] < \infty,$$

for a positive integer p. A channel is α any-time reliable for a sequential coding scheme if $P(\hat{m}^{t-d}(t) \neq m^{t-d}(t)) \leq K2^{-\alpha d}$ for all t, d. Here m^{t-d} is the message transmitted at time $t - d$, estimated at time t. One interesting aspect of an any-time decoder is the independence from the delay, with a fixed encoder policy. Sahai and Mitter [331] states that for a system driven by bounded noise, stabilization is possible if the maximum rate for which an any-time reliability of $2\log_2(|\lambda|)$ is satisfied and is greater than $\log_2(|\lambda|)$, where λ is the unstable pole of the scalar linear system.

In a related context, [258, 260, 266, 331] considered the relevance to Shannon capacity. Martins et al. [258] observed that when the moment coefficient goes to zero, Shannon capacity provides the right characterization on whether a channel is sufficient or insufficient, when noise is bounded. A parallel argument is provided in [331], observing that in the limit when $p \to 0$, capacity should be the right measure for the objective of establishing *stability in probability*. Their discussion was for bounded noise signals. Matveev and Savkin [266] presented a parallel discussion, again for bounded noise signals.

With a departure from the bounded noise assumption, Matveev [260] considered a more general model of multidimensional systems driven by an unbounded noise process, considering again *stability in probability*. Matveev [260] also showed that when the discrete noisy channel has capacity less than $\log_2(|a|)$, there exists no stabilizing scheme, and if the capacity is strictly greater than this number, there exists a stabilizing scheme in the sense of stability in probability.

Many network applications and networked control applications require the access of control and sensor information to be observed intermittently. Toward generating a solution for such problems, [416, 439] developed random-time state-dependent drift conditions, leading to the existence of an invariant distribution possibly with moment constraints, extending the earlier deterministic state-dependent results in [272]. Using drift arguments, [432] considered noisy channels (both discrete and continuous alphabet), [419] considered noiseless channels, and [439] considered erasure channels for the following stability criteria: the existence of an invariant distribution, and the existence of an invariant distribution with finite moments. Yüksel [422] considered discrete noisy channels, possibly with memory, with noiseless feedback.

Minero et al. [273] considered erasure channels and obtained necessary and sufficient time-varying rate conditions for control over such channels. Coviello et al. [104] considered second moment stability over a class of Markov channels with feedback and developed necessary and sufficient conditions for systems driven by unbounded noise. Gurt and Nair [182] considered stability of the state and quantizer parameters paralleling the results of [439].

Elia [130], Martins and Dahleh [256], and Martins et al. [257] considered general channels (possibly with memory) and, establishing connections with Jensen's formula and Bode's sensitivity integral, developed achievable rates for stabilization under various networked control settings. Elia [131] considered a general setting where channels are between the controller and the plant, as well as between the sensor and the controller.

For more traditional information theoretic settings where the source is revealed at the beginning of the transmission, and for cases where causality and delay are not important, the separation principle for source and channel-coding results are applicable for ergodic sources and information stable channels. The separation principle for more general setups has been considered in [376], among others. Bansal and Başar [41] has shown that for a scalar discrete-time linear Gaussian system controlled over a Gaussian channel, the encoder and the controllers with noiseless causal feedback which jointly minimize a quadratic objective functional are all linear. Walrand and Varaiya [385] and Witsenhausen [396] studied the optimal causal coding problem over, respectively, a noiseless channel and a noisy channel with noiseless feedback. Unknown sources have been considered in [91]. We also note that, when noise is bounded, binning, based strategies, inspired from Wyner-Ziv and Slepian-Wolf coding, schemes are applicable. This type of consideration has been applied in [178, 331, 432]. Finally time-invariant quantizer design for noiseless or bounded noise systems for control over noiseless channels includes [132, 135, 206, 207] which induce logarithmic quantization structures. Channel-coding algorithms for control systems have been presented in [301, 349]. Silva et al. [341] has also considered stabilization over noiseless channels with tight rate conditions. On channels with memory, we note the work [92] for Gaussian channels with memory.

There has also been progress on coding for noisy channels for the transmission of sources with memory. Due to practical relevance, for communication of sources with memory over channels, particular emphasis has been placed on Gaussian channels with feedback. For such channels, the fact that real-time linear schemes are rate-distortion achieving has been observed in [53, 156], and [41] in a control theoretic context. Aside from such results (which involve matching between rate-distortion achieving test channels and capacity achieving source distributions [156]), capacity is known not to be a good measure of information reliability for channels for real-time (zero-delay or delay-sensitive) control and estimation problems [331, 385].

The error exponent is typically improved with feedback, unlike capacity of DMCs. However, the error exponent under fixed length block coding with noiseless feedback is not currently known. Some partial results have been reported in [118] (in particular, the sphere packing bound is optimal for a class of symmetric channels for rates above a critical number even with feedback); see also [54, 68, 107, 128, 190, 285, 446]. Particularly related to this chapter, [68] has considered the exponent maximization for a special message symbol, at rates close to capacity. In case feedback is not used, Gilbert exponent [297] for low-rate regions may provide better bounds than the random coding exponent. Also in the information theory literature, performance of information transmission schemes for channels with feedback has been a recurring avenue of research, for both variable length and fixed length coding schemes [83, 122, 199, 310, 330]. In such setups, the source comes from a fixed alphabet, except for the sequential setup in [122, 330].

Erasure channels are practically important and there has been a large literature on control over such channels. Typically, such channels are investigated in two forms: (i) a continuous (packetized) model where when a transmission is to take place with no erasure, a real signal is transmitted without any error (i.e., this model ignores the issues regarding quantization/encoding), and (ii) a discrete (quantized) model, which is the model considered in this chapter. The packetized model assumes that the Shannon capacity of the channel is infinite. On the discrete model, Yüksel and Meyn [439] considered stochastic stabilization over erasure channels and established positive Harris recurrent properties. You and Xie [415] studied the problem of control over an erasure channel in the absence of noise. The continuous model has been studied in many contributions. For linear systems, stability and optimization have been considered in [181, 188, 205, 335], among other references (see [196]). Quevedo et al. [313,315] have considered stabilization and optimization over erasure channels for nonlinear settings. As mentioned earlier in Chap. 6, the effects of randomness in delay for transmission of sensor or controller signals (see, e.g., [96, 196]) is an important application area where the results of Chap. 6 can be applied.

The zooming algorithm and its variations have been used in source-coding and networked control literatures; see, e.g., the earlier papers [81, 166, 214] (zooming algorithms) and the more recent ones [260, 261, 266, 280, 419, 421].

In the context of stabilization, logarithmic quantizers have been considered in several publications, including [132, 135, 207].

Part of this chapter is based on [260, 420, 422, 426, 432]. The results on erasure channels are primarily based on [416, 439].

Chapter 9
Stabilization of Decentralized Systems Over Communication Channels

9.1 Introduction

A fundamental result in control theory is that a controllable and observable single station (or centralized) linear time-invariant (LTI) system can be stabilized through an observer (equivalently, dynamic output) feedback, and its poles can be altered by static output feedback. These results do not generalize directly to decentralized systems with multiple sensors or controllers. Decentralization presents further intricacies in large part because of the issues related to signaling in view of information transmission through control actions.

In the previous chapters, we covered stochastic stabilization of single-sensor and single-controller systems over a variety of communication channels. In this chapter, we study the problem of stabilization of multi-sensor and multi-controller systems over communication channels. We will see that decentralization will present more stringent requirements on information transmission for stabilization of multi-controller systems. Multi-sensor systems, however, do not suffer from a rate loss in communication due to decentralization.

The next section presents the problem model and some preliminary material. Existence of decentralized stabilizing control policies under any class of admissible policies given the decentralized information structure is discussed in Sect. 9.3, followed by an analysis of information transmission requirements in Sect. 9.4. Section 9.5 considers the important special case of the multi-sensor setting with a single controller. Section 9.6 considers multi-sensor and multi-controller settings with noise, and Sect. 9.7 introduces the notion of binning in the context of multi-sensor systems.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_9,

9.2 Problem Formulation

Consider the class of multi-station n-dimensional discrete-time LTI systems, described by

$$\begin{aligned} x_{t+1} &= Ax_t + \sum_{j=1}^{L} B^j u_t^j\,, \\ y_t^i &= C^i x_t\,,\ \ 1 \le i \le L,\ t = 0, 1, \ldots, \end{aligned} \tag{9.1}$$

where it is assumed that the joint system is stabilizable and detectable, that is, $(A, [B^1|B^2|\ldots|B^L])$ is stabilizable and $(A, [(C^1)'|(C^2)'|\ldots|(C^L)']')$ is detectable, but the individual pairs (A, B^i) may not be stabilizable or (A, C^i) may not be detectable, for $1 \le i \le L$. Here, $x_t \in \mathbb{R}^n$ is the state of the system, $u_t^i \in \mathbb{R}^{m_i}$ is the control applied by station i, and $y_t^i \in \mathbb{R}^{p_i}$ is the observation available at station i, all at time t. Without any loss of generality, we assume the system matrix A to be in Jordan form. The initial state x_0 is unknown, but is known to be generated according to some probability distribution which is supported on a compact set $\mathcal{X}_0 \subset \mathbb{R}^n$.

The information available to station i at time t is

$$I_t^i = \{y_{[0,t]}^i, u_{[0,t-1]}^i\}, \tag{9.2}$$

where, as before, $u_{[0,t-1]}^i$ denotes $\{u_0^i, u_1^i, \ldots, u_{t-1}^i\}$ and $y_{[0,t]}^i = \{y_0^i, y_1^i, \ldots, y_t^i\}$. Hence, for each time t, each station has access to only its measurement and also has full memory on its past measurements and actions. If γ_t^i denotes the strategy (policy, control law) of station i at time t, we have $u_t^i = \gamma_t^i(I_t^i)$. We assume that the entire system dynamics is common information to all agents (stations).

Definition 9.2.1. The decentralized system described in (9.1) is stabilizable under the decentralized information structure if there exists a set of policies $\underline{\gamma}$ such that, with $\{x_t^{\underline{\gamma}},\ t \ge 0\}$ denoting the state trajectory under $\underline{\gamma}$, $|x_t^{\underline{\gamma}}| \to 0$ as $t \to \infty$ almost surely. $\diamond$

Let us introduce the notation

$$\begin{aligned} \mathcal{K}^i &:= \left[B^i\ AB^i\ \cdots\ A^{n-1}B^i\right], \\ \mathcal{O}^i &:= \left[(C^i)'\ (C^iA)'\ \cdots\ (C^iA^{n-1})'\right]' \end{aligned}$$

and let the controllable and unobservable subspaces at station i be denoted by K^i and N^i, respectively, where K^i is the range space of $\mathcal{K}^i$ and N^i is the null-space of $\mathcal{O}^i$. We will, by a slight abuse of notation, refer to the subspace orthogonal to N^i as the observable subspace at the ith station and will denote it by O^i.

We now introduce some basic ingredients and notation used throughout the chapter. Let $\mathbb{U} \subset \mathbb{R}^n$, $\mathbb{V} \subset \mathbb{R}^n$ be Euclidean subspaces. We adopt the following notation:

$$\mathbb{U} \cup \mathbb{V} = \{x : x = \alpha u + \beta v, u \in \mathbb{U}, v \in \mathbb{V}, \alpha, \beta \in \mathbb{R}\},$$
$$\mathbb{U} \cap \mathbb{V} = \{x : x \in \mathbb{U}, x \in \mathbb{V}\},$$
$$\mathbb{U} - \mathbb{V} = \{u : u \in \mathbb{U}, u'v = 0, \forall v \in \mathbb{V}\}.$$

With the above definitions, for a vector space $S \subset \mathbb{R}^n$, we have $S^C = \mathbb{R} - S$, as the orthogonal complement of S. We denote by $P_{\mathbb{U}}(x) : \mathbb{R}^n \to \mathbb{U}$, the orthogonal projection of a vector x onto the subspace $\mathbb{U} \subset \mathbb{R}^n$. We denote by

$$[v_1, v_2, \ldots, v_m] := \{\sum_{i=1}^{m} \alpha_i v_i, \alpha_i \in \mathbb{R}\},$$

the space spanned by the vectors $\{v_1, v_2, \ldots, v_m\}$.

For two sets Ψ and Ξ, $\Psi \setminus \Xi = \{\eta : \eta \in \Psi, \eta \notin \Xi\}$ is the standard set difference.

By *modes* of a linear system, we refer to the subspaces (eigenspaces) which are invariant in the absence of control; as such, when all the eigenvalues of the system matrix A are distinct, the eigenvectors uniquely identify the modes of the system. In case the geometric multiplicity of an eigenvalue is less than its algebraic multiplicity, generalized eigenvectors (generalized modes) span the eigenspace corresponding to a particular eigenvalue.

We next introduce some relevant graph-theoretic notions: A *directed graph* $\mathcal{G}$ consists of a set of vertices, $\mathcal{V}$, and a set of directed edges, $(a, b) \in \mathcal{E}$, such that $a, b \in \mathcal{V}$. A path in $\mathcal{G}$ of length d consists of a sequence of d directed edges such that each edge is connected. A graph in which there exists a path from any node to any other node is a strongly connected graph. We define the minimum distance between two sets of nodes $S_1, S_2 \subset \mathcal{G}$ as $d(S_1, S_2) = \sum_{i \in S_1} \min\{d(i, j), j \in S_2\}$, where $d(i, j)$ denotes the minimum number of paths between node i and j (if such a finite number exists), with the trivial case being $d(i, i) = 0$ for all nodes.

9.3 Existence of Decentralized Stabilizing Controllers and Time-Varying Linear Feedback Laws

In this section we develop necessary and sufficient conditions for the existence of stabilizing controllers for the system in (9.1), under the information structure (9.2).

One of the important notions in decentralized control is that of *decentralized fixed modes*. For such fixed modes, there are two common classifications. In one classification, Wang and Davison have introduced the notion of fixed modes under linear time-invariant control policies.

Theorem 9.3.1 ([386]). *Let*

$$\mathbb{K}_D = \{K : K = \text{diag}(K^1, K^2, \ldots, K^L), K^i \in \mathbb{R}^{m_i \times p_i}\}.$$

For the system (9.1), the set of decentralized fixed modes under linear time-invariant laws is given by

$$\Lambda = \bigcap_{K \in \mathbb{K}_D} \lambda\left(A + \sum_{i=1}^{L} B^i K^i C^i\right),$$

where $\lambda(.)$ denotes the set of eigenvalues of its argument. The system (9.1) is stable under linear time-invariant output feedback laws of the form

$$u_t^i = K^i y_t^i, \quad i = 1, 2 \ldots, L, \tag{9.3}$$

if and only if Λ contains only stable eigenvalues. ◇

The original result in [386] is for continuous-time systems. A counterpart has been provided by Khargonekar and Özgüler [213] for discrete-time systems.

Toward identifying fixed modes under LTI policies, we first present a definition, which is followed by a result due to Willems [392].

Definition 9.3.1. An LTI system (A,B,C) is complete if the matrix

$$\begin{bmatrix} \lambda I - A & B \\ -C & 0 \end{bmatrix}$$

has rank no smaller than n for all complex-valued $\lambda \in \mathbb{C}$. If this holds for all $|\lambda| \geq 1$, then the system is weakly complete. ◇

Theorem 9.3.2. *There does not exist a decentralized unstable fixed mode under linear time-invariant decentralized controllers (9.3) if and only if the joint system is stabilizable and detectable, and for every partitioning of the system into*

$$\mathcal{E}_1 = \{a_1, a_2, \ldots, a_k\}, \mathcal{E}_2 = \{b_1, b_2, \ldots, b_{L-k}\} = \{1, 2, \ldots, L\} \setminus \mathcal{E}_1,$$

the systems

$$(A, [B^{a_1} B^{a_2} \ldots B^{a_k}], [(C^{b_1})'(C^{b_2})' \ldots (C^{b_{L-k}})']'),$$

are weakly complete. ◇

The other notion of fixed modes is one that is independent of the control policy applied, and this arises due to the uncontrollability of the decentralized system. Thus, a mode is a fixed mode if no decentralized algorithm/policy leads to a change in the dynamics of the mode.

Time-invariant output feedback laws are a restrictive class of control laws. In general, as discussed in Sect. 3.3, it is possible for the controllers to communicate through the plant with the process known as *signaling* which can be used for communication of mode information among the decision makers. Hence, through signaling, the controllable subspace can be expanded and the unobservable subspace can be shrunk. Anderson and Moore [11] showed that decentralized stabilization in a multi-controller setting is possible via time-varying control laws, if the system is jointly controllable, jointly observable, and strongly connected: If every station can communicate with every other station, possibly via other stations, the system is said to be strongly connected. This condition will be made precise, and weakened, later in this section.

In the following, we revisit a stabilizability result. This result is for any class of admissible policies, that is, the controllers are not assumed to be of a particular form (such as linear).

Willems [392] and Khargonekar and Özgüler [213] have noted that if a system is complete, and if $CA^lB = 0$ for all $l \geq 0$, then the controllable subspace of (A, B) needs to be identical to the unobservable subspace of (A, C) for stabilizability. This observation is related to the information theoretic approach we will take for decentralized stabilization.

Theorem 9.3.3 ([213]). *There exists a periodic-time-varying decentralized controller, that is, with time-dependent matrices K^i in (9.3), if and only if the joint system is stabilizable and detectable, and for every partitioning of the system into*

$$\mathcal{E}_1 = \{a_1, a_2, \ldots, a_k\}, \mathcal{E}_2 = \{b_1, b_2, \ldots, b_{L-k}\} = \{1, 2, \ldots, L\} \setminus \mathcal{E}_1,$$

such that if the system

$$(A, [B^{a_1} B^{a_2} \ldots B^{a_k}], [(C^{b_1})'(C^{b_2})' \ldots (C^{b_{L-k}})']')$$

has a zero transfer function, then it is weakly complete. ◇

In particular, through lifting (i.e., by viewing the system as a higher-dimensional linear system in a sampled setting), one can apply linear time-invariant policies. In the following, we state a theorem on the universality of linear time-varying controllers for decentralized stabilization. To facilitate the argument, we introduce the notion of a quotient system [217]. Toward this end, we first discuss the notion of connectivity. For two stations i and j, if $K^i \not\subseteq N^j$, then there exist control signals generated at station i which are observed at station j. We denote this by $i \to j$. This is equivalent to the condition that $C^j A^l B^i \neq 0$, for some $0 \leq l \leq n-1$, as we prove in the following:

Lemma 9.3.1. *$i \to j$ if and only if $C^j(A)^l B^i \neq 0$, for at least one l, $1 \leq l \leq n$.* ◇

Proof. The proof proceeds in two steps. (i) Suppose that $C^j(A)^l B^i \neq 0$ for at least one l. This implies the existence of a control u such that $\mathcal{O}^j \mathcal{K}^i u \neq 0$.

(ii) The observation at station j, as affected by controls from station i, is

$$y_t^j = C^j(A)^t x_0 + \sum_{k=0}^{t-1} C^j(A)^{t-k-1} B^i u_k^i.$$

If all the terms $C^j(A)^l B^i$ are zero, for $1 \leq l \leq n$, then via the Cayley–Hamilton theorem, $C^j A^l B^i = 0$ for all $l \in \mathrm{Z}^+$. Thus, we have $y_t^j = C^j(A)^t x_0$. Hence, the control of station i does not affect the observation of station j. □

Lemma 9.3.2 ([431]). *If $k \to m$, then station m can recover the message sent by station k in at most n time stages.* ◇

The proof of this lemma follows from Lemma 9.3.1 and the following result, which let station m extract the message signal.

Lemma 9.3.3. *Before signaling takes place, station m can compute $C^m A^n x_n$, at time n.* ◇

Proof. Suppose that there is no control action until time n. Then,

$$\begin{aligned} C^m A^n x_n &= C^m(\sum_{i=0}^{n-1} \alpha_i A^i) x_n = C^m(\sum_{i=0}^{n-1} \alpha_i A^i) A^n x_0 \\ &= C^m(\sum_{i=0}^{n-1} \alpha_i A^{i+n}) x_0 = C^m(\sum_{i=0}^{n-1} \alpha_i' A^i) x_0 = \sum_{i=0}^{n-1} \alpha_i' y_i^l, \end{aligned}$$

where $\alpha_i, 1 \leq i \leq n$, and $\alpha_i', 1 \leq i \leq n$, can be obtained by the Cayley–Hamilton theorem. This completes the proof. □

The above ensure that station i can send information to station j through control actions $\{u_t^i\}$. One may construct a directed communication graph given the above relationship, by considering also the possibility that two stations may be connected through other stations. In the following, we will, by an abuse of notation, use $i \to j$ to indicate that there is a path from station i to station j, possibly through other stations.

If every station is connected to every other station in the sense above, the system is said to be *strongly connected*. If the network of stations is not strongly connected, it can be uniquely represented as a disjoint union of strongly connected subsystems.

As a consequence of signaling, each subsystem can be effectively regarded as a single centralized subsystem. Let N be the number of strongly connected subsystems in the system. Let us order the strongly connected subsystems by an order of precedence, that is, subsystems i and j satisfy $i < j$ only if station i cannot signal to station j. This relation provides the direction of information flow among the subsystems. By concatenating the system as a composition of the modes controlled by strongly connected subsystems and regarding each subsystem as a single block, and using Kalman's canonical decomposition (see [93]) the system matrix can then be transformed to a block-upper-triangular form as follows (see [101] or [392]):

$$A = \begin{bmatrix} \tilde{A}^1 & * & * \cdots & * \\ 0 & \tilde{A}^2 & * \cdots & * \\ \vdots & \vdots & \vdots \cdots & \vdots \\ 0 & 0 & 0 \cdots & \tilde{A}^N \end{bmatrix}, \quad B = \begin{bmatrix} \tilde{B}^1 & * & * \cdots & * \\ 0 & \tilde{B}^2 & * \cdots & * \\ \vdots & \vdots & \vdots \cdots & \vdots \\ 0 & 0 & 0 \cdots & \tilde{B}^N \end{bmatrix},$$

$$C = \begin{bmatrix} \tilde{C}^1 & * & * \cdots & * \\ 0 & \tilde{C}^2 & * \cdots & * \\ \vdots & \vdots & \vdots \cdots & \vdots \\ 0 & 0 & 0 \cdots & \tilde{C}^N, \end{bmatrix} \tag{9.4}$$

where each of the subsystems $(\tilde{A}^k, \tilde{B}^k, \tilde{C}^k)$ is strongly connected and $*$ denotes some possibly nonzero matrix. Now, the stabilizability of each of these subsystems, in a sequential manner, leads to the result that such a system is stabilizable if and only if all of the individual strongly connected subsystems are stabilizable (by effectively centralized policies in each strongly connected component).

Theorem 9.3.4 (Gong and Aldeen [165]). *The decentralized system described in (9.1) is stabilizable under the decentralized information structure (9.2) if and only if each of the strongly connected subsystems* $(\tilde{A}^k, \tilde{B}^k, \tilde{C}^k)$ *is stabilizable detectable.* ◇

Furthermore, linear time-varying policies do not bring in any loss, as is captured in the following result.

Theorem 9.3.5 ([165]). *The decentralized system described in (9.1) is stabilizable under the decentralized information structure (9.2) if and only if it is stabilizable by periodic linear time-varying controllers.* ◇

Remark 9.3.1. Even though the proof of Gong and Aldeen for Theorem 9.3.4 does not use information theoretic ideas, this result has an information theoretic flavor. In particular, the theorem suggests that the controllers which can stabilize a given mode should either have local information about that mode or be provided information about that mode through signaling by those stations which can signal to them. Such a signaling approach also suggests a constructive linear time-varying control structure for Theorem 9.3.5. We will develop a few key ingredients toward such an interpretation in the next section in the context of information transmission requirements. ◇

9.4 Decentralized Stabilization over Communication Channels

Suppose the controllers are connected to the plant over a discrete noiseless channel. In this case, the control signals u^i are coded and decoded over discrete noiseless channels with finite capacity. Hence, the applied control and transmitted messages follow a coding (i.e., binary representation) and a decoding process. We assume here fixed-rate encoding, that is, the rate is defined as the (base-2) logarithm of the number of symbols to be transmitted: The coding process of the controller at station i is a mapping measurable with respect to the sigma-algebra generated by I_t^i to $\mathcal{M}_t^i = \{1, 2, \ldots, |\mathcal{M}_t^i|\}$, which is the quantizer codebook at station i at time t, and $|\cdot|$ denotes the cardinality function. Hence, at each time t, station i sends $\log_2(|\mathcal{M}_t^i|)$ bits over the channel to the plant.

From an operational viewpoint, we assume that there is a decoder at the actuator who receives the corresponding controller commands and applies the actions. This can be regarded as an actuator located in the plant who receives the commands from the controller. Hence, the channel is between the controller and the actuator. One may view the system in Fig. 9.1.

In this section, we consider the following problem: Let $\mathcal{R}$ denote the set of average rates on L sensor and controller channels which lead to stabilization, that is,

$$\mathcal{R} = \Bigg\{(R^i, i \in 1, 2, \ldots, L) : \\ \exists \{u^1_{[0,\infty)}, u^2_{[0,\infty)}, \ldots, u^L_{[0,\infty)}\}, \lim_{T\to\infty} ||x_T||_\infty = 0\Bigg\},$$

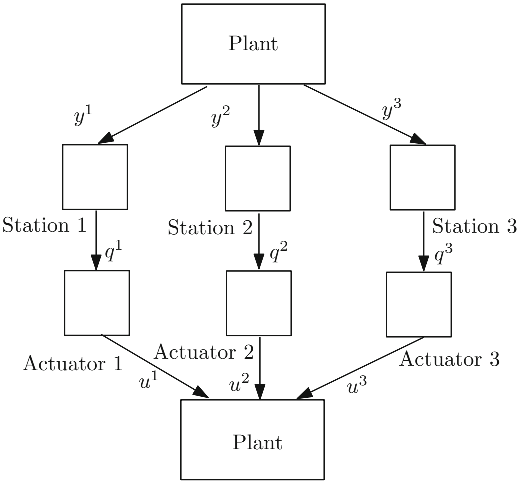

Fig. 9.1 Multi-station (multi-controller) system structure

where $R^i = \limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \log_2(|\mathcal{M}_t^i|)$. Then, what is the minimum *average total rate* $\mathbf{R} := \min_{\mathcal{R}}\{\sum_{i=1}^{L} R^i\}$, such that decentralized stabilization is possible?

We introduce two ingredients needed to address this question. First note that the observability of an LTI system (A,B,C) can be checked using the Hautus–Rosenbrock test: The pair (A,C) is observable if and only if for all $\lambda \in \mathbb{C}$, the following matrix is full rank:

$$\begin{bmatrix} \lambda I - A \\ C \end{bmatrix}.$$

Clearly, one needs to check the rank condition only for the eigenvalues of A. Likewise, controllability can be checked by replacing C with B' and A with A' above, that is, observability of the pair (A', B').

The following follows directly from this Hautus–Rosenbrock test and is stated without proof.

Proposition 9.4.1. *Consider (9.1), where the system is jointly controllable and jointly observable. Suppose that the system matrix A is in Jordan form, where each Jordan block admits distinct eigenvalues. Then, for each Jordan block, there exists at least one controller which can observe the entire eigenspace, and there exists at least one controller which can control the eigenspace.* ◇

Assumption 9.4.1. *The system matrix A is in Jordan form, where each Jordan block admits distinct eigenvalues.* ◇

Theorem 9.4.1. *Suppose that Assumption 9.4.1 holds. (i) Let A be such that the eigenvalues are real. Then, a tight lower bound on the total rate required,* $\mathbf{R}$, *between the controllers and the plant for stabilizability is given by*

$$\sum_{|\lambda_i|>1} (\eta_{M_i})\bigg(\log_2(|\lambda_i|)\bigg), \tag{9.5}$$

where

$$\eta_{M_i} = \min_{l,m\in\{1,2,\ldots,L\}} \bigg\{ D^*(l,m) : l \to m, [x^i] \subset O^i \cup O^m, \\ [x^i] \subset K^m, D^*(l,m) = d(l,m) + 1 \bigg\} \tag{9.6}$$

(ii) There exist stabilizing coding and control policies whose sum rate is arbitrarily close to the lower bound in (9.5). Hence, this bound is asymptotically achievable. ◇

Proof. See Sect. 9.8.2. □

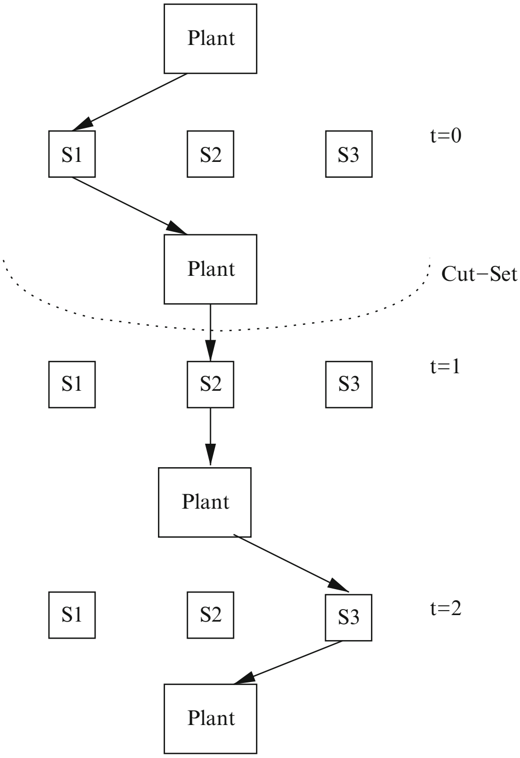

Fig. 9.2 A max-flow min-cut interpretation on a spatial and temporal graph. In the figure, station S1 observes a mode and relays it to station S2, which further relays it to station S3. Cut-set bounds in information theory are based on such information rates across cuts in a graph

We note that the first result of the theorem above admits a max-flow min-cut interpretation (Fig. 9.2) over a temporal graph [154]. One can approach the problem as information transfer over a network, where the rate of information flow across any cut is less than the mutual information between the inputs on either side of the cut conditioned on the inputs on the other side of the cut.

When Assumption 9.4.1 does not hold, the proof of Theorem 9.4.1 is still applicable: We can, instead of a Jordan form representation, invoke the representation in (9.4) and obtain a sequential stabilization characterization for each mode, considering also the updates in the observable modes as a result of sequential stabilization which sets the lower modes to 0. However, in this context, the combinatorial nature of the mode selection would require the expression to be more complicated. This combinatorial aspect is due to the selection of controller and observer sets and the communication paths between the stations.

To gain further insight on these settings, we consider two scenarios where Assumption 9.4.1 does not necessarily hold.

Case 1: Multiple Controllers Need Information on a Mode from Multiple Stations

Suppose that there is more than one controller which can control a single mode, yet their information is not sufficiently rich to recover the mode independently. Such a scenario can apply to modes with geometric multiplicities more than one. The following example captures this scenario:

$$x_{t+1} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_t^1 + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_t^2 + \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} u_t^3 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u_t^4,$$

$$y_t^1 = [0\ 0\ 0\ 1]\, x_t, \qquad y_t^2 = [0\ 0\ 1\ 0]\, x_t,$$
$$y_t^3 = [1\ 1\ 0\ 0]\, x_t, \qquad y_t^4 = [1\ -1\ 0\ 0]\, x_t.$$

In this case, the third and fourth stations send information to the first two controllers, which can control the first mode whose information is not enough to recover the mode independently.

Lemma 9.4.1. *The minimum average total information rate needed to be sent to the controllers for being able to control a mode with eigenvalue λ_i from those that can help recover the mode is lower bounded by*

$$\min_{\mathbb{K},\mathbb{L}:[x^i]\subset(\cup_{m\in\mathbb{K}}K^m)\cap(\cup_{j\in\mathbb{L},m\in\mathbb{K}}O^j\cup O^m)} d(\mathbb{L},\mathbb{K}) \max(0, \log_2(|\lambda_i|)), \tag{9.7}$$

where

$$d(\mathbb{L},\mathbb{K}) = \sum_{l\in\mathbb{L}} \min_{k\in\mathbb{K}} d(l,k).$$

◇

Proof. See Sect. 9.8.3. □

Case 2. Multiple Controllers Control a Given Mode Decentrally

Suppose that there is a number of controllers that can control a given mode, but only their joint information is sufficient to recover the mode. The following lemma, whose proof is given in the appendix, gives the information rate in that case.

Lemma 9.4.2. *The minimum average total information rate needed to be sent from the controllers to the plant for controlling a mode with eigenvalue λ_i is*

$$\min_{\mathbb{K}:[x^i]\subset(\cup_{m\in\mathcal{K}}K^m)\cap(\cup_{m\in\mathcal{K}}O^m)} \max(0, \log_2(|\lambda_i|))|\mathbb{K}|.$$

◇

Proof. See Sect. 9.8.4. □

To illustrate the results of the section presented above, we now consider two numerical examples.

Example 9.4.1. Consider the decentralized system below where the system matrix has an eigenvalue of multiplicity two and with two Jordan blocks.

$$x_{t+1} = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_t^1 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u_t^2 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} u_t^3 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} u_t^4,$$

$$y_t^1 = [0\ 1\ 0\ 1]\, x_t, \qquad y_t^2 = [1\ 0\ 0\ 0]\, x_t,$$
$$y_t^3 = [0\ 0\ 1\ 1]\, x_t, \qquad y_t^4 = [1\ 0\ 1\ \ -1]\, x_t.$$

◇

The average rate needed for stabilization is $2\log_2(4) + \log_2(2) + \log_2(2) + \log_2(2) = 7$ bits. As the first two modes are both observable and controllable by station 2, it only requires $2\log_2(4)$ bits to stabilize. For the third mode, there are two controllers which can control the mode, but they cannot observe the mode as the third station can observe $x^3 + x^4$ and the fourth station can observe $x^3 - x^4$. The third and the fourth stations can do the following: either take part in explicit signaling or apply a control decentrally. It turns out that both of these methods lead to the same rate $\log_2(2) + \log_2(2) = 2$ bits.

Example 9.4.2. For the system below, there is no stabilizing decentralized controller:

$$x_{t+1} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} x_t + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u_t^1 + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u_t^2 + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} u_t^3 + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_t^4,$$

$$y_t^1 = [0\ 0\ 0\ 1]\, x_t, \qquad y_t^2 = [0\ 1\ 0\ 0]\, x_t,$$
$$y_t^3 = [1\ 1\ 1\ 0]\, x_t, \qquad y_t^4 = [0\ 0\ 0\ 1]\, x_t.$$

◇

The reason is because the first controller cannot receive information about the mode it can control, and there is no other controller which can arrange information transmission for that control.

9.5 Multi-Sensor Structure with a Centralized Controller

We discuss in this section the case where there is a central controller which receives information from multiple sensors. Consider the following LTI system (see Fig. 9.3):

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t\,,\ t \geq 0, \\ y_t^i &= C^i x_t\,,\ t \geq 0\,,\ i = 1, \ldots, L\,, \end{aligned} \tag{9.8}$$

where (A, B) is stabilizable, $(A, [(C^1)' \ldots (C^L)']')$ is detectable, and the initial state x_0 is a random vector with a known continuous distribution over a compact support. The information received by the sensors is quantized and sent to the controller, which has access to only the information sent by the sensors. Suppose also that the controller is connected to the plant through a noiseless infinite capacity channel.

The system can be represented, using Kalman's canonical decomposition, in an upper-triangular form (see [209]) so that the system has the following form, where $*$ may or may not be zero and $(\tilde{A}^k, (\tilde{C}^k)', 1 \leq k \leq L)$ are detectable pairs:

$$A = \begin{bmatrix} \tilde{A}^1 & * & * \cdots & * \\ 0 & \tilde{A}^2 & * \cdots & * \\ \vdots & \vdots & \vdots \cdots & \vdots \\ 0 & 0 & 0 \cdots & \tilde{A}^L \end{bmatrix}, \qquad C = \begin{bmatrix} \tilde{C}^1 & * & * \cdots & * \\ 0 & \tilde{C}^2 & * \cdots & * \\ \vdots & \vdots & \vdots \cdots & \vdots \\ 0 & 0 & 0 \cdots & \tilde{C}^L \end{bmatrix}. \tag{9.9}$$

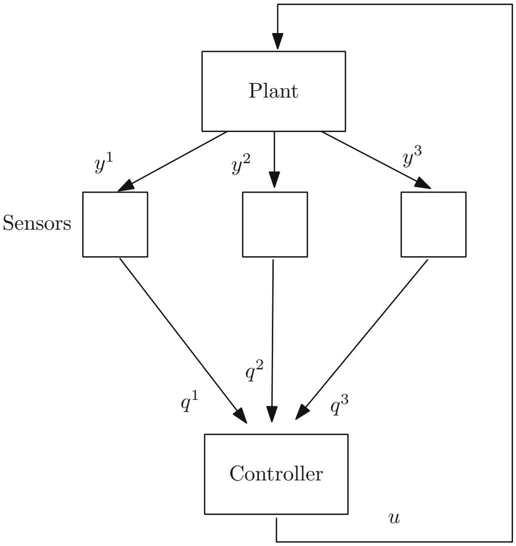

Fig. 9.3 Multi-sensor system structure

The controller can stabilize the components sequentially by receiving information from the sensors, as before, by moving up in the matrix. The following result then essentially follows from Theorem 9.4.1, through a sequential stabilization argument.

Theorem 9.5.1. *For the system considered in (9.8), for asymptotic stability, the average total rate for every stabilizing coding and control policy satisfies* $R \geq \sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$*, and this bound is tight in the sense that there exists a stabilizing coding and control policy for which the average total rate satisfies* $R > \sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$*.* ◇

9.6 Multi-Sensor and Multi-Controller Systems Driven by Noise

One further aspect of the difference between the multi-sensor case and the multi-controller case is that for the former there is no issue of signaling. This allows one to obtain tight results when there is both system and observation noise, as we discuss in the following.

9.6.1 Multi-Sensor Systems Driven by Unbounded Noise

In this subsection, we consider a multi-sensor system of the form

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t\,,\ t \geq 0, \\ y_t^i &= C^i x_t + v_t^i\,,\ t \geq 0\,,\ i = 1, \ldots, L\,, \end{aligned} \tag{9.10}$$

where (A, B) is stabilizable, $(A, [(C^1)' \ldots (C^L)']')$ is detectable, $\{v_t^i, w_t\}$ are i.i.d. noise processes with $E[|v_t^i|^{2+\epsilon}] < \infty$, $E[|w_t|^{2+\epsilon}] < \infty, \epsilon > 0$, and Assumption 9.6.1, to be introduced next, is applicable.

Assumption 9.6.1. *Suppose that* A *is in Jordan form. For every Jordan block in* A*, there exists a sensor which observes the entire block.* ◇

We note that here the Jordan blocks do not necessarily have distinct eigenvalues unlike Assumption 9.4.1.

Through the random-time state-dependent drift analysis considered in Chaps. 6 and 7, and obtaining an upper-diagonal structure as in (9.9), we can extend Theorem 7.4.1 to the multi-sensor case. However, there are two aspects which need to be addressed: coordination among the sensors for the stopping time analysis and the coupling between the *observable modes* of the sensors. In view of Theorem 7.4.1 and the construction in Sect. 7.4, we can let all the sensor quantizers zoom out and zoom in simultaneously by allowing a *zoom-out symbol* to be exchanged (see Fig. 9.4). Assumption 9.6.1 facilitates the separation of observable modes which partitions the (unstable portion of the) state space $\mathbb{R}^n$.

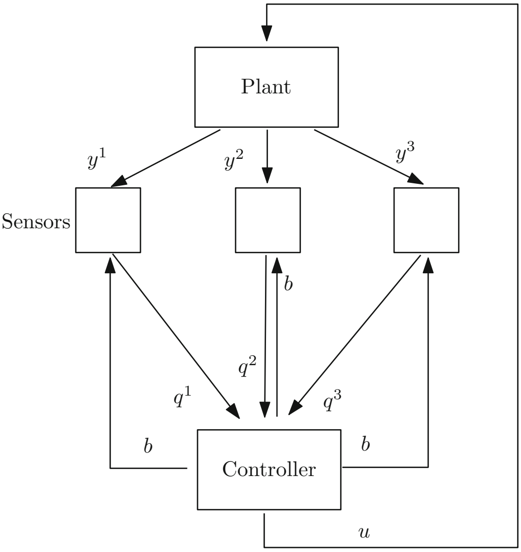

Fig. 9.4 Multi-sensor system structure with unbounded noise. Here b represents the one-bit coordination symbol for simultaneous zooming in and out

To facilitate the random-time drift analysis, with T denoting a sampling period in the system, a sequence of stopping times for zooming-in events can be defined as

$$\tau_0 = 0, \quad \tau_{z+1} = \inf\{k2nT > \tau_z : |h^i_{k2nT}| \leq 1, i \in \{1, 2, \ldots, n\}\}, \quad z \in \mathbb{Z}_+,$$

where $h^i_t = \frac{\bar{x}^i_t}{\Delta^i_t 2^{R'_i - 1}}$. Here Δ^i is the bin size of the quantizer in the direction of the eigenvector x^i, with rate used R'_i for a corresponding eigenvalue, and $\bar{x}^i_t$ is an estimate of the state component x^i_t, computed by a corresponding sensor.

Under update rules for the bin sizes as in (7.9), we can establish a finite geometric measure which dominates the stopping time distribution. In view of these, by taking T as a free parameter, we can establish the following.

Theorem 9.6.1 ([209]). *Consider the system in (9.10). To achieve asymptotic stability in the sense that*

$$\limsup_{t \to \infty} E[|x_t|_2^2] < \infty,$$

it suffices to have the average total rate satisfying $R > \sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$. ◇

The results can be extended to cases where Assumption 9.6.1 does not hold, for which case upper and lower bounds on the information rates can be provided to facilitate the drift arguments. In this case, the structure of the system in (9.9) plays an important role in the analysis. In particular, if the eigenvalues of the matrices $\tilde{A}^1, \cdots, \tilde{A}^L$ are ordered in decreasing magnitude, then Theorem 9.6.1 holds without Assumption 9.6.1. The reader is referred to [209] for details.

9.6.2 Multi-Controller Systems Driven by Unbounded Noise

In case one considers a noisy multi-controller counterpart of (9.10), the analysis is more involved due to the signaling aspects presented in Theorem 9.4.1: Consider a signaling phase where station k is signaling to station l. The observations at station l will be of the form:

$$y^l_{t+1} = C^l(Ax_t + B^k u^k_t + w_t) + v^l_t,$$

where station l tries to recover u^k_t without necessarily knowing x_t and the noise realizations. Signaling in such a setting corresponds to coding and decoding over an additive channel (with unequal side information on the *channel* due to unequal estimates regarding the state x_t at the controllers taking part in signaling; see [430] for a discussion). Unless there are no information transmission (power) constraints in the signaling phase, the analysis leads to a tedious communication problem. However, the essential construction of the signaling and control phases remains the same as that in the proof of Theorem 9.4.1.

9.7 Illustration of Binning and Its Use in Decentralized Stabilization

Before ending this chapter, we discuss and illustrate the useful information theoretic notion of binning, a coding method that exploits the available side information at the DMs which receive information.

Even though the discussion here is not related to the coverage earlier in this chapter and may be skipped by the reader, it is worth noting that binning is at the heart of multi-terminal information theoretic source and channel-coding problems (see [103, 312]), which, however, is also applicable to a class of real-time systems. This notion was discussed earlier in Chap. 3 in the context of the discrete Witsenhausen's counterexample and may be used to obtain practical design schemes for a class of decentralized problems. Binning will also be employed explicitly in minimum information exchange characterization for agreement problems in Chap. 12. The description in this section is meant to provide an instructive illustration on the use of binning.

Consider a 2-dimensional discrete-time noisy LTI system with a 2-dimensional control input, where each control component has a direct effect on a corresponding scalar state, that is,

$$x_{t+1} = Ax_t + Bu_t + w_t\ ,\ t \geq 1, \tag{9.11}$$

where A is an 2×2 matrix and $B = \text{diag}(b_1, b_2)$ is nonsingular. Suppose that the state x at time 1, x_1, is a continuous random vector with a known distribution with compact support, depicted in Fig. 9.5 with the square box, and w_t is an i.i.d. noise process whose distribution also has compact support.

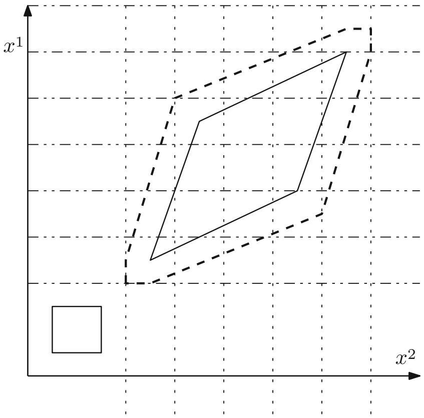

Fig. 9.5 A typical evolution of the uncertainty in a system of the form (9.11). Note that the evolution is parallel to the directions of the eigenvectors, in addition to the bounded noise effects. The encoders thus send correlated data and can exploit this dependency

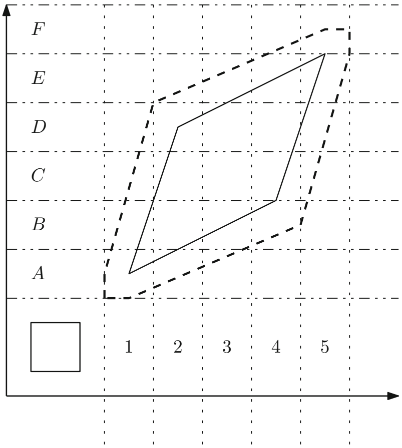

Fig. 9.6 The evolution of the uncertainty, with the same matrix in Fig. 9.5 being used. If the encoders do not collaborate, then each of their corresponding sensors will send information for 5 symbols. We encode $\log_2(30)$ bits if we do not let controllers cooperate (exchange their information)

Let x^1, x^2 be the scalar components of the system and let two sensors observe these variables. The goal is to send information to a receiver.

For the stabilization of such a system, a typically suboptimal (note that, we already discussed optimal transmission schemes earlier in the chapter), yet practical approach would be to encode the observations using a time-invariant scheme by exploiting the side information available at the controller. Figures 9.5–9.8 depict the essential ingredients of such an approach using ideas from information theory and

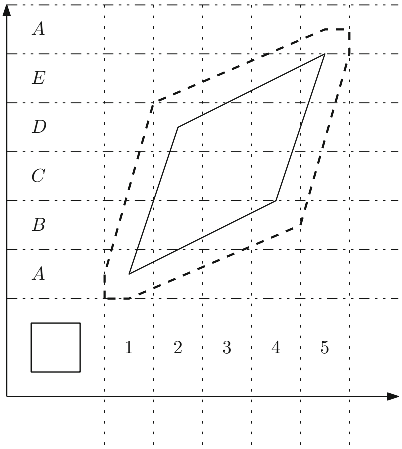

Fig. 9.7 For any level that sensor 2 has to send, there are only 5 bins, and not 6, that sensor 1 needs to send. We would need $\log_2(25)$ bits to be transmitted

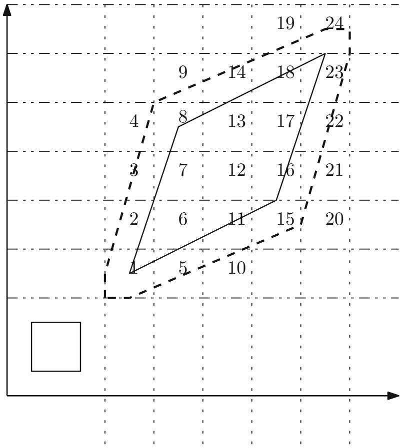

Fig. 9.8 Joint encoding outperforms independent encoding. If the sensors are allowed to share their information as in a centralized encoding scheme, there would only be $\log_2(24)$ bits to be transmitted

binning. Figure 9.5 depicts the evolution of the uncertainty, that is, the support set of the state random variable given that at time 1, the state's support was a unit box and the support set for time 2 is depicted. Figure 9.6 represents the rate conditions if the sensors do not employ binning. Figure 9.7 reveals the benefit of binning leading to an improvement in rate requirements, which, however, may not perform as well as a centralized scheme where joint encoding is allowed (see Fig. 9.8).

For an analysis of such constructions, we refer the reader to [312, 428, 429].

9.8 Appendix: Proofs

9.8.1 A Supporting Lemma

First we state and prove a lemma which will be used in the proof of Theorem 9.4.1.

Lemma 9.8.1. *Let* $\{k_t, t = 1, 2, \dots\}$ *be a sequence of positive numbers such that* $\lim_{t\to\infty} k_t = 0$. *For a sequence of scalar almost surely bounded random variables* $\{v_t, t = 1, 2, \dots\}$ *such that for all* $t \in \mathbb{N}_+$ $P(|v_t| \leq k_t) = 1$, *the entropy sequence has the property that* $\limsup_{t\to\infty} h(v_t) = -\infty$. ◇

Proof. The lemma follows from the fact that the entropy of a random variable with a bounded support is maximized by a uniform distribution. We now prove this argument. Let $p_t(dv)$ denote the probability measure of a random variable v_t, which we assume admits a density function, and denote the density also by $p_t(v_t)$. If this assumption does not hold, by constructing a sequence of discrete random variables appropriately converging to the discrete or a mixed random variable (see [171]), a similar argument as below is applicable. By Jensen's inequality below, we have that

$$h(v_t) = -E[\log_2(p_t(v_t))] = -\int_{-k_t}^{k_t} p_t(v_t)\log_2(p_t(v_t))dv_t$$

$$\leq \log_2(E[\frac{1}{p_t(v_t)}]) \leq \log_2(2k_t) \to -\infty.$$

□

9.8.2 Proof of Theorem 9.4.1

(i) Let $\mathbf{x}^\lambda$ be the eigenspace of λ, i.e., the space spanned by the eigenvectors corresponding to eigenvalue λ. Let us order the eigenvalues in an upper-triangular matrix as $x^k \in \mathbf{x}^\lambda$ with the smallest k being the lowermost mode and order them with increasing index in a Jordan block starting from the lowermost mode. Let the ith mode live in an ith eigenspace, with eigenvalue λ_i. Then, the dynamics of the ith mode can be written as

$$x_t^i = \lambda_i^t x_0^i + (\sum_k 1_{\{x^k \in \mathbf{x}^\lambda, k \leq i-1\}} f(x_0^k)) + B^{S_i}(u_{[0,t]}^{S_i}),$$

where $f(\cdot)$ is some function that depends on the upper-diagonal form of the system matrix, and the S_i denotes the set of stations that can control mode i, i.e., $x_0^i \in K^m$ for station m in S_i, $u_{[0,t]}^{S_i}$ denotes the control sequence applied by such stations, and $B^{S_i}(\cdot)$ denotes the mapping from the set of applied controls to the mode. The evolution equation above follows from the fact that modes sharing a similar eigenvalue are decoupled from other modes. We now have

$$
\begin{aligned}
h(x_t^i) &= h(\lambda_i^t x_0^i + (\sum_k 1_{\{x^k \in \mathbf{x}^\lambda, k \leq i-1\}} f(x_0^k)) + B^{S_i} u^{S_i}_{[0,t]}) \\
&\geq h\Bigg(\lambda_i^t x_0^i + (\sum_k 1_{\{x^k \in \mathbf{x}^\lambda, k \leq i-1\}} f(x_0^k)) \\
&\qquad + B^{S_i} u^{S_i}_{[0,t]} | u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_0^\lambda, k \leq i-1\Bigg) \\
&= h(\lambda_i^t x_0^i | u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_0^\lambda, k \leq i-1) \\
&= t \log_2(|\lambda_i|) + h(x_0^i | u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_0^\lambda, k \leq i-1). \qquad (9.12)
\end{aligned}
$$

The first inequality above follows from the fact that conditioning does not increase the entropy and the equalities follow from standard properties of the entropy function. It now follows from (9.12) that

$$h(x_0^i) - h(x_0^i | u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_0^\lambda, k \leq i-1) \geq h(x_0^i) + t \log_2(|\lambda_i|) - h(x_t^i).$$

Since the initial state has finite entropy, the sequence $\{x_t^i\}$ converges to zero, and by Lemma 9.8.1, $\limsup_{t \to \infty} h(x_t^i) = -\infty$, and it follows that

$$
\begin{aligned}
&\liminf_{t \to \infty} \frac{1}{t} \{h(x_0^i) - h(x_0^i | u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_{\mathbf{0}}^\lambda, k \leq i-1)\} \\
&\qquad \geq \liminf_{t \to \infty} \frac{1}{t} \{t \log_2(|\lambda_i|) + h(x_0^i) - h(x_t^i)\} \\
&\qquad \geq \log_2(|\lambda_i|). \qquad (9.13)
\end{aligned}
$$

Now, let $Z^m = \{k : k \to m\}$ be the set of stations which can communicate to station m. We have that $u^{S_i}_{[0,t]}$ is a causal function of the information available, $\{y^m_{[0,t]}, y^{Z^m}_{[0,t]}, m \in S_i, \{x_0^k\} \in \mathbf{x}_{\mathbf{0}}^\lambda, k \leq i-1\}$, and thus from the data-processing inequality it follows that

$$I(x_0^i; u^{S_i}_{[0,t]}, x_0^k \in \mathbf{x}_{\mathbf{0}}^\lambda, k \leq i-1) \leq I(x_0^i; \{y^m_{[0,t]}, y^{Z^m}_{[0,t]}, m \in S_i, x_0^k \in \mathbf{x}_{\mathbf{0}}^\lambda, k \leq i-1\})$$

and

$$\liminf_{t \to \infty} \frac{1}{t} I(x_0^i; \{y^m_{[0,t]}, y^{Z^m}_{[0,t]}, m \in S_i, x_0^k \in \mathbf{x}_{\mathbf{0}}^\lambda, k \leq i-1\}) \geq \log_2(|\lambda_i|).$$

Note that the above holds even if one allows nonlinear control policies, as long as these are causally measurable functions of the observations.

Hence, the information rate requirement is that, from the set of stations which can observe a mode to a station which can control a mode, the above rate requirement is necessary.

(ii) According to the system modes, the controllers who can control them can be identified. By Lemma 9.3.3, any station l can compute $C^l A^n x_n$, at time n. There exists at least one station, station l, that can control a mode i and at least one station, station k, that can observe the mode x^i and $k \to l$. Then, the information on mode x^i is to be transmitted to station l through the plant (where without any loss we assume that these stations are unique; otherwise, the most rate-efficient pair can be identified). Suppose that information on x_0^i is to be transmitted to station l. Sensor k recovers x_0^i at a time no later than n. It then quantizes x_0^i uniformly. Station k sets $u_t^k = Q_t(x_0^i)$, where Q_t denotes the quantization function at time t. In this case,

$$\begin{aligned} x_{n+1} &= Ax_n + B^k Q_n(x_0^i), \\ y_{n+1}^l &= C^l(Ax_{n+1} + B^k Q_n(x_0^i)). \end{aligned}$$

Assembling the observations $\{y_{[n+1,2n]}^l\}$ and using the fact that $C^l(A)^m B^k \neq 0$ for at least one $m, 1 \leq m \leq n$, and Lemma 9.3.3, the quantized output $Q_n(x_0^i)$ can be recovered at a time no later than $2n$. Sensor l can recover the quantized information $Q_n(x_0^i)$, which it subsequently sends to station l. Via this information, the estimate at time $2n$, $\hat{x^i}_0(2n)$, can be computed. Let $p > 0$ be an integer. If an average quantization rate of $R = n \log_2 |\lambda_i| + \epsilon$, for some $\epsilon > 0$, is used, then the estimation error $x_0^i - \hat{x^i}_0(pn)$ approaches zero at a rate faster than $1/(|\lambda_i|)^{pn}$.

The plant undoes the signaling, since it is assumed to know the control protocol. We assume (without any loss of generality) that after signaling takes place, the plant can undo it in the sense of canceling the effects of communication, since it is assumed to know the control protocol: The actions live in the controllable subspace, and these can be undone/cancelled by the same controller in at most n time stages, such that the system behaves as if it has been operating open loop by the end of $t + n$ time stages, where t is the time when signaling ends.

The controller can then drive the estimated value to zero in at most n time stages. Finally, we need to consider multiple transmissions. The remaining controllers can be designed to be idle, while a particular mode is being relayed by the plant. Such a sequential scheme ensures convergence, where ϵ can be taken to be arbitrarily small by adjusting the time stages. $\square$

9.8.3 *Proof of Lemma 9.4.1*

Let $\mathbb{K}$ be the set of controllers which can (possibly jointly) control the mode, that is, $x^i \in \cup_{m \in \mathbb{K}} K^m$. Following Lemma 9.8.1, we require $h(x_t) \to -\infty$ and $\liminf_{t \to \infty} \frac{1}{t} I(x_0^i; z_{[0,t]}^j, j \in \mathbb{K}) \geq \log_2(|\lambda_i|)$.

Without any loss of generality, suppose the controllers themselves do not have the information needed to recover the state of the mode to be controlled, for otherwise the rate needed for signaling would be zero. Let the information to be sent to station $1 \in \mathbb{K}$ at time t be z^1_t. It follows from the fact that the entropy of a discrete-valued random variable is nonnegative and conditioning does not increase the entropy of a random variable that

$$\begin{aligned}
\lim_{t\to\infty}\frac{1}{t}H(z^1_{[0,t]}) &\geq \lim_{t\to\infty}\frac{1}{t}\bigg\{H\bigg(z^1_{[0,t]}\bigg) - H\bigg(z^1_{[0,t]}\bigg|\lambda_i^t x_0^i, z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg)\bigg\} \\
&= \lim_{t\to\infty}\frac{1}{t}\bigg\{I\bigg(z^1_{[0,t]}; \lambda_i^t x_0^i, z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg)\bigg\} \\
&= \lim_{t\to\infty}\frac{1}{t}\bigg(I(z^1_{[0,t]}; \lambda_i^t x_0^i\bigg|z^j_{[0,t]}, j \in \mathbb{K}-\{1\})\bigg) \\
&\qquad + I\bigg(z^1_{[0,t]}; z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg)\bigg\} \\
&= \lim_{t\to\infty}\frac{1}{t}\bigg\{h\bigg(\lambda_i^t x_0^i\bigg|z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg) \\
&\qquad - h\bigg(\lambda_i^t x_0^i\bigg|z^j_{[0,t]}, j \in \mathbb{K}-\{1\}, z^1_{[0,t]}\bigg) \\
&\qquad + I\bigg(z^1_{[0,t]}; z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg)\bigg\} \\
&= \lim_{t\to\infty}\frac{1}{t}\bigg\{t\log_2(|\lambda_i|) + h\bigg(x_0^i\bigg|z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg) \\
&\qquad - h\bigg(\lambda_i^t x_0^i\bigg|z^j_{[0,t]}, j \in \mathbb{K}-\{1\}, z^1_{[0,t]}\bigg) \\
&\qquad + I\bigg(z^1_{[0,t]}; z^j_{[0,t]}, j \in \mathbb{K}-\{1\}\bigg)\bigg\} \\
&= \log_2(|\lambda_i|) + \eta \geq \log_2(|\lambda_i|),
\end{aligned}$$

where

$$\eta = \lim_{t\to\infty}\frac{1}{t}I(z^1_{[0,t]}; z^j_{[0,t]}, j \in \mathbb{K}-\{1\}) \geq 0.$$

Likewise, for the second message $z^2_{[0,t]}$ we have the same result and for all modes in $\mathbb{K}$. As such, a rate equal to the sum of these is necessary.

All stations which are designed to control the mode need to be sent information with at least an equal rate. □

9.8.4 Proof of Lemma 9.4.2

Proof is almost identical to that of Lemma 9.4.1. Let stations in the set $\mathbb{K}$ be able to control a mode and be sent the sufficient information to recover the mode. Let z_t^1 be the message sent by station $1 \in \mathbb{K}$ to the plant at time t. It then follows that

$$\begin{aligned}\lim_{t\to\infty} \frac{1}{t} H(z_{[0,t]}^1) &\geq \lim_{t\to\infty} \frac{1}{t} \Bigg\{ H\left(z_{[0,t]}^1\right) - H\left(z_{[0,t]}^1 \middle| \lambda_i^t x_0^i, z_{[0,t]}^j, j \in \mathbb{K} - \{1\}\right)\Bigg\} \\ &= \lim_{t\to\infty} \frac{1}{t} \Bigg(t \log_2(|\lambda_i|) + h\left(x_0^i \middle| z_{[0,t]}^j, j \in \mathbb{K} - \{1\}\right) \\ &\qquad - h\left(\lambda_i^t x_0^i \middle| z_{[0,t]}^j, j \in \mathbb{K} - \{1\}, z_{[0,t]}^1\right) \\ &\qquad + I\left(z_{[0,t]}^1 ; z_{[0,t]}^j, j \in \mathbb{K} - \{1\}\right)\Bigg\} \\ &\geq \log_2(|\lambda_i|) + \eta \geq \log_2(|\lambda_i|), \end{aligned} \tag{9.14}$$

where $\eta = \lim_{t\to\infty} \frac{1}{t} I(z_{[0,t]}^1 ; z_{[0,t]}^j, j \in \mathbb{K} - \{1\}) \geq 0$, following the same arguments as in the proof of Lemma 9.4.1. Hence, with the same arguments for the other control signals from stations in $\mathbb{K}$, $\log_2(|\lambda_i|)|\mathbb{K}|$ is a necessary amount of average rate. Any other coding scheme will require at least this rate. □

9.9 Concluding Remarks

In this chapter, we have provided a characterization for the existence of stabilizing controllers for multi-station systems under any control policy admissible under the given decentralized information structure.

Extension to the noisy case is a challenging issue when signaling is present. The signaling problem is equivalent to coding between two decision makers with nonidentical side information about the channel, as different stations have different information on the states they wish to control, and they signal communications over. The practicality of the problem requires further applicable and scalable design schemes.

9.10 Bibliographic Notes

In addition to the references listed in the chapter on decentralized stabilization, we should mention the following early efforts in the literature. We refer the reader to the comprehensive books by [252, 340, 342] for general references for decentralized control theory. Corfmat and Morse [101] provided conditions for decentralized stabilization with time-invariant, output feedback controllers when a leader is picked to control the entire system. If a leader is selected, by restricting the other agents to use time-invariant or time-varying linear laws, the leader might be able to control the entire system under strong connectivity conditions. Wang and Davison [386] proved that unless unstable fixed modes are present, a decentralized system can be stabilized by linear time-invariant controllers. Anderson and Moore [11] provided algebraic conditions for existence of decentralized fixed modes under linear time-invariant policies; the reader is also referred to [278]. Kobayashi et al. [217] presented a graph-theoretic discussion for the case where the decentralized system can be expressed as a set of strongly connected subsystems. They proved that the system is stabilizable by a linear controller if and only if there is no fixed mode between the decentralized systems composed of the strongly connected subsystems. Anderson and Moore [11] showed that decentralized stabilization in a multi-controller setting is possible via time-varying control laws, if the system is jointly controllable, jointly observable, and strongly connected. Further related references are the works of Wang [387], Willems [392], Khargonekar and Özgüler [213], and Gong and Aldeen [165] which further studied time-varying control laws for stabilization. Khargonekar and Özgüler [213] studied the necessary and sufficient requirements for stabilization via time-varying controllers in terms of input-output mappings. The conditions they provided are algebraic and further corroborate the fact that strong connectivity does ensure decentralized stabilizability under the assumption of joint controllability and observability. Gong and Aldeen [165] considered the decentralized stabilization problem and obtained the characterization for stabilizability along similar algebraic conditions. Özgüner and Davison [303] used a sampling technique to eliminate fixed modes resulting from time-invariant policies.

The characterization of minimum information requirements for multi-sensor and multi-controller linear systems with an arbitrary topology of decentralization has been discussed in various publications [263–266, 281, 282, 305, 353, 354, 429, 431]. In particular, references [263, 266, 305, 431] considered signaling in networked control problems with information theoretic coding perspectives. Gupta et al. [180] considered stabilization for multi-sensor systems over erasure channels.

The issue of complexity of decentralized computation is another important aspect of decentralized control applications. [404] studied the communication complexity of decentralized control, building on the notion of *communication complexity of computation* in [412]. In the information theory literature, distributed function computation with minimum information exchange is another important area, with some notable results being reported in [300], which does not consider a real-time setup, but an information theoretic setup, which considers an infinite copy of messages to be encoded and functions to be computed, extending the results in Csiszar and Körner ([107], Theorem. 4.6) to a computation setting. These will be considered further in the context of optimization of dynamic teams in Chap. 12.

Some of the results of this chapter are based on [209, 429, 431].

Part III

Optimization in Networked Control: Design of Optimal Policies Under Information Constraints

Chapter 10
Optimization of Real-Time Coding and Control Policies: Structural and Existence Results

10.1 Introduction

In Part II of this book we addressed the problem of *stabilization* of networked control systems. In this chapter, and Part III overall, we move beyond stabilization and study *optimization* of such systems, from the points of view of both encoding and control policies.

The chapter considers the optimal causal encoding/quantization problem for networked control systems. It presents structural results on optimal causal coding of Markov sources in a large class of settings: fully observed and partially observed Markov sources as well as multi-sensor systems and systems driven by control. For the optimal causal coding of a fully observed or a partially observed Markov source, the structure of optimal causal coders are obtained, which feature a separation structure. It is also shown that real-time decentralized coding of a partially observed i.i.d. source admits a memoryless optimal solution. Such a result does not, in general, extend to decentralized coding of partially observed Markov sources. We also establish in the chapter the existence of optimal control and quantization policies under appropriate technical conditions. Linear systems with quadratic cost will also be considered.

The contents of the chapter are as follows: In Sect. 10.2, we introduce the problem structure while also revisiting the setup of Sect. 5.2.2. We then present, in Sect. 10.3, structural results on optimal encoders, more precisely on optimal real-time coding of Markov sources when there is only one encoder. We study both fully observed and partially observed settings, as well as systems driven by control. Section 10.4 considers the existence of optimal quantization policies. In Sect. 10.5, we move to a decentralized setting, and show through a counterexample the difficulty one encounters in obtaining structural results for decentralized coding in the absence of a controlled Markov state construction, while providing a separation result when the source is memoryless. We discuss, in Sect. 10.6.2, the case of a partially observed Gaussian source and establish the optimality of a separation structure of estimation/filtering and quantization of the filtering output. We also investigate

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_10,

the optimal quantization and control problem for linear-quadratic-Gaussian (LQG) problems, and building on the developments in Chap. 4, we establish the existence of optimal quantizers and control policies. Finally, Sect. 10.7 considers the structure of optimal coding policies for the case with noisy channels and noiseless feedback. An appendix to the chapter includes proofs of the main results.

For background reading on Markov Decision Processes as well as for a review of the LQG control problem and Kalman filtering, we refer the reader to Appendix D.

10.2 Policies and Action Spaces for Encoding

We consider a typical causal encoding/quantization setup of the type introduced earlier in Chap. 5 (Sect. 5.2.2). For simplicity in exposition, but without much loss of conceptual generality, we consider the case of only two encoders and within this context introduce the causality and measurability constraints in quantizer design for decentralized systems.

Consider first a control-free partially observed Markov process, defined on a probability space, $(\varOmega, \mathcal{F}, P)$, and generated by the following scalar discrete-time equations for $t \geq 0$:

$$x_{t+1} = f(x_t, w_t), \tag{10.1}$$

$$y_t^i = g^i(x_t, v_t^i), \tag{10.2}$$

for (Borel) measurable functions $f, g^i, i = 1, 2$, with $\{w_t, v_t^i, i = 1, 2\}$ zero-mean noise processes with finite second moments, which are independent across time and space. We further have $x_t \in \mathbb{X}$, and $y_t^i \in \mathbb{Y}^i$, where $\mathbb{X}, \mathbb{Y}^i$ are Polish spaces. Let an encoder, Encoder i, be located at one end of a measurement channel characterized by (10.2), this being so for $i = 1, 2$. The encoders transmit their information to a receiver (see Fig. 10.1), over a discrete noiseless channel with finite capacity, and hence, they have to quantize their input.

We let, as before in Sect. 5.2.2, $\varPi^{comp,i}$ denote a *composite quantization policy* for Encoder i, defined as a sequence of functions $\{Q_t^{comp,i}, t \geq 0\}$ which are causal such that the quantization output at time t, q_t^i, under $\varPi^{comp,i}$ is generated by a function of its local information, that is, a mapping measurable with respect to the sigma-algebra generated by

$$I_t^i = \{y_{[0,t]}^i, q_{[0,t-1]}^i, z_{[0,t-1]}^i\}, \quad t \geq 1,$$

and $I_0^i = \{y_0^i\}$, with image space $\mathcal{M}_t^i$, where $\mathcal{M}_t^i := \{1, 2, \ldots, |\mathcal{M}_t^i|\}$, for $0 \leq t \leq T-1$ and $i = 1, 2$. Here z^i denotes some additional side information available, such as feedback from the receiver.

Let $\mathbb{I}_t^i$ denote the space I_t^i belongs to; hence

$$Q_t^{comp,i} : \mathbb{I}_t^i \to \mathcal{M}_t^i.$$

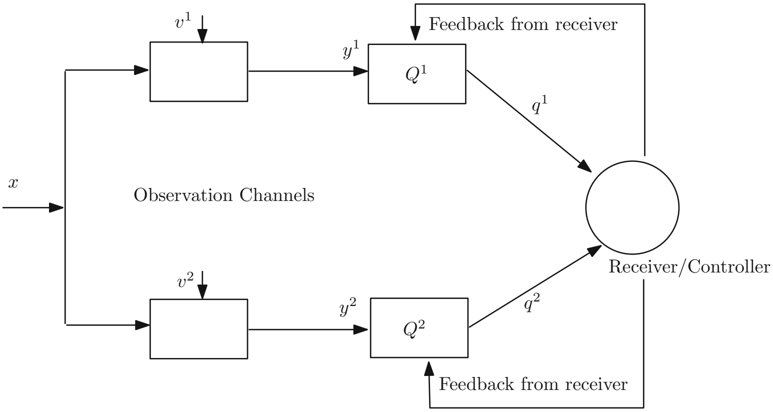

Fig. 10.1 Partially observed source under a decentralized structure

As discussed in Sect. 5.2.2, equivalently, we can express the policy $\Pi^{comp,i}$ as a composition of a *quantization policy* $\underline{\gamma}^i$ and a *quantizer*: A quantization policy of Encoder i, $\underline{\gamma}^i$, is a sequence of functions $\{\gamma_t^i\}$, such that for each $t \geq 0$, γ_t^i is a mapping from the information space $\mathbb{I}_t^i$ to the space of quantizers $\mathbb{Q}_t^i$. A quantizer is subsequently used to generate the quantizer output. Without any loss of generality, a quantizer action will be generated based on the common information at the encoders and the receiver, and the quantizer will map the relevant private information at the encoder to the quantization output. Let the information at the receiver at time t be $I_t^r = \{q^1_{[0,t-1]}, q^2_{[0,t-1]}\}$, for $t \geq 1$. Let the common information, under feedback, at the encoders and the receiver be I_t^c. Thus, we can express any measurable composite quantization policy as

$$Q_t^{comp,i}(I_t^i) = (\gamma_t^i(I_t^c))(I_t^i \setminus I_t^c), \tag{10.3}$$

mapping the information space to $\mathcal{M}_t^i$.

Viewing each encoder as an agent or a decision maker (DM), we let DMi have policy $\underline{\gamma}^i$ and under this policy generate quantizers $\{Q_t^i, t \geq 0\}$, $Q_t^i \in \mathbb{Q}_t^i$. Under action Q_t^i, the encoder generates q_t^i, as the *quantization output* at time t.

The receiver (or the controller), upon receiving the information from the encoders, generates its decision at time t, also causally: An admissible causal receiver policy is a sequence of measurable functions $\underline{\gamma}^0 = \{\gamma_t^0\}$ such that

$$\gamma_t^0 : \prod_{s=0}^{t} \left(\mathcal{M}_s^1 \times \mathcal{M}_s^2\right) \to \mathbb{U}, \qquad t \geq 0,$$

where $\mathbb{U}$ denotes the decision set for the receiver.

With the above formulation, one typical objective functional for the decision makers would be the following:

$$\inf_{\mathbf{\Pi}^{comp}} \inf_{\underline{\gamma}^0} E_{\nu_0}^{\mathbf{\Pi}^{comp},\underline{\gamma}^0}[\sum_{t=0}^{T-1} c(x_t, u_t)], \tag{10.4}$$

with initial condition distribution ν_0. Here $c(\cdot,\cdot)$, is a nonnegative, measurable function and $u_t = \gamma_t^0(\mathbf{q}_{[0,t]})$ (with $\mathbf{q} = (q^1, q^2)$) for $t \geq 0$.

Before concluding this section, it may be worth emphasizing the operational nature of causality, as different approaches could be adopted. The encoders at any given time can only use their local information to generate the quantization outputs. The receiver, at any given time, can only use its local information to generate its decision/estimate. These happen with zero delay, that is, if there is a common clock at the encoders and the receiver, the receiver at time t needs to make its decision before the realizations $x_{t+1}, y_{t+1}^1, y_{t+1}^2$ have taken place. This corresponds to the *zero-delay* coding schemes of, for example, Witsenhausen [396] and Linder and Lugosi [236].

10.3 Single Terminal Case: Optimal Causal Coding of a Partially Observed Markov Source

10.3.1 Single Terminal, Fully Observed Case

We first consider the single-encoder, fully observed case: In this setup, (10.1)–(10.2) hold with one encoder, that is,

$$x_{t+1} = f(x_t, w_t), \qquad y_t = x_t, t = 0, 1, \ldots . \tag{10.5}$$

Let $\mathcal{P}(\mathbb{X})$ denote the space of probability measures on $\mathcal{B}(\mathbb{X})$ under the topology of weak convergence and define $\pi_t \in \mathcal{P}(\mathbb{X})$ to be the regular conditional probability measure given by $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$, that is,

$$\pi_t(A) = P(x_t \in A | q_{[0,t-1]}), \quad A \in \mathcal{B}(\mathbb{X}).$$

We first state the following theorem on the structure of optimal causal quantization policies, due to Witsenhausen [396].

Theorem 10.3.1 (Witsenhausen [396]). *For system (10.5) and optimization problem (10.4), any composite quantization policy can be replaced, without any loss in performance, by one which only uses x_t and $q_{[0,t-1]}$ at time $t \geq 1$.* ◇

Proof. See Sect. 10.8.1. □

The following result is essentially due to Walrand and Varaiya [385]; however, the form below is more general since the spaces considered are not necessarily finite.

Theorem 10.3.2 ([425]). *For system (10.5) and optimization problem (10.4), any composite quantization policy can be replaced, without any loss in performance, by one which only uses the conditional probability measure $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$, the state x_t, and the time information t, at time $t \geq 1$.* ⋄

Proof. See Sect. 10.8.2. □

Remark 10.3.1. The difference between the structural results above is the following: In the setup of Theorem 10.3.1, the encoder's memory space is not fixed and keeps expanding as the decision horizon in the optimization, $T - 1$, increases. In the setup of Theorem 10.3.2, the memory space of an optimal encoder is fixed. In general, the space of probability measures is a very large one; however, it may be the case that different quantization outputs may lead to the same conditional probability measure on the state process, leading to a reduction in the required memory. Furthermore, Theorem 10.3.2 allows one to apply the theory of Markov Decision Processes, an aspect which we will elaborate on further in this chapter. ⋄

As we observed in Remark 4.7.2, the set [see (4.10)]

$$\Theta := \{\zeta \in P(\mathbb{R}^n \times \mathcal{M}) : \zeta = PQ, Q \in \mathcal{Q}\},$$

(with $\mathcal{Q}$ denoting the set of $|\mathcal{M}|$-cell quantizers) is the Borel measurable set of the extreme points of the set of probability measures on $\mathbb{R}^n \times \mathcal{M}$ with a fixed input marginal P. In view of this observation and that the class of quantization policies which admit the structure suggested in Theorem 10.3.2 is an important one, we henceforth define

$$\begin{aligned} \Pi_W := \Big\{ \Pi^{comp} = \{Q_t^{comp}, t \geq 0\} : \quad & \exists \gamma_t^1 : \mathcal{P}(\mathbb{X}) \to \mathcal{Q} \\ & Q_t^{comp}(I_t) = (\gamma_t^1(\pi_t))(x_t), \forall I_t \Big\}, \end{aligned} \tag{10.6}$$

to represent this class of policies. Here, the input measure is time varying and is given by π_t.

10.3.2 Partially Observed Markov Source

We consider here the setup of (10.1)–(10.2) but with a single encoder. Thus, the system considered is a discrete-time scalar system described by

$$x_{t+1} = f(x_t, w_t), \qquad y_t = g(x_t, v_t), t = 0, 1, ... \tag{10.7}$$

where x_t, $\{w_t, v_t\}$ are as introduced earlier. Let the quantizer, as described earlier, map its information to a finite set $\mathcal{M}_t$. At any given time, the receiver generates a quantity u_t as a function of its received information, that is, as a function of $\{q_0, q_1, \ldots, q_t\}$. The goal is to obtain a solution to (10.4) subject to constraints on the number of quantizer bins in $\mathcal{M}_t$ and the causality restriction in encoding and decoding.

Now, define $\tilde{\pi}_t \in \mathcal{P}(\mathbb{X})$ to be the regular conditional probability measure (whose existence for every realization of observation variables follows from the fact that both the state and the observation spaces are Polish) given by $P(dx_t|y_{[0,t]})$, that is,

$$\tilde{\pi}_t(A) = P(x_t \in A|y_{[0,t]}), \quad A \in \mathcal{B}(\mathbb{X}).$$

Under the topology of weak convergence for $\mathcal{P}(\mathbb{X})$, $\{\tilde{\pi}_t\}$ evolves according to a nonlinear filtering equation (see (10.40); see also [347]) and is itself a Markov process. Let us also define $\Xi_t \in \mathcal{P}(\mathcal{P}(\mathbb{X}))$ as the regular conditional measure

$$\Xi_t(A) = P(\tilde{\pi}_t \in A|q_{[0,t-1]}), \quad A \in \mathcal{B}(\mathcal{P}(\mathbb{X})).$$

The following are the main results of this subsection.

Theorem 10.3.3 ([425]). *For system (10.7) and optimization problem (10.4) with c bounded, any composite quantization policy can be replaced, without any loss in performance, by one which only uses $\{\tilde{\pi}_t, q_{[0,t-1]}\}$ as a sufficient statistic for $t \geq 1$. This can be expressed as a quantization policy which only uses $q_{[0,t-1]}$ to generate a quantizer, where the quantizer uses $\tilde{\pi}_t$ to generate the quantization output at time $t \geq 1$.* ◇

Proof. See Sect. 10.8.3. □

Theorem 10.3.4 ([425]). *For system (10.7) and optimization problem (10.4) with c bounded, any composite quantization policy can be replaced, without any loss in performance, by one which only uses $\{\Xi_t, \tilde{\pi}_t, t\}$ for $t \geq 1$. This can be expressed as a quantization policy which only uses $\{\Xi_t, t\}$ to generate a quantizer, where the quantizer uses $\tilde{\pi}_t$ to generate the quantization output at time $t \geq 1$.* ◇

Proof. See Sect. 10.8.4. □

A number of remarks are now in order.

Remark 10.3.2. From the proof of Theorem 10.3.4, we will see that (Ξ_t, Q_t) forms a controlled Markov chain. Defining the actions as the quantizers allows one to define a Markov Decision Problem with well-defined cost functions, and state and action spaces. ◇

Remark 10.3.3. The results above can be viewed as direct extensions of the ones in the previous subsection with perfect state measurements. In fact, once one recognizes the fact that $\{\tilde{\pi}_t\}$ forms a Markov source and the cost function can be expressed as $\tilde{c}(\tilde{\pi}, u)$, for some function $\tilde{c} : \mathcal{P}(\mathbb{X}) \times \mathbb{U} \to \mathbb{R}$, one could almost directly apply Theorems 10.3.1 and 10.3.2 to recover the structural results above. $\diamond$

The results of Theorems 10.3.3 and 10.3.4 are also generalizable to settings where (a) the source is Markov of order $m > 0$, (b) a finite delay d is allowed at the decoder, and (c) the observation process depends also on past source outputs in a sense described in (10.8) below. For these cases, we consider the following generalization of the source by expanding the state space.

Suppose that the partially observed source is such that either the source is Markov of order m or there is a finite delay $d > 0$ which is allowed at the decoder. Then we can augment the source to obtain $z_t = \{x_{[t-\max(d+1,m)+1,t]}\}$. Note that $\{z_t\}$ is Markov. We can thus consider the following representation:

$$z_{t+1} = \tilde{f}(z_t, \tilde{w}_t), \qquad y_t = \tilde{g}(z_t, \tilde{v}_t), \tag{10.8}$$

for some $\tilde{f}, \tilde{g}$, and where $z_t = \{x_{[t-\max(d+1,m)+1,t]}\} \in \mathbb{X}^{\max(d+1,m)}$, and $\tilde{w}_t, \tilde{v}_t$ are mutually independent, i.i.d. processes.

Any per-stage cost function of the form $c(x_t, u_t)$ can be written as for some $\tilde{c}$: $\tilde{c}(z_t, u_t)$. For the finite delay case, the cost per stage can further be specialized as $\tilde{c}(x_{t-d}, u_t)$. For the Markov case with memory, the cost function per stage writes as $\tilde{c}(x_{[t-m+1,t]}, u_t)$.

Now, by replacing $\mathbb{X}$ with $\mathbb{X}^{\max(d+1,m)}$, let $\tilde{\pi}_t \in \mathcal{P}(\mathbb{X}^{\max(d+1,m)})$ be given by

$$\tilde{\pi}_t(A) = P(z_t \in A | y_{[0,t]}), \quad A \in \mathcal{B}(\mathbb{X}^{\max(d+1,m)})$$

and $\Xi_t \in \mathcal{P}(\mathcal{P}(\mathbb{X}^{\max(d+1,m)}))$ be the regular conditional measure defined by

$$\Xi_t(A) = P(\tilde{\pi}_t \in A | q_{[0,t-1]}), \quad A \in \mathcal{B}(\mathcal{P}(\mathbb{X}^{\max(d+1,m)})).$$

Hence, we have the following result, which is a direct extension of Theorems 10.3.3 and 10.3.4. We assume that c is bounded.

Theorem 10.3.5. *Suppose that the partially observed source is such that either the source is Markov of order m or there is a finite delay $d > 0$ which is allowed at the decoder. With $z_t = \{x_{[t-\max(d+1,m)+1,t]}\}$ and y_t generated by (10.8), the following holds:*

(i) Any (causal) composite quantization policy *can be replaced, without any loss in performance, by one which only uses $\{\tilde{\pi}_t, q_{[0,t-1]}\}$ as a sufficient statistic for $t \geq 1$. This can be expressed as a quantization policy which only uses $q_{[0,t-1]}$ to generate a quantizer, where the quantizer uses $\tilde{\pi}_t$ to generate the quantization output at time $t \geq 1$.*

(ii) Any (causal) composite quantization policy *can be replaced, without any loss in performance, by one which only uses $\{\Xi_t, \tilde{\pi}_t, t\}$ for $t \geq 1$. This can be*

expressed as a quantization policy which only uses $\{\Xi_t, t\}$ *to generate a quantizer, where the quantizer uses* $\tilde{\pi}_t$ *to generate the quantization output at time* $t \geq 1$. ◇

For a further case where the decoder's memory is limited or imperfect, the results apply by replacing the full information considered so far at the receiver with the limited one with additional assumptions on the decoder's update of its memory (in particular, (10.42) in the proof of Theorem 10.3.4 does not apply in general). However, an equivalent result of Theorem 10.3.3 applies also for the limited memory setting. Such memory settings have been considered in [248, 385, 396].

10.3.3 Structural Results for Systems with Control

Theorem 10.3.2 applies also for Markov sources driven by control. That is, instead of (10.1)–(10.2), consider a system described by the following equations:

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, w_t), \\ y_t &= x_t. \end{aligned} \tag{10.9}$$

Suppose that the goal is the minimization of (10.4), with the information restrictions stated in Sect. 10.2.

For this system, we have the following result (which extends the finite state-action space analysis in [250, 384]).

Theorem 10.3.6 ([423]). *(i) For system (10.9) and optimization problem (10.4), any composite quantization policy (with a given control policy) can be replaced, without any loss in performance, by one which only uses* x_t *and* $q_{[0,t-1]}$ *at time* $t \geq 1$*, while keeping the control policy unaltered. This can be expressed as a quantization policy which only uses* $q_{[0,t-1]}$ *to generate a quantizer, where the quantizer uses* x_t *to generate the quantization output at time* t.

(ii) For system (10.9) and optimization problem (10.4), any composite quantization policy can be replaced, without any loss in performance, by one which only uses the conditional probability measure $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$*, the state* x_t*, and the time information* t*, at time* t*. This can be expressed as a quantization policy which only uses* $\{\pi_t, t\}$ *to generate a quantizer, where the quantizer uses* x_t *to generate the quantization output at time* t.

◇

Proof. See Sect. 10.8.5. □

The result also applies to the partially observed case with the conditional probability replacing the state as in Theorem 10.3.4. The proof follows from those of Theorems 10.3.4 and 10.3.6.

10.4 Existence of Optimal Zero-Delay Quantizers

We now discuss the problem of existence of optimal composite quantization policies, given the structural results for a fully observed setting. We assume that the source to be quantized is an $\mathbb{R}^n$-valued Markov source. The goal is to minimize the cost

$$J_{\pi_0}(\Pi^{comp}, \underline{\gamma}^0, T) := E_{\pi_0}^{\Pi^{comp}, \underline{\gamma}^0}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{10.10}$$

for some $T \geq 1$, where $c : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_+$ is a (measurable) stagewise cost function where $\mathbb{U}$ is an action set.

We have the following assumptions on the source $\{x_t\}$ and the cost function:

Assumption 10.4.1.

(i) *The evolution of the Markov source $\{x_t\}$ in (10.1)–(10.2) is given by*

$$\begin{aligned} x_{t+1} &= f(x_t) + w_t, \quad t \geq 0 \\ y_t &= x_t, \end{aligned} \tag{10.11}$$

where $\{w_t, t \geq 0\}$ is an i.i.d. Gaussian noise sequence and $f : \mathbb{R}^n \to \mathbb{R}^n$ is measurable and bounded.

(ii) *The cost function $c : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_+$ is continuous and bounded.*

(iii) *The initial probability measure π_0 is Gaussian.*

(iv) *$\mathbb{U}$ is compact (the compactness condition will be relaxed for LQG problems in Sect. 10.6.3).*

◇

In this section, we will assume that the number of bins for the quantizers is constant for every time stage such that $|\mathcal{M}_t| = M$ for all t. As discussed in Sect. 4.7, a quantizer Q with cells $\{B_1, \ldots, B_M\}$ can be characterized as a stochastic kernel Q from $\mathbb{R}^n$ to $\{1, \ldots, M\}$ defined by

$$Q(i|x) = 1_{\{x \in B_i\}}, \quad i = 1, \ldots, M.$$

We endow the quantizers with a topology induced by such a stochastic kernel interpretation as in Sect. 4.7. If P is a probability measure on $\mathbb{R}^n$ and Q is a stochastic kernel from $\mathbb{R}^n$ to $\mathcal{M}$, then PQ denotes the resulting joint probability measure on $\mathbb{R}^n \times \mathcal{M}$. That is, a quantizer sequence Q_n converges to Q weakly at P ($Q_n \to Q$ weakly at P) if $PQ^n \to PQ$ weakly. Similarly, Q_n converges to Q in total variation at P ($Q_n \to Q$ at P in total variation at P) if $PQ^n \to PQ$ in total variation.

Suppose we adopt a quantizer policy which is in Π_W, that is, it admits the form suggested by Theorem 10.3.2. Properties of conditional probability leads to the following expression for $\pi_t(dx_t) = P(dx_t|q_{[0,t-1]})$ for $t \geq 1$:

$$\pi_t(dx_t) = \frac{\int_{x_{t-1}} \pi_{t-1}(dx_{t-1})P(q_{t-1}|\pi_{t-1}, x_{t-1})P(dx_t|x_{t-1})}{\int_{x_{t-1}} \int_{x_t} \pi_{t-1}(dx_{t-1})P(q_{t-1}|\pi_{t-1}, x_{t-1})P(dx_t|x_{t-1})}.$$

Let $\mathcal{P}$ be the set of probability measures on $\mathbb{R}^n$ endowed with the topology of weak convergence. The following is a consequence of Theorem 10.3.4 and Remark 10.3.2.

Theorem 10.4.1. *The sequence of conditional measures and the sequence of quantizers,* (π_t, Q_t)*, form a joint Markov process in* $\mathcal{P} \times \mathcal{Q}$*.* ◇

Now, under any quantization policy in Π_W and for any $T \geq 1$, by optimizing the receiver policy given a composite quantization policy in (10.10), we can define

$$J_{\pi_0}(\Pi^{comp}, T) = E_{\pi_0}^{\Pi^{comp}}\left[\sum_{t=0}^{T-1} \tilde{c}(\pi_t, Q_t)\right],$$

where, with $B_i = Q_t^{-1}(i)$, $i = 1, \ldots, M$ denoting the cells of Q_t, we have

$$\begin{aligned} &\tilde{c}(\pi_t, Q_t) \\ &= \sum_{i \in \mathcal{M}} P(q_t = i|q_{[0,t-1]}) \inf_{u \in \mathbb{U}} \left(\int P(dx_t|q_{[0,t-1]}, q_t = i)c(x_t, u)\right) \\ &= \sum_{i \in \mathcal{M}} \inf_{u \in \mathbb{U}} \int_{B_i} \pi_t(dx)c(x, u). \end{aligned} \tag{10.12}$$

As in Sect. 4.7, we restrict the set of quantizers considered by only allowing quantizers having convex quantization bins (cells) B_i, $i = 1, \ldots, M$.

Assumption 10.4.2. *The quantizers have convex codecells with at most a given number of cells, that is, the quantizers live in* $\mathcal{Q}_c(M)$*, the collection of k-cell quantizers with convex cells where* $1 \leq k \leq M$*.* ◇

Let Π_W^C denote the set of all composite quantization policies Π_W [defined in (10.6)] which in addition satisfy the condition that all quantizers Q_t, $t \geq 0$ have convex cells (i.e., $Q_t \in \mathcal{Q}_c$ for all $t \geq 0$).

We have the following result on the existence of optimal quantizers.

Theorem 10.4.2 ([437]). *For any* $T \geq 1$ *and arbitrary initial condition* π_0*, under Assumptions 10.4.1 and 10.4.2, there exists a policy in* Π_W^C *such that*

$$\inf_{\Pi^{comp} \in \Pi_W^C} \inf_{\gamma^0} J_{\pi_0}(\Pi^{comp}, \gamma^0, T) \tag{10.13}$$

is achieved. Letting $J_T^T(\cdot) = 0$ *and*

$$J_0^T(\pi_0) := \min_{\Pi^{comp} \in \Pi_W^C, \gamma^0} J_{\pi_0}(\Pi^{comp}, \gamma^0, T),$$

the dynamic programming recursion

$$J_t^T(\pi_t) = \min_{Q \in \mathcal{Q}_c} \left(c(\pi_t, Q_t) + E[J_{t+1}^T(\pi_{t+1}) | \pi_t, Q_t] \right) \tag{10.14}$$

holds for all $t = 0, 1, \ldots, T-1$. ◇

Proof. See Sect. 10.8.6. □

10.5 Multiterminal (Decentralized) Setting

10.5.1 Memoryless Sources

Let us first consider a special, but important, case of (10.1)–(10.2) when $\{x_t, t \geq 0\}$ is an i.i.d. sequence. Further, suppose that the observations are generated by

$$y_t^i = g^i(x_t, v_t^i), \tag{10.15}$$

for measurable functions $g^i, i = 1, 2$, with $\{v_t^1, v_t^2\}$ (across time) an i.i.d. noise process. We do not require that v_t^1 and v_t^2 are independent for a given t. We note that the results presented here are also applicable when the process $\{v_t^1, v_t^2\}$ is only independent (across time), but not necessarily identically distributed. One difference with the general setup considered earlier is that we require the observation spaces $\mathbb{Y}^i, i = 1, 2$, to be finite spaces; $\mathbb{X}$ is Polish.

Suppose the goal is again the minimization

$$\inf_{\mathbf{\Pi}^{comp}} \inf_{\underline{\gamma}^0} E_{\nu_0}^{\mathbf{\Pi}^{comp}, \underline{\gamma}^0} [\sum_{t=0}^{T-1} c(x_t, u_t)]. \tag{10.16}$$

Toward this end, we introduce the class of nonstationary memoryless team policies, given by

$$\begin{aligned} \Pi^{NSM} := \Big\{ \mathbf{\Pi}^{comp} : P(\mathbf{q}_t | \mathbf{y}_{[0,t]}) &= P(q_t^1 | y_t^1, t) P(q_t^2 | y_t^2, t) \\ &= 1_{\{q_t^1 = Q_t^1(y_t^1)\}} 1_{\{q_t^2 = Q_t^2(y_t^2)\}}, \\ Q_t^1 : \mathbb{Y}^1 \to \mathcal{M}_t^1, \qquad Q_t^2 &: \mathbb{Y}^2 \to \mathcal{M}_t^2, \quad t \geq 0 \Big\}, \end{aligned} \tag{10.17}$$

where $\{Q_t^1, Q_t^2\}$ are arbitrary measurable functions.

Theorem 10.5.1 ([425]). *Consider the minimization problem of (10.16). An optimal composite quantization policy over all causal policies exists, and it is an element of* Π^{NSM}. ◇

Proof. See Sect. 10.8.7. □

The result says that an optimal composite quantization policy only uses the product form admitted by a nonstationary memoryless team policy. It ignores the past observations and past quantization outputs without any loss. We note that this result applies also to the case when the source is memoryless, but not necessarily i.i.d.

Remark 10.5.1. If there is an entropy constraint on the quantizer outputs, memory in the encoders might be useful for finite horizon problems as it provides common randomness, which cannot be achieved by time-sharing in a finite horizon problem. Neuhoff and Gilbert [292] noted that randomization of two scalar quantizers (operationally achievable through time-sharing) is optimal in causal coding of an i.i.d. source subject to entropy constraints. On the other hand, for the zero-delay setting, when one considers the distortion minimization problem subject to an entropy constraint, György and Linder [184] observed that the distortion-entropy curve is non-convex (leading to a benefit of common randomness which can be used to expand the set of achievable rate and distortion pairs) as we elaborated on in Sect. 5.4. ◇

10.5.2 Markov Sources: Nonclassical Information Structure and a Counterexample Under Signaling

We now consider general Markov sources and show that a separation result of the type seen in the single-terminal case may not hold when there are multiple terminals.

We have the following (negative) result for the two-encoder setup, where the encoders have access to the feedback from the receiver (Fig. 10.1).

Proposition 10.5.1 ([425]). *Consider the setup in (10.1)–(10.2), and let* $\tilde{\pi}^i_t(A) = P(x_t \in A | y^i_{[0,t]})$, $i = 1, 2$, *and* $A \in \mathcal{B}(\mathbb{X})$. *An optimal composite quantization policy cannot, in general, be replaced by a policy which uses only* $\{\mathbf{q}_{[0,t-1]}, \tilde{\pi}^i_t\}$ *to generate* q^i_t *for* $i = 1, 2$. ◇

Proof. It suffices to produce an instance where an optimal policy cannot admit the separated structure. Toward this end, let z_1, z_2, z_3 be uniformly distributed, independent, binary numbers; and let x_0, x_1 be defined by

$$x_0 = \begin{bmatrix} z_1 \; z_2 \; 0 \; 0 \end{bmatrix}', \quad x_1 = \begin{bmatrix} 0 \; 0 \; z_2 \; z_3 \end{bmatrix}',$$

such that $x_0(1) = z_1, x_0(2) = z_2, x_0(3) = x_0(4) = 0$. Let the observations be given as follows:

$$y_t^1 = g^1(x_t) = x_t(1) \oplus x_t(3) \oplus x_t(4), \qquad y_t^2 = g^2(x_t) = x_t(1) \oplus x_t(2), \quad t = 0, 1.$$

That is,

$$y_0^1 = \left[z_1\right], \qquad y_0^2 = \left[z_1 \oplus z_2\right],$$

where $\oplus$ is the x-or operation, and

$$y_1^1 = \left[z_2 \oplus z_3\right], \qquad y_1^2 = \left[0\right],$$

Let the cost be

$$E\left[(x_0(4) - E[x_0(4)|\mathbf{q}_{[0]}])^2 + (x_1(4) - E[x_1(4)|\mathbf{q}_{[0,1]}])^2\right].$$

That is, the cost is $E[(z_3 - E[z_3|\mathbf{q}_{[0,1]}])^2]$, where q_t^i are the information bits sent to the decoder for $t = 0$ and 1.

We further restrict the information rates to satisfy $|\mathcal{M}_0^1| = |\mathcal{M}_1^1| = |\mathcal{M}_1^2| = 2$, $|\mathcal{M}_0^2| = 1$. That is, the encoder 2 may only send information at time $t = 1$.

Under arbitrary causal composite quantization policies, a cost of zero can be achieved as follows: If the encoder 1 sends the value z_1 to the receiver and, at time 1, encoder 1 transmits $z_2 \oplus z_3$ and encoder 2 transmits z_2 (or $z_1 \oplus z_2$), the receiver can uniquely identify the value of z_3, for every realization of the random variables.

For such a source, an optimal composite policy cannot be written in the separated form, that is, an optimal policy of encoder 2 at time 1 cannot be written as $h_1(\mathbf{q}_0, \tilde{\pi}_1^2)$, for some measurable function h_1. To see this, note the following: The conditional distribution of x_1 at encoder 2 at time 1 is such that the conditional measure on (z_2, z_3) is uniform and independent, that is, $P(z_2 = a, z_3 = b|z_1 \oplus z_2) = (1/4)$ for all values of a, b. If a policy of the structure of h_1 is adopted, then it is not possible for encoder 2 to recall its past observation to extract the value of z_2. This is because $\tilde{\pi}_1^2$ will be a distribution only on z_2 and z_3, which will be uniform and independent, given $z_1 \oplus z_2$. Thus, the information y_0^2 will not be available in the memory and the receiver will have access to at most $z_2 \oplus z_3$ and z_1 and $P(z_2, z_3|z_1 \oplus z_2)$ (the last variable containing no useful information). The optimal estimator will be $E[z_3] = 1/2$, leading to a cost of $1/4$. □

Discussion: Connections with Team Decision Theory

Here, we interpret the results of this section in view of optimization for dynamic teams. With the characterization of information structures for dynamic teams provided in Chap. 3, every lossy coding problem is nonclassical, since a receiver cannot recover the information available at the encoder fully, while its information is clearly affected by the coding policy of the encoder. However, in an encoding problem, the problem itself is the transmission of information. We suggest the following: *Signaling in a coding problem is the policy of an encoder to use the quantizers/encoding functions to transmit a message to other decision makers or to itself to be used in future stages, through the information sent to the receiver.*

We have seen in Chap. 3 that in decentralized decision-making problems, when the information structure is nonclassical, the decision makers might benefit from communicating via their control actions, that is, by *signaling*. We also note that, in the information theory literature, signaling has been employed in coding for multiple-access channels with feedback, where active information transmission allows for coordination between encoders (see [82, 102, 378]).

The reason for the negative conclusion in Proposition 10.5.1 is that in general for an optimal policy,

$$P(q_t^i | \pi_t^i, \mathbf{q}_{[0,t-1]}, y^i_{[0,t-1]}) \neq P(q_t^i | \pi_t^i, \mathbf{q}_{[0,t-1]}), \tag{10.18}$$

when the encoders have engaged in signaling (in contrast with what we will have in the proof of the separation results). The encoders may benefit from using the received past observation variables explicitly.

As we will discuss in detail in Chap. 12, separation results for such dynamic team problems typically require information sharing between the encoders (decision makers), where the shared information is used to establish a sufficient statistic living in a fixed state space and which admits a controlled Markov recursion (hence, such a sufficient statistics can serve as a *state* for the decentralized system). For the proof of Theorem 10.3.4, we see that Ξ_t forms such a state. For the proof of Theorem 10.5.1, we see that information sharing is not needed for the encoders to agree on a sufficient statistic, since the source considered is memoryless. Furthermore, for the multiterminal setting with a Markov source, a careful analysis of the proof of Theorem 10.5.1 reveals that if the encoders agree on $P(dx_t | \mathbf{y}_{[0,t-1]})$ through sharing their beliefs for all $t \geq 1$, then a separation result involving this joint belief can be obtained. See Chap. 12 for further discussion on this topic and a discussion on the *belief sharing information pattern*.

10.6 Simultaneous Optimization of LQG Coding and Control Policies: Optimal Quantization and Control

In this section, we consider an important application of the results presented so far. We study a LQG setup, where a sensor encodes its noisy information to a controller/estimator. First, we discuss the case without control. The case with control will be considered subsequently.

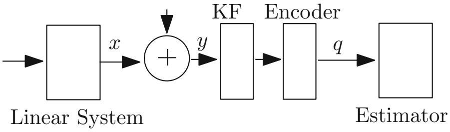

Fig. 10.2 Separation of estimation and quantization: When the source is Gaussian, generated by the linear system (10.19), the cost is quadratic, and the observation channel is Gaussian, the separated structure of the encoder above is optimal. That is, first the encoder runs a Kalman filter and then causally encodes its estimate

10.6.1 Application to the LQG Setup: Separation of Estimation and Quantization

Consider a control-free LQG setup, where a sensor is connected to an estimator over a discrete noiseless channel. Let $x_t \in \mathbb{R}^n, y_t \in \mathbb{R}^m$, and the evolution of the source be given by

$$\begin{aligned} x_{t+1} &= Ax_t + w_t, \\ y_t &= Cx_t + v_t, \end{aligned} \tag{10.19}$$

where $\{w_t, v_t\}$ is a mutually independent zero-mean Gaussian noise sequence with $E[w_t w_t'] =: W, E[v_t v_t'] =: V$, and A, C are matrices of appropriate dimensions. The goal is to obtain a solution to the minimization problem

$$\inf_{\Pi^{comp}} \inf_{\underline{\gamma}^0} E_{\nu_0}^{\Pi^{comp}, \underline{\gamma}^0} [\sum_{t=0}^{T-1} (x_t - u_t)' Q (x_t - u_t)], \tag{10.20}$$

with ν_0 denoting a Gaussian distribution for the zero-mean initial state, and $Q > 0$ a positive-definite matrix.

The conditional distribution $\tilde{\pi}_t(\cdot) = P(x_t \in \cdot | y_{[0,t]})$ is Gaussian for all time stages, which is characterized uniquely by its mean and covariance matrix for all time; thus $\tilde{\pi}_t$ can be uniquely characterized by an element of $\mathbb{R}^{\frac{n^2+3n}{2}}$. Furthermore, the nonlinear filter equation described in (D.4) admits a simpler recursion known as the Kalman filter (see Sect. B.4). We have the following result (see Fig. 10.2).

Theorem 10.6.1. *For the minimization of the cost in (10.20), any composite quantization policy can be replaced, without any loss in performance, by one which only uses the output of the Kalman filter and the information available at the receiver.* ◇

Proof. The result can be proven by considering a direct approach, rather than as an application of Theorems 10.3.3 and 10.3.4 (which require bounded costs; however, this assumption can be relaxed for this case), exploiting the specific quadratic nature of the problem. Let, again, $x_t \in \mathbb{R}^n$ and $|\cdot|_Q$ denote the norm generated by an inner product of the form $\langle x, y\rangle_Q = x^T Q y$ for $x, y \in \mathbb{R}^n$ for positive-definite $Q > 0$. The projection theorem for Hilbert spaces implies that the random variable $x_t - E[x_t|y_{[0,t]}]$ is orthogonal (see Sect. B.4) to the random variables $\{y_{[0,t]}, q_{[0,t]}\}$, where $q_{[0,t]}$ is included due to the Markov chain condition that $P(dx_t|y_{[0,t]}, q_{[0,t]}) = P(dx_t|y_{[0,t]})$. We thus obtain the following identity:

$$\begin{aligned} E[|x_t - E[x_t|q_{[0,t]}]|_Q^2] &= E[|x_t - E[x_t|y_{[0,t]}]|_Q^2] \\ &+ E\left[\left|E[x_t|y_{[0,t]}] - E\left[E[x_t|y_{[0,t]}]\middle|q_{[0,t]}\right]\right|_Q^2\right]. \end{aligned} \tag{10.21}$$

The second term is to be minimized through the choice of the quantizers. Hence, the term $\bar{m}_t := E[x_t|y_{[0,t]}]$, which is computed through a Kalman filter, is to be quantized (see Fig. 10.2). Recall that by the Kalman filter (see Sect. D.2), with

$$\Sigma_{0|-1} = E[x_0 x_0']$$

and for $t \geq 0$

$$\Sigma_{t+1|t} = A\Sigma_{t|t-1}A' + W - (A\Sigma_{t|t-1}C')(C\Sigma_{t|t-1}C' + V)^{-1}(C\Sigma_{t|t-1}A'),$$

the following recursion holds for $t \geq 0$ and with $\bar{m}_{-1} = 0$:

$$\bar{m}_t = A\bar{m}_{t-1} + \Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(CA(x_{t-1} - \bar{m}_{t-1}) + v_t).$$

Thus, the pair $(\bar{m}_t, \Sigma_{t|t-1})$ is a Markov source, where the evolution of $\Sigma_{t|t-1}$ is deterministic. Even though the cost to be minimized is not bounded, since $\bar{m}_t$ itself is a fully observed process, Theorem 10.3.1 can be used to develop the structural result that any causal encoder can be replaced with one which uses $(\bar{m}_t, \Sigma_{t|t-1})$ and the past quantization outputs. Likewise, the proof of Theorem 10.3.2 shows that, for the fully observed Markov source $(\bar{m}_t, \Sigma_{t|t-1})$, any causal coder can be replaced with one which only uses the conditional probability on $\bar{m}_t$ and the realization $(\bar{m}_t, \Sigma_{t|t-1}, t)$ at time t. □

10.6.2 Optimal LQG Coding and Control Policies and Separation Results

Here, we consider an LQG setup with control, where a sensor encodes its noisy information to a controller. Let $x_t \in \mathbb{R}^n$ and the evolution of the system be given by the following:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= x_t. \end{aligned} \tag{10.22}$$

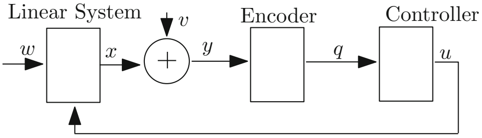

Fig. 10.3 Joint LQG optimal design of coding and control

Here, $\{w_t\}$ is a mutually independent, Gaussian noise sequence, $\{u_t\}$ is an $\mathbb{R}^m$-valued control action sequence, and A, B are matrices of appropriate dimensions. We assume that the initial state distribution is also Gaussian, denoted by ν_0.

As depicted in Fig. 10.3, we will follow the framework of Sect. 10.2 (in particular, see Theorem 10.3.6).

Suppose that the goal is the computation of

$$\inf_{\Pi^{comp}} \inf_{\underline{\gamma}^0} J(\Pi^{comp}, \underline{\gamma}^0, T), \tag{10.23}$$

where

$$J(\Pi^{comp}, \underline{\gamma}^0, T) := \frac{1}{T} E_{\nu_0}^{\Pi^{comp}, \underline{\gamma}^0} [\sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t].$$

Here, $Q \geq 0$, a positive semi-definite matrix, and $R > 0$, a positive-definite matrix.

Separation of Estimation Error and Control and Dual Effect

We first note that, by Theorem 10.3.6, an optimal composite quantization policy will be within the class Π_W.

Toward a solution, we adopt a dynamic programming approach and establish that the optimal controller is linear in its estimate [225]. This fact applies naturally for the terminal time stage control. That this also applies for the previous time stages applies from dynamic programming as we see in the following.

First consider the terminal time $t = T - 1$. For this time stage, to minimize $E[x_t' Q x_t + u_t' R u_t]$, the optimal control is $u_{T-1} = 0$ a.s.

To obtain a solution for $t = T - 2$, we look for a solution to

$$\min_{\gamma_t^0} E\left[\left(x_t' Q x_t + u_t' R u_t + E[(Ax_t + Bu_t + w_t)' Q (Ax_t + Bu_t + w_t) | \mathcal{I}_t^c, u_t]\right) \middle| \mathcal{I}_t^c\right].$$

By completing the squares and using the *orthogonality principle* (see Sect. B.4), we obtain that the optimal policy is linear and is given by

$$u_{T-2} = L_{T-2} E[x_{T-2} | q_{[0,T-2]}],$$

with

$$L_{T-2} = -R^{-1}B'QA.$$

For $t < T - 2$, to obtain the solutions, we will first establish that the estimation errors are uncorrelated. Toward this end, define for $1 \leq t \leq T - 1$:

$$\mathcal{I}_t^c = \{q_{[0,t]}, u_{[0,t-1]}\},$$

and note that

$$\tilde{m}_{t+1} := E[x_{t+1}|\mathcal{I}_{t+1}^c] = E[Ax_t + Bu_t + w_t|\mathcal{I}_{t+1}^c].$$

It then follows that

$$\begin{aligned}
\tilde{m}_{t+1} &= E[x_{t+1}|\mathcal{I}_{t+1}^c] \\
&= E[x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c] + E[x_{t+1}|\mathcal{I}_t^c]|\mathcal{I}_{t+1}^c] \\
&= E[x_{t+1}|\mathcal{I}_t^c] + E[x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]|\mathcal{I}_{t+1}^c] \\
&= E[Ax_t + Bu_t + w_t|\mathcal{I}_t^c] + E[x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]|\mathcal{I}_{t+1}^c] \\
&= A\tilde{m}_t + Bu_t + \Big(E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]\Big) \\
&= A\tilde{m}_t + Bu_t + \bar{w}_t, \qquad (10.24)
\end{aligned}$$

with

$$\bar{w}_t = (E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]). \qquad (10.25)$$

Now, $\bar{w}_t$ is orthogonal to the control action variable u_t, as control actions are determined by the past quantizer outputs and iterated expectation leads to the result that conditioned on $\mathcal{I}_t^c$, $\bar{w}_t$ is zero mean and is orthogonal to $\mathcal{I}_t^c$.

Now, for going into earlier time stages, the dynamic programming recursion for linear systems driven by an uncorrelated noise process would normally apply, since the estimate process $\{\tilde{m}_t\}$ is driven an uncorrelated noise (though, not necessarily independent) process $E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]$. However, this lack of independence may be important. Using the completion of the squares method, we can establish that the optimal controller at time t will be linear, provided that the random variable $\bar{w}_t' Q \bar{w}_t$ does not depend on $u_k, k \leq t$ under any policy. A sufficient condition for this is that the encoder is a predictive one. We state this formally as follows (see [283] for a similar, but not identical, construction):

Definition 10.6.1. A predictive quantizer policy is one where for each time stage t, the quantization has the form that the quantizer at all time stages subtracts the effect of the past control terms, that is, at time t it has the form $Q_t(x_t - \sum_{k=0}^{t-1} A^{t-k-1}Bu_k)$, and the past control terms are added at the receiver. Hence,

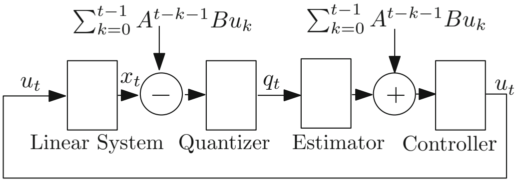

Fig. 10.4 For the LQG problem, a predictive encoder is without loss

the encoder quantizes a control-free process, defined by

$$\bar{x}_{t+1} = A\bar{x}_t + w_t,$$

and the receiver generates the quantized estimate and adds $\sum_{k=0}^{t-1} A^{t-k-1}Bu_k$ to compute the estimate of the state at time t. ◇

A predictive encoder is depicted in Fig. 10.4. We have the following key lemma.

Lemma 10.6.1 ([423]). *For problem (10.23), for any quantizer policy in class Π_W (which is without any loss as a result of Theorem 10.3.6), there exists a quantizer which satisfies the form of a predictive quantizer (see Definition 10.6.1) and attains the same performance under an optimal control policy.* ◇

Proof. See Sect. 10.8.8. □

Remark 10.6.1. We note that the structure in Definition 10.6.1 separates the estimation from control process in the sense that the estimation errors are independent of the control actions or policies. Hence, there is no *dual effect* of the control actions, in the sense that the estimation error at any given time is independent of the past applied control actions. ◇

As a consequence of the lack of dual effect, the cost function becomes

$$J(\Pi^{comp}, \underline{\gamma}^0, T) := \frac{1}{T} E_{\nu_0}^{\Pi^{comp}, \underline{\gamma}^0} [\sum_{t=0}^{T-1} \tilde{m}_t' Q \tilde{m}_t' Q x_t + u_t' R u_t + (x_t - \tilde{m}_t)' Q (x_t - \tilde{m}_t)].$$

Theorem 10.6.2. *For the minimization problem (10.23), the optimal control policy is given by $u_t = L_t E[x_t | q_{[0,t]}]$, where*

$$L_t = -(R + B' P_{t+1} B)^{-1} B' K_{t+1} A,$$

and

$$P_t = A_t' K_{t+1} B (R + B' K_{t+1} B)^{-1} B' K_{t+1} A,$$
$$K_t = A_t' K_{t+1} A_t - P_t + Q,$$

with $K_T = P_{T-1} = 0$. ◇

Given the optimal control policy, the following result is obtained after some analysis.

Theorem 10.6.3. *For the minimization problem (10.23), under an optimal control policy, the optimal cost is given by* $\frac{1}{T} J_0(\Pi^{comp}, T)$*, where*

$$J_0(\Pi^{comp},T)=E[x_0'K_0x_0]+E[(x_0-E[x_0|\mathcal{I}_0^c])'(Q+A'K_1A)(x_0-E[x_0|\mathcal{I}_0^c])]$$
$$+\sum_{t=1}^{T-1} E[(x_t - E[x_t|\mathcal{I}_t^c])'(Q + A'K_{t+1}A - K_t)(x_t - E[x_t|\mathcal{I}_t^c])]$$
$$+\sum_{t=0}^{T-1} E[w_t'K_{t+1}w_t]. \tag{10.26}$$

◇

Proof. See Sect. 10.8.9. □

We have thus established the solution to the optimal control problem. We address the optimal quantization problem in the following subsection.

10.6.3 Existence of Optimal Quantization Policies

Now that we have separated the costs due to control and quantization, under any such composite policy and $T \in \mathbb{N}$, we can define a cost to be minimized by a composite quantizer policy as

$$J(\Pi^{comp},T) = E_{\nu_0}^{\Pi^{comp}}[\frac{1}{T}(x_0'K_0x_0 + \sum_{t=0}^{T-1} c_t(\pi_t, Q_t))],$$

where

$$c_t(\pi_t, Q_t) = \sum_{i\in\mathcal{M}} \inf_{\gamma_t^0(i)} \int_{\mathbb{R}^n} 1_{\{q_t=i\}}\pi_t(d\bar{x})(\bar{x}_t - \gamma_t^0(i))'P_t(\bar{x}_t - \gamma_t^0(i)),$$

where now $\underline{\gamma}^0 = \{\gamma_t^0, t \geq 0\}$ denotes a receiver policy and $P_t = (Q + A'K_{t+1}A - K_t)$, by (10.59) and $P_0 = Q + A'K_1A$.

We note that here the process $\bar{x}_t$ is the control-free process given by $\bar{x}_{t+1} = A\bar{x}_t + w_t$.

Therefore, we consider the setting where in (10.22), $u_t = 0$ and the quantizer is designed for this system. We note that as a result of the decoupling from the control actions by the predictive quantization policy (see Definition 10.6.1), the separation results presented in Theorem 10.3.4 directly apply in this context.

In the analysis, we will restrict the quantizers to have convex codecells (see Assumption 10.4.2). As elaborated in Chap. 4, a quantizer can be characterized as a stochastic kernel Q from $\mathbb{X}$ to $\{1,\dots,M\}$) defined by

$$Q(i|x) = 1_{\{x \in B_i\}}, \quad i = 1,\dots,M.$$

We endow the quantizers by a topology induced by such a stochastic kernel interpretation. In view of the results of Sect. 4.4.1, we have the following.

Let $\pi_t(\cdot) = P(x_t \in \cdot | q_{[0,t-1]})$. Recall that the properties of conditional probability lead to the filtering expression in (D.4).

Hence, with $\mathcal{P}(\mathbb{R}^n)$ denoting the set of probability measures on $\mathcal{B}(\mathbb{R}^n)$ under weak convergence, the conditional density process and the quantization process $(\pi_t(x), Q_t)$ form a joint Markov process in $\mathcal{P}(\mathbb{R}^n) \times \mathcal{Q}_c(M)$, as in Theorem 10.4.1.

We have the following result on the existence of optimal quantizers for the finite horizon setting.

Theorem 10.6.4. *For any $T \geq 1$ and arbitrary initial condition π_0, there exists a policy in Π_W^C such that*

$$\inf_{\Pi^{comp} \in \Pi_W^C} J_{\pi_0}(\Pi^{comp}, T) \tag{10.27}$$

is achieved. Letting $J_T^T(\cdot) = 0$ and

$$J_0^T(\pi_0) := \min_{\Pi^{comp} \in \Pi_W^C, \gamma^0} J_{\pi_0}(\Pi^{comp}, \gamma^0, T),$$

the dynamic programming recursion

$$TJ_t^T(\pi_t) = \min_{Q \in \mathcal{Q}_c(M)} \Big(c(\pi_t, Q_t) + TE[J_{t+1}^T(\pi_{t+1}) | \pi_t, Q_t] \Big) \tag{10.28}$$

holds for all $t = 0, 1, \dots, T-1$. ◇

Proof. See Sect. 10.8.10. □

Note that the optimal control policy is linear in the conditional estimate and is given in Theorem 10.6.2.

10.6.4 Partially Observed Case

We consider now the setup in (10.22) and Fig. 10.3, with a partially observed state, that is, with

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \\ y_t &= Cx_t + v_t, \end{aligned} \tag{10.29}$$

where, different from (10.22), here $y_t \in \mathbb{R}^m$, v_t is Gaussian and C a matrix.

Define $\bar{m}_t := E[x_t|y_{[0,t]}]$, which is computed through a Kalman filter. With

$$\Sigma_{0|-1} = E[x_0 x_0']$$

and for $t \geq 0$,

$$\begin{aligned}\Sigma_{t+1|t} &= A\Sigma_{t|t-1}A' + W \\ &\quad -(A\Sigma_{t|t-1}C')(C\Sigma_{t|t-1}C' + V)^{-1}(C\Sigma_{t|t-1}A'),\end{aligned}$$

the following recursion holds for $t \geq 0$ and with $\tilde{m}_{-1} = 0$:

$$\begin{aligned}\bar{m}_t &= A\bar{m}_{t-1} + Bu_{t-1} \\ &+\Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(CA(x_{t-1} - \tilde{m}_{t-1}) + v_t).\end{aligned}$$

Now, note that the cost

$$\inf_{\Pi^{comp}} \inf_{\gamma} J(\Pi^{comp}, \gamma, T) \tag{10.30}$$

with

$$J(\Pi^{comp}, \gamma, T) = \frac{1}{T} E_{\nu_0}^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} x_t' Q x_t + u_t' R u_t]$$

can be written equivalently as

$$\begin{aligned}J(\Pi^{comp}, \gamma, T) &= \frac{1}{T} E_{\nu_0}^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} \bar{m}_t' Q \bar{m}_t + u_t' R u_t] \\ &\quad + \frac{1}{T} E_{\nu_0}[\sum_{t=0}^{T-1} (x_t - \bar{m}_t)' Q (x_t - \bar{m}_t)]\end{aligned}$$

since the quadratic error $(x_t - \bar{m}_t)'Q(x_t - \bar{m}_t)$ is independent of the coding or the control policy.

Thus, we have that the processes $(\bar{m}_t, \Sigma_{t+1|t})$ and u_t form a controlled Markov chain and we can invoke Theorem 10.3.6: Any causal quantizer policy can, without any loss, be replaced with one in Π_W (where the state is now $(\bar{m}_t, \Sigma_{t+1|t})$ instead of x_t) as a consequence of Theorem 10.3.6. Furthermore, any quantizer in Π_W can be replaced without any loss with a predictive quantizer with the new state $\bar{m}_t$, as a consequence of Lemma 10.6.1 applied to the new state with identical arguments: Observe that the past control actions do not affect the evolution of $\Sigma_{t+1|t}$.

We have the following result.

Theorem 10.6.5. *For the minimization problem (10.30), the optimal control policy is given by* $u_t = L_t E[x_t|q_{[0,t]}]$, *where*

$$L_t = -(R + B'P_{t+1}B)^{-1}B'K_{t+1}A,$$

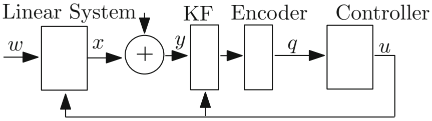

Fig. 10.5 Separation of estimation and quantization: When the source is Gaussian, generated by the linear system (10.29), the cost is quadratic, and the observation channel is Gaussian, the separated structure of the encoder above (with a predictive encoder) is optimal, where KF denotes the Kalman filter

and

$$P_t = A_t' K_{t+1} B(R + B' K_{t+1} B)^{-1} B' K_{t+1} A,$$
$$K_t = A_t' K_{t+1} A_t - P_t + Q,$$

with $K_T = P_{T-1} = 0.$ ◇

Given the optimal control policy, the following result is obtained.

Theorem 10.6.6. *For the minimization problem (10.30), the optimal cost it given by* $\frac{1}{T} J_0(\Pi^{comp}, T)$, *where*

$$J_0(\Pi^{comp},T)=E[x_0' K_0 x_0]+E[(x_0-E[x_0|\mathcal{I}_0^c])'(Q+A'K_1A)(x_0-E[x_0|\mathcal{I}_0^c])]$$
$$+\sum_{t=1}^{T-1} E[(x_t - E[x_t|\mathcal{I}_t^c])'(Q + A'K_{t+1}A - K_t)(x_t - E[x_t|\mathcal{I}_t^c])]$$
$$+\sum_{t=0}^{T-1} E[(x_t - \bar{m}_t)' Q(x_t - \bar{m}_t) + w_t' K_{t+1} w_t]. \quad (10.31)$$

◇

Now, that the cost has been separated; the following is a result of Theorem 10.6.1.

Theorem 10.6.7. *For the minimization of the cost in (10.30), any composite quantization policy can be replaced, without any loss in performance, by one which only uses the output of the Kalman filter and the information available at the receiver.* ◇

Thus, the optimality of Kalman filtering allows the encoder to only use the conditional estimate and the error covariance matrix without any loss of optimality (see Fig. 10.5), and the optimal quantization problem also has an explicit formulation.

10.7 Case with Noisy Channels and Noiseless Feedback

The results and the general program presented in this chapter apply also to coding over discrete memoryless (noisy) channels (DMCs) with feedback. In this context, consider the setup in Sect. 5.2.2, with one encoder and with $y_t = x_t$ and with the channel being a DMC. The equivalent results of Theorems 10.3.1 and 10.3.2 apply with q'_t terms replacing q_t, if q'_t is the output of a DMC at time t, as we state in the following.

In this context, let again $\mathcal{P}(\mathbb{X})$ denote the space of probability measures on $\mathcal{B}(\mathbb{X})$ under the topology of weak convergence and define $\pi_t \in \mathcal{P}(\mathbb{X})$ to be the regular conditional probability measure given by $\pi_t(\cdot) = P(x_t \in \cdot | q'_{[0,t-1]})$, where q'_t is the channel output when the input is q_t. That is, $\pi_t(A) = P(x_t \in A | q'_{[0,t-1]}), \quad A \in \mathcal{B}(\mathbb{X})$. The goal is the minimization

$$\inf_{\Pi^{comp}} \inf_{\underline{\gamma}^0} E_{\nu_0}^{\mathbf{\Pi}^{comp},\underline{\gamma}^0}[\sum_{t=0}^{T-1} c(x_t, u_t)], \tag{10.32}$$

with initial condition distribution ν_0. Here $c(\cdot,\cdot)$, is a nonnegative, measurable function and $u_t = \gamma_t^0(q'_{[0,t]})$. We state the following.

Theorem 10.7.1. *Any composite encoding policy can be replaced, without any loss in performance, by one which only uses x_t and $q'_{[0,t-1]}$ at time $t \geq 1$ to generate the channel input q_t.* ◇

Theorem 10.7.2. *Any composite quantization policy can be replaced, without any loss in performance, by one which only uses the conditional probability measure $\pi_t(\cdot) = P(x_t \in \cdot | q'_{[0,t-1]})$, the state x_t, and the time information t, at time $t \geq 1$ to generate the channel input q_t.* ◇

The proof of these results follow from those of Theorems 10.3.1 and 10.3.2 with almost identical steps with q_t being replaced with q'_t in the information available at the receiver and the encoder.

Likewise, for a partially observed setup, extensions of Theorems 10.3.3 and 10.3.4 also apply to this case.

Remark 10.7.1. When there is no feedback from the controller or when there is noisy feedback, the analysis requires a Markov chain construction in a larger state space provided memory restrictions are imposed on the decoders. We refer the reader to Teneketzis [361] and Mahajan and Teneketzis [248, 249] for a class of such settings. ◇

10.8 Appendix: Proofs

10.8.1 Proof of Theorem 10.3.1

At time $t = T-1$, the per-stage cost function can be written as follows, where γ_t^0 denotes a fixed receiver policy:

$$E[c(x_t, \gamma_t^0(q_{[0,t]}))|q_{[0,t-1]}] = E[F(x_t, q_{[0,t-1]}, q_t)|q_{[0,t-1]}]$$

where, $F(x_t, q_{[0,t-1]}, q_t) = c(x_t, \gamma_t^0(q_{[0,t]}))$.

This is equivalent to, by the smoothing property of conditional expectation, the following:

$$E\left[E[F(x_t, q_{[0,t-1]}, q_t)|x_t, q_{[0,t-1]}]\middle|q_{[0,t-1]}\right].$$

Now, we will apply Witsenhausen's two-stage lemma [396], to show that we can obtain a lower bound for the double expectation by picking q_t as a result of a measurable function of $x_t, q_{[0,t-1]}$. Thus, we will find a composite quantization policy which only uses $(x_t, q_{[0,t-1]})$ which performs as well as one which uses the entire memory available at the encoder. To make this precise, let us fix the decision function γ_t^0 at the receiver corresponding to a given composite quantization policy at the encoder Q_t^{comp}, let $t = T - 1$, and define for every $k \in \mathcal{M}_t$:

$$\beta_k := \left\{x_t, q_{[0,t-1]} : F(x_t, q_{[0,t-1]}, k) \leq F(x_t, q_{[0,t-1]}, q'), \forall q' \neq k, q' \in \mathcal{M}_t\right\}.$$

These sets are Borel, by the measurability of F on $\mathbb{X}$. Such a construction covers the domain set consisting of $(x_t, q_{[0,t-1]})$ but with overlaps. It covers the elements in $\mathbb{X} \times \prod_{t=0}^{T-2} \mathcal{M}_t$, since for every element in this product set, there is a minimizing $k \in \mathcal{M}_t$ ($\mathcal{M}_t$ is finite). To avoid the overlaps, we adopt the following technique which was introduced in Witsenhausen [396]. Let there be an ordering of the elements in $\mathcal{M}_t$ as $1, 2, \ldots, |\mathcal{M}_t|$, and for $k \geq 1$ in this sequence define a function $Q_t^{comp,*}$ as

$$q_t = Q_t^{comp,*}(x_t, q_{[0,t-1]}) = k, \quad \text{if}(x_t, q_{[0,t-1]}) \in \beta_k - \cup_{i=1}^{k-1} \beta_i,$$

with $\beta_0 = \emptyset$. Thus, for any random variable q_t appropriately defined on the probability space,

$$\begin{aligned} &E\left[E[F(x_t, q_{[0,t-1]}, q_t)|x_t, q_{[0,t-1]}]\middle|q_{[0,t-1]}\right] \\ &\qquad \geq E\left[E[F(x_t, q_{[0,t-1]}, Q_t^{comp,*}(x_t, q_{[0,t-1]}))|x_t, q_{[0,t-1]}]\middle|q_{[0,t-1]}\right]. \end{aligned} \tag{10.33}$$

Thus, the new composite policy performs at least as well as the original composite coding policy even though it has a restricted structure.

As such, if there is an optimal policy, it can be replaced with one which uses only $\{x_t, q_{[0,t-1]}\}$ without any loss of performance while keeping the receiver decision function γ_t^0 fixed.

We have thus obtained the structure of the encoder for the last stage. We iteratively proceed to study the other time stages. In particular, since $\{x_t\}$ is Markov, we could proceed as follows (in essence using Witsenhausen's three-stage lemma [396]): For a three-stage cost problem, the cost at time $t = 2$ can be written as, for measurable functions c_2, c_3:

$$E\left[c_2(x_2, \gamma_2^0(q_1, q_2), q_1, q_2) \right.$$
$$\left. +E[c_3(x_3, \gamma_3^0(q_1, q_2, Q_3^{comp,*}(x_3, q_2, q_1))|x_3, q_2, x_2, q_1, x_1)] \middle| x_2, x_1, q_2, q_1 \right].$$

Since

$$P(dx_3, q_2, q_1|x_2, x_1, q_2, q_1) = P(dx_3, q_2, q_1|x_2, q_2, q_1)$$

and since under $Q_3^{comp,*}$, q_3 is a function of x_3 and q_1, q_2, the expression above is equal to, for some measurable $F_2(.)$, $F_2(x_2, q_2, q_1)$. By a similar argument as above, a composite quantization policy at time 2 which uses x_2 and q_1 and which performs at least as good as the original policy can be constructed. By similar arguments, an encoder at time t, $1 \leq t \leq T-1$ only uses $(x_t, q_{[0,t-1]})$ can be constructed. The encoder at time $t = 0$ uses x_0, where $x_0 = \nu_0$ is the prior distribution on the initial state.

Now that we have obtained the restricted structure for a composite quantization policy which is without any loss, we can express this as

$$Q_t^{comp}(x_t, q_{[0,t-1]}) = Q^{q_{[0,t-1]}}(x_t), \quad \forall x_t, q_{[0,t-1]}$$

such that the quantizer action $Q^{q_{[0,t-1]}} \in \mathbb{Q}(\mathbb{X}; \mathcal{M}_t)$ is generated using only $q_{[0,t-1]}$ and the quantizer outcome is generated by evaluating $Q^{q_{[0,t-1]}}(x_t)$ for every x_t. □

10.8.2 Proof of Theorem 10.3.2

At time $t = T-1$, the per-stage cost function can be written as

$$E[c(x_t, v_t(q_{[0,t]}))|q_{[0,t-1]}] = E[\int_{\mathbb{X}} P(dx_t|q_{[0,t-1]}, q_t)c(x_t, v_t(q_{[0,t-1]}, q_t))].$$

Thus, at time $t = T-1$, an optimal receiver (which is deterministic without any loss of optimality) will use $P(dx_t|q_{[0,t]})$ as a sufficient statistic for an optimal decision (or any receiver can be replaced with one which uses this sufficient statistic without any loss). Let us fix a receiver policy which only uses the posterior $P(dx_t|q_{[0,t]})$ as its sufficient statistic. Let us further note that

$$\begin{aligned} P(dx_t|q_{[0,t]}) &= \frac{P(q_t, dx_t|q_{[0,t-1]})}{\int_{x_t} P(q_t, dx_t|q_{[0,t-1]})} \\ &= \frac{P(q_t|x_t, q_{[0,t-1]})P(dx_t|q_{[0,t-1]})}{\int_{x_t} P(q_t|x_t, q_{[0,t-1]})P(dx_t|q_{[0,t-1]})}. \end{aligned} \tag{10.34}$$

The term $P(q_t|x_t, q_{[0,t-1]})$ is determined by the quantizer action Q_t (this follows from Theorem 10.3.1). Furthermore, given Q_t, the relation (10.34) is measurable on $\mathcal{P}(\mathbb{X})$ (i.e., in $\pi_t(\cdot) = P(x_t \in \cdot|q_{[0,t-1]})$) under weak convergence.

To prove this technical argument, consider the numerator in (10.34) and note that the function $\kappa_B : \mathcal{P}(\mathbb{X}) \to \mathbb{R}$ defined as $\kappa_B(\pi) = \pi(B)$ is measurable under weak convergence topology as a consequence of Theorem B.2.1, for each $B \in \mathcal{B}(\mathbb{X})$. By Theorem B.2.2, this implies that the relation in (10.34) is measurable on $\mathcal{P}(\mathbb{X})$.

Let us denote the quantizer applied, given the past realizations of quantizer outputs as $Q_t^{q_{[0,t-1]}}$. Note that q_t is deterministically determined by (x_t, Q_t) and the optimal receiver function can be expressed as $\gamma_t^0(P(dx_t|q_{[0,t-1]}), q_t)$, given $Q_t^{q_{[0,t-1]}}$. The cost at time $t = T-1$ can be expressed, given the quantizer $Q_t^{q_{[0,t-1]}}$, for some Borel function G, as $G(P(dx_t|q_{[0,t-1]}), Q_t^{q_{[0,t-1]}})$, where

$$G(P(dx_t|q_{[0,t-1]}), Q_t^{q_{[0,t-1]}})$$

$$= \int_{\mathbb{X}} P(dx_t|q_{[0,t-1]}) \sum_{\mathcal{M}_t} \left(1_{\{q_t = Q_t^{q_{[0,t-1]}}(x_t)\}} \eta^{Q^{q_{[0,t-1]}}}(P(dx_t|q_{[0,t-1]}), q_t)) \right),$$

with $\eta^{Q^{q_{[0,t-1]}}}(P(dx_t|q_{[0,t-1]}), q_t)) = c(x_t, \gamma_t^0(P(dx_t|q_{[0,t-1]}), q_t))$.

Now, one can construct an equivalence class among the past $q_{[0,t-1]}$ sequences, which induce the same π_t, and can replace the quantizers in this class with one, which induces a lower cost among the finitely many elements in each class for the final time stage. An optimal quantization output thus may be generated using $\pi_t(\cdot) = P(x_t \in \cdot|q_{[0,t-1]})$ and x_t, by extending Witsenhausen's argument used earlier in the proof of Theorem 10.3.1 for the terminal time stage. Since there are only finitely many past sequences and finitely many π_t, this leads to a Borel measurable selection of x_t for every π_t, leading to a quantizer and a measurable selection in π_t, x_t. Hence, the final stage cost can be expressed as $F_t(\pi_t)$ for some F_t, without any performance loss.

The same argument applies for all time stages: At time $t = T-2$, the sufficient statistic both for the immediate cost and the cost-to-go is $P(dx_{t-1}|q_{[0,t-1]})$, and thus for the cost impacting the time stage $t = T-1$ as a result of the optimality result for Q_{T-1}. To show that the separation result generalizes to all time stages, it suffices to prove that $\{(\pi_t, Q_t)\}$ has a controlled Markov chain form, if the encoders use the structure above.

Now, for $t \geq 1$, for all $B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$,

$$P\left(P(dx_t|q_{[0,t-1]}) \in B \,\middle|\, P(dx_s|q_{[0,s-1]}), Q_s, s \leq t-1 \right)$$

$$= P\left(\int_{x_{t-1}} P(dx_t, dx_{t-1}|q_{[0,t-1]}) \in B \,\middle|\, P(dx_s|q_{[0,s-1]}), Q_s, s \leq t-1 \right)$$

$$= P\Bigg(\bigg\{\frac{\int_{x_{t-1}} P(dx_t|x_{t-1})P(q_{t-1}|x_{t-1},q_{[0,t-2]})P(dx_{t-1}|q_{[0,t-2]})}{\int_{x_{t-1},x_t} P(dx_t|x_{t-1})P(q_{t-1}|x_{t-1},q_{[0,t-2]})P(dx_{t-1}|q_{[0,t-2]})}\bigg\} \in B$$

$$\bigg| P(dx_s|q_{[0,s-1]}), Q_s, s \leq t-1\Bigg)$$

$$=P\Bigg(\bigg\{\frac{\int_{x_{t-1}} P(dx_t|x_{t-1})P(q_{t-1}|x_{t-1},q_{[0,t-2]})P(dx_{t-1}|q_{[0,t-2]})}{\int_{x_{t-1},x_t} P(dx_t|x_{t-1})P(q_{t-1}|x_{t-1},q_{[0,t-2]})P(dx_{t-1}|q_{[0,t-2]})}\bigg\} \in B$$

$$\bigg| P(dx_{t-1}|q_{[0,t-2]}), Q_{t-1}\Bigg) \quad (10.35)$$

$$=P\bigg(\int_{x_{t-1}} P(dx_t, dx_{t-1}|q_{[0,t-1]}) \in B\bigg| P(dx_{t-1}|q_{[0,t-2]}), Q_{t-1}\bigg)$$

$$= P\bigg(P(dx_t|q_{[0,t-1]}) \in B\bigg| P(dx_{t-1}|q_{[0,t-2]}), Q_{t-1}\bigg). \quad (10.36)$$

In the above derivation, (10.35) uses the fact that the term $P(q_{t-1}|x_{t-1}, q_{[0,t-2]})$ is uniquely identified by $P(dx_{t-1}|q_{[0,t-2]})$ and Q_{t-1}, which in turn is uniquely identified by $q_{[0,t-2]}$ and Q_{t-1}. Furthermore, (10.36) defines a regular conditional probability measure since for all $B \in \mathcal{B}(\mathbb{X})$,

$$\pi_t(B) = P(x_t \in B|q_{[0,t-1]})$$
$$= \int_{x_{t-1}} P(x_t \in B, dx_{t-1}|q_{[0,t-1]})$$
$$= \int_{x_{t-1}} P(x_t \in B|x_{t-1})P(dx_{t-1}|q_{[0,t-1]})$$

is measurable in π_{t-1}, given Q_{t-1} (as a consequence of the measurability of (10.34) in π_t). As a consequence the conditional probability $\pi_t(B)$, $B \in \mathcal{B}(\mathbb{X})$, is a measurable function of π_{t-1}, given Q_{t-1}. By Theorem B.2.2, we conclude that for any measurable function F_t of $P(dx_t|q_{[0,t-1]})$

$$E[F_t(P(dx_t|q_{[0,t-1]}))|P(dx_s|q_{[0,s-1]}), Q_s, s \leq t-1]$$
$$= E[F_t(P(dx_t|q_{[0,t-1]}), Q_t)|P(dx_{t-1}|q_{[0,t-2]}), Q_{t-1}], \quad (10.37)$$

for every given Q_{t-1}. Once again an equivalence relationship between the finitely many past quantizer outputs, based on the equivalence of the conditional measures $P(dx_{t-1}|q_{[0,t-2]})$ they induce, can be constructed and as a consequence the conditional probability measure π_t is measurable in $\{P(dx_{t-1}|q_{[0,t-2]}), Q_{t-1}\}$, given Q_{t-1}. With the controlled Markov structure, we can essentially follow the same argument for earlier time stages. As such, it suffices that the encoder uses only $(P(dx_t|q_{[0,t-1]}),t)$ as its sufficient statistic for all time stages, to generate the optimal quantizer. An optimal quantizer uses x_t to generate the optimal quantization outputs. □

10.8.3 Proof of Theorem 10.3.3

We transform the problem into a real-time coding problem involving a fully observed Markov source. At time $t = T - 1$, the per-stage cost function can be written as follows, where γ_t^0 denotes a fixed receiver policy:

$$
\begin{aligned}
&E[c(x_t,\gamma_t^0(q_{[0,t]}))|q_{[0,t-1]}]\\
&=\sum_{\mathcal{M}_t} P(q_t=k|q_{[0,t-1]})(\int_{\mathbb{X}} P(dx_t|q_{[0,t-1]},k)c(x_t,\gamma_t^0(q_{[0,t-1]},k)))\\
&=\int_{\mathbb{X}}\sum_{\mathcal{M}_t} P(dx_t,q_t=k|q_{[0,t-1]})c(x_t,\gamma_t^0(q_{[0,t-1]},k))\\
&=\int_{\mathcal{P}(\mathbb{X})}\int_{\mathbb{X}}\sum_{\mathcal{M}_t} P(dx_t,q_t{=}k,d\tilde{\pi}_t|q_{[0,t-1]})c(x_t,\gamma_t^0(q_{[0,t-1]},k))\\
&=\int_{\mathcal{P}(\mathbb{X})}\int_{\mathbb{X}}\sum_{\mathcal{M}_t} P(d\tilde{\pi}_t|q_{[0,t-1]})P(dx_t|\tilde{\pi}_t)P(q_t{=}k|\tilde{\pi}_t,q_{[0,t-1]})c(x_t,\gamma_t^0(q_{[0,t-1]},k))\\
&=\int_{\mathcal{P}(\mathbb{X})}\sum_{\mathcal{M}_t}P(d\tilde{\pi}_t|q_{[0,t-1]})P(q_t{=}k|\tilde{\pi}_t,q_{[0,t-1]})\int_{\mathbb{X}}P(dx_t|\tilde{\pi}_t)c(x_t,\gamma_t^0(q_{[0,t-1]},k))\\
&=E[F(\tilde{\pi}_t,q_{[0,t-1]},q_t)|q_{[0,t-1]}],
\end{aligned}
\tag{10.38}
$$

where $\tilde{\pi}_t(\cdot){=}P(x_t\in\cdot|y_{[0,t]})$ and $F(\tilde{\pi}_t,q_{[0,t-1]},q_t){=}\int_{\mathbb{X}}\tilde{\pi}_t(dx)c(x,\gamma_t^0(q_{[0,t-1]},q_t))$.

In the above derivation, the fourth equality follows from the property that

$$x_t \leftrightarrow P(dx_t|y_{[0,t]}) \leftrightarrow q_{[0,t]}.$$

We note that $F(\cdot,\gamma_t^0(q_{[0,t-1]},q_t))$ is measurable by Theorem B.2.1 and the fact that the cost is bounded.

As in the proof of Theorem 10.3.1, one may define q_t as a random variable on the probability space such that the joint distribution of $(q_t,\tilde{\pi}_t,q_{[0,t-1]})$ matches the characterization that $q_t = Q_t^{comp}(y_{[0,t]},q_{[0,t-1]})$, since

$$P(q_t|\tilde{\pi}_t,q_{[0,t-1]}) = \int_{\mathbb{Y}^{t+1}} P(q_t|y_{[0,t]},q_{[0,t-1]}))P(y_{[0,t]}|\tilde{\pi}_t,q_{[0,t-1]}).$$

The cost at the final stage is thus written as $E[F(\tilde{\pi}_t, q_{[0,t-1]}, q_t)|q_{[0,t-1]}]$, which is equivalent to, by the smoothing property of conditional expectation, to the following:

$$E\left[E[F(\tilde{\pi}_t, q_{[0,t-1]}, q_t)|\tilde{\pi}_t, q_{[0,t-1]}]\middle| q_{[0,t-1]}\right].$$

Now, we will apply Witsenhausen's two-stage lemma [396], to show that we can obtain a lower bound for the double expectation by picking q_t to be a measurable function of $\tilde{\pi}_t, q_{[0,t-1]}$. Thus, we will find a composite quantization policy which only uses $(\tilde{\pi}_t, q_{[0,t-1]})$ which performs as well as one which uses the entire memory available at the encoder. Let us fix the receiver function γ_t^0 at the receiver corresponding to a given composite quantization policy at the encoder Q_t^{comp}, let $t = T-1$, and define for every $k \in \mathcal{M}_t$:

$$\beta_k := \left\{\tilde{\pi}_t, q_{[0,t-1]} : F(\tilde{\pi}_t, q_{[0,t-1]}, k) \leq F(\tilde{\pi}_t, q_{[0,t-1]}, q'), \forall q' \neq k, q' \in \mathcal{M}_t\right\}.$$

These sets are Borel, by the measurability of F on $\mathcal{P}(\mathbb{X})$. Such a construction covers the domain set consisting of $(\tilde{\pi}_t, q_{[0,t-1]})$ but with overlaps. It covers the elements in $\mathcal{P}(\mathbb{X}) \times \prod_{t=0}^{T-2} \mathcal{M}_t$, since for every element in this product set, there is a minimizing $k \in \mathcal{M}_t$ ($\mathcal{M}_t$ is finite). To avoid the overlaps, we adopt the following technique which was introduced in Witsenhausen [396]. Let there be an ordering of the elements in $\mathcal{M}_t$ as $1, 2, \ldots, |\mathcal{M}_t|$, and for $k \geq 1$ in this sequence define a function $Q_t^{comp,*}$ as

$$q_t = Q_t^{comp,*}(\tilde{\pi}_t, q_{[0,t-1]}) = k, \quad \text{if} (\tilde{\pi}_t, q_{[0,t-1]}) \in \beta_k - \cup_{i=1}^{k-1} \beta_i,$$

with $\beta_0 = \emptyset$. Thus, for any random variable q_t appropriately defined on the probability space,

$$\begin{aligned} &E\left[E[F(\tilde{\pi}_t, q_{[0,t-1]}, q_t)|\tilde{\pi}_t, q_{[0,t-1]}]\middle| q_{[0,t-1]}\right] \\ &\quad \geq E\left[E[F(\tilde{\pi}_t, q_{[0,t-1]}, Q_t^{comp,*}(\tilde{\pi}_t, q_{[0,t-1]}))|\tilde{\pi}_t, q_{[0,t-1]}]\middle| q_{[0,t-1]}\right]. \end{aligned} \tag{10.39}$$

Thus, the new composite policy performs at least as well as the original composite coding policy even though it has a restricted structure.

As such, if there is an optimal policy, it can be replaced with one which uses only $\{\tilde{\pi}_t, q_{[0,t-1]}\}$ without any loss of performance, while keeping the receiver decision function γ_t^0 fixed. It should now be noted that $\{\tilde{\pi}_t\}$ is a Markov process. Further note that

$$P(dx_t|dy_{[0,t]}) = \frac{\int_{x_{t-1}} P(dy_t|x_t)P(dx_t|x_{t-1})P(dx_{t-1}|dy_{[0,t-1]})}{\int_{x_{t-1},x_t} P(dy_t|x_t)P(dx_t|x_{t-1})P(dx_{t-1}|dy_{[0,t-1]})}$$

and that $P(dy_t|\tilde{\pi}_s, s \le t-1) = \int_{x_t} P(dy_t, dx_t|\tilde{\pi}_s, s \le t-1) = P(dy_t|\tilde{\pi}_{t-1})$. These imply that the following is a Markov kernel:

$$P(d\tilde{\pi}_t|\tilde{\pi}_s, s \le t-1) = P(d\tilde{\pi}_t|\tilde{\pi}_{t-1}). \tag{10.40}$$

We have thus obtained the structure of the encoder for the last stage. We iteratively proceed to study the other time stages. In particular, since $\{\tilde{\pi}_t\}$ is Markov, we could proceed as follows (in essence using Witsenhausen's three-stage lemma [396]): For a three-stage cost problem, the cost at time $t = 2$ can be written as, for measurable functions c_2, c_3,

$$E\Bigg[c_2(\tilde{\pi}_2, \gamma_2^0(q_1,q_2), q_1, q_2) \\ + E[c_3(\tilde{\pi}_3, \gamma_3^0(q_1,q_2,Q_3^{comp}(\tilde{\pi}_3,q_2,q_1))|\tilde{\pi}_3,q_2,\tilde{\pi}_2,q_1,\tilde{\pi}_1)]\Bigg|\tilde{\pi}_2,\tilde{\pi}_1,q_2,q_1\Bigg].$$

Since

$$P(d\tilde{\pi}_3, q_2, q_1|\tilde{\pi}_2, \tilde{\pi}_1, q_2, q_1) = P(d\tilde{\pi}_3, q_2, q_1|\tilde{\pi}_2, q_2, q_1)$$

and since under $Q_3^{comp,*}$, q_3 is a function of $\tilde{\pi}_3$ and q_1, q_2, the expectation above is equal to, for some measurable $F_2(.)$, $E[F_2(\tilde{\pi}_2, q_2, q_1)]$. Measurability follows since

$$E\Bigg[c_2(\tilde{\pi}_2, \gamma_2^0(q_1,q_2), q_1, q_2) \\ + E[c_3(\tilde{\pi}_3, \gamma_3^0(q_1,q_2,Q_3^{comp,*}(\tilde{\pi}_3,q_2,q_1))|\tilde{\pi}_3,q_2,\tilde{\pi}_2,q_1,\tilde{\pi}_1)]\Bigg|\tilde{\pi}_2,\tilde{\pi}_1,q_2,q_1\Bigg]$$

is measurable. Thus, a composite quantization policy at time 2 which uses $\tilde{\pi}_2$ and q_1 and which is without any loss in comparison with the original policy can be constructed.

By a similar argument, an optimal encoder at time t, $1 \le t \le T-1$ only uses $(\tilde{\pi}_t, q_{[0,t-1]})$. The encoder at time $t = 0$ uses $\tilde{\pi}_0$, where $\tilde{\pi}_0 = \nu_0$ is the prior distribution on the initial state.

Now that we have obtained the restricted structure for a composite quantization policy which is without any loss, we can express this as

$$Q_t^{comp}(\tilde{\pi}_t, q_{[0,t-1]}) = Q^{q_{[0,t-1]}}(\tilde{\pi}_t), \quad \forall \tilde{\pi}_t, q_{[0,t-1]}$$

such that the quantizer action $Q^{q_{[0,t-1]}} \in \mathbb{Q}(\mathcal{P}(\mathbb{X}); \mathcal{M}_t)$ is generated using only $q_{[0,t-1]}$ and the quantizer outcome is generated by evaluating $Q^{q_{[0,t-1]}}(\tilde{\pi}_t)$ for every $\tilde{\pi}_t$. □

10.8.4 Proof of Theorem 10.3.4

At time $t = T - 1$, an optimal receiver will use $P(dx_t|q_{[0,t]})$ as a sufficient statistic for an optimal decision (or any receiver can be replaced with one which uses this sufficient statistic without any loss). As in the proof of Theorem 10.3.2, let us fix a receiver policy which only uses the posterior $P(dx_t|q_{[0,t]})$ as its sufficient statistic. We now note that

$$P(dx_t|q_{[0,t]}) = \int_{\tilde{\pi}_t} P(dx_t|\tilde{\pi}_t)P(d\tilde{\pi}_t|q_{[0,t]}). \tag{10.41}$$

Let us note that

$$\begin{aligned} P(d\tilde{\pi}_t|q_{[0,t]}) &= \frac{P(q_t, d\tilde{\pi}_t|q_{[0,t-1]})}{\int_{\tilde{\pi}_t} P(q_t, d\tilde{\pi}_t|q_{[0,t-1]})} \\ &= \frac{P(q_t|\tilde{\pi}_t, q_{[0,t-1]})P(d\tilde{\pi}_t|q_{[0,t-1]})}{\int_{\tilde{\pi}_t} P(q_t|\tilde{\pi}_t, q_{[0,t-1]})P(d\tilde{\pi}_t|q_{[0,t-1]})}. \end{aligned} \tag{10.42}$$

The term $P(q_t|\tilde{\pi}_t, q_{[0,t-1]})$ is determined by the quantizer action Q_t (this follows from Theorem 10.3.3). Furthermore, given Q_t, the relation (10.42) is measurable on $\mathcal{P}(\mathcal{P}(\mathbb{X}))$ (i.e., in $\Xi_t(\cdot) = P(\tilde{\pi}_t \in \cdot|q_{[0,t-1]})$) under weak convergence.

This argument, as in the proof of Theorem 10.3.2, follows from the observation that in the numerator of (10.42) the function $\kappa_B : \mathcal{P}(\mathcal{P}(\mathbb{X})) \to \mathbb{R}$ defined as $\kappa_B(\Xi) = \Xi(B)$ is measurable under weak convergence topology as a consequence of Theorem B.2.1, for each $B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$. By Theorem B.2.2, this implies that the relation in (10.42) is measurable on $\mathcal{P}(\mathcal{P}(\mathbb{X}))$.

Let us denote the quantizer applied, given the past realizations of quantizer outputs as $Q_t^{q_{[0,t-1]}}$. Note that q_t is deterministically determined by $(\pi_t, Q_t^{q_{[0,t-1]}})$ and the optimal receiver function can be expressed as $\gamma_t^0(\Xi_t, q_t)$ (as a measurable function), given $Q_t^{q_{[0,t-1]}}$. The cost at time $t = T - 1$ can be expressed, given the quantizer $Q_t^{q_{[0,t-1]}}$, for some Borel function G, as $G(\Xi_t, Q_t^{q_{[0,t-1]}})$, where

$$\begin{aligned} &G(\Xi_t, Q_t^{q_{[0,t-1]}}) \\ &= \int_{\mathcal{P}(\mathbb{X})} \Xi_t(d\tilde{\pi}_t) \sum_{\mathcal{M}_t} 1_{\{q_t = Q_t^{q_{[0,t-1]}}(\pi_t)\}} \eta^{Q^{q_{[0,t-1]}}}(\Xi_t, q_t)), \end{aligned}$$

with $\eta^{Q^{q_{[0,t-1]}}}(\Xi_t, q_t)) = \int \tilde{\pi}_t(dx_t)c(x_t, \gamma_t^0(\Xi_t, q_t))$. As in the proof of Theorem 10.3.2, one can construct an equivalence class among the past $q_{[0,t-1]}$ sequences which induce the same Ξ_t and can replace the quantizers $Q_t^{q_{[0,t-1]}}$ for each class with one which induces a lower cost among the finitely many elements in each such class, for the final time stage. Thus, an optimal quantization output may be generated using $\Xi_t(\cdot) = P(\tilde{\pi}_t \in \cdot|q_{[0,t-1]})$ and $\tilde{\pi}_t$. Since there are only finitely many past sequences and finitely many Ξ_t, this leads to a Borel measurable selection of $\tilde{\pi}_t$ for every Ξ_t, leading to a quantizer and a measurable selection in $\Xi_t, \tilde{\pi}_t$.

Since such a selection for Q_t only uses Ξ_t, an optimal quantization output may be generated using $\Xi_t(\cdot) = P(\tilde{\pi}_t \in \cdot|q_{[0,t-1]})$ and $\tilde{\pi}_t$. Hence, $G(\Xi_t, Q_t^{q_{[0,t-1]}})$ can be replaced with $F_t(\Xi_t)$ for some F_t, without any performance loss.

The same argument applies for all time stages: At time $t = T - 2$, the sufficient statistic both for the immediate cost and the cost-to-go is $P(d\tilde{\pi}_{t-1}|q_{[0,t-1]})$, and thus for the cost impacting the time stage $t = T - 1$, as a result of the optimality result for Q_{T-1}. To show that the separation result generalizes to all time stages, it suffices to prove that $\{(\Xi_t, Q_t)\}$ is a controlled Markov chain, if the encoders use the structure above.

Now, for $t \geq 1$, for all $B \in \mathcal{B}(\mathcal{P}(\mathcal{P}(\mathbb{X})))$,

$$P\bigg(P(d\tilde{\pi}_t|q_{[0,t-1]}) \in B \bigg| P(d\tilde{\pi}_s|q_{[0,s-1]}), Q_s, s \leq t-1\bigg)$$

$$= P\bigg(\int_{\tilde{\pi}_{t-1}} P(d\tilde{\pi}_t, d\tilde{\pi}_{t-1}|q_{[0,t-1]}) \in B \bigg| P(d\tilde{\pi}_s|q_{[0,s-1]}), Q_s, s \leq t-1\bigg)$$

$$= P\bigg(\bigg\{\frac{\int_{\tilde{\pi}_{t-1}} P(d\tilde{\pi}_t|\tilde{\pi}_{t-1})P(q_{t-1}|\tilde{\pi}_{t-1}, q_{[0,t-2]})P(d\tilde{\pi}_{t-1}|q_{[0,t-2]})}{\int_{\tilde{\pi}_{t-1},\tilde{\pi}_t} P(d\tilde{\pi}_t|\tilde{\pi}_{t-1})P(q_{t-1}|\tilde{\pi}_{t-1}, q_{[0,t-2]})P(d\tilde{\pi}_{t-1}|q_{[0,t-2]})}\bigg\} \in B \bigg| P(d\tilde{\pi}_s|q_{[0,s-1]}), Q_s, s \leq t-1\bigg)$$

$$=P\bigg(\bigg\{\frac{\int_{\tilde{\pi}_{t-1}} P(d\tilde{\pi}_t|\tilde{\pi}_{t-1})P(q_{t-1}|\tilde{\pi}_{t-1}, q_{[0,t-2]})P(d\tilde{\pi}_{t-1}|q_{[0,t-2]})}{\int_{\tilde{\pi}_{t-1},\tilde{\pi}_t} P(d\tilde{\pi}_t|\tilde{\pi}_{t-1})P(q_{t-1}|\tilde{\pi}_{t-1}, q_{[0,t-2]})P(d\tilde{\pi}_{t-1}|q_{[0,t-2]})}\bigg\} \in B \bigg| P(d\tilde{\pi}_{t-1}|q_{[0,t-2]}), Q_{t-1}\bigg) \quad (10.43)$$

$$= P\bigg(\int_{\tilde{\pi}_{t-1}} P(d\tilde{\pi}_t, d\tilde{\pi}_{t-1}|q_{[0,t-1]}) \in B \bigg| P(d\tilde{\pi}_{t-1}|q_{[0,t-2]}), Q_{t-1}\bigg)$$

$$= P\bigg(P(d\tilde{\pi}_t|q_{[0,t-1]}) \in B \bigg| P(d\tilde{\pi}_{t-1}|q_{[0,t-2]}), Q_{t-1}\bigg). \quad (10.44)$$

Here, (10.43) uses the fact that $P(q_{t-1}|\tilde{\pi}_{t-1}, q_{[0,t-2]})$ is identified by $\{\Xi_{t-1}, Q_{t-1}\}$, which in turn is uniquely identified by $q_{[0,t-2]}$ and Q_{t-1}. Furthermore, the expression in (10.44) defines a regular conditional probability measure, since for all $B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$,

$$\begin{aligned}\Xi_t(B) &= P(\tilde{\pi}_t \in B | q_{[0,t-1]}) \\ &= \int_{\tilde{\pi}_{t-1}} P(\tilde{\pi}_t \in B, d\tilde{\pi}_{t-1} | q_{[0,t-1]}) \\ &\quad = \int_{\tilde{\pi}_{t-1}} P(\tilde{\pi}_t \in B | \tilde{\pi}_{t-1}) P(d\tilde{\pi}_{t-1} | q_{[0,t-1]})\end{aligned}$$

is measurable in Ξ_{t-1}, given Q_{t-1} (as a consequence of the measurability of (10.42) in Ξ_t). Hence, by Theorem B.2.2, we conclude that for any measurable function F_t of Ξ_t

$$E[F_t(\Xi_t)|\Xi_{[0,t-1]}, Q_{[0,t-1]}] = E[F_t(\Xi_t), Q_t)|\Xi_{t-1}, Q_{t-1}],$$

for every given Q_{t-1}. Now, once again an equivalence relationship between the finitely many past quantizer outputs, based on the equivalence of the conditional measures Ξ_{t-1} they induce, can be constructed. With the controlled Markov structure, we can follow the same argument for earlier time stages. Therefore, it suffices that the encoder uses only (Ξ_t, t) as its sufficient statistic for all time stages, to generate the optimal quantizer. An optimal quantizer uses $\tilde{\pi}_t$ to generate the optimal quantization outputs. □

10.8.5 Proof of Theorem 10.3.6

We note that the analysis in (10.38)–(10.33) apply identically for the case with control, by replacing a receiver policy with a fixed control policy. We can thus obtain the structure of the optimal encoder for the last stage. We iteratively proceed to study the other time stages. The only difference here is that, with control, $\{x_t\}$ is no longer Markov, but $\{x_t, u_t\}$ forms a controlled Markov chain. For a three-stage cost problem, the cost at time $t = 2$ can be written as, for measurable functions c_2, c_3,

$$c_2(x_2, u_2(q_1, q_2)) + E[c_3(x_3, u_3(q_1, q_2, Q_3^{comp}(x_3, q_2, q_1))|q_2, x_2, u_2, q_1, x_1)].$$

Since

$$P(dx_3, q_2, q_1|x_2, u_2, x_1, q_2, q_1) = P(dx_3, q_2, q_1|x_2, u_2, q_2, q_1),$$

and u_2 is a function of q_1, q_2 (with the control policy fixed, as mentioned earlier), and since under $Q_3^{comp,*}$, q_3 is a function of x_3 and q_1, q_2, the expectation above is equal to $F_2(x_2, q_2, q_1)$ for some measurable $F_2(.)$. Thus, an optimal composite quantization policy at time 2 uses x_2 and q_1. The proof follows identically for other time stages. Hence, we have established an analogue of Theorem 10.3.1.

By observing that $P(dx_t|q_{[0,t]})$ is a sufficient statistic for an optimal control policy and the construction of a controlled Markov chain as in (10.35), it follows that the discussion in Theorem 10.3.2 also applies in this case. $\square$

10.8.6 *Proof of Theorem 10.4.2*

We show that the measurable selection hypothesis (see Sect. D.1.5) applies in the set of states which are visited with probability 1. In particular, the elements in the set of reachable probability measures admit densities which furthermore satisfy the equicontinuity condition in view of Lemma 10.8.1 below.

The following is a key lemma.

Lemma 10.8.1. *For all $t \geq 1$, $\pi_t(dx)$ is absolutely continuous with respect to the Lebesgue measure, i.e., it has a probability density function, which we will also denote by π_t by an abuse of notation. The density function π_t is uniformly continuous for every t and the sequence $\{\pi_t\}$ is a uniformly bounded and uniformly equicontinuous family.* ◇

Proof. Let ϕ denote the common density of the Gaussian noise variables w_t. Since $x_t = f(x_{t-1}) + w_t$ and w_t is independent of $q_{[0,t-1]}$, it is easy to see that the pdf π_t of the conditional measure $P(dx_t|q_{[0,t-1]})$ is given by

$$\pi_t(z) = \int_{\mathbb{R}^n} \phi(z - f(x_{t-1}))P(dx_{t-1}|q_{[0,t-1]}), \quad z \in \mathbb{R}^n.$$

The uniform boundedness of $\{\pi_t\}$ is immediate. Since ϕ is a Gaussian density, there is a $C > 0$ such that $\left|\frac{\partial}{\partial z_j}\phi(z)\right| \leq C$, $j = 1, \dots, n$. A standard application of the dominated convergence theorem implies that the partial derivatives of π_t exist and they also satisfy $\left|\frac{\partial}{\partial z_j}\pi_t(z)\right| \leq C$, $j = 1, \dots, n$. Since C does not depend on t, the sequence of densities $\{\pi_t\}$ is uniformly equicontinuous. $\square$

The following lemma is important for the proof.

Lemma 10.8.2 ([437]).

(a) Let $\{\mu_n\}$ be a sequence of density functions on $\mathbb{R}^n$ which are uniformly equicontinuous and uniformly bounded and assume $\mu_n \to \mu$ weakly. Then $\mu_n \to \mu$ in total variation.

(b) Let $\{Q_n\}$ be a sequence in $\mathcal{Q}_c$ such that $Q_n \to Q$ weakly at P. If P admits a density, then $Q_n \to Q$ in total variation at P.

(c) Let $P_nQ_n \to PQ$ weakly, where $\{Q_n\}$ is a sequence in $\mathcal{Q}_c$. Suppose further that $P_n \to P$ in total variation where P admits a density. Then $P_nQ_n \to PQ$ in total variation.

(d) *Let $P_n \to P$ in total variation where P admits a density function. Let $\{Q_n\}$ be a sequence in $\mathcal{Q}_c$. Then, there exists some subsequence such that $P_{k_n} Q_{k_n} \to PQ$ for some $Q \in \mathcal{Q}_c$.*

◇

Proof. (a) Since $\mu_n \to \mu$ weakly, the sequence $\{\mu_n\}$ is tight. Then, using a minor modification of Lemma 4.6.3, one can show that μ has a density and that the equicontinuity and uniform boundedness of the $\{\mu_n\}$ imply that along some subsequence $\mu_{k_n}(x) \to \mu(x)$ pointwise for all x. By Scheffe's theorem [70], μ_{k_n} converges to μ in L_1, which is equivalent to convergence in total variation. This convergence holds for the original sequence μ_n as well, since there cannot be a subsequence which does not converge in L_1: Suppose otherwise. Then, there exists a subsequence such that μ_{m_n} converges to μ weakly, but μ_{m_n} does not converge in L_1. Then for some $\epsilon > 0$ and a further subsequence with index m'_n, $\|\mu_{m'_n} - \mu\|_{TV} \geq \epsilon$. But, along a further subsequence $\mu_{m''_n}$ will converge to μ in total variation (by the arguments above), leading to a contradiction.

(b) It was shown in (4.15) that

$$\|PQ_n - PQ\|_{TV} \leq \sum_{i=1}^{M} P(B_i^n \triangle B_i),$$

where $B_1^n, \ldots, B_M^n$ and $B_1, \ldots, B_M$ are the cells of Q_n and Q, respectively, and $B_i^n \triangle B_i = (B_i^n \setminus B_i) \cup (B_i \setminus B_i^n)$. Since Q has convex cells, the boundary ∂B_i of each cell B_i has zero Lebesgue measure, so $P(\partial B_i) = 0$ because P has a density. Since $\partial(B_i \times \{j\}) = \partial B_i \times \{j\}$ and

$$P(A \times \{j\}) = P(A \cap B_j),$$

we have

$$PQ(\partial(B_i \times \{j\})) = P(\partial B_i \cap B_j) = 0,$$

for all i and j. Thus if $PQ^n \to PQ$ weakly, then $PQ^n(B_i \times \{j\}) \to PQ(B_i \times \{j\})$ by the Portmanteau theorem, which is equivalent to

$$P(B_i \cap B_j^n) \to P(B_i \cap B_j)$$

for all i and j. Since $B_1^n, \ldots, B_M^n$ and $B_1, \ldots, B_M$ are both partitions of $\mathbb{R}^n$, this implies $P(B_i^n \triangle B_i) \to 0$ for all i, which in turns proves that $PQ^n \to PQ$ in total variation via (4.15).

(c) For any measurable $A \subset \mathbb{R}^n \times \mathcal{M}$ we have

$$\begin{aligned} |PQ_n(A) - PQ(A)| \leq\ & |PQ_n(A) - P_n Q_n(A)| \\ & + |P_n Q_n(A) - PQ(A)|. \end{aligned}$$

It is relatively easy to see that $P_n \to P$ in total variation implies $|P_n Q_n(A) - PQ_n(A)| \to 0$. This follows from the observation that for $A_1 = \{x : (x, y) \in A\}$ and for $x \in A_1$, $A_2(x) = \{y : (x, y) \in A\}$,

$$|P_nQ_n(A) - PQ_n(A)| \leq \int_{A_1} |P_n(dx) - P(dx)| \bigg(\int_{A_2(x)} Q_n(dy|x) \bigg)$$

and that

$$\int_{A_1} |P_n(dx) - P(dx)| \bigg(\int_{A_2(x)} Q_n(dy|x) \bigg) \leq \|P_n - P\|_{TV} \to 0.$$

On the other hand, for any A with $PQ(\partial A) = 0$, we have $|P_nQ_n(A) - PQ(A)| \to 0$ since $P_nQ_n \to PQ$ weakly. This proves that $PQ_n \to PQ$ weakly. But since P admits a density, part (b) now implies that $Q_n \to Q$ in total variation.

Then we have

$$\begin{aligned}
&\|P_nQ_n - PQ\|_{TV} \\
&= \sup_{f:\|f\|_\infty \leq 1} \bigg| \sum_{i=1}^{M} \bigg(\int_{\mathbb{R}^n} f(x,i)Q_n(i|x)P_n(dx) - \int_{\mathbb{R}^n} f(x,i)Q(i|x)P(dx) \bigg) \bigg| \\
&\leq \sup_{f:\|f\|_\infty \leq 1} \bigg| \sum_{i=1}^{M} \bigg(\int_{\mathbb{R}^n} f(x,i)Q_n(i|x)P_n(dx) - \int_{\mathbb{R}^n} f(x,i)Q_n(i|x)P(dx) \bigg) \bigg| \\
&\quad + \sup_{f:\|f\|_\infty \leq 1} \bigg| \sum_{i=1}^{M} \bigg(\int_{\mathbb{R}^n} f(x,i)Q_n(i|x)P(dx) - \int_{\mathbb{R}^n} f(x,i)Q(i|x)P(dx) \bigg) \bigg| \\
&= \sup_{f:\|f\|_\infty \leq 1} \bigg| \bigg(\int_{\mathbb{R}^n} (P_n(dx) - P(dx)) \sum_{i=1}^{M} f(x,i)Q_n(i|x) \bigg) \bigg| \\
&\quad + \sup_{f:\|f\|_\infty \leq 1} \bigg| \sum_{i=1}^{M} \bigg(\int_{\mathbb{R}^n} f(x,i)Q_n(i|x)P(dx) \\
&\quad - \int_{\mathbb{R}^n} f(x,i)Q(i|x)P(dx) \bigg) \bigg| \to 0
\end{aligned}$$

since $P_n \to P$ in total variation and $Q_n \to Q$ in total variation at P.

(d) By (b) above, there exists a subsequence Q_{m_n} such that PQ_{m_n} converges to PQ for some Q. Since $P_n \to P$, we have that

$$\|P_{m_n}Q_{m_n} - PQ\|_{TV} \leq \|P_{m_n}Q_{m_n} - PQ_{m_n}\|_{TV} + \|PQ_{m_n} - PQ\|_{TV} \to 0,$$

and the result follows. □

A Measurable Selection Condition and the Proof of Theorem 10.4.2

We now provide a relaxation of the measurable selection conditions considered in [194, Theorem 3.3.5] (see also Appendix D). Let $\mathcal{S} \subset \mathcal{P}(\mathbb{R}^n)$ be the set of reachable states for π under any composite coding policy. Note that by Lemma 10.8.1, the set of densities in $\mathcal{S}$ is uniformly bounded and equicontinuous.

Condition D

(i) The cost function $c(\pi, Q)$ is continuous on $\mathcal{S} \times \mathcal{Q}_c$ in the sense that $P_n Q_n \to PQ$ implies that $c(P_n, Q_n) \to c(P, Q)$.
(ii) $\mathcal{Q}_c$ is compact in total variation at any input π admitting a density.
(iii)

$$\int_{\mathcal{P}(\mathbb{R}^n)} J_{t+1}^T(\pi) P(d\pi_{t+1}|\pi_t, Q_t)$$

is a continuous function on $\mathcal{S} \times \mathcal{Q}_c$, for the value function J_{t+1}^T at time t defined recursively as

$$J_t^T(\pi_t) = \min_{Q_t \in \mathcal{Q}_c} c(\pi_t, Q_t) + E[J_{t+1}^T(\pi_{t+1})|\pi_t = \pi, Q_t = Q],$$

with $J_T^T = 0$.

The proof of the following theorem follows essentially from dynamic programming equation itself. This is related to (but weaker than) the Measurable Selection Condition 3.3.2 (with a,b,c1) and the subsequent Theorem 3.3.5 in [194] since here we directly consider the value function.

Theorem 10.8.1. *Under Condition D, there exists an optimal (Borel measurable) policy in Π_W achieving* (10.13). ◇

In view of the preceding theorem, to prove Theorem 10.4.2 it suffices to show that Condition D holds. We note that (ii) in Condition D directly follows from Theorem 4.7.4.

An important supporting lemma is the following.

Lemma 10.8.3. *Let for $\pi_n \in \mathcal{S}$, $Q_n \in \mathcal{Q}_c$*

$$\pi'(m, \pi_n, Q_n)(C) := P(x_{n+1} \in C|\pi_n, Q_n, q_n = m)$$
$$= \frac{1}{\pi_n(B_m^n)} \int_{z \in C} \left\{ \int \pi_n(dx) 1_{\{x \in B_m^n\}} \phi(z - f(x)) \right\} dz. \quad (10.45)$$

As $(\pi_n, Q_n) \to (\pi, Q)$ in total variation, for every $m \in \{1, \cdots, M\}$,

$$\|\pi'(m, \pi_n, Q_n) - \pi'(m, \pi, Q)\|_{TV} \to 0,$$

provided that $\pi(B_m) > 0$. ◇

The following lemma is a minor generalization of Theorem 4.5.4.

Lemma 10.8.4. *Let* $P_nQ_n \to PQ$ *in total variation. Then,*

$$\inf_\gamma \int P_n(dx)Q_n(q|x)c(x,\gamma(q)) \to \inf_\gamma \int P(dx)Q(q|x)c(x,\gamma(q)). \tag{10.46}$$

◇

The next result establishes (i) in Condition D, the continuity of $c(\pi_t, Q_t)$ on $\mathcal{S} \times \mathcal{Q}_c$.

Theorem 10.8.2. $c(\pi, Q)$ *is continuous on* $\mathcal{S} \times \mathcal{Q}_c$. ◇

Proof. Let $\{(\pi_n, Q_n)\}$ be a sequence in $\mathcal{S} \times \mathcal{Q}_c$ such that $\pi_n Q_n \to \pi Q$ weakly. It follows from Lemma 10.8.2(a) that $\pi_n \to \pi$ in total variation and from Lemma 10.8.2(c) that $\pi_n Q_n \to \pi Q$ in total variation. Then Lemma 10.8.4 implies that $c(\pi_n, Q_n) \to c(\pi, Q)$. □

Next we establish (iii) in Condition D. We wish to prove that $E[J_t^T(\pi_t)|\pi, Q]$ is continuous in π, Q. We apply backward induction. At $t = T-1$, let

$$J_{T-1}^T(\pi_{T-1}) = \min_Q c(\pi_{T-1}, Q_{T-1}).$$

By Theorem 10.8.2 and the compactness of the set of quantizers (by Theorem 4.7.5), there exists an optimal quantizer, Q_{T-1}^*. Furthermore, the following holds:

Lemma 10.8.5. *Let* $F : \mathcal{S} \times \mathcal{Q}_c \to \mathbb{R}$ *be continuous on* $\mathcal{S} \times \mathcal{Q}_c$ *in the sense that* $P_nQ_n \to PQ$ *implies that* $F(\pi_n, Q_n) \to F(\pi, Q)$. *Then, the function* $\min_Q F(\pi, Q)$ *is continuous in* π. ◇

Proof. Let $\pi_n \to \pi$, Q_n be optimal for π_n, and Q be optimal for π. Then

$$|\min_Q F(\pi_n, Q) - \min_Q F(\pi, Q)| \leq \max\Big(F(\pi_n, Q) - F(\pi, Q), F(\pi, Q_n) - F(\pi_n, Q_n)\Big). \tag{10.47}$$

The first term above converges since F is continuous in π, Q. The second converges also. Suppose otherwise. Then, for some $\epsilon > 0$, there exists a subsequence such that

$$F(\pi, Q_{k_n}) - F(\pi_{k_n}, Q_{k_n}) \geq \epsilon.$$

Consider the sequence (π_{k_n}, Q_{k_n}). There exists a subsequence such that $(\pi_{k'_n}, Q_{k'_n})$ which converges to π, Q' for some Q', by Lemma 10.8.2(d). Hence, for this subsequence, we have convergence of $F(\pi_{k'_n}, Q_{k'_n})$ as well as $F(\pi, Q_{k'_n})$, leading to a contradiction. □

As a consequence, $J^T_{T-1}(\pi_{T-1})$ is continuous in π_{T-1}. Consider now $t = T - 2$ and we wish to see if there is a solution to the following equality:

$$J^T_{T-2}(\pi_{T-2}) = \min_{Q} \left(c(\pi_{T-2}, Q_{T-2}) + E[J^T_{T-1}(\pi_{T-1})|\pi_{T-2}, Q_{T-2}] \right). \tag{10.48}$$

Note that

$$\begin{aligned} &E[J^T_{T-1}(\pi_{T-1})|\pi_{T-2}, Q_{T-2}] \\ &= \sum_{m=1}^{M} P(\pi'(m, \pi_{T-2}, Q_{T-2})|\pi_{T-2}, Q_{T-2}) J^T_{T-1}(\pi'(m, \pi_{T-2}, Q_{T-2}), \end{aligned}$$

where

$$P(\pi'(m, \pi_{T-2}, Q_{T-2})|\pi_{T-2}, Q_{T-2}) = P(q_{T-2} = m|\pi_{T-2}, Q_{T-2})$$

and

$$\pi'(m, \pi_{T-2}, Q_{T-2})(dz) = \frac{\int \pi_{T-2}(dx_{T-2}) P(q_{t-2}=m|\pi_{T-2}, x_{T-2}) P(dz|x_{T-2})}{\int\int \pi_{T-2}(dx_{T-2}) P(q_{t-2}=m|\pi_{T-2}, x_{T-2}) P(dz|x_{T-2})},$$

or

$$\pi'(m, \pi, Q)(C) = \frac{1}{\pi(B_m)} \int_{z \in C} \left\{ \pi(dx) 1_{\{x \in B_m\}} \phi(z - f(x)) \right\} dz. \tag{10.49}$$

Lemma 10.8.3 shows that as $\pi Q_n \to \pi Q$, $\|\pi'(m, \pi, Q_n) - \pi'(m, \pi, Q)\|_{TV} \to 0$ whenever $\pi'(m, \pi, Q) > 0$. If $\pi'(m, \pi_{T-2}, Q_{T-2}) = 0$, by the boundedness of the cost, it follows that

$$P(\pi'(m, \pi_{T-2}, Q_{T-2,n})|\pi_{T-2}, Q_{T-2,n}) J^T_{T-1}(\pi'(m, \pi_{T-2}, Q_{T-2,n}) \to 0,$$

since $P(\pi'(m, \pi_{T-2}, Q_{T-2,n})|\pi_{T-2}, Q_{T-2,n}) \to 0$.

As a consequence, we have that $E[J^T_{T-1}(\pi_{T-1})|\pi_{T-2}, Q_{T-2}]$ is continuous in (π_{T-2}, Q_{T-2}), since both of the expressions in (10.48) are continuous. Hence,

$$J^T_{T-2}(\pi_{T-2}) = \min_{Q_{T-2}} \left(c(\pi_{T-2}, Q_{T-2}) + E[J^T_{T-1}(\pi_{T-1})|\pi_{T-2}, Q_{T-2}] \right)$$

exists and by Lemma 10.8.5 is continuous. The recursion applies for all time stages.

This concludes the proof of Theorem 10.4.2. □

10.8.7 Proof of Theorem 10.5.1

The proof is in three steps: (i), (ii), and (iii) below.

Step (i) In decentralized dynamic decision problems where the decision makers have the same objective (i.e., in team problems), more information provided to the decision makers does not lead to any degradation in performance, since the decision makers can always choose to ignore the additional information (as we saw in Sect. 3.5.2, in view of *expansion of information structures*). In view of this, let us relax the information structure in such a way that the decision makers now have access to all the previous observations; that is, the information available at the encoders 1 and 2 are

$$I_t^i = \{y_t^i, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}\} \quad t \geq 1, \quad i = 1, 2.$$

$$I_0^i = \{y_0^i\}, \quad i = 1, 2.$$

The information pattern among the encoders is now the *one-step delayed observation sharing pattern*. We will show that the past information can be eliminated altogether, to prove the desired result.

Step (ii) The second step uses the following technical lemma.

Lemma 10.8.6. *Under the relaxed information structure in step* (i) *above, any decentralized quantization policy at time t,* $1 \leq t \leq T-1$*, can be replaced, without any loss in performance, with one which only uses* $(\pi_t, \mathbf{y}_t, \mathbf{q}_{[0,t-1]})$*, satisfying the following form:*

$$\begin{aligned} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]}) &= P(q_t^1|y_t^1, \mathbf{q}_{[0,t-1]})P(q_t^2|y_t^2, \mathbf{q}_{[0,t-1]}) \\ &= 1_{\{q_t^1 = \bar{Q}^1(y_t^1, \mathbf{q}_{[0,t-1]})\}} 1_{\{q_t^2 = \bar{Q}^2(y_t^2, \mathbf{q}_{[0,t-1]})\}}, \end{aligned} \tag{10.50}$$

for measurable functions $\bar{Q}^1$ *and* $\bar{Q}^2$. ◇

Proof. Let us fix a composite quantization policy $\mathbf{Q}^{comp}$. At time $t = T-1$, the per-stage cost function can be written as

$$E[\int_{\mathbb{X}} P(dx_t|\mathbf{q}_{[0,t]})c(x_t, u_t)|\mathbf{q}_{[0,t-1]}]. \tag{10.51}$$

For this problem, $P(dx_t|\mathbf{q}_{[0,t]})$ is a sufficient statistic for an optimal receiver. Hence, at time $t = T-1$, an optimal receiver will use $P(dx_t|\mathbf{q}_{[0,t]})$ as a sufficient statistic for an optimal decision as the cost function conditioned on $\mathbf{q}_{[0,t]}$ is written as $\int P(dx_t|\mathbf{q}_{[0,t]})c(x_t, u_t)$, where u_t is the decision of the receiver. Now, let us fix this decision policy at time t. We now observe that

$$P(dx_t|\mathbf{q}_{[0,t]}) = \sum_{\mathbb{Y}^{t+1}} P(dx_t, \mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) = \frac{\sum_{\mathbb{Y}^{t+1}} P(dx_t, \mathbf{q}_t, \mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t-1]})}{P(\mathbf{q}_t|\mathbf{q}_{[0,t-1]})}$$
$$= \frac{\sum_{\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)P(dx_t|\mathbf{y}_{[0,t-1]})P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}{\int_{\mathbb{X},\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)P(dx_t|\mathbf{y}_{[0,t-1]})P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}$$
$$= \frac{\sum_{\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)\pi(dx_t)P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}{\int_{\mathbb{X},\mathbb{Y}^{t+1}} P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|x_t)\pi(dx_t)P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})}. \tag{10.52}$$

The term $P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})$ is determined by the composite quantization policies:

$$P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})$$
$$= P(q_t^1|y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})P(q_t^2|y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})$$
$$= 1_{\{q_t^1 = Q_t^{comp,1}(y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}} 1_{\{q_t^2 = Q_t^{comp,2}(y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}}.$$

In (10.52), we use the relation $P(dx_t|\mathbf{y}_{[0,t-1]}) = P(dx_t) =: \pi(dx_t)$, where $\pi(\cdot)$ denotes the marginal probability on x_t (recall that the source is memoryless). The term $P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})$ in (10.52) is determined by the composite quantization policies:

$$P(\mathbf{q}_t|\mathbf{y}_{[0,t]}, \mathbf{q}_{[0,t-1]})$$
$$= P(q_t^1|y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})P(q_t^2|y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})$$
$$= 1_{\{q_t^1 = Q_t^{comp,1}(y_t^1, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}} 1_{\{q_t^2 = Q_t^{comp,2}(y_t^2, \mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]})\}}.$$

The above is valid since each encoder knows the past observations of both encoders.

As such, $P(dx_t|\mathbf{q}_{[0,t]})$ can be written as, for some function $\varUpsilon$,

$$\varUpsilon(\pi, \mathbf{q}_{[0,t-1]}, \mathbf{Q}_t^{comp}(.)).$$

Note that $\mathbf{q}_{[0,t-1]}$ appears due to the term $P(\mathbf{y}_{[0,t-1]}|\mathbf{q}_{[0,t-1]})$. Now, consider the space of joint mappings at time t, denoted by $\mathcal{G}_t$:

$$\mathcal{G}_t = \{\mathbf{\Psi}_t : \mathbf{\Psi}_t = \{\Psi_t^1, \Psi_t^2\}, \Psi_t^i : \mathbb{Y}^i \to \mathcal{M}_t^i, \quad i = 1, 2\}.$$

For every composite quantization policy there exists a distribution P' on random variables $(\mathbf{q}_t, \pi, \mathbf{q}_{[0,t-1]})$ such that

$$P'(\mathbf{q}_t|\pi, \mathbf{q}_{[0,t-1]}) = \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} P(\mathbf{q}_t, \mathbf{y}_{[0,t]}|\pi, \mathbf{q}_{[0,t-1]})$$
$$= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} \Bigg(P(q_t^1|\mathbf{y}_{[0,t-1]}, y_t^1, \mathbf{q}_{[0,t-1]}, \pi)$$
$$P(q_t^2|\mathbf{y}_{[0,t-1]}, y_t^2, \mathbf{q}_{[0,t-1]}, \pi)P(y_t^1, y_t^2)P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]})\Bigg). \tag{10.53}$$

Furthermore, with every composite quantization policy and every realization of $\mathbf{y}_{[0,t-1]}, \mathbf{q}_{[0,t-1]}$, we can associate an element in the space $\mathcal{G}_t$, $\boldsymbol{\Psi}_{\mathbf{y}_{[0,t-1]},\mathbf{q}_{[0,t-1]}}$, such that the induced stochastic relationship in (10.53) can be obtained:

$$P'(\mathbf{q}_t|\pi, \mathbf{q}_{[0,t-1]}) = \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} P(\mathbf{q}_t, \mathbf{y}_{[0,t]}|\pi, \mathbf{q}_{[0,t-1]})$$

$$= \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t+1}} 1_{\{\boldsymbol{\Psi}_{\mathbf{y}_{[0,t-1]},\mathbf{q}_{[0,t-1]}}(y_t^1, y_t^2) = (q_t^1, q_t^2)\}} P(y_t^1, y_t^2) P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]}).$$

We can thus express the cost, for some measurable function F in the following way:

$$E[F(\pi, \mathbf{q}_{[0,t-1]}, \boldsymbol{\Psi})|\pi, \mathbf{q}_{[0,t-1]}],$$

where

$$P(\boldsymbol{\Psi}|\pi, \mathbf{q}_{[0,t-1]}) = \sum_{(\mathbb{Y}^1 \times \mathbb{Y}^2)^{t}} 1_{\{\boldsymbol{\Psi} = \boldsymbol{\Psi}_{\mathbf{y}_{[0,t-1]},\mathbf{q}_{[0,t-1]}}\}} P(\mathbf{y}_{[0,t-1]}|\pi, \mathbf{q}_{[0,t-1]}).$$

Now let $t = T - 1$ and define for every possible realization $\boldsymbol{\Psi}_t = (\varPsi_t^1, \varPsi_t^2) \in \mathcal{G}_t$ (with the decision policy considered earlier fixed):

$$\beta_{\boldsymbol{\Psi}_t} := \Big\{ \pi, \mathbf{q}_{[0,t-1]} : F(\pi, \mathbf{q}_{[0,t-1]}, \boldsymbol{\Psi}_t) \leq F(\pi, \mathbf{q}_{[0,t-1]}, \boldsymbol{\Psi}'_t) \quad \forall ((\varPsi_t^1)', (\varPsi_t^2)') \in \mathcal{G}_t \Big\}.$$

As we had observed in the proof of Theorem 10.3.3, such a construction covers the domain set consisting of $(\pi, \mathbf{q}_{[0,t-1]})$ but possibly with overlaps. Note that for every $(\pi, \mathbf{q}_{[0,t-1]})$, there exists a minimizing function in $\mathcal{G}_t$, since $\mathcal{G}_t$ is a finite set. In this sequence, let there be an ordering of the finitely many elements in $\mathcal{G}_t$ as $\{\boldsymbol{\Psi}_t(1), \boldsymbol{\Psi}_t(2), \ldots, \boldsymbol{\Psi}_t(k), \ldots\}$, and define a function $\mathbf{T}_t^*$ as

$$\boldsymbol{\Psi}_t(k) = \mathbf{T}_t^*(\pi, \mathbf{q}_{[0,t-1]}), \text{if} \quad \Big(\pi, \mathbf{q}_{[0,t-1]}\Big) \in \beta_{\boldsymbol{\Psi}_t(k)} - \cup_{i=0}^{k-1} \beta_{\boldsymbol{\Psi}_t(i)},$$

with $\beta_{\boldsymbol{\Psi}_t(0)} = \emptyset$.

Thus, we have constructed a policy which performs at least as well as the original composite quantization policy. It has a restricted structure in that it only uses $(\pi, \mathbf{q}_{[0,t-1]})$ to generate the team action and the local information y_t^1, y_t^2 to generate the quantizer outputs.

Now that we have obtained the structure of the optimal encoders for the last stage, we can sequentially proceed to study the other time stages. Note that given a fixed π, $\{(\pi, \mathbf{y}_t)\}$ is i.i.d. and hence Markov. Now, define $\pi'_t = (\pi, \mathbf{y}_t)$. For a three-stage cost problem, the cost at time $t = 2$ can be written as, for measurable functions c_2, c_3,

$$c_2(\pi'_2, v_2(\mathbf{q}_{[1,2]})) + E[c_3(\pi'_3, v_3(\mathbf{q}_{[1,2]}, Q_3(\pi'_3, \mathbf{q}_{[1,2]})))|\pi'_{[1,2]}, \mathbf{q}_{[1,2]}].$$

Since $P(d\pi'_3, \mathbf{q}_{[1,2]}|\pi'_2, \pi'_1, \mathbf{q}_{[1,2]}) = P(d\pi'_3, \mathbf{q}_{[1,2]}|\pi'_2, \mathbf{q}_{[1,2]})$, the expression above is equal for some $F_2(\pi'_2, \mathbf{q}_2, \mathbf{q}_1)$ for some measurable F_2. By a similar argument, an optimal composite quantizer at time t, $1 \leq t \leq T-1$ only uses $(\pi, \mathbf{y}_t, \mathbf{q}_{[0,t-1]})$. An optimal (team) policy generates the quantizers Q^1_t, Q^2_t using $\mathbf{q}_{[0,t-1]}, \pi$, and the quantizers use $\{y^i_t\}$ to generate the quantizer outputs at time t for $i = 1, 2$. □

Step(iii) The final step will complete the proof. At time $t = T - 1$, an optimal receiver will use $P(dx_t|\mathbf{q}_{[0,t]})$ as a sufficient statistic for the optimal decision. We now observe that

$$\begin{aligned} P(dx_t|\mathbf{q}_{[0,t]}) &= \sum_{\mathbb{Y}^{t+1}} P(dx_t|\mathbf{y}_{[0,t]})P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) \\ &= \sum_{\mathbb{Y}^{t+1}} P(dx_t|\mathbf{y}_t)P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) = \sum_{\mathbb{Y}} P(dx_t|\mathbf{y}_t)\sum_{\mathbb{Y}^t} P(\mathbf{y}_{[0,t]}|\mathbf{q}_{[0,t]}) \\ &= \sum_{\mathbb{Y}} P(dx_t|\mathbf{y}_t)P(\mathbf{y}_t|\mathbf{q}_{[0,t]}). \end{aligned}$$

Thus, $P(dx_t|\mathbf{q}_{[0,t]})$ is a function of $P(\mathbf{y}_t|\mathbf{q}_{[0,t]})$. Now, let us note that

$$\begin{aligned} P(\mathbf{y}_t|\mathbf{q}_{[0,t]}) &= \frac{P(\mathbf{q}_t, \mathbf{y}_t|\mathbf{q}_{[0,t-1]})}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t, \mathbf{y}_t|\mathbf{q}_{[0,t-1]})} \\ &= \frac{P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|\mathbf{q}_{[0,t-1]})}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t|\mathbf{q}_{[0,t-1]})} \\ &= \frac{P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t)}{\sum_{\mathbf{y}_t} P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})P(\mathbf{y}_t)}, \end{aligned} \tag{10.54}$$

where the term $P(\mathbf{q}_t|\mathbf{y}_t, \mathbf{q}_{[0,t-1]})$ is determined by the quantizer team action $\mathbf{Q}^{comp}_t$. As such, the cost at time $t = T - 1$ can be expressed as a measurable function $G(P(\mathbf{y}_t), \mathbf{Q}_t)$. Thus, it follows that, an optimal quantizer policy at the last stage, $t = T - 1$ may only use $P(\mathbf{y}_t)$ to generate the quantizers, where the quantizers use the local information y^i_t to generate the quantization output. The rest of the proof follows the arguments in the proof of Theorem 10.3.4: At time $t = T - 2$, the sufficient statistic for the cost function is $P(dx_{t-1}|\mathbf{q}_{[0,t-1]})$ both for the immediate cost and the cost-to-go, that is, the cost impacting the time stage $t = T - 1$, as a result of the optimality result for Q_{T-1} and the memoryless nature of the source dynamics. The same argument applies for all time stages.

Hence, any policy without loss can be replaced with one in Π^{NSM}. Since there are finitely many policies in this class, an optimal composite quantization policy exists. □

10.8.8 *Proof of Lemma 10.6.1*

We apply dynamic programming. Let for the final stage, $t = T - 1$, $f_t(q_{[0,t-1]}) := \sum_{k=0}^{t-1} A^{t-k-1}Bu_k$ and $x_t = \bar{x}_t + f_t(q_{[0,t-1]})$. If the policy is in Π_W, the composite quantization policy is of the form

$$Q_t(\bar{x}_t + \sum_{k=0}^{t-1} A^{t-k-1}Bu_k, P(\bar{x}_t + \sum_{k=0}^{t-1} A^{t-k-1}Bu_k \in \cdot|q_{[0,t-1]})).$$

For this time stage, let there be an optimal decoder and controller for which a sufficient statistic for the optimal control policy is $E[x_t|q_{[0,t]}]$. Observe that

$$\begin{aligned} E[\bar{x}_t + f_t(q_{[0,t-1]})|q_{[0,t]}] &= E[\bar{x}_t|q_{[0,t]}] + f_t(q_{[0,t-1]}) \\ &= E[\bar{x}_t|q_{[0,t-1]}, q_t] + f_t(q_{[0,t-1]}). \end{aligned} \tag{10.55}$$

The quantization output q_t represents the bin information for x_t. By shifting the quantizer bins by $f_t(q_{[0,t-1]})$, a new quantizer which quantizes $\bar{x}_t$ can generate the same bin information on x_t through q_t. Hence, there is no information loss due to the elimination of the past control actions. Therefore, this new quantizer, by adding $f_t(q_{[0,t-1]})$ to the output, generates the same conditional estimate of the state as the original quantizer. Thus, there exists a quantizer of the form $\tilde{Q}_t(\bar{x}_t, P(\bar{x}_t \in \cdot|q_{[0,t-1]}))$ with the following property: The estimation error realization and hence the estimation is the same almost surely and as a consequence of the structure of the cost and linearity in the system, the conditional estimate is a sufficient statistic, the cost realization is identical almost surely. Furthermore, $\bar{w}_t$ is independent of the control actions applied earlier (due to the separated structure).

Consequently, for $t = T - 3$, since u_{T-2} is independent of $\bar{w}_{T-2}$ and $\bar{w}_{T-1}$, an optimal controller will use $E[x_t|q_{[0,t]}]$ as a sufficient statistic given the structural result above for u_{T-1}, u_{T-2} and the encoder policies. Hence, the analysis above applies for $t = T-4$ and by induction, for all time stages until $t = 0$. The estimation error is independent of the control actions under an optimal coding and control policy without any loss. □

10.8.9 Proof of Theorem 10.6.3

As a consequence of Theorem 10.6.2, we obtain that for $t \geq 0$, the unnormalized value function to be given by

$$J_t(\mathcal{I}_t^c) = E[x_t' K_t x_t | \mathcal{I}_t^c] + \sum_{k=t}^{T-1} \Bigg(E[(x_k - E[x_k|\mathcal{I}_k^c])' Q (x_k - E[x_k|\mathcal{I}_k^c])] \\ + E[\bar{w}_k' K_{k+1} \bar{w}_k] \Bigg), \tag{10.56}$$

where the effective noise process is $\bar{w}_t = E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]$ with

$$J(\Pi^{comp}, \underline{\gamma}^0, T) = \frac{1}{T} J_0(\mathcal{I}_0^c).$$

Given a positive-definite matrix Λ define an inner product as

$$\langle z_1, z_2 \rangle_\Lambda = z_1' \Lambda z_2.$$

and the norm generated by this inner product as $|z|_\Lambda = \sqrt{z' \Lambda z}$. We now note the following:

$$E\Bigg[|E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c]|_\Lambda^2 \Bigg] \\ = E\Bigg[|\Bigg((E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}) + (x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]) \Bigg)|_\Lambda^2 \Bigg] \\ = E[|(E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1})|_\Lambda^2] + E[|(x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c])|_\Lambda^2] \\ +2E\Bigg[\langle (E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}), (x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]) \rangle_\Lambda \Bigg].$$

Note that

$$E\Bigg[\langle (E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}), (x_{t+1} - E[x_{t+1}|\mathcal{I}_t^c]) \rangle_\Lambda \Bigg] \\ = E\Bigg[- \langle (E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}), (E[x_{t+1}|\mathcal{I}_t^c]) \rangle_\Lambda \\ + \langle (E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}), (x_{t+1}) \rangle_\Lambda \Bigg]$$

$$= E\Bigg[\langle (E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1}), (x_{t+1}) \rangle_\Lambda \Bigg] \tag{10.57}$$

$$= -E[|(E[x_{t+1}|\mathcal{I}_{t+1}^c] - x_{t+1})|_\Lambda^2], \tag{10.58}$$

where (10.57)–(10.58) follow from the orthogonality property of minimum mean-square estimation and that $E[x_{t+1}|\mathcal{I}_t^c]$ is measurable on $\sigma(\mathcal{I}_{t+1}^c)$, the sigma-field generated by $\mathcal{I}_{t+1}^c$. Therefore, we have

$$\begin{aligned}
&E\Bigg[|(E[x_{t+1}|\mathcal{I}_{t+1}^c] - E[x_{t+1}|\mathcal{I}_t^c])|^2_{K_{t+1}}\Bigg]\\
&= E\Bigg[|(x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}^c])|^2_{K_{t+1}} + E[|(Ax_t + w_t - AE[x_t|\mathcal{I}_t^c])|^2_{K_{t+1}}\Bigg]\\
&\qquad -2E[|(x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}^c])|^2_{K_{t+1}}]\\
&= -E\Bigg[(x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}^c])'(K_{t+1})(x_{t+1} - E[x_{t+1}|\mathcal{I}_{t+1}^c])\\
&\qquad +E[(x_t - E[x_t|\mathcal{I}_t^c])'(A'K_{t+1}A)(x_t - E[x_t|\mathcal{I}_t^c])] + E[w'K_{t+1}w]\Bigg].
\end{aligned}$$

Thus, the finite horizon cost could be written as, since $K_T = 0$,

$$\begin{aligned}
J_t(\mathcal{I}_t^c) &= E[x_t'K_tx_t|\mathcal{I}_t^c]\\
&\quad + \sum_{k=t}^{T-1} E[(x_k - E[x_k|\mathcal{I}_k^c])'(Q + A'K_{k+1}A)(x_k - E[x_k|\mathcal{I}_k^c])]\\
&\quad - \sum_{k=t}^{T-1} E[(x_{k+1} - E[x_{k+1}|\mathcal{I}_{k+1}^c])'(K_{k+1})(x_{k+1} - E[x_{k+1}|\mathcal{I}_{k+1}^c])]\\
&\quad + \sum_{k=t}^{T-1} E[w_k'K_{k+1}w_k]\\
&= E[x_t'K_tx_t|\mathcal{I}_t^c]+\sum_{k=t}^{T-1}E[(x_k-E[x_k|\mathcal{I}_k^c])'(Q+A'K_{k+1}A)(x_k-E[x_k|\mathcal{I}_k^c])]\\
&\quad - \sum_{k=t+1}^{T-1} E[(x_k - E[x_k|\mathcal{I}_k^c])'(K_k)(x_k - E[x_k|\mathcal{I}_k^c])]\\
&\quad + \sum_{k=t}^{T-1} E[w_k'K_{k+1}w_k].\\
&=E[x_t'K_tx_t|\mathcal{I}_t^c]+E[(x_t-E[x_t|\mathcal{I}_t^c])'(Q+A'K_{t+1}A)(x_t-E[x_t|\mathcal{I}_t^c])]\\
&\quad + \sum_{k=t+1}^{T-1} E[(x_k - E[x_k|\mathcal{I}_k^c])'(Q + A'K_{k+1}A - K_k)(x_k - E[x_k|\mathcal{I}_k^c])]\\
&\quad + \sum_{k=t}^{T-1} E[w_k'K_{k+1}w_k].
\end{aligned} \tag{10.59}$$

Now, with a fixed horizon T and for $t < T-1$,

$$\begin{aligned} J_t(\mathcal{I}_t^c) &= E[x_t' K_t x_t | \mathcal{I}_t^c] + E[(x_t - E[x_t|\mathcal{I}_t^c])'(Q + A'K_{t+1}A)(x_t - E[x_t|\mathcal{I}_t^c])] \\ &\quad + \sum_{k=t+1}^{T-1} E[(x_k - E[x_k|\mathcal{I}_k^c])'(Q + A'K_{k+1}A - K_k)(x_k - E[x_k|\mathcal{I}_k^c])] \\ &\quad + \sum_{k=t}^{T-1} E[w_k' K_{k+1} w_k]. \end{aligned} \tag{10.60}$$

Letting $t = 0$ completes the proof. □

10.8.10 Proof of Theorem 10.6.4

We will show that *Condition D* used in the proof of Theorem 10.4.2 applies. We need to modify the proof of Theorem 10.4.2 only in view of the unboundedness of the cost, which appears in two contexts. One is with regard to the continuity of $c(\pi, Q)$ and the other is the weak continuity of the expected value function in the transition kernel. We address both below. To economize the notation, we take $P_t = I$ in the following.

(a) *Continuity of* $c(\pi, Q)$

Continuity under total variation can be extended for unbounded functions, provided there is a uniform integrability condition as follows:

$$\lim_{L\to\infty} \sup_{Q_n} \inf_{\gamma} E_\pi^{Q_n}[(x - Q_n(x))'(x - Q_n(x)) 1_{\{(x-\gamma(Q_n(x)))'(x-\gamma(Q_n(x)))\geq L\}}] = 0,$$

where by an abuse of notation, the infimization $\inf_\gamma$ is not for the truncated expression

$$E_\pi^{Q_n}[(x - Q_n(x))'(x - Q_n(x)) 1_{\{(x-\gamma(Q_n(x)))'(x-\gamma(Q_n(x)))\geq L\}}],$$

but for the original cost

$$E_\pi^{Q_n}[(x - Q_n(x))'(x - Q_n(x))].$$

Now, by the parallelogram law

$$(x - Q(x))'(x - Q(x)) \leq 2x'x + 2|\sup_x Q(x)|^2.$$

As a consequence, for any Q_n, π_n, we have that for some sequence D_n

$$\sup_{\pi_n, Q_n} \inf_{\gamma} \int \pi_n(dx)(x-\gamma(Q_n(x))'(x-\gamma(Q_n(x))1_{\{(x-\gamma(Q_n(x))'(x-\gamma(Q_n(x))\geq L\}}$$
$$\leq \sup_{\pi_n} \inf_{\gamma} \int \pi_n(dx)(2x'x + 2D_n)1_{\{2x'x \geq L-2D_n\}}. \tag{10.61}$$

For every π, Q and every sequence π_n, Q_n converging to π, Q, since the bins converge setwise, so does the minimizing quantizer reconstruction levels in the sense that

$$\int_{B_k^n} \pi_n(dx)x \to \int_{B_k} \pi(dx)x, \quad 1 \leq k \leq M.$$

Hence, for some $D < \infty$

$$\sup_{\pi_n, Q_n} \inf_{\gamma} \int \pi_n(dx)(x - \gamma(Q_n(x))'(x - \gamma(Q_n(x))1_{\{(x-\gamma(Q^(x))'(x-\gamma(Q^(x))\geq L\}}$$
$$\leq \sup_{\pi_n, Q_n} \inf_{\gamma} \int \pi_n(dx)(2x'x + 2D_n)1_{\{2x'x \geq L-2D_n\}}$$
$$\leq \sup_{\pi_n} \inf_{\gamma} \int \pi_n(dx)(2x'x + 2D)1_{\{2x'x \geq L-2D\}}. \tag{10.62}$$

Hence, one needs to prove that $\{\pi_n\}$ itself is uniformly integrable. If this holds then, for every ϵ, there exists an L such that for all $(\pi_n, Q_n) \to (\pi, Q)$, it follows that

$$\Big| E^{Q}_{\pi_n}[(x - Q_n(x))'(x - Q_n(x))1_{\{(x-\gamma(Q_n(x)))'(x-\gamma(Q_n(x)))\leq L\}}]$$
$$- E^{Q_n}_{\pi_n}[(x - Q_n(x))'(x - Q_n(x))]\Big| \leq \epsilon/2$$

and that for sufficiently large n, given L,

$$\Big| E^{Q_n}_{\pi_n}[(x - Q_n(x))'(x - Q_n(x))1_{\{(x-\gamma(Q_n(x)))'(x-\gamma(Q_n(x)))\leq L\}}] \tag{10.63}$$
$$- E^{Q}_{\pi}[(x - Q(x))'(x - Q(x))1_{\{(x-\gamma(Q(x)))'(x-\gamma(Q(x)))\leq L\}}]\Big| \leq \epsilon/2.$$

Hence, for every $\epsilon > 0$ there exists n_0 such that for all $n \geq n_0$,

$$|E^{Q}_{\pi_n}[(x - Q_n(x))'(x - Q_n(x))] - E^{Q}_{\pi}[(x - Q(x))'(x - Q(x))]| \leq \epsilon.$$

Thus, continuity is established under the uniform integrability condition. The following technical lemma addresses the uniform integrability of π_n.

Lemma 10.8.7 ([434]). *Let* $\pi_{t,n} \to \pi_t$ *be a uniformly integrable sequence. Then,* $\pi'(m, \pi_{t,n}, Q_{t,n})$ *[defined in (10.45)] is uniformly integrable for* $(\pi_{t,n}, Q_{t,n}) \to \pi_t, Q_t$. ◇

(b) *Continuity of the Value Function in the Quantizer*
We apply backward induction. Let

$$J^T_{T-1}(\pi_{T-1}) = \min_Q c(\pi_{T-1}, Q_{T-1}).$$

We observed in part (a) above that for this case the optimal cost function is continuous in π_{T-1}, Q_{T-1}, provided π_{T-1} varies along a uniformly integrable sequence, and hence by Lemma 10.8.5, the value function J^T_{T-1} is continuous in π_{T-1}. Now, we wish to see if

$$J^T_{T-2}(\pi_{T-2}) = \min_Q \left(c(\pi_{T-2}, Q_{T-2}) + E[J^T_{T-1}(\pi_{T-1}) | \pi_{T-2}, Q_{T-2}] \right) \tag{10.64}$$

is continuous in (π_{T-2}, Q_{T-2}) (along a uniformly integrable sequence). Lemma 10.8.7 suggests that for $(\pi_{T-2,n}, Q_{T-2,n})$ a uniformly integrable converging sequence, converging to (π_{T-2}, Q_{T-2}), the term $J^T_{T-1}(\pi_{T-1}(m, \pi_{T-2,n}, Q_{T-2,n}))$ also converges for every m, and hence continuity is established. It can be shown that by Lemma 10.8.3, as $Q_n \to Q$, $\|\pi'(m, \pi, Q_n) - \pi'(m, \pi, Q)\|_{TV} \to 0$ for any quantizer Q with M cells of positive measure [434]. Continuity can be established even if the number of cells is less then M by a bounding argument: if a bin probability decreases to zero, so does the value function (note that an optimal quantizer cannot have less than M cells since by splitting a given cell into two leads to the existence of another cell, yet the value function strictly decreases). Thus, in the update equation

$$\begin{aligned} &E[J^T_{T-1}(\pi_{T-1}) | \pi_{T-2}, Q_{T-2}] \\ &= \sum_{m=1}^{M} P(\pi'(m, \pi_{T-2}, Q_{T-2}) | \pi_{T-2}, Q_{T-2}) J^T_{T-1}(\pi'(m, \pi_{T-2}, Q_{T-2}), \end{aligned}$$

$E[J^T_{T-1}(\pi_{T-1}) | \pi_{T-2}, Q_{T-2}]$ is continuous in π_{T-2}, Q_{T-2}. By the continuity of $c(\pi_{T-2}, Q_{T-2})$, we have that $c(\pi_{T-2}, Q_{T-2}) + E[J^T_{T-1}(\pi_{T-1}) | \pi_{T-2}, Q_{T-2}]$ is continuous in π_{T-2}, Q_{T-2}; the value function is continuous by Lemma 10.8.5. Thus, (10.64) admits a solution, and

$$J^T_{T-3}(\pi_{T-3}) = \min_Q \left(c(\pi_{T-3}, Q_{T-3}) + E[J^T_{T-2}(\pi_{T-2}) | \pi_{T-3}, Q_{T-3}] \right),$$

is continuous in π_{T-3}.

Continuing the same reasoning for the previous time stages, continuity and the existence of an optimal policy follows. □

10.9 Concluding Remarks

This chapter presented structural results on optimal causal coding of Markov sources in a large class of settings. The structural results are shown to feature a separation structure. For the optimal causal coding of a partially observed Markov source, the structure of the optimal causal coders is obtained and is shown to admit a separation structure. We observed in particular that *separation of estimation (conditional probability computation) and quantization (of this probability)* applies in such a setup. We also observed that optimal real-time decentralized coding of a partially observed i.i.d. source admits separation. Such a separation result does not, in general, extend to decentralized coding of partially observed Markov sources. The chapter has also established the existence of optimal control and quantization policies under some technical conditions.

The joint optimization of encoding and control policies for the LQG problem has also been studied in the chapter, and it has been shown that separation of estimation and control applies, an optimal quantizer exists under some technical assumptions on the space of policies considered, and the optimal control policy is linear in its conditional estimate.

The separation result presented in this chapter will likely find many applications in sensor networks and networked control problems where sensors have imperfect observation of a plant to be controlled. One direction still to explore is to find explicit results on the optimal policies using computational tools. One promising approach is expert-based systems, which are very effective once one imposes a structure on the designs; see [187] for details.

Theorem 10.3.4 motivates the problem of optimal quantization of probability measures. This remains an interesting problem to be investigated in a real-time coding context, with important practical consequences in control and economics applications. Toward this direction, Graf and Luschgy, in [167, 168], have studied the optimal quantization of probability measures.

10.10 Bibliographic Notes

Related papers on real-time coding include the following: [292] established that the optimal causal encoder minimizing the data rate subject to a distortion for an i.i.d. sequence is memoryless. If the source is kth-order Markov, then the optimal causal fixed-rate coder minimizing any measurable distortion uses only the last k source symbols, together with the current state at the receiver's memory [396]. Walrand and Varaiya [385] considered the optimal causal coding problem of finite-state Markov sources over noisy channels with feedback. Teneketzis [361] and Mahajan and Teneketzis [249] considered optimal causal coding of Markov sources over noisy channels without feedback. Mahajan and Teneketzis [248] considered the optimal causalcoding over a noisy channel with noisy feedback. Linder and Zamir [237] considered the causal coding of more general sources, stationary sources, under a high-rate assumption. An earlier reference on quantizer design is [108]. Relevant discussions on optimal quantization, randomized decisions, and optimal quantizer design can be found in [149, 438].

Borkar et al. [74] have studied a related problem of coding of a partially observed Markov source. This work also regarded actions as the *quantizer functions*. Nayyar and Teneketzis [289] considered within a multiterminal setup decentralized coding of correlated sources when the encoders observe conditionally independent messages given a finitely valued random variable and obtained separation results for the optimal encoders. Their paper also considers noisy channels. Some related studies include optimal control with multiple sensors and sequential decentralized hypothesis testing problems [375] and multi-access communications with feedback [8].

Existence of optimal quantizers for a one-stage cost problem has been investigated by Abaya and Wise [1], Pollard [309], and Yüksel and Linder [438]. For dynamic vector quantizers, Borkar, Mitter, and Tatikonda [74] obtained existence results for an infinite horizon setting. Mahajan and Teneketzis [250], Teneketzis [361], and Yüksel [418] considered zero-delay coding of Markov sources under various setups. Tatikonda et al. [358] considered general channels in the context of sequential rate distortion and established the result that uniform quantization is asymptotically optimal in the limit of large rates for quadratic distortion criteria. A similar discussion can be found in [427]. Linder and Zamir [237] considered causal coding of stationary sources in the limit of low distortion. Matveev and Savkin [262] established the existence of optimal coding and quantizer policies for the LQG setup under the assumption that the controller is memoryless.

There is a large literature on jointly optimal quantization for the LQG problem dating back to early 1960s (see, e.g., [108, 232]). References [42, 73, 139, 147, 262, 283, 358, 423] considered the optimal LQG quantization and control, with various results on the optimality or the lack of optimality of the separation principle. We also note that [425] provides a discussion for optimal quantization of control-free linear Gaussian systems. The LQG system analysis in this chapter builds primarily on [423].

Weissman and Merhav [389] considered optimal causal variable-rate coding under side information and [433] considered optimal variable-rate causal coding under distortion constraints.

In this chapter, we also presented structural results for optimal decentralized coding of i.i.d. sources, considered in [425]. There are algorithmic and asymptotic results available in the literature when the encoders satisfy the optimal structure obtained in the chapter; important contributions in this direction include [141, 172, 390].

A parallel line of consideration which is of a rate-distortion theoretic nature is the *sequential-rate distortion* proposed in [358] and the *feedforward* setup, which has been investigated in [129, 377].

This chapter is also related to Witsenhausen's indirect rate distortion problem [397] (see also [119]). Further related papers include [20, 39, 53, 119, 138, 201].

Related papers considering multiterminal information theory problems in a team theoretic angle include [99, 289].

Theorems 10.3.3–10.3.5 and 10.5.1 follow from [418,425]. Some of these results generalize the approaches in [385, 396]. Theorem 10.4.2 is due to [437]. Fischer [139], Nair et al. [283], and Tatikonda et al. [358] considered the optimal LQG quantization over general channels and established separation results; Theorem 10.6.2 has essentially appeared in [283, 423]. Lemma 10.6.1 is due to [423].

Chapter 11
Optimal Coding and Control for Linear Gaussian Systems Over Gaussian Channels Under Quadratic Cost

11.1 Introduction

In this chapter, we consider linear systems driven by Gaussian noise, which are controlled over Gaussian channels. We study the optimization of encoders and controllers for the minimization of quadratic costs for such linear and Gaussian systems (known as linear quadratic Gaussian (LQG) problems). We devote an entire chapter to such problems because Gaussian source and channel models are widely used in practice. Gaussian systems give rise to some of the most popular and easily implementable control and filtering algorithms in view of the LQG theory and Kalman filtering. Gaussian source and channel models are effective at capturing robustness, for a Gaussian distribution has unbounded support, an attribute that proved to play an important role in Part II of the book. Quadratic costs are practically important because such costs penalize unstable behavior.

We present in the chapter information theoretic results on communicating a Gaussian source over a Gaussian channel and obtain structural results for optimal encoders and controllers. We obtain conditions under which linear encoding and control policies are optimal and identify situations where such an optimality does not hold. We saw in Chap. 3, in the context of Witsenhausen's counterexample, that even LQG systems with nonclassical information structures may fail to admit linear policies as their optimal solutions. Yet, we also saw that a variation on the Witsenhausen's counterexample, with the cost function modified but still with nonclassical information, the Gaussian test channel problem, admits linear optimal policies. Two further important examples are presented in this chapter, which are similar to the spirit of the Witsenhausen's counterexample, exhibiting that optimality of linear encoding policies is a rather rare event.

In the chapter, Sect. 11.2 provides information theoretic ingredients for optimal joint source-channel coding over a Gaussian channel and discusses the notion of matching in the context of Gaussian systems. Sect. 11.3 considers the problem of jointly optimal LQG control and coding schemes for optimization, and Sect. 11.4 studies stabilization over Gaussian channels in various settings. Section 11.5

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_11,

presents the two examples where linear policies are suboptimal, and Sect. 11.6 discusses information theoretic bounds for system performance.

For background reading, we refer the reader to Sect. D.2 for a review of the LQG control problem and Kalman filtering.

11.2 Gaussian Source-Channel Pairs and Optimality of Linear Policies

In this section, we consider Gaussian channels and networks of Gaussian channels motivated by LQG problems.

11.2.1 Optimality of Linear Coding Policies over a Gaussian Channel with Matching Between the Source and the Channel

Converse to Shannon's Channel Coding Theorem

As seen earlier in Sect. 5.3.3, the *rate-distortion function*, $R(D)$, of a source is the minimum average amount of information that can be transmitted to satisfy a given level of average distortion D. The capacity, $C(P)$, of a channel with cost (power) constraint P, on the other hand, is the maximum amount of information which can be transmitted reliably per channel use given an average cost (power) constraint at the channel input.

These two notions lead to the following classical converse theorem in information theory [103]:

Theorem 11.2.1. *Let $\{x_t\}$ be a real ergodic information source and let a distortion function $\rho : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_+$ be given. This information source can be transmitted over a memoryless channel with capacity $C(P)$ with an average distortion less than or equal to D, that is, there exists a coding policy with*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=0}^{n-1} \rho(x_t, \hat{x}_t) \leq D,$$

where $\hat{x}_t$ is the decoder output, only if

$$R(D) < C(P).$$

◇

As a consequence of the above result, we obtain a lower bound on the achievable distortion over a channel with capacity $C(P)$ as $D \geq R^{-1}(C(P))$, where R^{-1}

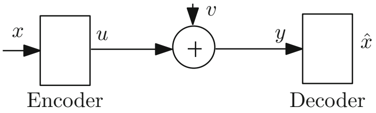

Fig. 11.1 The Gaussian source-channel pair belongs to a special class of pairs where source-channel matching occurs

denotes the inverse rate-distortion function (known also as the distortion-rate function; see Sect. 5.3.3).

The bound is in fact achievable arbitrarily closely, provided that the encoder and the decoder are allowed to operate with delay, that is, block codes are allowed. This is known as the source-channel separation theorem in information theory. However, in the context of delay-sensitive real-time systems, the above is typically a very loose bound, as discussed in Sect. 5.4. There are very few source and channel pairs which allow the above to be tight. The Gaussian source-channel pair is one such pair; we elaborate on this next.

11.2.2 The Gaussian Pair: Gaussian Sources and Channels

Optimality and Suboptimality of Linear Policies for Estimation over a Gaussian Channel

In Sect. 3.3.3, we considered the generalized Gaussian test channel. One of the findings was that for a scalar Gaussian source transmitted over a scalar Gaussian channel, an optimal encoder is linear, a derivation of which essentially utilized Theorem 11.2.1 above. Let x be a zero-mean real-valued Gaussian random variable transmitted over a scalar Gaussian channel with capacity C and let $y = u + v$, $u = \gamma_1(x)$, with u being the channel input and v a zero-mean Gaussian channel noise, independent of x and $\hat{x} = \gamma_2(y)$ (see Fig. 11.1).

Let the Gaussian channel have input power constraint P, such that $E[u^2] \leq P$ and the Gaussian channel noise have variance σ_v^2. The capacity then is given by $C = \frac{1}{2}\log_2(1 + (P/\sigma_v^2))$. In Chap. 3, we saw that the minimum estimation error variance D of a Gaussian source transmitted over a Gaussian channel with capacity C satisfies the following set of inequalities:

$$\begin{aligned} C(P) = \frac{1}{2}\log_2(1 + (P/\sigma_v^2)) \geq I(u;y) \\ \geq I(x;\hat{x}) \geq R(D) = \frac{1}{2}\log_2(E[x^2]/D). \end{aligned} \tag{11.1}$$

Furthermore, the above is tight and the minimum attainable distortion is given by $D = E[x^2]/2^{2C}$. This is achieved by a linear scaling of the input. Thus, what is known as the *matching principle* is applicable in this context:

Lemma 11.2.1 ([156]). *The equality $R(D) = C(P)$ holds if and only if:*

(a) The distribution P_u of $u = \gamma_1(x)$ achieves the capacity of the channel $P(dy|u)$ with input cost constraint $E[\eta(u)] \leq P$ (for some cost function η).
(b) The conditional distribution $P(d\hat{x}|x)$ with $\hat{x} = \gamma_2(y)$ given x achieves the rate-distortion function of the source x at distortion $D = E[\rho(x, \hat{x})]$.
(c) γ_1 and γ_2 are such that $I(u; y) = I(x; \hat{x})$.

◇

For the Gaussian source and channel case, with $\rho(\cdot, \cdot)$ a quadratic error function and η a quadratic function, we can take γ_1 and γ_2 to be linear. In this sense, we say that a Gaussian source is matched to a Gaussian channel. We summarize the result for the scalar Gaussian source-channel pair in the following.

Theorem 11.2.2 ([211]). *Consider the minimization of $E[(x - E[x|y])^2]$, subject to a power constraint on the input $E[u^2] \leq P$ over a Gaussian channel considered in Fig. 11.1. The optimal encoder-decoder pair is linear, and the minimum achievable cost is given by $E[x^2]/2^{2C}$. The encoder applies $z = \alpha x$, with $\alpha^2 = P/E[x^2]$. The optimal decoder policy is*

$$E[x|y] = \frac{P}{P + \sigma_v^2} \frac{1}{\alpha} y.$$

◇

11.2.3 Multi-Dimensional Source and Channels

We now consider a multidimensional source-channel pair. Consider the same setup as above, but now with $u \in \mathbb{R}^m$, $v \in \mathbb{R}^m$ with $u = (u^1, u^2, \cdots, u^m)$. Suppose further that the channel input is constrained to satisfy $E[u'u] \leq P$. The capacity for such a multidimensional Gaussian channel is given by what is known as the *water-filling* scheme. Toward that end, let us first define a multidimensional (parallel) Gaussian channel as a set of channels:

$$y^j = u^j + v^j, j = 1, 2, \ldots, m,$$

where $\{v^j\}$ is an independent set of random variables, with each v^j being zero-mean Gaussian with variance N_j. The problem of maximization of the mutual information $I(u; y)$ subject to the power constraint $E[u'u] \leq P$ admits a clean solution, given in the following theorem.

Theorem 11.2.3 ([103]). *The capacity of a multidimensional Gaussian channel is given by*

$$\sum_{i=1}^{m} \frac{1}{2} \log_2(1 + \frac{P_i}{N_i}),$$

with

$$P_i = \max(\lambda - N_i, 0),$$

where λ *is chosen such that* $\sum_{i=1}^{m} P_i = P$. ◇

Furthermore, the rate-distortion function for a multidimensional source also admits what is known as a *reverse water-filling* solution, as given in the following theorem.

Theorem 11.2.4 ([103]). *Let* x *be an* n*-dimensional, zero-mean Gaussian random vector with independent components, where component* i, x^i, *has a Gaussian distribution with variance* σ_i^2. *Suppose that the distortion criterion is* $\rho(x, \hat{x}) = \sum_{i=1}^{n}(x^i - \hat{x}^i)^2$, *where* $\hat{x}^i$ *denotes the conditional expectation of* x^i. *The rate-distortion function is given by*

$$R(D) = \sum_{i=1}^{n} \frac{1}{2} \log_2(\frac{\sigma_i^2}{D_i}),$$

where

$$D_i = \lambda 1_{\{\lambda < \sigma_i^2\}} + \sigma_i^2 1_{\{\lambda \geq \sigma_i^2\}}$$

and λ *is chosen such that* $\sum_{i=1}^{n} D_i = D$. ◇

Remark 11.2.1. In general, for such multidimensional Gaussian source-channel pairs, matching conditions do not hold. Even when the Gaussian source and channel dimensions are the same, matching conditions may not be realized. What is required is that $R(D) = C(P)$ and that the water-filling solutions giving rise to optimal input distributions are such that the rate-distortion achieving test channel is the effective channel from the estimator output to the source (after the encoding and the decoder mappings are performed) itself. To see this aspect, note that linear encoding may well achieve the capacity, but the effective channel from the source to the decoder output also needs to satisfy the rate-distortion constraint, which may not hold. One special case where such a matching holds is the case when the noise and signal power levels are identical in every channel and the distortion criterion is identical for all scalar components [308]. ◇

11.3 Joint Optimization of Encoder and Controllers for Linear Systems Controlled over Gaussian Channels

11.3.1 Problem Setup

We now consider a controlled scalar linear system of the form:

$$x_{t+1} = a x_t + b u_t + w_t, \tag{11.2}$$

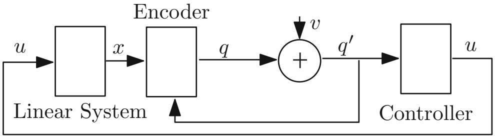

Fig. 11.2 Control over a Gaussian channel

where $\{w_t\}$ is an i.i.d. sequence of zero-mean Gaussian variables and x_0 is an independent Gaussian random variable.

The encoder transmits its information to a receiver/controller, over an additive Gaussian channel. We assume that the encoder has access to past channel outputs, that is, there is noiseless feedback. See Fig. 11.2.

In this context, we mean by a composite encoding policy Π^{comp} a sequence of functions $\{Q_t^{comp}, t \geq 0\}$ which are causal such that the coding output at time t, q_t, under Π^{comp} is generated by a causally measurable function of its local information, that is, a mapping measurable with respect to the sigma-algebra generated by $\mathcal{I}_t^e = \{x_{[0,t]}, q'_{[0,t-1]}\}, \quad t \geq 1$, and $\mathcal{I}_0^e = \{x_0\}$, to $\mathbb{R}$. The setup follows that of Chap. 10 (see Sect. 10.2), except that here the channel is a Gaussian channel. The channel is such that

$$q'_t = q_t + v_t, \tag{11.3}$$

where $\{v_t\}$ is an i.i.d. sequence of zero-mean Gaussian variables with variance σ_v^2. The controller's information at time t is $\mathcal{I}_t^c = \{q'_{[0,t]}\}$. The goal is to minimize

$$E[\sum_{t=0}^{T-1} Qx_t^2 + Ru_t^2], \tag{11.4}$$

with $R > 0, Q > 0$, subject to the constraint that $E[q_t^2] \leq P$, for some power constraint P. We note that if for all time stages, the encoder applies

$$q_t = \alpha_t(x_t - aE[x_{t-1}|\mathcal{I}_{t-1}^c]), \tag{11.5}$$

with $\alpha_t = \frac{P}{E[(x_t - aE[x_{t-1}|\mathcal{I}_{t-1}^c])^2]}$, the estimation error satisfies the recursion

$$E[(x_t - E[x_t|\mathcal{I}_t^c])^2] = \frac{E[a^2(x_{t-1} - E[x_{t-1}|\mathcal{I}_{t-1}^c])^2] + \sigma_w^2}{1 + (P/\sigma_v^2)}.$$

Upon recognizing the capacity expression in the denominator, $C = \frac{1}{2}\log_2(1 + (P/\sigma_v^2))$, we obtain the following: If $C > \log_2(|a|)$, stabilization in the sense of having $\lim_{t\to\infty} E[(x_t - E[x_t|\mathcal{I}_t^c])^2] < \infty$ is possible. This result will be refined in Theorem 11.3.2.

We next consider the optimization problem.

11.3.2 Optimality of Linear Policies

We first note that in the current context there is no dual effect of control, that is, the control actions do not affect the estimation errors for future time stages. The process x_t, conditioned on the past applied control actions, is always Gaussian. What is not clear in general, however, is whether conditioning on the past channel outputs leads to a Gaussian measure for the state process under optimal policies. We will consider this in this subsection.

Before we proceed, we present three important lemmas. The first lemma is the following.

Lemma 11.3.1. *Let x be zero-mean Gaussian and $z_{[0,t]}$ be any given collection of random variables for some $t \in \mathbb{Z}_+$. Then,*

$$E[(x - E[x|z_{[0,t]}])^2] \geq E[x^2]2^{-2I(x;z_{[0,t]})}.$$

This inequality is tight if $z_{[0,t]}$ is a Gaussian collection. ◇

Proof. By the data-processing inequality (see Lemma 5.3.1) and the fact that conditioning does not increase entropy, it follows that

$$\begin{aligned} I(x;z_{[0,t]}) &\geq I(x;E[x|z_{[0,t]}]) = h(x) - h(x|E[x|z_{[0,t]}]) \\ &= h(x) - h\left(x - E[x|z_{[0,t]}] \middle| E[x|z_{[0,t]}]\right) \geq h(x) - h(x - E[x|z_{[0,t]}]) \\ &\geq (1/2)\log_2\left(\frac{E[x^2]}{E[(x-E[x|z_{[0,t]}])^2]}\right), \end{aligned} \tag{11.6}$$

where the last inequality follows from the fact that among all random variables with a given variance, the Gaussian random variable has the largest differential entropy. Tightness for the Gaussian case follows from the following:

$$\begin{aligned} I(x;z_{[0,t]}) &= h(x) - h(x - E[x|z_{[0,t]}]|z_{[0,t]}) \\ &= h(x) - h(x - E[x|z_{[0,t]}]) \qquad (11.7) \\ &= \frac{1}{2}\log_2\left(\frac{E[x^2]}{E[(x-E[x|z_{[0,t]}])^2]}\right). \qquad (11.8) \end{aligned}$$

Here, (11.7) follows since for a Gaussian collection, $x - E[x|z_{[0,t]}]$ is also Gaussian and independent of $z_{[0,t]}$. □

The second lemma is the following (see [39, 440]).

Lemma 11.3.2. *Consider the scheme in Fig. 11.3, with $u = \gamma^e(y)$ for a measurable γ^e such that $E[u^2] \leq P$, x is zero-mean Gaussian and v, w are zero-mean Gaussian noise variables with variances σ_v^2, σ_w^2, respectively. Then $I(x;z)$ is maximized with*

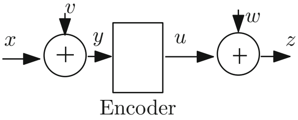

Fig. 11.3 Leak in information through a noisy observation channel

$u = \gamma^{e,*}(y) = \alpha y$ such that $E[u^2] = P$. Furthermore, under any coding policy γ^e,

$$E[(x - E[x|z])^2] \geq E[x^2]2^{-2\tilde{C}},$$

where

$$\tilde{C} = \frac{1}{2}\log_2(1 + \frac{\alpha^2 E[x^2]}{\alpha^2\sigma_v^2 + \sigma_w^2}).$$

Under $u = \gamma^{e,*}(y) = \alpha y$, the inequality is an equality. ◇

The proof can be established by noting that for any measurable γ^c

$$\begin{aligned} E[(x - \gamma^c(z))^2] &\geq E[(x - E[x|z])^2] \\ &= E[(x - E[x|y])^2] + E[(E[x|y] - E[x|z])^2], \end{aligned} \tag{11.9}$$

where the equality follows from similar arguments as in (10.21). The term $E[(E[x|y] - E[x|z])^2]$ is minimized by a linear policy in view of (11.1), since $E[x|y]$ is itself a Gaussian random variable. The final important ingredient is the following.

Lemma 11.3.3. *For the linear Gaussian system (11.2) controlled over a Gaussian channel (11.3), the cost (11.4) is equivalent to*

$$E\left[\sum_{t=0}^{T-1} \tilde{R}_t(u_t - \tilde{a}_t x_t)^2\right] + K_0 E[x_0^2] + \sum_{t=0}^{T-1} K_{t+1}\sigma_w^2, \tag{11.10}$$

where for $t = 0, \ldots, T-1$,

$$K_t = Q + a^2 K_{t+1} - \frac{ba^2 K_{t+1}}{K + b^2 K_{t+1}}, \quad \tilde{a}_t = -\frac{baK_{t+1}}{R + b^2 K_{t+1}}, \quad \tilde{R}_t = R + b^2 K_{t+1},$$

with $K_T = 0$. ◇

Hence, the problem reduces to a state estimation problem. Given these lemmas, one can show that optimal coding and control policies are linear, as we state in the following.

Theorem 11.3.1 ([41]). *Consider the linear Gaussian system (11.2) controlled over a Gaussian channel (11.3). For the minimization of (11.4), an optimal encoding policy is of the form*

$$q_t = \alpha_t(x_t - aE[x_{t-1}|\mathcal{I}_{t-1}^c]),$$

with

$$\alpha_t = \frac{P}{E[(x_t - aE[x_{t-1}|\mathcal{I}_{t-1}^c])^2]},$$

and the optimal control policy is given by

$$u_t = \tilde{a}_t E[x_t|\mathcal{I}_t^c],$$

where $\tilde{a}_t$ is given in (11.10). ◇

Proof. See Sect. 11.7.1. □

Theorem 11.3.1 does not generalize directly to multidimensional sources or channels since source-channel matching needs to take place at every time stage when the channel is used. This is a very restrictive condition in a practical setting. Partial results along this direction have been presented in [358].

In the appendix, we provide a proof for the following result, which is related to Theorem 11.3.1.

Theorem 11.3.2. *For the linear Gaussian system $x_{t+1} = ax_t + u_t + w_t$, with $\{w_t\}$ a zero-mean Gaussian i.i.d. sequence with variance σ_w^2, to satisfy*

$$\lim_{t\to\infty} E[x_t^2] \leq d,$$

over a memoryless (discrete or continuous) communication channel with noiseless feedback under some policy, the channel capacity must satisfy

$$C \geq \frac{1}{2}\log_2(\frac{a^2 d}{d - \sigma_w^2}).$$

These inequalities become equalities for a Gaussian channel and when optimal linear coding and control policies are adopted. ◇

Structural Results for Optimal Encoders for Partially Observed LQG Problems

In Sect. 10.6.1, we considered optimal coding and control over a discrete noiseless channel. Provided that noiseless feedback is available, the analysis there carries over to the case with Gaussian channels. Specifically, let $x_t \in \mathbb{R}^n, y_t \in \mathbb{R}^m$, and the evolution of the process be given by the following:

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad y_t = Cx_t + v_t^0. \tag{11.11}$$

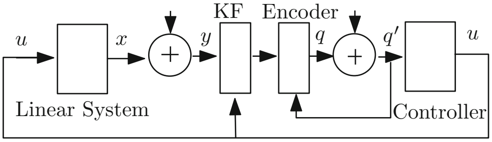

Fig. 11.4 LQG control of a partially observed linear Gaussian system over a Gaussian channel. Here, the channel encoder has access to the Gaussian channel output at the controller and KF denotes the Kalman filter. Separation of estimation and coding is optimal as in Theorem 11.3.3

Here, $\{w_t, v_t^0\}$ is a mutually independent, white zero-mean Gaussian noise sequence with $W = E[w_t w_t'], V = E[v_t^0 v_t^{0\prime}]$, A, B, C are matrices of appropriate dimensions. Suppose the goal is the minimization problem

$$\inf_{\Pi^{comp}} \inf_{\gamma} E_{\nu_0}^{\Pi^{comp},\gamma}[\sum_{t=0}^{T-1} x_t' Q x_t + u_t' Q u_t], \tag{11.12}$$

over all admissible coding and control policies, with ν_0 denoting a Gaussian distribution for the initial state, and $Q > 0, R > 0$, where now Π^{comp} generates the real-valued channel inputs using the Gaussian channel outputs in a causal manner under a given power constraint for every time stage. Building on Theorem 10.6.1, we have the following result (see Fig. 11.4).

Theorem 11.3.3. *For the minimization of the cost in (11.12) over all coding policies, a coder which uses the Kalman filter output and the information available at the receiver is as good as any other causal coder.* ◇

A particular application of this result is the encoding problem depicted in Fig. 11.3. Here the encoder uses its estimate without any loss, and as a consequence of (11.9), the optimal encoder is linear in its estimate. One further message here is that a partially observed Gaussian source is essentially not different from a fully observed Gaussian source.

11.4 Stabilization over Gaussian Channels and Sufficiency of Shannon Capacity Conditions

Consider an $\mathbb{R}^n$-valued linear Gaussian system of the form

$$x_{t+1} = Ax_t + Bu_t + w_t, \qquad y_t = x_t, \tag{11.13}$$

where we take B to be diagonal. The stabilizability of such a system over a Gaussian channel in the sense of ergodicity or the AMS property follows directly from Chap. 8: By the discussions in Sect. 5.3.1, the capacity of a Gaussian channel can be approximated arbitrarily well by capacities of a sequence of discrete channels obtained through quantization. As a consequence of the results in Sect. 8.4, Shannon capacity being greater than $\sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$ is a sufficient condition for the AMS stabilization of the linear system (11.13), where λ_i are the eigenvalues of A. However, not only the AMS property but also the positive Harris recurrence can be established in the special setting of scalar sources and scalar channels through the analysis in the previous section.

In this subsection, we see that this capacity condition is also sufficient for quadratic stability for a large class of settings. In particular, for stabilization across a scalar Gaussian channel with minimum capacity requirements, source-channel matching is not required. This argument is not surprising since the scalar Gaussian channel with noiseless feedback has its error exponent as infinite; see [133] and [334]. In fact, in these papers, Elias [133] and Schalkwijk and Kailath [334] use variations of the linear innovation coding scheme given in (11.5), discussed in the previous section. Thus, as a consequence of Theorem 8.4.4 and arguments in Chap. 8, the infinite exponent implies the existence of coding and control policies which lead to quadratic stability.

However, instead of the constructions in Chap. 8, simpler coding schemes (such as linear time-invariant, linear time-varying, or memoryless nonlinear policies) may also be sufficient in the context of linear Gaussian systems. In the following, we will show that this is the case. Toward this goal, we observe in the following that for a multidimensional system controlled over a scalar Gaussian channel, it is sufficient to use linear time-varying coding policies for sequential transmission of scalar components in a Jordan form representation of the system. The proof of the following theorem builds on first stabilizing the lower modes in a Jordan block and then regarding these as noise variables for the upper modes, once one moves up in the Jordan matrix.

Theorem 11.4.1 ([443]). *Consider the n-dimensional Gaussian linear system (11.13) to be controlled over a scalar Gaussian channel with noiseless feedback. If the channel capacity is greater than* $\sum_{|\lambda_i|>1} \log_2(|\lambda_i|)$*, then there exist coding and control schemes such that*

$$\limsup_{t\to\infty} E[|x_t|^2] < \infty.$$

Furthermore, a linear time-varying policy is sufficient through sequential linear encoding of scalar components. ◇

The above theorem can be generalized to a class of multidimensional sources and multidimensional channels through time-varying (periodic) linear policies. What is essential in such a construction is that every scalar mode of a linear system is to be transmitted over at most a single scalar channel at any given time: Consider m parallel Gaussian channels, each channel with individual rates $C_i, 1 \leq i \leq m$, so that the capacity of the parallel channel is $\sum_{i=1}^{m} C_i$. We have the following.

Theorem 11.4.2 ([445]). *The system (11.13) can be mean square stabilized over m parallel Gaussian channels (having independent noise variables) with feedback, using a linear time-varying policy if there exist $f_{ij} \in \mathbb{Q}$ such that $f_{ij} \geq 0$, $\sum_{j=1}^{m} f_{ij} \leq 1$, $\sum_{i=1}^{n} f_{ij} \leq 1$, and for all i*

$$\log(|\lambda_i|) < \sum_{j=1}^{m} f_{ij} C_j,$$

where λ_i are eigenvalues of the system matrix A in (11.13) and C_i is the capacity of the ith channel. ◇

In Theorem 11.4.2, the rational terms $\{f_{ij}\}$ dictate the period of the *time-sharing* policy for each of the modes.

One question of practical interest is the following: Can one achieve stabilization with time-invariant policies? In general, the answer is clearly negative as the following simple counterexample shows. Consider the system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t, \qquad t \geq 0, \\ y_t &= \gamma^e(x_t) + v_t\,, \end{aligned} \tag{11.14}$$

where the second equation describes the joint map from the encoder and the channel noise to the output. Consider now, with $|a| > 1$,

$$A = \text{diag}(a, a), \quad \gamma^e(x_t) = Cx_t, \quad C = [\alpha \quad \beta].$$

For this system, independent of the values of α and β, the corresponding noise-free system is unobservable, and therefore the system cannot be stabilized, regardless of the capacity of the channel. For this system a time-varying encoding scheme can be utilized to achieve stability, by transmitting different components of the state at alternating times such that $C_{2t} = [0 \quad \alpha_t]$, $C_{2t+1} = [\beta_t \quad 0]$ for $t \in \mathbb{Z}_+$, $\alpha_t, \beta_t \in \mathbb{R}$. If the channel is one dimensional and there is feedback, at every channel use, the channel can carry independent information. Furthermore, the encoder can adjust the signals to be transmitted in a time-varying fashion to ensure that the receiver's estimation error for every (open-loop) unstable mode is stable.

Remark 11.4.1. In Sect. 8.6.4 (see Theorem 8.6.7), stabilization of a linear system over a Gaussian forward channel with a noisy Gaussian feedback channel was considered. We observed that if the forward and the reverse channel capacities satisfy

$$2^{-2C_f} + 2^{-2C_r} - 2^{-2C_f - 2C_r} < 1/a^2,$$

then the steady-state variance is finite under linear memoryless policies. In particular, if the reverse channel is noise-free, $C_r = \infty$, the condition reduces to $C_f > \log_2(|a|)$. ◇

11.5 Two Counterexamples on Sub-optimality of Linear Policies

In the following, we will see that separation results involving optimality of linear policies do not necessarily extend to decentralized settings. We present two counterexamples for Gaussian source-channel pairs. Note that we have already seen that for multidimensional channels, linear policies may not be optimal.

In particular, the results obtained for scalar sources and channels do not extend to settings where there are two sensors and to settings when there is a relay in the system. These two are perhaps the most immediate extensions of the scalar setup considered. These, together with Witsenhausen's counterexample considered in Sect. 3.3.2, reveal that optimality of linear policies does not hold for a general class of Gaussian systems, and typically one needs to go beyond the linear structure to find optimal policies.

11.5.1 Gaussian Relay Channels with Two Encoders: Person-by-Person-Optimality of Linear Policies and Lack of Convexity of the Team Problem

Consider the transmission of a Gaussian source over a Gaussian relay channel, as depicted in Fig. 11.5. The relay is an intermediate encoder/decision maker which helps transmit the message from the source to the receiver.

We wish to minimize $E[(x-\hat{x})^2]$ over source encoder and relay encoder policies: We assume that the source x is Gaussian with zero mean and variance σ_x^2. The encoder mapping is γ_e, with $s_e = \gamma_e(x)$ satisfying $\mathbb{E}[s_e^2] \leq P_S$. The transmitted signal s_e is then observed with noise by the relay node as $y = s_e + v_e$, where v_e is a zero-mean independent Gaussian noise of variance N_e. The relay node applies a measurable mapping γ_r on the received signal to produce s_r under the following average relay power constraint: $\mathbb{E}[s_r^2] \leq P_R$. The signal s_r is then transmitted over a Gaussian channel. Accordingly the destination node receives $r = s_r + v_r$, where $\{v_r\}$ is a zero-mean white Gaussian noise with variance N_r. The decoder generates $\hat{x} = g(r)$.

In Lemmas 11.3.1 and 11.3.2, we saw that if the source encoder is restricted to be linear, the optimal relay policy is also linear. On the other hand, if the relay encoder policy is fixed to be linear and memoryless, then the problem becomes equivalent to the transmission of a Gaussian source over a Gaussian channel subject to an average

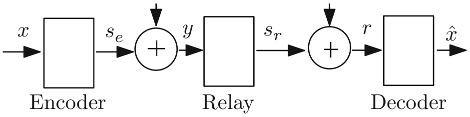

Fig. 11.5 Gaussian relay channel

power constraint, for which we already know that linear encoding is optimal. Hence, there exists a person-by-person optimal linear solution to the problem considered in Fig. 11.5 (existence can be seen easily by noting that with encoder, relay, and decoder all restricted to linear policies, the best policy in that class can be obtained by minimization of a continuous function on a closed and bounded subset of $\mathbb{R}^3$, and this globally optimal solution over the linear class is person-by-person optimal over the general class).

We have seen in Chap. 2 that in decentralized team optimization problems, person-by-person optimal solutions are globally optimal if the cost function is convex in the policies of the decision makers and the cost function satisfies appropriate differentiability conditions on the policies. For the problem at hand, however, such a structural result does not hold, as we discuss next. Let P be an observation channel from the input variable x at source encoder to the channel output variable r such that $P(\cdot|x)$ is a probability measure on the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ on $\mathbb{R}$ for every $x \in \mathbb{R}$, and $P(A|\cdot) : \mathbb{R} \to [0,1]$ is a Borel measurable function for every $A \in \mathcal{B}(\mathbb{R})$. Similarly we define P_1 as an observation channel from the variable x to the variable y and P_2 as an observation channel from the variable y to the variable r. From Chap. 4 (see Sect. 4.3), it follows that the distortion is concave in the joint observation channel $P(A|x) = \int_{\mathbb{R}} P_2(A|y)P_1(dy|x)$ for every $A \in \mathcal{B}(\mathbb{R})$, where the individual channels P_1 and P_2 are induced by the source and the relay encoding policies. This implies that person-by-person optimal encoding policies do not guarantee team optimality. We also note that, even under linear policies, the problem is not convex (see [302]).

In view of the discussion above, we now provide a simple counterexample for the problem depicted in Fig. 11.5 to show that linear policies are not optimal for causal transmission of a Gaussian source over the given relay channel (see [444]). First note that optimal linear policies lead to a cost of

$$D_L^\star = \sigma_x^2 \left(1 - \frac{P_S P_R}{(P_S + N_e)(P_R + N_r)}\right).$$

Consider now the following policies at the source encoder and the relay encoder, respectively:

$$\gamma_e(x) = \begin{cases} a, & \text{for} \quad x > m_1, \\ 0, & \text{for} \quad |x| \le m_1, \\ -a, & \text{for} \quad x < -m_1, \end{cases} \qquad \gamma_r(y) = \begin{cases} b, & \text{for} \quad y > m_2, \\ 0, & \text{for} \quad |y| \le m_2, \\ -b, & \text{for} \quad y < -m_2, \end{cases}$$

where the scalars $a, b, m_1, m_2 \in \mathbb{R}_+$ are to be specified. Under these policies, the signals observed at the relay and the destination are respectively given by

$$y = \begin{cases} a + v_e, & \text{for} \quad x > m_1, \\ v_e, & \text{for} \quad |x| \le m_1, \\ -a + v_e, & \text{for} \quad x < -m_1, \end{cases} \qquad r = \begin{cases} b + v_r, & \text{for} \quad y > m_2, \\ v_r, & \text{for} \quad |y| \le m_2, \\ -b + v_r, & \text{for} \quad y < -m_2, \end{cases}$$

The nonlinear policies above have to satisfy the average transmit power constraints. This leads to

$$a \le \sqrt{\frac{P_S}{2E(\frac{m_1}{\sigma_x})}}, \qquad b \le \sqrt{\frac{P_R}{2\kappa(m_1,m_2,a,\sigma_x,N_e)}},$$

where

$$\kappa(m_1, m_2, a, \sigma_x, N_e)$$
$$= \left(1-2E_q\left(\frac{m_1}{\sigma_x}\right)\right) E_q\left(\frac{m_2}{\sqrt{N_e}}\right) + E_q\left(\frac{m_1}{\sigma_x}\right)\left(E_q\left(\frac{m_2-a}{\sqrt{N_e}}\right) + E_q\left(\frac{m_2+a}{\sqrt{N_e}}\right)\right)$$

and

$$E_q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{\tau^2}{2}} d\tau.$$

Numerical computation [444] leads to the following: Letting $\sigma_x^2 = P_S = P_R = 1$, $N_e = N_d = 4$, $m_1 = 2.45$, and $m_2 = 6.84$, the performance of the given nonlinear policy is $D_{NL} = 0.926$, whereas the performance of the best linear policy is $D_L^\star = 0.96$.

Hence, for the setup in Fig. 11.5, we can state the following:

- If the relay is restricted to be linear, the optimal encoder is linear from information theoretic arguments. If the source encoder is restricted to be linear, the best relay encoding is linear (by Lemmas 11.3.1 and 11.3.2).
- The problem is non-convex when the encoders are viewed as stochastic kernels (as a consequence of Theorem 4.3.1). Hence, person-by-person optimality above does not necessarily imply optimality of linear policies.
- Indeed, policies optimal in the linear class are not globally optimal.

11.5.2 A Decentralized Sensing Problem over Vector Gaussian Channels

As a different setting where optimality of linear policies may not hold, this subsection studies a distributed sensing problem. Consider a two-sensor, single-controller, LTI system:

$$\begin{aligned} x_{t+1} &= ax_t + u_t + w_t\,, \\ y_t^i &= x_t + v_t^i\,, \quad i = 1, 2. \end{aligned} \tag{11.15}$$

Here, $x_t \in \mathbb{R}$ is the state of the system with the initial state x_0 a zero-mean Gaussian random variable with variance σ_x^2, $u_t \in \mathbb{R}$ is the control signal, and $y_t^i \in \mathbb{R}$ is the observation available at sensor station i at time t. Here $\{w_t, v_t^1, v_t^2\}$

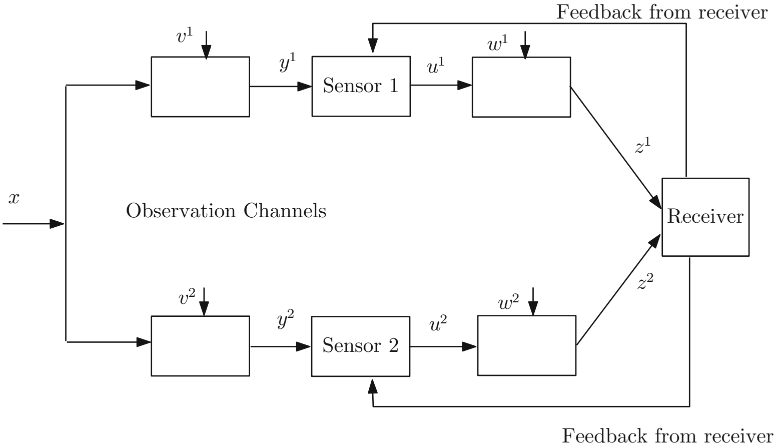

Fig. 11.6 Controller has access to (z_1, z_2)

are mutually independent, i.i.d. zero-mean Gaussian disturbance processes with variances $\{\sigma_w^2, \sigma_{v^1}^2, \sigma_{v^2}^2\}$.

The sensors transmit their signals at time t, which we denote by $\{u_t^1, u_t^2\}$, over two noisy Gaussian channels where the channel outputs

$$z_t^i = u_t^i + w_t^i, \quad i = 1, 2,$$

are received by the controller. Here w_t^i are zero-mean Gaussian random variables with variances $\{\sigma_{w^i}^2, i = 1, 2\}$. Upon observing the channel outputs, the controller generates its control input u_t.

In the following, we discuss the information structures under which the sensor signals $\{u_t^1, u_t^2\}$ and the control signal u_t are generated. A pictorial description for the sequence of events for a single stage is presented in Fig. 11.6. Here, the sensor action u_t^i $(i = 1, 2)$ is the output of a mapping γ^i measurable with respect to the sigma-field generated by $\mathcal{I}_t^i = \{y_t^i, \mathbf{z}_t\}$ and is a mapping to $\mathbb{R}$. Here,

$$\mathbf{z}_t = \{\mathbf{y}_{[0,t-1]}, \mathbf{z}_{[0,t-1]}\}$$

is the common past information, where z_t^1, z_t^2 are the information received by the controller. The controller policy γ^0 is measurable with respect to the sigma-algebra generated by $\mathcal{I}_t^c = \{\mathbf{z}_{[0,t-1]}\}$ and is a mapping to $\mathbb{R}$. We also have a power constraint on the sensors: $E[(u_t^i)^2] \le P^i, \quad i = 1, 2$. The objective is to minimize the following cost function:

$$J(\gamma^1, \gamma^2, \gamma^0) = E[\sum_{t=0}^{T} x_t^2 + Q(u_t)^2], \tag{11.16}$$

for some $Q > 0$, over all admissible sensor and control policies under the information structure presented above.

Theorem 11.5.1 ([441]). *For the minimization of the cost (11.16) under the specified information structure, linear policies are not necessarily optimal. In particular, there exists a nonlinear coding scheme which outperforms an optimal linear sensing set of policies for a particular instance of the problem.* ◇

Proof. See Sect. 11.7.3. □

11.6 Looseness of Information Theoretic (Cut-Set) Bounds for Gaussian Networks

We provide in this section a general discussion on the looseness of the bounds derived from the source-channel separation theorem considered in Theorem 11.2.1. Let us consider the configuration of 11.3, where v, w are Gaussian noise variables with strictly positive variances. We observed in Lemmas 11.3.1 and 11.3.2 that linear policies are optimal for quadratic error minimization. We note that under such policies

$$I(x;y) - I(x;z) = -h(x|y) + h(x|z) = -h(x|y,z) + h(x|z) = I(x;y|z) > 0.$$

Hence, it follows that the end to end mutual information is less than the mutual information in the first channel, unless the Markov chain condition holds

$$x \leftrightarrow z \leftrightarrow y,$$

which implies that there is no information loss with regard to the message at the second channel. Furthermore, it is evident then that

$$I(x;y) < \min(C_1, C_2),$$

where C_1 is the capacity of the additive first channel (where there is no encoding) and C_2 is the capacity of the channel between u and z. As a consequence,

$$E[(x - E[x|z])^2] > E[x^2]/2^{2\min(C_1,C_2)}.$$

Likewise, if we consider an encoder in the first channel as in the setup considered in Fig. 11.5, with strictly positive noise variances, we arrive at the following.

Theorem 11.6.1. *The mutual information satisfies*

$$\max_{P(s_e)} I(s_e;r) \leq \min(\max_{P(s_e)} I(s_e;y), \max_{P(s_r)} I(s_r;r)) < \min(C_3, C_4),$$

where C_3, C_4 *denote the Shannon capacities of the additive Gaussian channels between* s_e, y *and* s_r, r*, respectively, and* $P(s)$ *denotes the probability measure on the variable* s*. Hence, there is a leak in the end-to-end information transmission when the channels are noisy and*

$$E[(x - E[x|r])^2] > \frac{E[x^2]}{2^{2\min(C_1,C_2)}}.$$

◇

The above holds despite the fact that the information theoretic capacity of a block code is still equal to $C = \min(C_1, C_2)$.

The message in this subsection is that such information theoretic bounds (which are also known as *cut-set* bounds) may only be used as lower bounds on system performance in general. As with the case of optimality of linear policies, tightness of such bounds is also a rare phenomenon and is related to the matching discussion in Lemma 11.2.1.

11.7 Appendix: Proofs

11.7.1 Proof of Theorem 11.3.1

Given Lemma 11.3.3, the goal is to minimize the following expression in (11.10):

$$E\left[\sum_{t=0}^{T-1} \tilde{R}_t(u_t - \tilde{a}_t x_t)^2\right]. \tag{11.17}$$

Note that due to the lack of a dual effect, the control actions do not affect the estimation errors. Therefore, the optimal control policies are of the form

$$u_t = \tilde{a}_t E[x_t|\mathcal{I}_t^c].$$

As a consequence, (11.17) can be regarded as the sum of per-stage estimation errors. Since the control actions do not affect the estimation errors, we can instead consider a control-free process $\bar{x}_t = a\bar{x}_t + w_t$ and let u_t be the estimation of $\tilde{a}_t\bar{x}_t$. Hence, the goal is the minimization of

$$E\left[\sum_{t=0}^{T-1} \tilde{R}_t\left(\tilde{a}_t E[\bar{x}_t|\mathcal{I}_t^c] - \tilde{a}_t\bar{x}_t\right)^2\right].$$

We will first consider a two-stage problem. Note that

$$I(\bar{x}_1; q'_{[0,1]}) = I(\bar{x}_1; q'_1|q'_0) + I(\bar{x}_1; q'_0).$$

Properties of mutual information lead to

$$I(\bar{x}_1;q_1'|q_0') = h(q_1'|q_0') - h(q_1'|\bar{x}_1,q_0') = h(q_1'|q_0') - h(q_1'|\bar{x}_1,q_0',Q_1^{comp}(\bar{x}_1,q_0'))$$

$$= h(q_1'|q_0') - h(q_1'|Q_1^{comp}(\bar{x}_1,q_0')) \leq h(q_1') - h(q_1'|Q_1^{comp}(\bar{x}_1,q_0')) \tag{11.18}$$

$$\leq \frac{1}{2}\log_2(1+\frac{P}{\sigma_v^2}) =: C, \tag{11.19}$$

where C is the Gaussian channel capacity.

Furthermore, $I(\bar{x}_1;q_0')$ is also maximized by a linear policy at time 0: The relation $\bar{x}_1 = a\bar{x}_0 + w_0$ leads to $E[\bar{x}_0|\bar{x}_1] = \beta_1\bar{x}_1$ for some scalar β_1 so that we can write $z_0 = \beta_1\bar{x}_1 + \bar{v}_0$ where $\bar{x}_1 \leftrightarrow z_0 \leftrightarrow q_0'$ forms a Markov chain, $\bar{v}_0$ is Gaussian independent of $\bar{x}_1$, and z_0 has the same distribution as $\bar{x}_0$. Therefore, the setup reduces to the setting considered in Lemma 11.3.2 and a linear policy at time 0 also maximizes the mutual information $I(\bar{x}_1;q_0')$.

As a result, by Lemma 11.3.2, it follows that $I(\bar{x}_1;q_{[0,1]}') \leq C_1$, where $C_1 = C + \max I(\bar{x}_1;q_0')$. This upper bound is tight when the optimal encoding policy is of the form

$$q_t = \alpha_t(x_t - aE[x_{t-1}|\mathcal{I}_{t-1}^c]),$$

since Lemma 11.3.2 applies and (11.18)–(11.19) hold with equality since q_1' is independent of q_0' and is Gaussian.

By writing

$$I(\bar{x}_2;q_{[0,2]}') = I(\bar{x}_2;q_2'|q_{[0,1]}') + I(\bar{x}_2;q_{[0,1]}') \leq C + \max I(\bar{x}_2;q_{[0,1]}'),$$

the same reasoning applies. Hence, we can obtain an upper bound on $I(\bar{x}_t;q_{[0,t]}') = I(\bar{x}_t;q_t'|q_{[0,t-1]}') + I(\bar{x}_t;q_{[0,t-1]}')$, and the recursion is established for all time stages. Hence, the mutual information is maximized for all time stages.

Finally, by Lemma 11.3.1, we observe that the process $\{\bar{x}_t\}$ is zero mean and

$$E[(\bar{x}_t - E[\bar{x}_t|q_{[0,t]}'])^2] \geq E[\bar{x}_t^2]2^{-2I(\bar{x}_t;q_{[0,t]}')},$$

leading to a tight lower bound on the estimation errors which is achieved by linear policies. □

11.7.2 Proof of Theorem 11.3.2

Let $\sigma^2 = E[w^2]$ and $D_t = E[(x_t + \frac{b}{a}u_t)^2]$ and $d_t = E[x_t^2]$. It follows that $D_t = (d_{t+1} - \sigma^2)/a^2$ for $t \geq 0$. Consider a memoryless channel with inputs q_t and outputs q_t' and any deterministic control policy. Then,

$$\begin{aligned}
&\frac{1}{T}\sum_{t=0}^{T-1} I(x_t; q'_t|q'_{[0,t-1]}) = \frac{1}{T}\sum_{t=0}^{T-1} h(x_t|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \\
&= \frac{1}{T}\sum_{t=0}^{T-1} h(ax_{t-1} + bu_{t-1} + w_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \\
&= \frac{1}{T}\sum_{t=0}^{T-1} h(ax_{t-1} + w_{t-1}|q'_{[0,t-1]}) - h(x_t|q'_{[0,t]}) \\
&= \frac{1}{T}\bigg(\sum_{t=1}^{T} h(ax_{t-1} + w_{t-1}|q'_{[0,t-1]}) - h(x_{t-1}|q'_{[0,t-1]})\bigg) \\
&\qquad + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&\geq \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(2^{2h(ax_{t-1}|q'_{[0,t-1]})} + 2^{2h(w_{t-1})}) - h(x_{t-1}|q'_{[0,t-1]})\bigg) \\
&\qquad + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&= \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(2^{2h(a(x_{t-1}+(b/a)u_{t-1})|q'_{[0,t-1]})} + 2^{2h(w_{t-1})}) - h(x_{t-1}+(b/a)u_{t-1}|q'_{[0,t-1]})\bigg) \\
&\qquad + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&\geq \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(2^{2h(a(x_{t-1}+(b/a)u_{t-1}))} + 2^{2h(w_{t-1})}) - h(x_{t-1} + (b/a)u_{t-1})\bigg) \\
&\qquad + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&\geq \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(2\pi e(a^2 D_{t-1} + \sigma^2)) - \frac{1}{2}\log_2(2\pi e D_{t-1})\bigg) \\
&\qquad + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&= \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(\frac{(a^2 D_{t-1}+\sigma^2)}{D_{t-1}})\bigg) + \frac{1}{T}(-h(ax_{T-1}+w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&= \frac{1}{T}\bigg(\sum_{t=1}^{T} \frac{1}{2}\log_2(a^2 + \frac{\sigma^2}{D_{t-1}})\bigg) + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)).
\end{aligned}$$

Here, the first inequality follows from a conditional version of the entropy-power inequality (see Lemma 5.3.2) and the second and the third ones by the following argument: $\frac{1}{2}\log_2(a^2e^s + b) - s$, with $b > 0$, is a decreasing function of s; hence the entropy term

$$h(x_t|q'_{[0,t]}) = h(x_t + \frac{b}{a}u_t|q'_{[0,t]}) \leq h(x_t + \frac{b}{a}u_t)$$

can be replaced by its maximum, which is achieved by a zero-mean Gaussian random variable with the same second moment: D_t. Hence, for every $T \in \mathbb{N}$, we have that

$$\begin{aligned}
&\frac{1}{T}\sum_{t=0}^{T-1} I(x_t; q'_t|q'_{[0,t-1]}) \\
&\geq \frac{1}{T}\left(\sum_{t=1}^{T}\frac{1}{2}\log_2(a^2 + \frac{\sigma^2}{D_{t-1}})\right) + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) \\
&\geq \frac{1}{2}\log_2\left(a^2 + \frac{\sigma^2}{(\frac{1}{T}\sum_{t=1}^{T} D_{t-1})}\right) + \frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)),
\end{aligned}$$

where the last line follows from the convexity of $\log(1 + \frac{1}{x})$ in x. Observe now that

$$\lim_{T\to\infty}\frac{1}{T}(-h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) + h(x_0)) = 0,$$

since

$$\begin{aligned}
h(w_{T-1}) &= h(ax_{T-1} + w_{T-1}|x_{T-1}, q'_{[0,T-1]}) \\
&\leq h(ax_{T-1} + w_{T-1}|q'_{[0,T-1]}) \leq h(x_T) < M < \infty, \qquad (11.20)
\end{aligned}$$

for some $M \in \mathbb{R}_+$ (by the assumption that a limit exists for $E[x_t^2]$ and hence a uniform upper bound exists for the entropy sequence $\{h(x_t)\}$). Now, following the analysis in the proof of Theorem 8.5.2 (see (8.21)), a memoryless Gaussian channel is such that its capacity C satisfies

$$\begin{aligned}
C &\geq \limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1} I(x_t; q'_t|q'_{[0,t-1]}) \\
&\geq \limsup_{T\to\infty}\frac{1}{2}\log_2\left(a^2 + \frac{\sigma^2}{(\frac{1}{T}\sum_{t=0}^{T-1} D_t)}\right) \geq \frac{1}{2}\log_2\left(a^2 + \frac{\sigma^2}{D}\right), \qquad (11.21)
\end{aligned}$$

where the last inequality follows from the continuity of $\log_2(1 + \frac{1}{x})$ for $x > 0$.

Hence, a lower bound on the capacity is expressed as $C \geq \frac{1}{2}\log(\frac{(a^2D+\sigma^2)}{D})$ for the second moment to be less than d. This is tight for a Gaussian channel by linear encoding and control policies. □

11.7.3 Proof of Theorem 11.5.1

We take $T = 1$ and seek to minimize $E[\sum_{t=0}^{1} x_t^2 + Qu_t^2]$, under the given information structure. The cost is

$$E[Q(u_1)^2 + (ax_0 + u_0 + w_0)^2 + Q(u_0)^2 + (x_0)^2].$$

Clearly $u_1 = 0$, and using completion of squares,

$$\begin{aligned} J(\gamma^1, \gamma^2, \gamma^0) = E\Bigg[&Q(u_1)^2 + (\frac{a}{\sqrt{1+Q}}x_0 + \sqrt{1+Q}u_0)^2 \\ &+(a^2(1 - \frac{1}{1+Q}) + 1)(x_0)^2 + (w_0)^2\Bigg]. \end{aligned} \tag{11.22}$$

Observing the fact that the estimation error is orthogonal to the best estimate at the controller, the optimal control at time $t = 0$ can be evaluated as

$$u_0 = -\frac{a}{1+Q}E[x_0|z_0^1, z_0^2].$$

Hence, the total cost for $T = 1$ can be written as

$$\begin{aligned} J(\gamma^1, \gamma^2, \gamma^0) = &\frac{a^2}{1+Q}E_{\mathbf{z}_0}[(x_0 - E[x_0|\mathbf{z}_0])^2] \\ &+E[(x_0)^2](a^2(1 - \frac{1}{1+Q}) + 1) + E[(w_0)^2]. \end{aligned} \tag{11.23}$$

As such, the remaining issue becomes the minimization of the estimation error variance $E_{\mathbf{z}_0}[(x_0 - E[x_0|\mathbf{z}_0])^2]$.

In the following, we first compute the performance of the best linear sensing policies and introduce an alternative sensing scheme. We conclude the proof by a comparison of the performances under the two schemes.

The performance of the optimal linear coding and decoding can be computed from the performance of a minimum mean square error (MMSE) decoder ([311]): In particular, the estimation error reduces to

$$E[x^2] - [E[x^2]E[x^2]]A^{-1}[E[x^2]E[x^2]]', \tag{11.24}$$

where

$$A = \begin{bmatrix} E[x^2] + \sigma^2_{w'^1} & E[x^2] \\ E[x^2] & E[x^2] + \sigma^2_{w^{2'}} \end{bmatrix}$$

with the induced channel noise variances

$$\sigma^2_{w'^1} = \sigma^2_{v^1} + \frac{\sigma^2_{w^1}}{\frac{P_1}{E[x^2]+\sigma^2_{v^1}}}, \qquad \sigma^2_{w'^2} = \sigma^2_{v^2} + \frac{\sigma^2_{w^2}}{\frac{P_2}{E[x^2]+\sigma^2_{v^2}}}.$$

We now provide an alternative sensing scheme, which will lead to a better performance than the optimal linear policy. Toward obtaining the sensing scheme, we first revisit a relevant result from source-channel coding literature to guide the construction of the alternative coding scheme. In an information theoretic setup, for a distributed joint source-channel code to be optimal in the sense of minimum mean-square error, the following two conditions are sufficient (e.g., see [326]): (i) All channels send independent information and (ii) all channels utilize the capacity (source-channel needs to be matched).

Hence, one characteristic of an optimal code (in the information theoretic setup) is the transmission of independent information over the channels (in the linear case, the transmitted signals are inevitably correlated). Building on this insight, in the scheme that will be provided, one of the coders will transmit the magnitude of the signals they received at time 0 (we suppress the time index hereafter), and the other will transmit the sign of the signal. These two are independent but note that they do not satisfy the matching conditions with the channel (only a Gaussian source is matched to a Gaussian channel). The approach is to express the signal to be transmitted as $x = \tilde{u}^2 u^1$, with $\tilde{u}^2$ denoting the magnitude and u_1 denoting the sign of the random variable x. To minimize the power of the transmitted signal at the sensor, we write $\tilde{u}^2 = u^2 + \eta$, where η is $E[|x|]$ and only transmit u^2. As such, the decoder policy $\zeta : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ will be as follows:

$$\hat{x}(z^1, z^2) = \zeta(z^1, z^2) = (\hat{u}^2(z^2) + \eta)\hat{u}^1(z^1)$$

with $\hat{u}^2(z^2)$ being the best linear decoder estimate of the shifted magnitude of the source and $\hat{u}^1(z^1)$ the information regarding the sign of the source. We can write the corresponding estimation error variance as

$$\begin{aligned} E[(x-\hat{x})^2] = {} & Pr(u^1 \neq \hat{u}^1) E[(u^2 - \hat{u}^2)^2 | u^1 \neq \hat{u}^1] \\ & + Pr(u^1 = \hat{u}^1) E[(u^2 - \hat{u}^2)^2 | u^1 = \hat{u}^1]. \end{aligned} \tag{11.25}$$

Picking the values as $a = 1.2, Q = 0.005, P_1 = 10$, $P_2 = 40$, $E[x^2] = 5$, $E[(v^1)^2] = 0$, $E[(v^2)^2] = 0$, $E[(w^2)^2] = E[(w^1)^2] = 2, E[w^2] = 0.2$ one checks that with these values an upper bound on the cost J with the alternative coding scheme is 5.36, whereas the optimal linear scheme leads to a cost of 5.51. Hence, there exists a nonlinear scheme that outperforms the best linear policy. □

11.8 Concluding Remarks

This chapter has identified conditions for stabilization of a linear system over Gaussian channels and the conditions for optimality in quadratic cost minimization.

We have presented counterexamples to optimality of linear policies for two basic classes of Gaussian channels. The message is that when one departs from the setting of the transmission of scalar Gaussian sources over scalar Gaussian channels, linearity of optimal policies is a rather rare event unless *matching* occurs. We also observed that cut-set type bounds can, in general, only be used as lower bounds for system performance in real-time problems.

11.9 Bibliographic Notes

Regarding Gaussian source-channel pairs with noiseless feedback, optimality of linear coding policies for the scalar setting has been known since 1960s; see [133, 211] and [447]. The transmission over scalar Gaussian channels has been considered also in [133,334] and [152], where the error exponents have been shown to be unbounded (and the error probability has been shown to decrease at least doubly exponentially in the block length). Such a strong result, however, does not hold when there is noisy feedback [216]. The case where there is a mismatch in the source and the channel is a challenging one. Partial results are known regarding optimal zero-delay policies. Regarding the source-channel matching principle, we refer the reader to [52,107,156,355] and [358] for related discussions. For Gaussian source and channel pairs, multidimensional source and multichannel systems have been considered in [228] and [308], and optimal linear encoders have been studied in [21, 23, 228] and [5]. Partially observed settings have been considered in [20, 53, 119], and [138]. Gastpar [155] has considered various settings of multi-access and broadcast channels. A comprehensive discussion and literature review is available at [443].

On the relationship between the minimum mean-square estimation error and mutual information, we further mention the following results (which exhibit that the relation goes beyond the matching principle considered earlier). Duncan [126] has observed that for a general continuous-time stochastic process x_t (not necessarily Gaussian), observed through an additive Gaussian channel in the form $dy_t = x_t dt + dw_t$, with $\{w_t\}$ denoting a standard Brownian process, the estimation error satisfies a relationship of the form

$$(1/2)\int_0^T E[(x_t - E[x_t|y_{[0,t]}])^2]dt = \int_0^T I(x_{[0,t]}; y_{[0,t]})dt,$$

with the integration involving a continuous-time generalization of mutual information. Weissman, Kim, and Permuter [388] considered an extension of this result involving settings where the process x_t is affected by feedback and showed that the result holds provided that directed information replaces mutual information.

Along similar lines, Başar and Bansal [27] have considered the continuous-time version of the optimal coding and control problem for quadratic cost functions, establishing an analogous result to Theorem 11.3.1 and the optimality of linear policies for scalar Gaussian source-channel pairs.

Functional and topological properties of MMSE decoding have been considered in [409]. References [339,358] and [92] discussed this for vector sources in a control setting. References [92] and [39] consider the partially observed formulation.

For stabilization of vector sources over scalar and parallel channels, linear time-invariant encoding schemes have been considered in [80, 339] for the case without system/process noise and it is established that for scalar channels such policies are stabilizing for a large class of settings in view of rate requirements provided that the system is observable under a time-invariant policy. Stronger results on linear time-invariant policies have been established for scalar systems in [145]. For noise-free settings, [220] and [221] have considered nonlinear policies to establish the sufficiency of capacity bounds for quadratic stabilization.

A counterexample to optimality of linear policies over relay channels was presented by Lipsa and Martins [238] for a sequence of relay channels with the number of relays being greater than 2. Zaidi et al. [444] generalized this also to the case of a single relay. When there is no relay, linear policies are optimal as the problem is one involving the Gaussian test channel. Sufficient conditions for stabilization of linear systems over relay channels has been considered also in [222].

We observed that optimal (encoding) policies may not be linear among all possible policies when the source or channels are not scalar. Nonetheless, it may be of practical interest to obtain the best linear policies. A dynamic programming over the space of matrices can be carried out, leading to a family of nonlinear optimization problems. Such an analysis has been carried out in [27] for continuous-time systems.

Part of this chapter is based on [39, 41, 441, 443, 444], and [445].

Chapter 12
Agreement in Teams and the Dynamic Programming Approach Under Information Constraints

12.1 Introduction

As discussed extensively in Chaps. 2 and 3, at the heart of a decentralized decision problem are the delineation of an information structure, identification of a loss functional, and a probabilistic description of the unknown quantities (the so-called states of nature).

One of the important messages in those two chapters was that some information structures lead to computationally efficient derivations or programs for generation of optimal decision policies regardless of the nature of loss functions or probabilistic structure of the underlying system, but for some other information structures loss functions and probabilistic structures also play a role in tractability. Particularly, with nonclassical information structures, one may run into intractable problems as far as the derivation of optimal decision rules goes, with one prime example being Witsenhausen's counterexample.

In this chapter, we will further elaborate on the issue of tractability and show that in a dynamical decentralized system, agreement on certain variables through communication and availability of some common information are essential for the development of a systematic program to obtain optimal solutions. Agreement on common information can be useful for a team policy to be generated, or a hierarchically higher decision maker can organize the overall system according to an acceptable or optimal performance. For example, in the context of stochastic control theory (see Appendix D), for a large class of decentralized decision problems, all one needs to do is construct a related Markov decision problem (MDP) (with an appropriate Markovian state in a stage-independent state space and a cost function) and work with such a chain if the information structure allows for a recursive (iterative) derivation and computation using dynamic programming. The intractability of some (but not all) decentralized control problems stems from the fact that it is not possible to construct such a chain in a general setting. An important relevant question entails the communication requirements for generating such an MDP which leads to a tractable computation for optimal policies. This chapter addresses such problems.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4_12,

In this chapter, Sect. 12.2 discusses the concepts of common knowledge and agreement. Section 12.3 introduces the dynamic programming approach to team decision problems, and Sect. 12.4 introduces the belief sharing information pattern in view of developing optimal team policies for a dynamic setting and obtains information requirements for such a pattern. Section 12.5 introduces a team cost-rate function, which captures the trade-off between the information rate and the optimal cost attained under an optimal policy (or the infimum cost over all policies given a total information rate constraint).

12.2 Common Knowledge and Agreement

12.2.1 Common Knowledge

We begin our discussion by revisiting an important notion introduced by Aumann [19]. Let $(\Omega, \mathcal{F}, P)$ be a probability space, that is, Ω a sample space, $\mathcal{F}$ a σ-field of subsets of Ω, and P a probability measure, where we first assume that Ω is a finite space. Let $E \in \mathcal{F}$ denote an event and further let two information variables measurable on this probability space be available to two decision makers DM 1 and DM 2 such that $\mathcal{I}^1$ is available at DM 1 and $\mathcal{I}^2$ is available to DM 2.

Consider a finite space setting for Ω. Let $\mathcal{F}^i = \sigma(\mathcal{I}^i)$ denote the sigma-field generated by $\mathcal{I}^i$, $i = 1, 2$ (hence $\mathcal{F}^i$ is generated by a partition of Ω). Let us say that DM i knows $E \in \mathcal{F}$ at ω (and denote this by $\{\omega \in E\} \in \mathcal{F}^1$) if there exists a set $B^i \in \mathcal{F}^i$ such that $\omega \in B^i \subset E$. Assume that the partitions of each decision maker (induced by the local information variables) are known by both decision makers. Aumann introduced the notion of an event E being *common knowledge* between DM 1 and DM 2 at $\omega \in E$ as follows: Whenever $\omega \in E$ happens, DM 1 knows E, DM 2 knows E, DM 1 knows that DM 2 knows E, DM 2 knows that DM 1 knows E, and so on. An event E is common knowledge if it is common knowledge for all $\omega \in E$.

Aumann made the following characterization of common knowledge based on information fields: If $E \in \mathcal{F}^1 \cap \mathcal{F}^2$, then it is common knowledge. This is not only sufficient but also necessary: Suppose that E is common knowledge and some $\omega \in E$ takes place. As stated earlier, the event that DM 1 knows E at ω means that there exists some set $B \in \mathcal{F}^1$ such that $\omega \in B \subset E$. Now, if E is common knowledge, it must be that $\{\omega : \{\omega \in E\} \in \mathcal{F}^1\} \in \mathcal{F}^2$, that is, DM 2 knows that DM 1 knows that E took place, for all $\omega \in E$. For this to happen for every $\omega \in E$, it must be that there exists a finite index set T_2 such that $E = \cup_{t \in T_2} E^2_t$, where $E^2_t \in \mathcal{F}^2, t \in T_2$ and for all $t \in T_2$, $E^2_t \subset E$ (for, otherwise, there would exist some $\omega \in E$ such that DM 2 could not know if DM 1 knows whether E happened or not). A parallel argument applies for DM 1 for another finite index set. Hence, $E \in \mathcal{F}^1 \cap \mathcal{F}^2$. Thus, an event $E \in \mathcal{F}$ is common knowledge *only if* $E \in \sigma(\mathcal{I}^1) \cap \sigma(\mathcal{I}^2)$. The *if* direction follows from the observation that $\{\{\omega \in E\} \in \mathcal{F}^1\} \in \mathcal{F}^2$ (since $\{\omega : \{\omega \in E\} \in \mathcal{F}^1\} = E$ and $E \in \mathcal{F}^2$) and $\{\omega : \omega \in \{\{\omega \in E\} \in \mathcal{F}^1\} \in \mathcal{F}^2\} \in \mathcal{F}^1$ and so on, for every such finite iteration.

Hence, an event $E \in \mathcal{F}$ is common knowledge if and only if $E \in \sigma(\mathcal{I}^1) \cap \sigma(\mathcal{I}^2)$. We note that this relation does not depend on the particular probability measure P.

This equivalence between the definition of common knowledge and the sigma-field characterization can be extended to cases where Ω is an uncountable space. However, one technical issue is that sets (of one knows that the other knows) of the form $\{\{\omega \in E\} \in \sigma(\mathcal{I}^1)\}$ may not be an event in $\mathcal{F}$, since this may involve uncountable unions. By incorporating the probability measure P, Nielsen [293] has provided an approach by defining an equivalence class among the elements in $\mathcal{F}$ by excluding null events, allowing a generalized definition of Aumann to apply for uncountable settings. Brandenburger and Dekel [79] have considered *common knowledge with probability 1*, through a completion of sigma-fields by adding null events to the information sets of decision makers (note that if one partitions Ω into measurable sets with positive probability, the partition is a countable one).

Hereafter, we will adopt the definition of common knowledge based on the sigma-field characterization. That is, an event $E \in \mathcal{F}$ is common knowledge if $E \in \sigma(\mathcal{I}^1) \cap \sigma(\mathcal{I}^2)$. In most cases, however, common knowledge with probability 1 is sufficient since null events do not affect the expected costs.

Aumann's main result is that if decision makers have a common prior measure and if their posterior probability measures on some event is common knowledge, then the posteriors should be equal. In the following subsection, we will discuss this result in a more general context.

12.2.2 Asymptotic Agreement with Common Priors but Different Posteriors

If the posteriors of two DMs are not common knowledge and yet the DMs have a common prior, then a common posterior can be obtained through communication, a topic which we will consider further in this chapter. In control and economics literatures, agreement on such a common knowledge has typically been obtained through iterative exchange of decisions.

Let X be an integrable random variable (so that $E[|X|] < \infty$), such as 1_A, with A denoting an event. Suppose two decision makers DM 1 and DM 2 have access to some local random variables defined on a common probability space and correlated with X, and exchange their conditional probability measures over time. Suppose further that:

- The information σ-fields at each decision maker is increasing: $\mathcal{F}_t^i \subset \mathcal{F}_{t+1}^i$.
- For all $n \in \mathbb{N}$, there exists $m > n$ such that $\mathcal{F}_m^i$ contains information on $E[X|\mathcal{F}_n^j]$, $i, j = 1, 2$. That is, the decision makers exchange their estimates (but not their raw data) infinitely often.

Note that by the smoothing property of conditional expectation,

$$E[E[X|\mathcal{F}^i_{n+1}|\mathcal{F}^i_n] = E[X|\mathcal{F}^i_n],$$

and that by Jensen's inequality $\sup_n E[|E[X|\mathcal{F}^i_n]|] \leq E[E[|X||\mathcal{F}^i_n]] = E[|X|] < \infty$. Hence, $(E[X|\mathcal{F}^i_n], \mathcal{F}^i_n)$ is a martingale sequence and the submartingale convergence theorem (Theorem C.2.2) leads to the conclusion that $\lim_{n\to\infty} E[X|\mathcal{F}^i_n]$ exists almost surely for $i = 1, 2$. Let $\mathcal{F}^i_\infty$ denote the smallest sigma-algebra containing $\mathcal{F}^i_n$ for all $n \in \mathbb{N}$. Hence, the almost sure limit is $E[X|\mathcal{F}^i_\infty]$. Furthermore, $E[X|\mathcal{F}^i_n]$ is a uniformly integrable martingale sequence since X is integrable (see Theorem 3.3.2 in [70]) and as a consequence one also has that $E[|E[X|\mathcal{F}^i_n] - E[X|\mathcal{F}^i_\infty]|] \to 0$.

Since $E[X|\mathcal{F}^j_n]$ is $\mathcal{F}^i_m$ measurable, it is also $\mathcal{F}^i_\infty$ measurable and likewise for all n. Hence $E[X|\mathcal{F}^j_\infty]$ is $\mathcal{F}^i_\infty$ measurable, and $E[X|\mathcal{F}^i_\infty]$ is measurable on both $\mathcal{F}^1_\infty$ and $\mathcal{F}^2_\infty$. Thus, it is measurable on $\cap_{i=1,2}\mathcal{F}^i_\infty$. Hence the expectations are asymptotically equal and are common knowledge. A similar reasoning applies for a larger number of decision makers as well. Such a reasoning has been considered by Borkar and Varaiya [75] (see also Geanakoplos and Polemarchakis [159], Tsitsiklis and Athans [367], Li and Başar [233], and Teneketzis and Varaiya [362], among other contributions and approaches in the literature).

At this point, it is also worth stating that exchanging expectations is information theoretically inefficient. We will discuss this issue further in the following subsections.

12.2.3 Inconsistent Priors (Probability Models), Lack of Agreement and Merging

Common knowledge is a notion which is crucially important in team problems. In a team setting, if a prior measure on the event space is not common knowledge, the problem ceases to be a team problem, since the decision makers effectively solve different optimization problems, transforming the problem to a game problem: To make this more explicit, let Q denote a measurement channel as in Chap. 4 and define the following cost:

$$E^\gamma_\pi[c(x,u)] = \int \pi(dx)Q(dy|x)c(x,\gamma(y)).$$

Now, given a fixed policy γ, the costs $E^\gamma_\pi[c(x,u)]$ and $E^\gamma_\mu[c(x,u)]$ lead to different values since the prior measures π and μ are different [24]. Therefore $E^\gamma_\cdot[c(x,u)]$ defines a (measurable) function on the space of priors (under weak convergence), leading to possibly different values for different priors given a fixed policy γ. In particular, with multiple DMs present, a cost function to be minimized will not be a common cost when viewed as a function of team policies.

This then brings up the question of what the DMs can do in the absence of such common priors.

Consider a two-DM setup. If the priors are different, *the fact that they are different* may or may not be common knowledge. The DMs may run some agreement protocol to exchange information. In such a case, the following may take place:

- The DMs may run an agreement protocol assuming that the other DM has the same prior as himself [362] or have a probabilistic belief on the prior of the other DM [89]. In these cases, the DMs may realize that their probability models are inconsistent and yet reach an agreement or may not realize that the models are inconsistent but reach an agreement or reach different conclusions. In other scenarios the DMs may *agree to disagree*, that is, they may realize that their models are different, and a communication exchange protocol will not provide any further information that would change the agreement status or lack of it (see Teneketzis and Varaiya [362] and Teneketzis and Castanon [89] for further discussions).
- The DMs may run Bayesian estimators and asymptotically achieve common knowledge in the more restricted sense of agreement for events in the future (which Blackwell and Dubins [63] refer to as *merging*): When DMs have inconsistent priors, they can update their probability measures for events taking place as a function of future random variables with increasing information, and in some settings the learned information can overrule the priors leading to a form of consistency. Blackwell and Dubins [63] and Kalai and Lehrer [212] have observed that a sufficient condition for agreement on such conditional probabilities in the case of strictly common (knowledge) observations is that the decision makers' prior beliefs are absolutely continuous with respect to each other (i.e., each of the DMs assigns a positive prior probability to an event on which the other DM also has a positive prior probability). However, in the cases with controlled observations, such a condition has been shown to be too restrictive for many applications [143, 277]. The reader is also referred to Diaconis and Freedman [117] on asymptotic consistency of inconsistent priors and the importance of prior selection for uncontrolled observations. For a related information theoretic angle, see Barron [47].

The notion of common knowledge, beyond agreement on a common probability measure, is a subtle notion. Before we end this discussion, we provide an intriguing example, considered by Littlewood (see Geanakoplos [158]): Suppose there are three students sitting in a circle, each wearing either a red hat or a white hat (each only knowing the colors of the other students' hats). Suppose that all the hats are actually red (and the students do not know that), and a teacher asks the students what color their hats are. No student can provide a conclusive answer since there is not sufficient information ruling out the possibility that their hats are not red, or not white. However, if the teacher reveals that there is at least one red hat in the circle, and reveals this to everyone, so that this is common knowledge, and sequentially asks the students whether their hats are red or not, the first two students cannot give a conclusive answer, but the third one, given the inconclusive answers of the first two students, can deduce that her hat is red by taking into account all the possible answers of the first two students.

In the remaining of the chapter, we will primarily consider settings where the DMs have common priors (and consistent probability models) but possibly different posteriors.

12.2.4 Agreement in Finite Time Over Noisy Channels

Even though asymptotic agreement results require in general an unbounded time for resolution, it may be the case that DMs in a team can agree in finite time over a communication channel by incorporating encoding and decoding rules.

Consider two decision makers who communicate over a noisy channel with feedback, with the goal of agreement on a certain variable which can take finitely many values. This variable may be, for example, synchronization of encoding times and agreement on zooming times considered earlier in Chaps. 8 and 9. Even though (almost sure) agreement over such a noisy channel typically takes an infinite amount of time, one may achieve agreement in finite expected time for a class of channels. If the following assumption holds, then such agreements are possible in finite expected time: The channel is such that there exist input letters x_1, x_2, x_3, x_4 where $D(P(\cdot|x_1)||P(\cdot|x_2)) = \infty$ and $D(P(\cdot|x_3)||P(\cdot|x_4)) = \infty$, where $D(\cdots||\cdots)$ denotes the divergence between two input probability measures (see Definition 5.3.4). Here, x_1 can be equal to x_4 and x_2 can be equal to x_3, and as an example, the erasure channel satisfies this property. The argument here essentially follows from [83].

If there is no feedback or if there is noisy feedback, achieving agreement in finite time requires conditions which are similar to having a strictly positive zero-error capacity [103], which is equivalent to the condition that at least two messages can be distinguished unambiguously for some finite block length n.

Hence, one may achieve common knowledge even through noisy channels provided that the channels are reliable enough in the context described above. These may, for example, allow for the application of the random-time state-dependent stochastic drift arguments (see Theorem 6.2.4) to settings where a system's stabilization is conditioned on agreement between various decision makers communicating over noisy channels. These may also allow for the dynamic programming principle, considered in the following section, to be applicable.

12.3 Common Knowledge as Information State and the Dynamic Programming Approach to Team Decision Problems

In a team problem, if all the random information at any given decision maker is common knowledge between all decision makers, then the system is essentially centralized. If only some of the system variables are common knowledge, the remaining unknowns may or may not lead to a computationally tractable program generating an optimal solution. A possible approach toward establishing a tractable program is through the construction of a controlled Markov chain where the controlled Markov state may now live in a larger state space (e.g., a space of probability measures) and the actions are elements in possibly function spaces. This controlled Markov construction may lead to a computation of optimal policies.

Such a *dynamic programming approach* has been adopted extensively in the literature (see, e.g., [4,18,95,287,413,417] and generalized in [288]) through the use of a team policy which uses common information to generate partial functions for each DM to generate their actions using local information. Thus, in the dynamic programming approach, a separation of team decision policies in the form of a two-tier architecture, *a higher-level controller* and a *lower-level controller*, can be established with the use of common knowledge.

In the following, we present the ingredients of such an approach, as formalized by Nayyar, Mahajan and Teneketzis [288] and termed *the common information approach*:

1. Elimination of irrelevant information at the DMs: In this step, irrelevant local information at the DMs, say DM k, is identified as follows. By letting the policy at other DMs to be arbitrary, the policy of DM k can be optimized as a best-response function, and irrelevant data at DM k can be removed. Theorems 10.3.1 and 10.3.3 are examples of this step.
2. Construction of a coordinated system: This step identifies the common information and local/private information at the DMs, after Step 1 above has been carried out. A *fictitious coordinator* (*higher-level controller*) uses the common information to generate team policies, which in turn dictates the (*lower-level*) DMs what to do with their local information. The construction of quantization policies in Chap. 5.2.2 and in (10.3) are examples of this step.
3. Formulation of the cost function as a partially observed Markov decision process (POMDP), in view of the coordinator's optimal control problem: A fundamental result in stochastic control is that the problem of optimal control of a partially observed Markov chain (with additive per-stage costs) can be solved by turning the problem into a fully observed one on a larger state space where the state is replaced by the "belief" on the state (see Appendix D). Theorems 10.3.2, 10.3.4, and 10.5.1 are examples of this step.

4. Solution of the POMDP leads to the structural results for the coordinator to generate optimal team policies, which in turn dictates the DMs what actions to take given their local information realizations. Theorem 10.5.1 is an example of this step.
5. Establishment of the equivalence between the solution obtained and the original problem and translation of the optimal policies. Any coordination strategy can be realized in the original system. Note that, even though there is no real coordinator, such a coordination can be realized implicitly, due to the presence of common information.

In addition to the examples in Chap. 10, we will provide a further explicit setting with such a recipe at work, in the context of the *k-stage periodic belief sharing pattern* in the next section. In particular, Lemmas 12.4.1 and 12.4.2 will highlight this approach.

When a given information structure does not allow for the construction of a controlled Markov chain even in a larger, but fixed for all time stages, state space, one question that can be raised is what information requirements would lead to such a structure. We will also investigate this problem in the context of the *one-stage belief sharing pattern* in the next section.

12.4 k-Stage Periodic Belief Sharing Pattern and Communication Requirements

In this section, we will use the term *belief* for a probability measure-valued random variable. This terminology has been used particularly in the artificial intelligence and computer science communities, which we adopt here. We will, however, make precise what we mean by such a belief process in the following.

12.4.1 k-Stage Periodic Belief Sharing Pattern

As mentioned earlier and discussed in detail in Appendix D, a fundamental result in stochastic control is that the problem of optimal control of a partially observed Markov chain can be solved by turning the problem into a fully observed one on a larger state space where the state is replaced by the belief on the state. Such an approach is very effective in the centralized setting; in a decentralized setting, however, the notion of a state requires further specification. In the following, we illustrate this approach under the k-step periodic belief sharing information pattern.

Consider a joint process $\{x_t, y_t, t \in \mathbb{Z}_+\}$, where we assume for simplicity that the spaces where x_t, y_t take values from are finite-dimensional real-valued or countable. They are generated by

$$x_{t+1} = f(x_t, u_t^1, \ldots, u_t^L, w_t),$$

$$y_t^i = g(x_t, v_t^i),$$

where x_t is the state, $u_t^i \in \mathbb{U}^i$ is the control action, and $(w_t, v_t^i, 1 \leq i \leq L)$ are second order, zero-mean, mutually independent, i.i.d. noise processes. We also assume that the state noise, w_t, either has a probability mass function or a probability measure with a density function. To minimize the notational clutter, $P(x)$ will denote the probability mass function for discrete-valued spaces or probability density function for continuous spaces.

Suppose that there is a common information vector $\mathcal{I}_t^c$ at some time t, which is available to all the decision makers. At times $ks-1$, with $k > 0$ fixed, and $s \in \mathbb{Z}_+$, the decision makers share all their information: $\mathcal{I}_{ks-1}^c = \{\mathbf{y}_{[0,ks-1]}, \mathbf{u}_{[0,ks-1]}\}$ and for $\mathcal{I}_0^c = \{P(x_0)\}$, that is, at time 0 the DMs have the same *a priori* belief on the initial state. Hence, at time t, DM i has access to $\{y_{[ks,t]}^i, \mathcal{I}_{ks-1}^c\}$.

Until the next *common* observation instant $t = k(s+1)-1$ we can regard the individual decision functions specific to DM i as $\{u_t^i = \gamma_s^i(y_{[ks,t]}^i, \mathcal{I}_{ks-1}^c)\}$; we let γ_s denote the ensemble of such decision functions and let $\underline{\gamma}$ denote the team policy.

It then suffices to generate γ_s for all $s \geq 0$, as the decision outputs conditioned on $y_{[ks,t]}^i$, under $\gamma_s^i(y_{[ks,t]}^i, \mathcal{I}_{ks-1}^c)$, can be generated. In such a case, we can define $\gamma_s(., \mathcal{I}_{ks-1}^c)$ to be the joint team decision rule mapping $\mathcal{I}_{ks-1}^c$ into a space of action vectors: $\{\gamma_s^i(y_{[ks,t]}^i, \mathcal{I}_{ks-1}^c), i \in \mathcal{L}=\{1, 2.., L\}, t\in\{ks, ks+1, \ldots, k(s+1)-1\}\}$.

Let $[0, T-1]$ be the decision horizon, where T is divisible by k. Let the objective of the decision makers be the joint minimization of

$$E_{x_0}^{\gamma^1,\gamma^2,\ldots,\gamma^L}[\sum_{t=0}^{T-1} c(x_t, u_t^1, u_t^2, \ldots, u_t^L)],$$

over all policies $\gamma^1, \gamma^2, \ldots, \gamma^L$, with the initial condition x_0 specified. The cost function

$$J_{x_0}(\underline{\gamma}) = E_{x_0}^{\underline{\gamma}} \sum_{t=0}^{T-1} c(x_t, \mathbf{u}_t)$$

can be expressed as

$$J_{x_0}(\underline{\gamma}) = E_{x_0}^{\underline{\gamma}}[\sum_{s=0}^{\frac{T}{k}-1} \bar{c}(\gamma_s(., \mathcal{I}_{ks-1}^c), \bar{x}_s)]$$

with

$$\bar{c}(\gamma_s(., \mathcal{I}_{ks-1}^c), \bar{x}_s) = E_{\bar{x}_s}^{\underline{\gamma}}[\sum_{t=ks}^{k(s+1)-1} c(x_t, \mathbf{u}_t)].$$

Lemma 12.4.1 ([417]). *Consider the decentralized system setup above. Let $\mathcal{I}_t^c$ be a common information vector supplied to the DMs regularly every k time stages, so* *that the DMs have common memory with a control policy generated as described above. Then,* $\{\bar{x}_s := x_{ks}, \gamma_s(\cdot, \mathcal{I}^c_{ks-1}), s \geq 0\}$ *forms a controlled Markov chain.* ◇

Proof. Let$\gamma_{[0,s]} = \{\gamma_m(., \mathcal{I}^c_{km-1}), 0 \leq m \leq s\}$. It follows that

$$\begin{aligned}
&P(\bar{x}_{s+1}|\gamma_{[0,s]}, \bar{x}_{[0,s]}) \\
&= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}} P(\bar{x}_{s+1}, x_{[ks,k(s+1)-1]}, \mathbf{y}_{[ks,k(s+1)-1]}|\gamma_{[0,s]}, \bar{x}_{[0,s]}) \\
&= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}} \Bigg(\prod_{m=ks}^{k(s+1)-1} P(x_{m+1}|x_m, \gamma_s(\mathbf{y}_{[ks,m]}, \mathcal{I}^c_{ks-1}), \bar{x}_s, \gamma_{[0,s-1]}, \bar{x}_{[0,s-1]}) \\
&\qquad\qquad P(y^1_m|x_m)P(y^2_m|x_m)\ldots P(y^L_m|x_m)\Bigg) \\
&= \sum_{(\mathbf{x},\mathbf{y})_{[ks,k(s+1)-1]}} \Bigg(\prod_{m=ks}^{k(s+1)-1} P(x_{m+1}|x_m, \gamma_s(\mathbf{y}_{[ks,m]}, \mathcal{I}^c_{ks-1})) \\
&\qquad\qquad P(y^1_m|x_m)P(y^2_m|x_m)\ldots P(y^L_m|x_m)\Bigg) \\
&= P(\bar{x}_{s+1}|\gamma_s(., \mathcal{I}^c_{ks-1}), \bar{x}_s).
\end{aligned}$$

□

In view of the above, we have the following separation result.

Lemma 12.4.2 ([417]). *Let* $\mathcal{I}^c_t$ *be a common information vector supplied to the DMs regularly every* k *time steps. There is no loss in performance if* $\mathcal{I}^c_{ks-1}$ *is replaced by* $P(\bar{x}_s|\mathcal{I}^c_{ks-1})$. ◇

Proof. As we observed earlier, the cost can be written as a function of additive costs:

$$J_{x_0}(\underline{\gamma}) = E^{\underline{\gamma}}_{x_0}[\sum_{s=0}^{\frac{T}{k}-1} \bar{c}(\gamma_s, \bar{x}_s)],$$

with

$$\bar{c}(\gamma_s, \bar{x}_s) = E^{\underline{\gamma}}_{\bar{x}_s}[\sum_{t=ks}^{k(s+1)-1} c(x_t, \mathbf{u}_t)].$$

For the minimization of a stage-additive cost in partially observed Markov chains, it suffices to transform the state to an equivalent state of conditional distributions as discussed in Appendix D. This follows from the construction of equivalence classes based on belief states as in the proof of Theorem 10.3.2. Hence $P(\bar{x}_s|\mathcal{I}^c_{ks-1})$ acts as a sufficient statistic. □

An essential issue for a tractable solution is to ensure a common information vector which will act as a sufficient statistic for future control policies. This can be done via sharing information at every stage or some structure possibly requiring larger but finite delay.

The above motivates us to introduce the following pattern.

Definition 12.4.1 (k-stage periodic belief sharing pattern). An information pattern in which the decision makers share their posterior beliefs to reach a joint belief about the system state is called a belief sharing information pattern. If the belief sharing occurs periodically every k-stages ($k > 1$), the DMs also share the control actions they applied in the last $k-1$ stages, together with intermediate belief information. In this case, the information pattern is called the k-stage periodic belief sharing information pattern. ◇

Remark 12.4.1. For $k > 1$, it should be noted that the exchange of the control actions is essential, as was also observed in performance-irrelevant signaling or stochastic nestedness discussion in Sect. 3.5. The decision makers also need to exchange information for intermediate beliefs. The following algorithmic discussion will make this clear. ◇ ◇

In accordance with the preceding remark, we now discuss how beliefs are shared sequentially. We proceed by induction. Suppose at time $ks - 1$, the DMs have an agreement on $P(\bar{x}_s|\mathcal{I}^c_{ks-1})$ and know the policies used by all the DMs. It then follows that

$$\pi_{s+1} := P(\bar{x}_{s+1}|\mathbf{y}_{[ks,k(s+1)-1]}, \mathbf{u}_{[ks,k(s+1)-1]}, \pi_s)$$

can be written as

$$\frac{P(\bar{x}_{s+1}, (\mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}{\sum_{\bar{x}_{s+1}} P(\bar{x}_{s+1}, (\mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}$$

$$= \frac{\sum_{x_{[ks,k(s+1)-1]}} P(\bar{x}_{s+1}, (x, \mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}{\sum_{x_{[ks,k(s+1)-1]},\bar{x}_{s+1}} P(\bar{x}_{s+1}, (x, \mathbf{y}, \mathbf{u})_{[ks,k(s+1)-1]}|\pi_s)}. \tag{12.1}$$

We now express the numerator in (12.1) more explicitly as

$$\sum_{x_{k(s+1)-1}} \Bigg(P(x_{k(s+1)}|x_{k(s+1)-1}, \mathbf{u}_{k(s+1)-1}) (\prod_{l=1}^{L} P(y^l_{k(s+1)-1}|x_{k(s+1)-1}))$$

$$\sum_{x_{k(s+1)-2}} \Bigg(P(x_{k(s+1)-1}|x_{k(s+1)-2}, \mathbf{u}_{k(s+1)-2}) (\prod_{l=1}^{L} P(y^l_{k(s+1)-2}|x_{k(s+1)-2}))$$

$$\dots \qquad \dots \qquad \dots$$

$$\sum_{x_{ks+1}} \Bigg(P(x_{ks+2}|x_{ks+1}, \mathbf{u}_{ks+1}) (\prod_{l=1}^{L} P(y^l_{ks+1}|x_{ks+1}))$$

$$\sum_{x_{ks}} \Bigg(P(x_{ks+1}|x_{ks}, \mathbf{u}_{ks}) (\prod_{l=1}^{L} P(y^l_{ks}|x_{ks})) P(x_{ks}|\mathcal{I}^c_{ks-1}) \Bigg) \Bigg) \dots \Bigg) \Bigg).$$

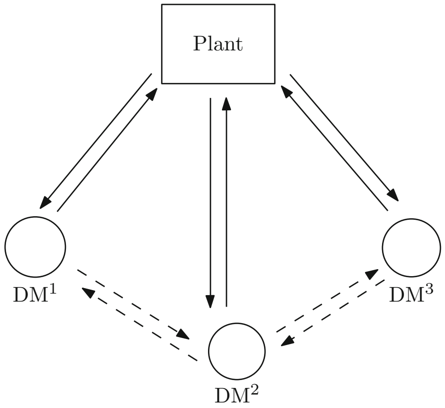

Fig. 12.1 Belief propagation converges to the true conditional measure in a finite number of iterations in a cycle-free network

Thus, if $k > 1$, then the DMs also need to share the control actions applied in the previous $k - 1$ time stages, as well as beliefs on individual states.

When the belief sharing occurs at every stage, then controls can be generated by each of the DMs, and hence the control actions need not be shared. We will discuss this further while considering the belief propagation algorithm in the following.

Iterative Belief Propagation Algorithm

We now present an iterative approach for the belief sharing pattern. This will be an extension of the belief propagation algorithm, which is a local message exchange algorithm among several remote DMs/sensors located at the vertices of a graph through the edges [146]. In a belief propagation algorithm, each DM has an *a priori* belief about the state of the system. With local observations, each DM generates an *a posteriori* belief and then exchanges this with other DMs. Belief propagation reaches the correct measure (one that would be achieved under a centralized information structure) if there are no cycles, for example, when the topology of the communication graph connecting the DMs forms a tree.

We discuss below how belief exchanges can be carried out to achieve the belief sharing pattern in view of (12.1) and (12.2). We first consider a cycle-free network of DMs. Consider the setting of Fig. 12.1. For $m \geq 0$, suppose the DMs have an agreement on x_{ks+m}, before they receive their local observations, that is, they all have access to $P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|\mathcal{I}^c_{ks+m-1}))$ for all x_{ks+m} values. Once DMs observe local measurements y^i_{ks+m}, first, DM 1 sends to DM 2 its belief on x_{ks+m}: $P(x_{ks+m}|y^1_{ks+m}, P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|\mathcal{I}^c_{ks+m-1})))$. Thus, DM 1 sends its belief about the state of the system at time $ks + m$ to DM 2, for all possible x_{ks+m} values. DM 2 then for all x_{ks+m} values generates its own belief on x_{ks+m}, using its local information (only using the belief sent by DM 1 and its own information, together with the prior belief), to obtain

$$P(x_{ks+m}|(y^1, y^2)_{ks+m}, P(x_{ks+m}|.))$$
$$= \frac{P(y^2|x_{ks+m})P\Big(x_{ks+m}|y^1_{ks+m}, P(\cdot|\cdot)\Big)}{\sum_{x_{ks+m}} P(y^2|x_{ks+m})P\Big(x_{ks+m}|y^1_{ks+m}, P(\cdot|\cdot)\Big)},$$

where $P(\cdot|\cdot)$ denotes $P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|\mathcal{I}^c_{ks+m-1}))$, which is the prior belief on x_{ks+m} using the information from the previous time stage. DM 2 then sends this information to DM 3, who upon receiving the information from DM 2 generates the final conditional probability measure

$$P(x_{ks+m}|(y^1, y^2, y^3)_{ks+m}, P(x_{ks+m}|\mathbf{u}_{ks+m-1}, P(x_{ks+m-1}|\mathcal{I}^c_{ks+m-1}))).$$

In the next iteration, DM 3 sends this belief information back to DM 2 and finally to DM 1. Upon such a forward and backward sweeping, all the DMs have access to the joint belief on the state of the system, which we denote by $P(x_{ks+m}|\mathcal{I}'^c_{ks+m})$.

Now that all the DMs have the same belief on the state, they share all their control actions that they applied at time $ks + m$. With this information, all of the DMs have access to $P(x_{ks+m+1}|\mathbf{u}_{ks+m}, P(x_{ks+m}|\mathcal{I}'_{ks+m}))$. This now acts as a common prior for the next iteration in this algorithm. DM 1 once again sends its belief on the state x_{ks+m+1} and the iteration continues. As such, upon such a double forward and backward sweeping (first for the beliefs, then for the control actions), all DMs have access to the correct joint belief on the state of the system. Once this convergence occurs for all time stages $ks, ks+1, ks+2\ldots, ks+s-1$, the DMs start the iteration for the next time stage, eventually agreeing on the conditional measure for the time stage $k(s + 1) - 1$ and subsequently having a common prior on $k(s + 1)$.

Consider now the setting of Fig. 12.2. In this case, as there is a cycle (loop) in the communication topology, it is essential that the DMs have a predefined route for information exchange for convergence to the actual joint belief: If all DMs talk to their neighbors, then one DM might have more than one path to send a message to another DM and hence her opinions might be incorrectly emphasized by the virtue of the cycle.

If the sharing of measurements occurs at every stage (that is, $k = 1$), then control action sharing would not be needed, since every DM can generate the others' control actions given the joint belief, under deterministic policies. In this case, we only need to apply the original belief sharing algorithm, which, in the cycle-free case, only requires one forward and backward message passing.

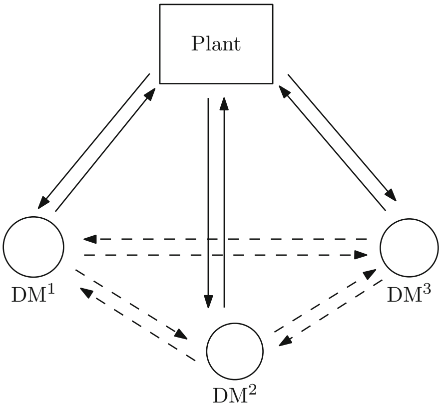

Fig. 12.2 When there are cycles in the communication topology, in our setting, one needs to pick an ordering *a priori* and avoid cycles in the ordering until convergence is reached

12.4.2 Minimum Communication for the Belief Sharing Pattern

In view of Lemmas 12.4.1 and 12.4.2, what needs to be exchanged is a sufficient amount of information such that the DMs have a common $P(\bar{x}_s|\mathcal{I}_s^c)$, so that their recursions can be based on this information. The question that we are interested in addressing in this subsection is the following: How much information exchange is needed between the DMs so that they have an agreement on the state of the system (i.e., the joint belief on the state) and a dynamic programming recursion is possible?

As was done earlier in the book, the information in this context will be measured by the number of bits; when the coding is variable rate, information is measured by the average number of bits that need to be exchanged among the DMs, whereas when the coding scheme is fixed rate, information is measured by the actual number of bits that are exchanged for any given time stage $t \geq 0$. The coding process of the controller at DM i is a mapping measurable with respect to the sigma-algebra generated by $\mathcal{I}_t^i$. DM i's coding policy for communicating to DM j maps the information $\mathcal{I}_t^i$ to $\mathcal{M}^{i,j}$ at time t; that is, the set of entropy-coder variable-rate or fixed-rate codewords for communication from DM i to DM j. Hence, at each time stage t, DM i sends $R^{i,j}(t)$ bits over an external channel to DM j. Let $\mathcal{R}_t = \{R^{i,j}(t), i \neq j, i, j \in 1, 2, \dots, L\}$ such that belief sharing is possible. Define $R_t := \inf_{\mathcal{R}_t}\{\sum_{i=1}^{L}\sum_{j=1, j\neq i}^{L} R_t^{i,j}\}$, such that belief sharing is realized for a given time stage t. Our aim is to obtain such R_t values.

In the following, we consider the case where the measurement spaces $\mathbb{Y}^i, i \in \mathcal{L}$, are discrete.

We consider the one-stage belief sharing pattern, first for a two-DM setup. In this case, the information needed at both controllers is such that they all need to exchange the relevant information on the state, and need to agree on $P(\bar{x}_t|\mathcal{I}_t^1, \mathcal{I}_t^2)$, where $\mathcal{I}_t^i$ denotes the information available at DM i. In the one-step belief sharing pattern, $\bar{x}_t = x_t$, since the period for information exchange is $k = 1$.

We note that, when control policies are deterministic, the actions can uniquely be identified by both DMs. As such, control signals need not be exchanged.

Theorem 12.4.1 ([417]). *To achieve the one-stage belief sharing information pattern, the following rate region is achievable using fixed-rate codes:*

$$\mathcal{R}(t) = \Big\{(R^{i,j}, R^{j,i}) : R^{i,j} = \lceil \log_2(|\mathcal{S}_t|) \rceil, R^{j,i} = \lceil \log_2(\sup_{\pi^i} |\mathcal{S}_{\pi^i,t}|) \rceil,$$

$$\mathcal{S}_t = \Big\{\pi^i = P\Big(x_t \Big| y_t^i = y^i, P(.|.)\Big) : P\Big(y_t^i = y^i \Big| P(.|.)\Big) > 0, y^i \in \mathbb{Y}^i\Big\},$$

$$\mathcal{S}_{\pi^i,t} = \Big\{P\Big(x_t \Big| y_t^j {=} y^j, \pi^i, P(.|.)\Big) : P\Big(y_t^j {=} y^j \Big| \pi^i, P(.|.)\Big) > 0, y^j {\in} \mathbb{Y}^j\Big\}\Big\},$$

where $P(.|.)$ stands for $P(x_t|\mathcal{I}^c_{t-1})$. ◇

Proof. The result follows from binning arguments used in the context of decentralized communication. First, DM 2 learns the conditional belief of DM 1 and then computes the joint belief using the algorithm discussed in the previous section. DM 2 then sends the set of all distinct possible *a posteriori* joint beliefs to DM 1 consistent with the belief at DM 1. □

The following is a counterpart of the above result when the communication rate is measured by the average number of bits:

Theorem 12.4.2 ([417]). *Suppose that the measurement variables are discrete valued, i.e., $\mathbb{Y}^i, i = 1, 2,$ are countable spaces. To allow for the belief sharing information pattern, a lower bound on the minimum average number of bits to be transmitted to DM i from DM j, $i, j \in \{1, 2\}, i \neq j$ is*

$$R^{j,i} \geq H\Big(P(x_t|\mathcal{I}^c_{t-1}, y_t^i, y_t^j) \Big| P(x_t|\mathcal{I}^c_{t-1}), y_t^i\Big),$$

$$R^{i,j} \geq H\Big(P(x_t|\mathcal{I}^c_{t-1}, y_t^i, y_t^j) \Big| P(x_t|\mathcal{I}^c_{t-1}), y_t^j, z_t^i\Big),$$

where z_t^i is the variable sent to DM i. ◇

Proof. Let z_t^1 be the random variable that is transmitted from DM 2 to DM 1. Then, we have (with $||.||$ denoting the total variation norm),

$$R^{2,1} \geq \inf\Big\{H(z_t^1) : ||P(x_t|\mathcal{I}^c_{t-1}, z_t^1, y_t^1) - P(x_t|\mathcal{I}^c_{t-1}, y_t^1, y_t^2)|| = 0\Big\}$$

$$\geq \inf\Big\{H\Big(z_t^1 \Big| P(x_t|\mathcal{I}^c_{t-1}), y_t^1\Big) : ||P(x_t|\mathcal{I}^c_{t-1}, z_t^1, y_t^1) {-} P(x_t|\mathcal{I}^c_{t-1}, y_t^1, y_t^2)|| {=} 0\Big\}$$

$$
\begin{aligned}
&\geq \inf \Bigg\{ H\Bigg(z_t^1 \Bigg| P(x_t|\mathcal{I}_{t-1}^c), y_t^1 \Bigg) - H\Bigg(z_t^1 \Bigg| P(x_t|\mathcal{I}_{t-1}^c, y_t^2, y_t^1), P(x_t|\mathcal{I}_{t-1}^c), y_t^1 \Bigg) \\
&\qquad\qquad : ||P(x_t|\mathcal{I}_{t-1}^c, z_t^1, y_t^1) - P(x_t|\mathcal{I}_{t-1}^c, y_t^1, y_t^2)|| = 0 \Bigg\} \\
&= \inf \Bigg\{ I\Bigg(z_t^1 ; P(x_t|\mathcal{I}_{t-1}^c, y_t^2, y_t^1) \Bigg| P(x_t|\mathcal{I}_{t-1}^c), y_t^1 \Bigg) \\
&\qquad\qquad\qquad : ||P(x_t|\mathcal{I}_{t-1}^c, z_t^1, y_t^1) - P(x_t|\mathcal{I}_{t-1}^c, y_t^1, y_t^2)|| = 0 \Bigg\} \\
&= \inf \Bigg\{ H\Bigg(P(x_t|\mathcal{I}_{t-1}^c, y_t^2, y_t^1) \Bigg| P(x_t|\mathcal{I}_{t-1}^c), y_t^1 \Bigg) \\
&\qquad\qquad\qquad - H\Bigg(P(x_t|\mathcal{I}_{t-1}^c, y_t^2, y_t^1) \Bigg| P(x_t|\mathcal{I}_{t-1}^c), z_t^1, y_t^1 \Bigg) \\
&\qquad\qquad\qquad : ||P(x_t|\mathcal{I}_{t-1}^c, z_t^1, y_t^1) - P(x_t|\mathcal{I}_{t-1}^c, y_t^1, y_t^2)|| = 0 \Bigg\} \\
&= H\Bigg(P(x_t|\mathcal{I}_{t-1}^c, y_t^2, y_t^1) \Bigg| P(x_t|\mathcal{I}_{t-1}^c), y_t^1 \Bigg).
\end{aligned}
$$

Here the last two steps follow from the observation that given z^1 the conditional measures are identical, and the constraint is independent of

$$H(P(x_t|\mathcal{I}_{t-1}^c, y_t^1, y_t^1)|P(x_t|\mathcal{I}_{t-1}^c), y_t^1).$$

A parallel discussion applies to the reverse direction. □

We note that the information rate needed is less than the one needed for achieving the centralized information pattern which would require all the decision makers to exchange all their observations. By the above argument, one would need $R^{i,j} \geq H(y_t^i|y_t^j, \mathcal{I}_{t-1}^c)$ for the centralized information pattern as a lower bound. The entropy of the conditional measure is at most as high as the entropy of the measured variable. This is because different outputs may lead to the same values for $P(y_t^2 = y|x_t, \mathcal{I}_{t-1}^c)$.

We may also obtain an upper bound on the communication rates in the two-DM setting for variable-rate schemes. The result is intuitive and is given below without a proof.

Proposition 12.4.1. *When the measurement space is discrete, to allow for the one-stage belief sharing information pattern, an upper bound on the minimum average amount of bits to be transmitted to DM i from DM j,* $i, j \in \{1, 2\}, i \neq j$, *is given by*

$$R^{j,i} \leq \min \left\{ H\left(\zeta(y_t^j, P(.|.)) \middle| P(.|.) \right): \\ P\left(x_t \middle| P(.|.), y_t^1, y_t^2 \right) = P\left(x_t \middle| P(.|.), y_t^i, \zeta(y_t^j, P(.|.)) \right) \right\},$$

where $P(.|.)$ *denotes* $P(x_t|\mathcal{I}_{t-1}^c)$. ◇

Remark 12.4.2. In the setup considered, the goal is for each DM to compute the joint belief. We note here the interesting discussion between decentralized computation and communication provided by Csiszar and Körner (see [107], Theorem. 4.6) and Orlitsky and Roche [300], as well as by Witsenhausen [395]. However, the settings of these works assumes a typical information theoretic setup of an infinite copy of messages to be encoded and functions to be computed, which is not applicable in a real-time setting. In particular, such a decentralized computation problem can be posed as a multiterminal source coding problem with a cost function aligned with the computation. See El Gamal and Kim [153] for a comprehensive treatment of information exchange requirements for computing. It is also important to point out that multi-round protocols typically reduce the average rate requirements. See also Sect. 5.4 for related discussions in the context of real-time coding. ◇ ◇

In the case of more than two DMs, one encounters a distributed coding with side information setting. In this case the DMs will send correlated information to another decision maker. This leads to the following lower bound, which we provide without a proof since it follows from the analysis made earlier for the two-DM case.

Proposition 12.4.2 ([417]). *To allow for the belief sharing information pattern, a lower bound on the minimum average amount of bits that needs to be supplied to any DM i,* $\{i = 1, 2, \ldots, L\}$ *is obtained by the relation*

$$R^i \geq H\left(P(x_t|\mathcal{I}_{t-1}^c, y_t^1, y_t^2, \ldots, y_t^L) \middle| P(x_t|\mathcal{I}_{t-1}^c), y_t^i \right).$$

◇

Consider the following special case with zero-capacity measurement channels.

Proposition 12.4.3. *Consider the case where the measurement channel for each of the DMs has zero capacity. Then, as*

$$P(\bar{y}_s^i = \eta|\bar{x}_s) = P(\bar{y}_s^i = \beta|\bar{x}_s)$$

for all η, β *values that the measurements can take, there is no further information that is needed for the belief sharing pattern.* ◇

Proof. By Theorem 12.4.2, a lower bound on the rate is

$$H\left(P(x_t|\mathcal{I}^c_{t-1}, y^1_t, y^2_t)\middle|P(x_t|\mathcal{I}^c_{t-1}), y^1_t\right) = 0.$$

This rate bound is tight, since by Theorem 12.4.1, the rate achievable by a fixed-rate scheme is zero as well. □

A similar result applies to the case of more than two DMs, by Proposition 12.4.2.

As such, there is no need for information exchange since there is no information generated by the measurement of each controller with regard to the state and no transmitted information transmission/exchange will be useful. Hence, the communication required for belief sharing pattern is zero if all of the information channels have zero-capacity. We note that, in such a case, the control actions do not need to be exchanged either, as there is already an agreement on the beliefs, based on the *a priori* belief, and the optimal team decisions can be generated in a decentralized fashion.

12.5 A Team Cost-Rate Function

Building on the results of the previous section, here we generalize the value of information exchange to general sequential team problems. Consider the general setting of the L-agent dynamic team problem of Chap. 2 with dynamics

$$x_{t+1} = f_t(x_t, u^1_t, \ldots, u^L_t; w_t)\,, \quad t \in \mathcal{T}, \tag{12.2}$$

and measurements

$$y^i_t = g^i_t(x_t, u^1_{t-1}, \ldots, u^L_{t-1}; v^i_t), \qquad i \in \mathcal{L}, \qquad t \in \mathcal{T}, \tag{12.3}$$

where $x_0, w_{[0,T-1]}, v^i_{[0,T-1]}, i \in \mathcal{L}$ are random variables with specified probability distributions. The information exchange is facilitated by an encoding protocol $\mathcal{E}$ which, in the usual setting, admits partial functions from each of the encoders measurable on the local information, described as follows. Let the information available to DM i at time t be

$$\mathcal{I}^i_t = \{y^i_{[1,t]}, u^i_{[1,t-1]}, z^{i,j}_{[0,t]}, z^{j,i}_{[0,t]}, j \in \mathcal{L}\},$$

where $z^{i,j}$ is the information transmitted from DM j to DM i such that for $t \geq 1$ with

$$\mathbf{z}^i_t = \{z^{i,j}_t, j \in \mathcal{L}\} = \mathcal{E}^i_t(\mathcal{I}^i_{t-1}, u^i_{t-1}, y^i_t) \tag{12.4}$$

and for $t = 0$,

$$\mathbf{z}_0^i = \{z_0^{i,j}, j \in \mathcal{L}\} = \mathcal{E}_0^i(y_0^i).$$

As before,

$$u_t^i = \gamma_t^i(\mathcal{I}_t^i),$$

for all DMs. Let $\underline{\mathcal{E}}^i = \{\mathcal{E}_t^i, t \geq 0\}$. Under an encoding policy $\underline{\mathcal{E}} = \{\underline{\mathcal{E}}^1, \underline{\mathcal{E}}^2, \ldots, \underline{\mathcal{E}}^L\}$ and control policies $\underline{\gamma} = \{\underline{\gamma}^1, \underline{\gamma}^2, \ldots, \underline{\gamma}^L\}$, let the induced cost be

$$E^{\underline{\gamma},\underline{\mathcal{E}}}[\sum_{t=0}^{T-1} c(x_t, \mathbf{u}_t)].$$

Definition 12.5.1. Given a decentralized control problem as above, *team cost-rate function* $C : \mathbb{R} \to \mathbb{R}$ is

$$C(R) := \inf_{\underline{\gamma},\underline{\mathcal{E}}} \left\{ E^{\underline{\gamma},\underline{\mathcal{E}}}[\sum_{t=0}^{T-1} c(x_t, \mathbf{u}_t)] : \frac{1}{T}\mathcal{R}(\mathbf{z}_{[0,T-1]}) \leq R \right\},$$

where $\mathcal{E}$ denotes the class of all causal encoder protocols/policies leading to the exchange of random variables $z^{i,j}, i, j \in \mathcal{L}$. ◇

Here, $\frac{1}{T}\mathcal{R}(\mathbf{z}_{[0,T-1]})$ denotes the average rate of communication. This rate can be measured under fixed-rate or variable-rate formulations (see Sect. 5.3.2).

The formulation here can be adjusted to include sequential (iterative) information exchange given a fixed ordering of actions, as opposed to a simultaneous (parallel) information exchange at any given time t. That is, instead of (12.4), we may have

$$\mathbf{z}_t^i = \{z_t^{i,j}, j \in \{1, 2, \ldots, L\}\} = \mathcal{E}_t^i(\mathcal{I}_{t-1}^i, u_{t-1}^i, y_t^i, \{z_t^{k,i}, k < i\}). \quad (12.5)$$

Proposition 12.5.1. *A sequential (iterative) communication protocol may perform strictly better than an optimal parallel communication protocol given a total rate constraint.* ◇

Proof. We provide an instance where a sequential transmission may lead to better performance. Consider the following setup with two DMs [164]. Let x^1, x^2, p be uniformly distributed binary random variables. Let

$$x = (p, x^1, x^2), \quad y^1 = p, \quad y^2 = (x^1, x^2),$$

and the loss function be

$$c(x, u^1, u^2) = 1_{\{p=0\}} c(x^1, u^1, u^2) + 1_{\{p=1\}} c(x^2, u^1, u^2),$$

with

$$c(s, u^1, u^2) = (s - u^1)^2 + (s - u^2)^2.$$

Suppose that we wish to compute the minimum cost subject to a total rate of 2 bits that can be exchanged.

Under a sequential scheme, if we allow DM 1 to encode y^1 to DM 2 with one bit, then a cost of 0 is achieved since DM 2 knows the relevant information that needs to be transmitted to DM 1, again with one bit: If $p = 0$, x^1 is relevant with an optimal policy $u^1 = u^2 = x^1$, and if $p = 1$, x^2 is relevant with an optimal policy $u^1 = u^2 = x^2$, and a cost of 0 is achieved. However, if the information exchange is parallel, then DM 2 does not know which state is the relevant one, and it can be shown that a cost of 0 cannot be achieved under any coding policy. □

We now make a few remarks.

Remark 12.5.1. Replacing the fixed-rate or variable-rate (entropy) constraint in Definition 12.5.1 with a mutual information constraint may lead to more desirable mathematical properties for $C(R)$; however, such a definition would not be operational in a real-time setting in the same spirit as discussed in Sect. 5.4. ◇

◇

Remark 12.5.2. In Sect. 5.4, we introduced the *distortion-constrained entropy minimization* problem [184,427]. The dual of the problem is the *entropy-constrained distortion minimization* problem, defined as

$$D_Q(R) = \inf\{D = E[\rho(x, Q(x))] : H(Q(x)) \leq R\}.$$

Note that, for the causal quantization problem, the entropy-constrained distortion minimization problem becomes a special case of the formulation in Definition 12.5.1, with the Markov source being a control-free source and the fixed-rate constraint being replaced with a variable-rate constraint. ◇

Remark 12.5.3. The formulation in Definition 12.5.1 can also be adjusted to allow for multiple rounds of communication per time stage. Keeping the total rate constant, having multiple rounds can enhance the performance for a class of team problems while keeping the total rate constant. For an information theoretic setup, see [245]. ◇

The discussion on the one-stage belief sharing pattern leads to the following.

Theorem 12.5.1. *If R is large enough to allow the realization of the one-stage belief sharing pattern, then the optimal performance that could have been achieved for a centralized system would be achieved.* ◇

Therefore, $C(R)$ can be expressed explicitly in a regime when the one-stage belief sharing pattern can be realized: The rate requirements for the one-stage belief sharing pattern were considered earlier in the chapter. If the available information rate cannot allow for belief sharing, however, the answer to the question on what an optimal coordination/agreement scheme would be is not clear for a dynamic team problem. Tsitsiklis [365] has observed that from a computational complexity viewpoint, variants of such a problem is non-tractable (NP-hard). However, explicit structural and existence results on optimal coding and control policies, extending the findings in Chap. 10, would be useful to obtain approximately optimal policies.

12.6 Concluding Remarks

This chapter developed further insights into the problem of *the value of information channels in decentralized stochastic control*. At the expense of communications, the controllers benefit from a reduced computational complexity. Further work is needed to obtain a trade-off between computational complexity and communications, although the chapter provided some discussion toward a unifying theory in this direction. It is perhaps counterintuitive first to think that more information leads to less complexity, but stochastic control theory shows that all one needs is to construct a Markov decision problem (with an appropriate controlled Markovian state, and a cost function). The intractability of some decentralized optimal control problems (but not all) stems from not being able to construct such a non-exploding chain in a general setting. Belief sharing is an attempt to provide a systematic, rate-efficient way to construct a tractable setting, via an appropriate Markov chain. This pattern also provides a partial answer to what needs to be exchanged in real time to obtain optimal performance subject to communication constraints when the communication allows an agreement on beliefs.

One related problem is how to optimally encode the control actions; such an analysis will intimately depend on the cost functions [367]. Furthermore, as shown in [368], solutions to such problems can be computationally difficult. Finally, the analysis of belief propagation algorithms for the belief sharing information pattern is a practically important problem which merits further analysis.

12.7 Bibliographic Notes

In case there is a prior agreement on a system model, common knowledge can be generated through iterative algorithms. Most of such results in the literature use either martingale methods or contraction methods. There has been considerable amount of research on agreement problems. A large number of results are based on iterative convergence and consensus algorithms. One of the earliest results is due to Borkar and Varaiya [75] and Tsitsiklis and Athans [367]. It is evident that almost sure agreement in finite time is impossible over general noisy channels, even though this may be possible over noiseless channels, or noiseless channels which satisfy infinite error exponents in finite expected time, such as those that can be used for Burnashev's exponent [83]. Erasure channels constitute one example of a channel where agreement in finite expected time is possible if the source has finite cardinality.

Convergence properties of consensus algorithms have been discussed in [97, 208, 296, 364, 365] among many other contributions in the literature. Regarding agreement under inconsistent priors under Bayesian and non-Bayesian (which may be through a *frequentist approach* or through other iterative methods) update laws, we refer the reader to Blackwell and Dubins [63], Diaconis and Freedman [117], and Acemoglu and Ozdaglar [3].

Coordinator-based policies have been considered by Athans [18], Yoshikawa [413], Aicardi et al. [4], Ooi and Wornell [299], Nayyar, Mahajan and Teneketzis [287], Mahajan [246], Yüksel [417], Singh [342], and hierarchical control literature including Mahmoud [252] and Tsitsiklis [366] who have discussed coordination. A detailed investigation on when a dynamic programming approach algorithm is applicable has been investigated in [286]. In such cases, common information can be used to generate a team decision policy, provided one can design an MDP typically in an uncountable topological state space. The periodic information sharing considered in this chapter has also been considered in [299], in addition to [417]. Several authors [4, 49, 286] have considered the use of common information for more general patterns. An extensive review is available in [288].

For the belief propagation algorithm and its convergence properties, we refer the reader to [275, 355] and the references therein. Algorithms for the computation of joint beliefs have been considered in [88, 417].

Related to the construction of agreement results by Nielsen [293] and Brandenburger and Dekel [79] is the probability dependency of a notion of a state of a system. A related approach is considered in Yüksel [417] in stochastic nestedness and in Mahajan and Yüksel [251] in the context of *measure and cost dependency* of information structures.

It is worth pointing out, in the context of the discussion in this chapter, the results provided in the information theory literature on decentralized computation and communication by Csiszár and Körner (see [107], Theorem. 4.6), Orlitsky and Roche [300], and Ma and Ishwar [245]. However, the setting presented in the mentioned works assumes an information theoretic setup with an infinite copy of messages to be encoded and functions to be computed, which is not applicable in a real-time setting (in the same spirit as discussed in Sect. 5.4). However, the analysis, as in the case of the bounds provided by the rate-distortion function for the real-time quantization problem, provides lower bounds on achievable encoding and agreement policies. In a control context, however, one method to achieve the information theoretic bounds is via binning [417]. Orlitsky and Roche [300] investigate coding rate requirements for computing and present connections with multiterminal source coding theory. The results are combinatorial generalizations of multiterminal source coding problems and provide lower and upper bounds on information transmission rates for communication protocols. Giridhar and Kumar [162] provide a comprehensive overview of results in distributed computation under information constraints. Distributed function computation through linear policies is considered in [350].

The information exchange requirements for decentralized optimization depend also on the structural properties of the cost functional to be minimized. For a class of static team problems, one might simply need to exchange a sufficient statistic needed for optimal solutions, such as the exchange of a Lagrange multiplier [77, 344].

References [160, 290], among many others, have considered static distributed convex optimization under information constraints modeled by communication graphs. Even further, for some problems, there may be no need for an exchange at all, if the sufficient statistics are already available, as in the case of *mean field equilibrium* problems when the number of decision makers is very large; see [200, 202, 226].

Optimization under local interaction and sparsity constraints and various criteria have been investigated in a number of publications including [38, 134, 210, 231, 324, 327]. A review for the literature on norm-optimal control as well as optimal stochastic dynamic teams is provided in [247].

The communication complexity literature (see, e.g., [412]) has considered the following problem: Consider two agents who have access to local variables x, y. Given a function f of variables (x, y), what is the maximum (of all input variables x, y) of the minimum amount of information exchange, one bit at a time, needed for at least one agent to compute the value of the function? We note that, in this setting, what is addressed is not the optimization of a given function, but its computation, and the problem is deterministic. We refer the reader to [291] for a comprehensive resource on information complexity for optimization problems. A further distribution-free interpretation of the amount of common information between two decision makers is given in [279]. Tsitsiklis and Luo [369] considered the communication complexity in a two-DM setup, for the minimization of a sum of convex functions, where only a local function is known by each DM. A sequential setting with an information theoretic approach to the formulation of communication complexity has been considered in [318], and Tsitsiklis and Athans [368] established intractability of cost-rate-type formulations from the point of view of communication complexity. Wong and Baillieul have developed an approach for communication cost complexity for control systems in [404, 405]. A comprehensive overview of communication and computation complexity aspects primarily in the economics literature is available in [372]. Finally, a formulation relevant to the one in Sect. 12.5 has been considered in [359] with mutual information constraints, where also the differences between information theoretic approaches and real-time formulations have been investigated.

Part of this chapter builds on [417].

Appendix A
Topological Notions and Optimization

This appendix provides some background material on those aspects of real analysis and optimization that are frequently used in the text; it also serves to introduce the reader to some of the notation and terminology used in the book. For a more detailed exposition on the topics covered here, two standard references are [109, 242].

A.1 Sets

A set S is a collection of elements. If s is a member (element) of S, we write $s \in S$; if s does not belong to S, we write $s \notin S$. If S contains a finite number of elements, it is called a *finite set*; otherwise it is called an *infinite set*. If the number of elements of an infinite set is countable (i.e., if there is a one-to-one correspondence between its elements and positive integers); we say that it is a *denumerable* (*countable*) set, otherwise it is a *nondenumerable* (*uncountable*) set. An example for the latter is the set of all real numbers, $\mathbb{R}$.

A set $\mathbb{X}$ with some specific structure attached to it is called a *space*, and it is called a linear (*vector*) *space* if this specific structure is of algebraic nature with certain well-known properties which we assume the reader is familiar with. If S is a vector space, a subset of S which is also a vector space is called a *subspace*. An example of a vector space is the *n-dimensional Euclidean space* (denoted as $\mathbb{R}^n$), each element of which is determined by n real numbers. An $x \in \mathbb{R}^n$ can either be written as a *row vector* $x = (x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are real numbers and denote the components of x, or as a *column vector* which is the "transpose" of $(x_1, \ldots, x_n)$ [written as $x = (x_1, \ldots, x_n)'$]. We shall adopt the latter convention in this text, unless indicated otherwise.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,

Linear Independence

Given a finite set of vectors $\{s_1, \ldots, s_n\}$, in a vector space S, we say that this set of vectors is *linearly independent* if the equation $\sum_{i=1}^{n} \alpha_i s_i = 0$ implies $\alpha_i = 0 \; \forall i = 1, \ldots, n$. Furthermore, if every element of S can be written as a linear combination of these vectors, we say that this set of vectors *generates* S. Now, if S is generated by such a linearly independent finite set (say, X), it is said to be *finite dimensional* with its unique "dimension" being equal to the number of elements of X; otherwise, S is *infinite dimensional.*

A.2 Vector Spaces

Normed Linear (Vector) Spaces

A *normed linear vector space* is a linear (vector) space S which has some additional structure of topological nature. This structure is induced on S by a real-valued function which maps each element $u \in S$ into a real number $||u||$ called the *norm* of u. The norm satisfies the following three axioms, where $\underline{o}$ is the *zero* element:

(1) $||u|| \geq 0 \;\; \forall u \in S$; $\;\; ||u|| = 0$ if, and only if, $u = \underline{o}$
(2) $||u + v|| \leq ||u|| + ||v||$, for each pair $u, v \in S$
(3) $||\alpha u|| = |\alpha| \cdot ||u|| \;\; \forall \alpha \in \mathbb{R}$ and for each $u \in S$

Convergent Sequences and Cauchy Sequence

Aninfinite sequence of vectors $\{s_1, s_2, \ldots\}$ in a normed vector space S is said to *converge* to a vector s if, given an $\epsilon > 0$, there exists an integer N such that $||s - s_i|| < \epsilon$ for all $i \geq N$. In this case, we write $s_i \to s$, or $\lim_{i \to \infty} s_i \to s$, and call s the *limit point* of an infinite sequence $\{s_i\}$. More generally, a point s is said to be a *limit point* of the infinite sequence $\{s_i\}$ if the latter has an infinite subsequence $\{s_i\}$ that converges to s.

An infinite sequence $\{s_i\}$ in a normed vector space is said to be a *Cauchy sequence* if, given an $\epsilon > 0$, there exists an integer N such that $||s_n - s_m|| < \epsilon$ for all n, $m \geq N$. A normed vector space S is said to be *complete*, or a Banach space, if every Cauchy sequence in S is convergent to an element of S.

Open, Closed, and Compact Sets

Let S be a normed vector space. Given an $s \in S$ and an $\epsilon > 0$, the set $N_\epsilon(s) = \{x \in S : \|x - s\| < \epsilon\}$ is said to be an ϵ- *neighborhood* of s. A subset X of S is open if, for every $x \in X$, there exists an $\epsilon > 0$ such that $N_\epsilon(x) \subset X$. A subset X of S is *closed* if its complement in S is open; equivalently, X is closed if every convergent sequence in X has its limit point in X. Given a set $X \subset S$, the largest subset of X which is open is called the *interior* of X.

A subset X of a normed vector space S is said to be *(sequentially) compact* if every infinite sequence in X has a convergent subsequence whose limit point is in X. If X is finite dimensional, compactness is equivalent to being closed and bounded.

Denseness and Separability

Let S be a normed linear space and X be a subset of S. X is said to be *dense* in S if given any $s \in S$ and $\epsilon > 0$, there exists $x \in X$ such that $\|x - s\| \leq \epsilon$. S is said to be *separable* if there exists a countable set $X \subset S$ which is dense in S.

Metric Spaces

A metric defined on a set X is a function $d : X \times X \to \mathbb{R}_+$ such that

$$\begin{aligned}
&(1)\ d(x,y) \geq 0, \quad \forall x, y \in X, \quad d(x,y) = 0 \text{ if, and only if, } x = y \\
&(2)\ d(x,y) = d(y,x), \quad \forall x, y \in X \\
&(3)\ d(x,y) \leq d(x,z) + d(z,y), \quad \forall x, y, z \in X
\end{aligned}$$

A metric space (X, d) is a set equipped with a metric d. A normed linear space is also a metric space, with metric $d(x,y) = \|x - y\|$. A metric space (X, d) is complete, if every Cauchy sequence in the space has a limit in X (where the definition of a Cauchy sequence is the same as that for a normed linear space with $\|\cdot\|$ replaced by $d(\cdot,\cdot)$). A complete, separable metric space is also called a Polish space.

Inner-Product Spaces

Given a linear vector space S, an *inner product* on S is a bilinear map $\langle \cdot, \cdot \rangle : S \times S \to \mathbb{R}$ which is positive definite and symmetric, i.e.,

(1) $\langle u, u\rangle \geq 0\ \forall u \in S$; $\langle u, u\rangle = 0$ if, and only if, $u = \underline{o}$, the *zero* element
(2) $\langle u, v\rangle = \langle v, u\rangle, \quad \forall u, v \in S$
(3) $\langle \alpha u + \beta v, \gamma w\rangle = \alpha\gamma\langle u, w\rangle + \beta\gamma\langle v, w\rangle \quad \forall \alpha, \beta, \gamma \in \mathbb{R}; u, v, w \in S$

An important property of the inner product is that it satisfies the Cauchy–Schwarz inequality:

$$|\langle u, v\rangle|^2 \leq \langle u, u\rangle \langle v, v\rangle$$

A vector space S, equipped with an inner product $\langle \cdot, \cdot\rangle$, is called an *inner-product space*. The inner product induces a natural norm, $\|u\| := \sqrt{\langle u, u\rangle}$, so that an inner-product space is also a normed vector space. Furthermore, if this inner-product space is also complete (i.e., a Banach space), it is called a real *Hilbert space*, which we denote by $\mathbf{H}$.

An important notion in a Hilbert space $\mathbf{H}$ is orthogonality. We say that $u \in \mathbf{H}$ and $v \in \mathbf{H}$ are *orthogonal* if $\langle u, v\rangle = 0$, in which case we write $u \perp v$. If S is a subspace of $\mathbf{H}$, then the *orthogonal complement* $S^{\perp}$ is defined by

$$S^{\perp} = \{u \in \mathbf{H} : \langle u, v\rangle = 0 \quad \forall v \in S\}.$$

An important property of $S^{\perp}$ is that it is a *closed linear subspace* of $\mathbf{H}$, and that given any $x \in \mathbf{H}$ there is a unique element in $S^{\perp}$ (say $x_{S^{\perp}}$) and a unique element in the closure of S ($\bar{S}$) (say $x_{\bar{S}}$) such that

$$x = x_{\bar{S}} + x_{S^{\perp}}.$$

Equivalently, $\mathbf{H}$ can be written as the direct sum of $\bar{S}$ and $S^{\perp}$:

$$\mathbf{H} = \bar{S} \oplus S^{\perp}.$$

In the above, $x_{S^{\perp}}$ is called the *orthogonal projection* of x *on* $S^{\perp}$, and we have the norm identity

$$\|x\|^2 = \|x_{\bar{S}}\|^2 + \|x_{S^{\perp}}\|^2,$$

which is also known as the *Pythagorean theorem.*

A fundamental result in functional analysis, which we will have occasion to use in the text, is the celebrated *orthogonal projection theorem*:

The Orthogonal Projection Theorem. Let $\mathbf{H}$ be a Hilbert space and S a closed subspace of it. Then, given $x \in \mathbf{H}$, there exists a unique $s_0 \in S$, such that

$$\|x - s_0\| = \min_{s \in S} \|x - s\|.$$

Furthermore, a necessary and sufficient condition for $s_0 \in S$ to be the minimizing vector is that $(x - s_0) \perp S$.

Transformations and Continuity

A mapping f of a vector space S into a vector space T is called a *transformation* or a *function* and is written symbolically $f : S \to T$ or $y = f(x)$, for $x \in S, y \in T$. f is said to be a *functional* if $T = \mathbb{R}$.

Let $f : S \to T$ where S and T are normed linear spaces. f is said to be *continuous at a point* $x_0 \in S$ if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in N_\epsilon(f(x_0))$ for every $x \in N_\delta(x_0)$, where the notation $N_\delta(z) := \{x : \|x - z\| \leq \delta\}$ denotes the ball with radius δ centered at z. If f is continuous at every point of S it is said to be *continuous everywhere* or, simply, *continuous*. f is said to be *uniformly continuous* if δ depends only on ϵ and not on x_0. Furthermore, f is said to be *linear* (or a *linear operator*) if $f(\alpha u + \beta v) = \alpha f(u) + \beta f(v)$, for all $u, v \in S$; $\alpha, \beta \in \mathbb{R}$.

Let S and T be as above and F be a family of functions from S to T. The family F is said to be *equicontinuous at a point* $x_0 \in S$ if, for every $\epsilon > 0$, there exists a $\delta > 0$ such that $f(x) \in N_\epsilon(f(x_0))$ for all $f \in F$ and for every $x \in N_\delta(x_0)$. The family F is said to be *equicontinuous* if it is equicontinuous at each $x \in S$. The family F is *uniformly equicontinuous* if δ depends only on ϵ and not on x_0.

Let S and T be Banach spaces, and $A : S \to T$ a linear operator. The quantity

$$\ll A \gg = \sup_{\|x\| \leq 1} \|Ax\|$$

is called the (*operator*) *norm* of A, and A is said to be *bounded* if there exists a scalar $k > 0$ such that $\ll A \gg \leq k$. A is said to be a *contraction operator* if this bound is less than 1, i.e., $\ll A \gg < 1$.

If $S = T$, another related quantity is the *spectral radius* of a bounded linear operator $A : S \to S$, defined by

$$\rho(A) = \lim_{k \to \infty} \sup \, [\ll A^k \gg]^{1/k}.$$

An important relationship between the norm and the spectral radius is the inequality

$$\rho(A) \leq \ll A \gg,$$

for $A : S \to S$.

Now again let S and T not necessarily be the same, but assume that each is a Hilbert space. Then, the adjoint of a linear bounded operator $A : S \to T$, denoted by A^*, is defined by the equality

$$< Au, v > = < u, A^* v >, \quad \forall u \in S, \ \forall \, v \in T.$$

Note that here the two inner products are not the same, unless S and T are identical. Now let $S = T =: \mathbf{H}$, a common Hilbert space, and introduce the notation $\mathcal{L}(\mathbf{H})$ to denote the space of all linear bounded operators from $\mathbf{H}$ into $\mathbf{H}$. Then we have the following useful terminology and properties:

1. $A \in \mathcal{L}(\mathbf{H})$ is said to be *self-adjoint* if $A^* = A$.
2. For a self-adjoint operator A,
$$\ll A \gg = [\rho(AA^*)]^{\frac{1}{2}} = [\rho(A^*A)]^{\frac{1}{2}}.$$
3. A self-adjoint operator A is *positive* (equivalently, *nonnegative*) if
$$\langle u, Au \rangle \ \geq \ 0, \quad \forall u \in \mathbf{H}.$$
A is *strongly positive* if there exists a scalar $\alpha > 0$ such that
$$\langle u, Au \rangle \ \geq \ \alpha \langle u, u \rangle, \quad \forall u \in \mathbf{H}.$$
4. The norm of a self-adjoint operator $A \in \mathcal{L}(\mathbf{H})$ is also given by
$$\ll A \gg = \ \sup_{\|x\| \leq 1} |\langle x, Ax \rangle|.$$
5. Every strongly positive operator $A \in \mathcal{L}(\mathbf{H})$ has a unique strongly positive square root $A^{\frac{1}{2}} \in \mathcal{L}(\mathbf{H})$, so that $A^{\frac{1}{2}} A^{\frac{1}{2}} = A$.
6. Given an $A \in \mathcal{L}(\mathbf{H})$, its inverse A^{-1} is defined by
$$A^{-1}A = AA^{-1} = I,$$
where $I \in \mathcal{L}(H)$ is the *identity* operator (i.e., $Ix = x$, $\forall x \in \mathbf{H}$). A^{-1} also belongs to $\mathcal{L}(\mathbf{H})$.
7. Let S be a closed linear subspace of $\mathbf{H}$. $P \in \mathcal{L}(H)$ is called a *projection operator* (into S) if $Px \in S$ for every $x \in \mathbf{H}$. Here S is called the *range space* of P. Two important properties of P are that (i) $P^2 = I$ (i.e., P is *idempotent*) and (ii) $\ll P \gg = 1$. (Projection operators are also defined on more general Banach spaces.)
8. Given two operators A, $B \in \mathcal{L}(\mathbf{H})$, the following inequalities hold:
$$\begin{aligned} &\ll AB \gg \leq \ll A \gg \ \ll B \gg \\ &\rho(AB) \qquad \leq \rho(A)\rho(B). \end{aligned}$$
(The first inequality is valid also on Banach spaces, S and T, where $A : M \to T$, $B : S \to M$, and M is another Banach space. Here the operator norm is the one induced by the norm defined on T.)

If a linear operator has finite-dimensional domain and range spaces, it is called a matrix, whose properties are further discussed in the following section.

A.3 Matrices

An $(m \times n)$ matrix A is a rectangular array of numbers, called *elements* or *entries*, arranged in m rows and n columns. The element in the ith row and jth column of A is denoted by a subscript ij, such as a_{ij} or $[A]_{ij}$, in which case we write $A = \{a_{ij}\}$.

A matrix is said to be *square* if it has the same number of rows and columns; an $(n \times n)$ square matrix A is said to be an *identity matrix* if $a_{ii} = 1, i = 1, \ldots, n$ and $a_{ij} = 0, i \neq j, i, j = 1, \ldots, n$. An $(n \times n)$ identity matrix will be denoted by I_n or, simply, by I whenever its dimension is clear from the context.

The *transpose* of an $(m \times n)$ matrix A is the $(n \times m)$ matrix A' with elements $a'_{ij} = a_{ji}$. A square matrix A is symmetric if $A = A'$; it is nonsingular if there is an $(n \times n)$ matrix called the inverse of A, denoted by A^{-1}, such that $A^{-1}A = I = AA^{-1}$.

Eigenvalues and Quadratic Forms

Corresponding to a square matrix A, a scalar (possibly complex valued) λ and a nonzero vector (also possibly complex valued) x satisfying the equation $Ax = \lambda x$ are said to be, respectively, an *eigenvalue* and an *eigenvector* of A.

A square symmetric matrix A has all its eigenvalues real. If they are all positive (respectively, nonnegative), then A is said to be *positive definite* (respectively, *nonnegative definite* or *positive semi-definite*). An equivalent definition is as follows. A symmetric $(n \times n)$ matrix A is said to be positive definite (respectively, nonnegative definite) if $x'Ax > 0$ (respectively, $x'Ax \geq 0$) for all nonzero $x \in \mathbb{R}^n$. The matrix A is said to be *negative definite* (respectively, *nonpositive definite*) if the matrix $(-A)$ is positive (respectively, nonnegative) definite. We symbolically write $A > 0$ (respectively, $A \geq 0$) to denote that A is positive (respectively, nonnegative) definite.

Norm and Spectral Radius

The *norm* of an $(m \times n)$ matrix A, consistent with the operator norm introduced in Sect. A.2, is given by

$$\ll A \gg = \max\{\lambda : \det(A'A - \lambda I) = 0\} ;$$

in other words, $\ll A \gg$ is the maximum eigenvalue of the symmetric matrix $A'A$ (or equivalently that of AA').

If A is a square matrix, its *spectral radius* is equal to

$$\rho(A) = \max\{|\lambda| : \det(A - \lambda I) = 0\},$$

that is, it is the magnitude of that eigenvalue of A with maximum absolute value.

A.4 Convex Sets and Functionals

A subset C of a vector space S is *convex* if for every $u, v \in C$ and every $\alpha \in [0, 1]$, we have $\alpha u + (1 - \alpha)v \in C$. A functional $f : C \to \mathbb{R}$ defined over a convex subset C of a vector space S is *convex* if, for every $u, v \in C$ and every scalar $\alpha \in [0, 1]$, we have $f(\alpha u + (1 - \alpha)v) \leq \alpha f(u) + (1 - \alpha)f(v)$. If this is a strict inequality for every $\alpha \in (0, 1)$, then f is *strictly convex*. The functional f is *concave* if $(-f)$ is convex and *strictly concave* if $(-f)$ is strictly convex.

A functional $f : \mathbb{R}^n \to \mathbb{R}$ is *continuously differentiable* if, with $x \in \mathbb{R}^n$, the partial derivatives of f with respect to the components of x exist and are continuous, in which case we write

$$\nabla f(x) = [\partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n].$$

$\nabla f(x)$ is the *gradient* of f at x and is a row vector. We shall also use the notation $f_x(x)$ or $df(x)/dx$ to denote the same quantity. If we partition x into two vectors y and z of dimensions n_1 and $n - n_1$, respectively, and are interested only in the partial derivatives of f with respect to the components of y, then we use the notation $\nabla_y f(y, z)$ or $\partial f(y, z)/\partial z$ to denote this partial gradient.

Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a vector-valued function whose components are continuously differentiable with respect to the components of $x \in \mathbb{R}^n$. Then, we say that $g(x)$ is differentiable, with the derivative $dg(x)/dx$ being an $(m \times n)$ matrix whose ijth element is $\partial g_i(x)/\partial x_j$. (Here g_i denotes the ith component of g.) The gradient $\nabla f(x)$ being a vector, its derivative (which is the second derivative of $f : \mathbb{R}^n \to \mathbb{R}$) will thus be an $(n \times n)$ matrix, assuming that $f(x)$ is twice continuously differentiable in terms of components of x. This matrix, denoted by $\nabla^2 f(x)$, is symmetric and is called the *Hessian matrix* of f at x.

Properties of Convex Functionals

1. A functional $f : \mathbb{R}^n \to \mathbb{R}$ which is twice continuously differentiable on $\mathbb{R}^n$ is *convex* (respectively, *strictly convex*) if, and only if, its Hessian matrix $\nabla^2 f(x)$ is *nonnegative definite* (respectively, *positive definite*) for all $x \in \mathbb{R}^n$.
2. Let C be a convex subset of a vector space S and $f : C \to \mathbb{R}$ be a *convex* (respectively, *strictly convex*) functional. Let u and v be two vectors in C and let

$[a, b]$ be a subinterval of $\mathbb{R}$ such that for all $\epsilon \in [a, b]$, $u + \epsilon v \in C$. Then, the function $h : \mathbb{R} \to \mathbb{R}$, defined by $h(\epsilon) := f(u + \epsilon v)$, is a *convex* (respectively, *strictly convex*) functional on $[a, b]$.

3. If $h : \mathbb{R} \to \mathbb{R}$ is convex on an interval $[a, b]$, then for c_1, c_2, c_3 such that $a \leq c_1 < c_2 < c_3 \leq b$,

$$\frac{h(c_2) - h(c_1)}{c_2 - c_1} \leq \frac{h(c_3) - h(c_1)}{c_3 - c_1}.$$

If h is strictly convex, then the inequality above is strict.

4. Let h, and (c_1, c_2) be as above, and let $\dot{h}_+(c_1)$ denote the right-hand derivative of h at the point c_1. Then,

$$\dot{h}_+(c_2) \leq \frac{h(c_2) - h(c_1)}{c_2 - c_1}$$

with the inequality being strict if h is strictly convex. Of course, if h is differentiable at c_1, then the derivative of h satisfies this bound.

A.5 Optimization of Functionals

Given a functional $f : S \to \mathbb{R}$, where S is a vector space, and a subset $X \subseteq S$, by the optimization problem

$$\text{minimize } f(x) \text{ subject to } x \in X,$$

we mean the problem of finding an element $x^* \in X$ (called a *minimizing element* or an *optimal solution*) such that

$$f(x^*) \leq f(x) \quad \forall x \in X.$$

This is sometimes also referred to as a *globally minimizing solution* in order to differentiate it from the other alternative—a *locally minimizing solution*. An element $x^\circ \in X$ is called a locally minimizing solution if there exists an $\epsilon > 0$ such that

$$f(x^\circ) \leq f(x) \quad \forall x \in N_\epsilon(x^\circ) \cap X,$$

i.e., we compare $f(x^\circ)$ with values of $f(x)$ in that part of a certain ϵ-neighborhood of x°, which lies in X.

For a given optimization problem, it is not necessary that an optimal solution exists; an optimal solution will exist if the set of real numbers $\{f(x) : x \in X\}$ is bounded below and there exists an $x^* \in X$ such that $\inf\{f(x) : x \in X\} = f(x^*)$, in which case we write

$$f(x^*) = \inf_{x \in X} f(x) = \min_{x \in X} f(x).$$

If such an x^* cannot be found, even though inf $\{f(x) : x \in X\}$ is finite, we simply say that an optimal solution does not exist, but we declare the quantity

$$\text{inf}\,\{f(x) : x \in X\} \quad \text{or} \quad \inf_{x \in X} f(x)$$

as the *optimal value* of the optimization problem. In this case, given any $\epsilon > 0$, one can find $x_\epsilon \in X$ with the property that $f(x_\epsilon) < \inf_{x \in X} f(x) + \epsilon$, that is, one can get arbitrarily close to $\inf_{x \in X} f(x)$ by picking an element out of X. If $\{f(x) : x \in X\}$ is not bounded below, i.e., $\inf_{x \in X} f(x) = -\infty$, then neither an optimal solution nor an optimal value exists.

An optimization problem which involves maximization instead of minimization may be converted into a minimization problem by simply replacing f by $-f$. Any optimal solution of this minimization problem is also an optimal solution for the initial maximization problem and the optimal value of the former. When a *maximizing element* $x^* \in X$ exists, then $\sup_{x \in X} f(x) = \max_{x \in X} f(x) = f(x^*)$.

Existence of Optimal Solutions

In the minimization problem formulated above, an optimal solution exists if X is a finite set, since then there is only a finite number of comparisons. When X is not finite, however, existence of an optimal solution is guaranteed if f is continuous (or lower semicontinuous) and X is compact—a result known as the *Weierstrass theorem*. For the special case when X is finite dimensional, we should recall that compactness is equivalent to being closed and bounded.

Necessary and Sufficient Conditions for Optimality

Let $S = \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. If X is an open set, a first-order necessary condition for an optimal solution to satisfy is

$$\nabla f(x^*) = 0.$$

If, in addition, f is twice continuously differentiable on $\mathbb{R}^n$, a second-order necessary condition is

$$\nabla^2 f(x^*) \geq 0.$$

The pair of conditions $\{\nabla f(x^*) = 0, \nabla^2 f(x^*) > 0\}$ is sufficient for $x^* \in X$ to be a locally minimizing solution. These conditions are also sufficient for global optimality if, in addition, X is a convex set and f is a convex functional on X.

These results, by and large, hold also for the case when S is infinite dimensional, but then one has to replace the gradient vector and the Hessian matrix by first and second Gateaux (or Frèchet) derivatives, and positive definiteness requirement is replaced by "strong positiveness" of an operator. See [242] for these extensions.

A.6 Contraction Mappings and Fixed-Point Theorems

Let S be a normed linear vector space, with norm $\|\cdot\|$. A mapping f of S into itself, is called a *contraction of* S (or equivalently, a *contraction mapping*) if there exists a scalar $p \in [0,1)$ such that

$$\|f(u) - f(v)\| \leq p\|u - v\|, \quad \forall u, v \in S.$$

It is possible to extend this definition to *metric* spaces where the "distance" function $\|u - v\|$ is replaced by a more general metric $d(u, v)$. Consider an equation of the form

$$f(u) = u$$

defined on S. Such an equation is known as a *fixed-point equation*, and any solution to it (i.e., a $u^\circ \in S$ such that $f(u^\circ) = u^\circ$) is called a *fixed point* of f.

Theorem A.6.1 (Banach). *If S is a Banach space under the norm $\|\cdot\|$ and $f : S \to S$ is a continuous mapping which is also a contraction, then f has a unique fixed point.* ◇

It is possible for f to have a fixed point without being a contraction mapping. A sufficient condition for such an existence is provided in the following theorem.

Theorem A.6.2 (Schauder). *Let S be a Banach space, X a nonempty convex set in S, and Y a compact subset of X. Let $f : X \to Y$ be a continuous map. Then there is a fixed point of f (not necessarily unique).* ◇

When specialized to finite-dimensional spaces, this result is known as *Brouwer's fixed-point theorem*, which of course predates the more general one:

Theorem A.6.3 (Brouwer). *Let $S \equiv \mathbb{R}^n$ and X be a closed and bounded subset of S. Let $f : X \to X$ be continuous. Then, f has a fixed point.* ◇

One way of computing the fixed point of a mapping $f : S \to S$, is to use the iteration

$$u_{(k+1)} = f(u_{(k)}), \quad k = 0, 1, \ldots$$

which is called "successive approximation." If S is a Banach space and f is contraction mapping, then the iteration above converges to the unique fixed point of f, for all starting points $u_{(0)} \in S$. If f is an affine mapping, then this result can be strengthened, as given in the following theorem.

Theorem A.6.4 (Successive Approximation). *Let S be a Banach space and A a linear bounded operator mapping S into itself. Consider the equation*

$$u = Au + b \tag{$*$}$$

defined on S, where $b \in S$ is given. Furthermore, consider the "successive approximation"

$$u_{(k+1)} = Au_{(k)} + b, \quad k = 0, 1, \dots \tag{$**$}$$

to the solution of $(*)$. Then, the sequence $\{u_{(k)}\}$ generated by $(**)$ converges to a unique element of S, for any starting point $u_{(0)} \in S$, if, and only if, the spectral radius of A is less than 1 (i.e., $\rho(A) < 1$). The limit of $\{u_{(k)}\}$ is the unique fixed point of $(*)$. ◇

Appendix B
Probability Theory and Stochastic Processes

This appendix presents some notions of probability theory and stochastic processes which are used in the text. For a more complete exposition the reader should consult with the standard texts on probability theory, such as [16, 137, 241, 304], and texts on stochastic processes, such as [402].

B.1 Probability

B.1.1 Measurable Spaces

Let $\mathbb{X}$ be a collection of points (elements) and $\mathcal{F}$ be a collection of subsets of $\mathbb{X}$ with the following properties:

- $\mathbb{X} \in \mathcal{F}$
- If $A \in \mathcal{F}$, then $\mathbb{X} \setminus A \in \mathcal{F}$
- If $A_k \in \mathcal{F}, k = 1, 2, 3, \ldots$, then their countable union is also in $\mathcal{F}$, *i.e.*,

$$\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$$

Then, $\mathcal{F}$ is said to be a σ-field. By De Morgan's laws, $\mathcal{F}$ has to be closed under countable intersections as well. Note, for example, that the full power set of any set is a σ-field. If the countable union above is replaced with finite union, $\mathcal{F}$ is called a field.

With the above, the pair $(\mathbb{X}, \mathcal{F})$ is called a *measurable space*.

If elements of $\mathbb{X}$ correspond to outcomes of a random experiment, then $\mathbb{X}$ is called a *sample space* and its subsets are called *events*. In this case one generally uses the notation Ω instead of $\mathbb{X}$.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,

Let $\mathcal{A}$ be a collection of sets and $\mathcal{J}$ be the smallest σ-field containing the sets in $\mathcal{A}$; in this case, we write $\mathcal{J} = \sigma(\mathcal{A})$. An important class of σ-fields is the Borel σ-field on a metric (or more generally topological) space, generated by open sets. We will denote the Borel σ-field on such a space $\mathbb{X}$ by $\mathcal{B}(\mathbb{X})$. A measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ is called *standard Borel*, if $\mathbb{X}$ is a Polish space.

Definition B.1.1. If $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ and $(\mathbb{Y}, \mathcal{B}(\mathbb{Y}))$ are two measurable spaces, we say a mapping $h : \mathbb{X} \to \mathbb{Y}$ is measurable if

$$h^{-1}(B) = \{x : h(x) \in B\} \in \mathcal{B}(\mathbb{X}), \quad \forall B \in \mathcal{B}(\mathbb{Y}).$$

◇

A positive measure μ on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ is a function from $\mathcal{B}(\mathbb{X})$ to $[0, \infty]$ which is *countably additive*, that is, for $A_k, A_j \in \mathcal{B}(\mathbb{X})$ and $A_k \cap A_j = \emptyset \quad \forall k, j \in \mathbb{N}$,

$$\mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k).$$

Definition B.1.2. μ is a probability measure if it is a positive measure and $\mu(\mathbb{X}) = 1$. ◇

Definition B.1.3. A measure μ is finite if $\mu(\mathbb{X}) < \infty$ and σ-finite, if there exists a collection of subsets $\{A_k\}$ such that $\mathbb{X} = \bigcup_{k=1}^{\infty} A_k$ with $\mu(A_k) < \infty$ for all k. ◇

Theorem B.1.1 (Carathéodory's Extension Theorem). *Let $\mathcal{M}$ be a collection of subsets, which is a field, and suppose that there exists a countably additive measure P on $\mathcal{M}$. Then, there exists a unique measure P' on the σ-field generated by $\mathcal{M}$, $\sigma(\mathcal{M})$, which is consistent with P on $\mathcal{M}$.* ◇

Theorem B.1.2 (Kolmogorov's Extension Theorem). *Let $\mathbb{X}$ be a complete, separable metric space and let μ_n be a probability measure on $\mathbb{X}^n$, the n product of $\mathbb{X}$, for each $n = 1, 2, \ldots$, such that*

$$\mu_n(A_1 \times A_2 \times \cdots \times A_n) = \mu_{n+1}(A_1 \times A_2 \times \cdots \times A_n \times \mathbb{X}),$$

for every n and every sequence of Borel sets A_k. Then, there exists a unique probability measure μ on $(\mathbb{X}^{\infty}, \mathcal{B}(\mathbb{X}^{\infty}))$ which is consistent with each of the μ_n's.

◇

The above result is useful because to determine whether two measures are equal it suffices to check if they are equal on the collection of sets which generate the σ-field, and not necessarily on the entire σ-field. For example, if the σ-field on a product space $\mathbb{X}^{\mathbb{Z}}$ is generated by a sequence of finite-dimensional distributions, one can define a measure on the product space which is consistent with the finite-dimensional distributions. In particular, the Borel σ-field defined on a product space is generated by the open sets in the product topology, which in turn allows one to uniquely define a probability measure on the infinite-dimensional space by its measures on cylinder sets of the form

$$\{x = (x_0, x_1, \cdots) \in \mathbb{X}^{\mathbb{Z}_+} : x_m \in A_m \in \mathcal{B}(\mathbb{X}), m \in I \subset \mathbb{Z}_+\}.$$

One could construct the Lebesgue measure on the Borel σ-field on $\mathbb{R}$ by restricting it to intervals only, since a field can be constructed by finitely many disjoint unions of intervals. The Lebesgue measure λ is defined on the Borel σ-field such that for $A = (a, b)$, $\lambda(A) = b - a$.

B.1.2 Integration

Let h be a nonnegative measurable function from $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. The Lebesgue integral of h with respect to a measure μ can be defined in three steps:

First, for $A \in \mathcal{B}(\mathbb{X})$, define $1_{\{x \in A\}}$ (or $1_{(x \in A)}$, or $1_A(x)$) as an indicator function for event $x \in A$, that is, the function that takes the value 1 if $x \in A$, and 0 otherwise. In this case, $\int_{\mathbb{X}} 1_{\{x \in A\}} \lambda(dx) =: \lambda(A)$.

Next, define simple functions such that for $A_1, A_2, \ldots, A_n$, all in $\mathcal{B}(\mathbb{X})$, and positive numbers $b_1, b_2, \ldots, b_n$, $h_n(x) = \sum_{k=1}^n b_k 1_{\{x \in A_k\}}$. Then, $\int_{\mathbb{X}} h_n(x) \lambda(dx) =: \sum_{k=1}^n b_k \lambda(A_k)$.

Finally, observe that for any given nonnegative measurable function h, there exists a sequence of simple functions h_n such that $h_n(x) \uparrow h(x)$ monotonically. Then, the Lebesgue integral is defined as

$$\int h(x)\mu(dx) := \lim_{n\to\infty} \int h_n(x)\mu(dx) = \lim_{n\to\infty} \sum_{k=1}^n b_k^n \lambda(A_k^n).$$

There are four important convergence theorems related to integration.

Theorem B.1.3 (Monotone Convergence Theorem). *If μ is a σ-finite positive measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ and $\{f_n, n \in \mathbb{Z}_+\}$ is a sequence of measurable functions from $\mathbb{X}$ to $\mathbb{R}$ which pointwise, monotonically, converges to f, that is, $0 \le f_n(x) \le f_{n+1}(x)$ for all n, and*

$$\lim_{n\to\infty} f_n(x) = f(x),$$

for μ-almost every x, then

$$\lim_{n\to\infty} \int_{\mathbb{X}} f_n(x)\mu(dx) = \int_{\mathbb{X}} f(x)\mu(dx).$$

◇

Theorem B.1.4 (Extended Monotone Convergence Theorem). *Let h be a measurable and integrable function from $\mathbb{X}$ to $\mathbb{R}$, μ be a σ-finite positive measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, and $\{f_n, n \in \mathbb{Z}_+\}$ be a monotonically nondecreasing sequence of measurable functions from $\mathbb{X}$ to $\mathbb{R}$ which pointwise converges to f such that $h(x) \le f_n(x) \le f_{n+1}(x)$ for all n, and*

$$\lim_{n\to\infty} f_n(x) = f(x),$$

for μ-almost every x, and $\int \mu(dx)h(x) > -\infty$. Then,

$$\lim_{n\to\infty} \int_{\mathbb{X}} f_n(x)\mu(dx) = \int_{\mathbb{X}} f(x)\mu(dx).$$

◇

Theorem B.1.5 (Fatou's Lemma). *(also Fatou–Lebesgue Theorem) If μ is a σ-finite positive measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ and $\{f_n, n \in \mathbb{Z}_+\}$ is a sequence of nonnegative measurable functions from $\mathbb{X}$ to $\mathbb{R}$, then*

$$\int_{\mathbb{X}} \liminf_{n\to\infty} f_n(x)\mu(dx) \leq \liminf_{n\to\infty} \int_{\mathbb{X}} f_n(x)\mu(dx).$$

◇

Theorem B.1.6 (Dominated Convergence Theorem). *If (i) μ is a σ-finite positive measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, (ii) $g(x) \geq 0$ is a Borel measurable function such that*

$$\int_{\mathbb{X}} g(x)\mu(dx) < \infty,$$

and (iii) $\{f_n, n \in \mathbb{Z}_+\}$ is a sequence of measurable functions from $\mathbb{X}$ to $\mathbb{R}$ which satisfy $|f_n(x)| \leq g(x)$ for μ-almost every x, and $\lim_{n\to\infty} f_n(x) = f(x)$, then

$$\lim_{n\to\infty} \int_{\mathbb{X}} f_n(x)\mu(dx) = \int_{\mathbb{X}} f(x)\mu(dx).$$

◇

B.1.3 Probability Spaces and Random Variables

Let $(\Omega, \mathcal{F})$ be a measurable space and $\mathbf{P}$ be a probability measure on this space. Then, the triple $(\Omega, \mathcal{F}, \mathbf{P})$ is known as a *probability space*.

If $\Omega = \mathbb{R}^n$, then its subsets of interest are the n-dimensional rectangles, and the smallest σ-algebra generated by these rectangles is called the n-dimensional *Borel σ-algebra* and is denoted by $\mathcal{B}(\mathbb{R}^n)$. Elements of $\mathcal{B}(\mathbb{R}^n)$ are *Borel sets*, and the pair $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is a *Borel (measurable) space*. A probability measure defined on this space is known as a *Borel probability measure*.

If Ω is a finite set (say, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$), we can assign probability weights on individual elements of Ω, instead of on subsets of Ω, in which case we write p_i to denote the probability of the single event ω_i. We call the n-tuple $\{p_1, p_2, \ldots, p_n\}$ a *probability mass function (pmf)* on Ω. Clearly, we have the restriction $0 \leq p_i \leq 1\ \forall i = 1, \ldots, n$, and $\sum_{i=1}^{n} p_i = 1$. The same convention applies when Ω is countable (i.e., $\Omega = \{\omega_1, \omega_2, \ldots, \omega_i, \ldots\}$), in which case we simply replace n by ∞.

Let $(\Omega, \mathcal{F})$ and $(E, \mathcal{E})$ be two measurable spaces, and $X : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ be a measurable map. We call X an *E-valued random variable*, which is in fact a function.[1] If $\mathbf{P}$ is a probability measure on $(\Omega, \mathcal{F})$, then the image under X constitutes a probability measure on $(E, \mathcal{E})$, called the law of X and denoted by $\mathbf{P}_X$. If $E = \mathbb{R}^N$, we call X an N-dimensional random vector. In this case we can restrict, $\mathbf{P}_X$, without any loss of generality, to semi-infinite open rectangles, $\{\xi \in \mathbb{R}^N : \xi_i < a_i, i = 1, \ldots, N\}$, in $\mathcal{B}(\mathbb{R}^N)$, resulting in a function $P_X(a_1, a_2, \ldots, a_N)$ of N variables, $\{a_1, a_2, \ldots, a_N\}$, which are real numbers. This function is known as the *cumulative (probability) distribution function (cdf)* of the random vector X. The relationship with the probability measure $\mathbf{P}$ on the original probability space is:

$$P_X(a_1, a_2, \ldots, a_N) = \mathbf{P}(\{\omega \in \Omega : X_1(\omega) < a_1, X_2(\omega) < a_2, \ldots, X_N(\omega) < a_N\})$$

where X_i is the ith component of X.

The σ-field generated by the events $\{\{\omega : X(\omega) \in A\}, A \in \mathcal{E}\}$ is called the σ-field generated by X and is denoted by $\sigma(X)$. This is the smallest sub-σ-field of $\mathcal{F}$ on which X is measurable.

Probability Density Function

Let $\mathbf{P}$ be a Borel probability measure on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ such that any element of $\mathcal{B}(\mathbb{R}^N)$ which is of Lebesgue measure zero is also of $\mathbf{P}$-measure zero; then we say that $\mathbf{P}$ is *absolutely continuous* with respect to the Lebesgue measure. Now, if $X : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathbf{P}_X)$ is random vector and if $\mathbf{P}_X$ is absolutely continuous with respect to the Lebesgue measure, there exists a nonnegative Borel function $p_X(\cdot)$ such that, for every $A \in \mathcal{B}(\mathbb{R}^N)$,

$$\mathbf{P}_X(A) = \int_A p_X(\xi) d\xi.$$

Such a function $p_X(\cdot)$ is called the *probability density function* (*pdf*) of the random vector X, and X is said to be a *continuous random vector*. In terms of the cdf P_X, the preceding relationship can be written as

$$\mathbf{P}_X(a_1, \ldots, a_N) = \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_N} p_X(\xi_1, \ldots, \xi_N) d\xi_1 \ldots d\xi_N$$

for every N scalars $a_1, \ldots, a_N$.

---

[1] It is common to use uppercase letters for a random variable and lowercase letters for its realization, that is, $X(\omega) = x$, for $\omega \in \Omega$. In the book, to conserve notation, we will occasionally use lowercase letters for the random variable itself as well, when there is no ambiguity from context.

Let $X : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathbf{P}_X)$ be a random vector and $f : (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N)) \to \mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$ be a nonnegative Borel function. Then, $f(X)$ can be considered as a random vector from $(\Omega, \mathcal{F})$ to $(\mathbb{R}^M, \mathcal{B}(\mathbb{R}^M))$, and its *average value* (*expected value*) is defined either by $\int_\Omega f(X(\omega))\mathbf{P}(d\omega)$ or by $\int_{\mathbb{R}^n} f(\xi)\mathbf{P}_X(d\xi)$ depending on which interpretation one adopts. Both of these integrals are well defined and are uniquely equal in value. If f changes signs, then we take $f = f^+ - f^-$ where both f^+ and f^- are nonnegative and write the expected value of f as

$$\begin{aligned} E_X f(X) \equiv E[f(X)] &= \int_{\mathbb{R}^N} f^+(\xi)\mathbf{P}_X(d\xi) - \int_{\mathbb{R}^N} f^-(\xi)\mathbf{P}_X(d\xi) \\ &=: \int_{\mathbb{R}^N} f(\xi)\mathbf{P}_X(d\xi), \end{aligned}$$

provided that at least one of the pair $E[f^+(X)]$ and $E[f^-(X)]$ is finite. Since, by definition, $\mathbf{P}_X(d\xi) = \mathbf{P}_X(\xi + d\xi) - \mathbf{P}_X(\xi) =: d\mathbf{P}_X(\xi)$, this integral can further be written as

$$E[f(X)] = \int_{\mathbb{R}^N} f(\xi) d\mathbf{P}_X(\xi),$$

which is a Lebesgue-Stieltjes integral. For the special case when $f(x) = x$, we have

$$E[X] := \int_{\mathbb{R}^N} \xi\mathbf{P}_X(d\xi) =: \bar{X} \equiv \mu_X,$$

which is known as the *mean* (*expected*) *value* of X. The *covariance* of the N-dimensional random vector X is defined as

$$E[(X - \bar{X})(X - \bar{X})'] = \int_{\mathbb{R}^N} (\xi - \bar{X})(\xi - \bar{X})'\mathbf{P}_X(d\xi) =: \text{cov}(X)$$

which is a nonnegative definite matrix of dimensions $N \times N$. Now, if $\mathbf{P}_X$ is absolutely continuous with respect to the Lebesgue measure, $E[f(X)]$ can equivalently be written, in terms of the corresponding density function p_X, as

$$E[f(X)] = \int_{\mathbb{R}^N} f(\xi) p_X(\xi) d\xi.$$

If Ω consists of only a finite number of disjoint events $\omega_1, \omega_2, \dots, \omega_n$, then the integrals are all replaced by the single summation

$$E[f(X(\omega))] = \sum_{i=1}^{n} f(X(\omega_i)) p_i,$$

where p_i denotes the probability of occurrence of event ω_i. For a countably infinite set Ω, we have the counterpart

$$E[f(X(\omega)) = \lim_{n\to\infty} \sum_{i=1}^{n} f(X(\omega_i))p_i.$$

A continuous N-dimensional random vector $X := (X_1, \ldots, X_N)$ is said to be *Gaussian distributed* (or simply, a *Gaussian random vector*) with parameters $(\underline{\mu}, \Sigma)$—$\underline{\mu}$ an N-vector, Σ a (symmetric) positive definite matrix of dimensions $N \times N$—if its *pdf* is given by

$$p_X(x) = \frac{1}{(2\pi)^{\frac{N}{2}}(\det \Sigma)^{\frac{1}{2}}} \exp\{-\frac{1}{2}(x-\underline{\mu})'\Sigma^{-1}(x-\underline{\mu})\},$$

where $x = (x_1, \ldots, x_N)$. In this case we use the notation $X \sim N(\underline{\mu}, \Sigma)$ to denote that X is Gaussian with the given mean vector and covariance matrix.

Let $\mathcal{L}$ be the class of all N-dimensional random vectors $X : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N), \mathbf{P}_X)$, with bounded second moments (i.e., $E[X'X] := \sum_{i=1}^{N} E[(X_i)^2] < \infty$). Then $\mathcal{L}$ is a *Hilbert space* (see Appendix A for a definition) under the *inner product* $\langle X, Y\rangle := E[X'Y]$, for $X, Y \in \mathcal{L}$. A consequence of this is the important (and useful) *Cauchy-Schwarz inequality*:

$$\mid E[X'Y] \mid^2 \leq E[X'X]E[Y'Y].$$

Sequences of Random Variables and Convergence

A *random sequence* $\{X_i\}_{i=1}^{\infty}$ is a denumerable family of random variables

$$X_i : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}, P_{X_i}), i = 1, 2, \ldots.$$

Given a random sequence $\{X_i\}$, we say that it *converges in probability* to $X : (\Omega, \mathcal{F}, \mathbf{P}) \to (\mathbb{R}, \mathcal{B}, \mathbf{P}_X)$ (and write $X_i \to_{\text{p}} X$) if, for every $\epsilon > 0$, Prob$\{\mid X_i - X \mid \geq \epsilon\} \to 0$ as $i \to \infty$. We say that $\{X_i\}$ *converges almost surely* (*a.s.*) to X (and write $X_i \to_{\text{as}} X$), if $X_i(\omega) \to X(\omega)$ pointwise, except possibly on a subset of Ω which receives *zero* probability measure (under $\mathbf{P}$).

B.2 Convergence of Probability Measures

Let $\mathcal{P}(\mathbb{R}^{\mathbb{N}})$ denote the family of all probability measures on $(\mathbb{R}^N, \mathcal{B}(\mathbb{R}^N))$ for some $N \in \mathbb{N}$. Let $\{\mu_n, n \in \mathbb{N}\}$ be a sequence in $\mathcal{P}(\mathbb{R}^N)$. It is said to converge to $\mu \in$

$\mathcal{P}(\mathbb{R}^N)$ *weakly* if

$$\int_{\mathbb{R}^N} c(x)\mu_n(dx) \to \int_{\mathbb{R}^N} c(x)\mu(dx)$$

for every continuous and bounded $c : \mathbb{R}^N \to \mathbb{R}$. On the other hand, $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^N)$ *setwise* if

$$\int_{\mathbb{R}^N} c(x)\mu_n(dx) \to \int_{\mathbb{R}^N} c(x)\mu(dx)$$

for every measurable and bounded $c : \mathbb{R}^N \to \mathbb{R}$. Setwise convergence can also be defined through pointwise convergence on Borel subsets of $\mathbb{R}^N$ (see, e.g., [195]), that is,

$$\mu_n(A) \to \mu(A), \quad \text{for all } A \in \mathcal{B}(\mathbb{R}^N),$$

since the space of simple functions is dense in the space of bounded and measurable functions under the supremum norm.

For two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^N)$, the *total variation* metric is defined by

$$\begin{aligned} \|\mu - \nu\|_{TV} &:= 2 \sup_{B \in \mathcal{B}(\mathbb{R}^N)} |\mu(B) - \nu(B)| \\ &= \sup_{f\colon \|f\|_\infty \le 1} \left| \int f(x)\mu(dx) - \int f(x)\nu(dx) \right|, \end{aligned} \tag{B.1}$$

where the supremum is over all measurable real f such that $\|f\|_\infty := \sup_{x\in\mathbb{R}^N} |f(x)| \le 1$. A sequence $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^N)$ in total variation if $\|\mu_n - \mu\|_{TV} \to 0$.

Setwise convergence is equivalent to pointwise convergence on Borel sets, whereas total variation requires uniform convergence on Borel sets. Thus these three convergence notions are in increasing order of strength: convergence in total variation implies setwise convergence, which in turn implies weak convergence.

We close this section with two useful measurability results. A proof of the first result can be found in [6] (see Theorem 15.13 in [6] or p. 215 in [66]).

Theorem B.2.1. *Let $\mathbb{S}$ be a Polish space and M be the set of all measurable and bounded functions $f : \mathbb{S} \to \mathbb{R}$. Then, for any $f \in M$, the integral*

$$\int \pi(dx)f(x)$$

defines a measurable function on $\mathcal{P}(\mathbb{S})$ under the topology of weak convergence. $\diamond$

This is a useful result since it allows us to define measurable functions in integral forms on the space of probability measures when we work with the topology of weak convergence.

The second useful result follows from Theorem B.2.1 and Theorem 2.1 of Dubins and Freedman [123] and Proposition 7.25 in Bertsekas and Shreve [57].

Theorem B.2.2. *Let $\mathbb{S}$ be a Polish space. A function $F : \mathcal{P}(\mathbb{S}) \to \mathcal{P}(\mathbb{S})$ is measurable on $\mathcal{B}(\mathcal{P}(\mathbb{S}))$ (under weak convergence), if for all $B \in \mathcal{B}(\mathbb{S})$, $(F(\cdot))(B) : \mathcal{P}(\mathbb{S}) \to \mathbb{R}$ is measurable under weak convergence on $\mathcal{P}(\mathbb{S})$, that is, for every $B \in \mathcal{B}(\mathbb{S})$, $(F(\pi))(B)$ is a measurable function when viewed as a function from $\mathcal{P}(\mathbb{S})$ to $\mathbb{R}$.* ◇

B.3 Conditional Expectation and Estimation

Definition B.3.1. Given a measurable space $(\Omega, \mathcal{F})$, let $\mathcal{G} \subset \mathcal{F}$ be a (sub-) sigma-field. A function $g : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is said to be *measurable with respect to* $\mathcal{G}$ if

$$\text{for every } B \in \mathcal{B}(\mathbb{R}),\ \{\omega \in \Omega : g(\omega) \in B\} \in \mathcal{G}$$

where $\mathcal{B}$ is the Borel field. In this case we say that the function g is $\mathcal{G}$*-measurable*. ◇

Fact. Suppose that X and Y are two real random variables on $(\Omega, \mathcal{F}, \mathbf{P})$, where X is $\sigma(Y)$-measurable. Then, there is a Borel function $g : \mathbb{R} \to \mathbb{R}$ such that $X = g(Y)$. ◇

Definition B.3.2. Let X be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ such that $E[|X|] < \infty$ and Y be another random variable (on the same probability space) with generated sigma-field $\sigma(Y)$. *Conditional expectation of* X *given* $\sigma(Y)$ is any $\sigma(Y)$-measurable random variable Z satisfying

$$\int_A Z(\omega)\mathbf{P}(d\omega) = \int_A X(\omega)\mathbf{P}(d\omega)$$

for every $A \in \sigma(Y)$. Such a Z is *essentially unique* and is written as $E[X \mid \sigma(Y)]$ and sometimes as $E[X \mid Y]$. ◇

Note. In the definition above, we can replace $\sigma(Y)$ by any sigma-field $\mathcal{G} \subset \mathcal{F}$, in which case we call $E[X \mid \mathcal{G}]$ the "conditional expectation of X given $\mathcal{G}$."

Some Properties

1. If Y_1 and Y_2 are two random variables, defined on the same probability space, then

$$\sigma(Y_1, Y_2) \supset \sigma(Y_1) \text{ and } \sigma(Y_1, Y_2) \supset \sigma(Y_2).$$

Hence, by the law of the iterated expectation,

$$E\{E[X \mid \sigma(Y_1, Y_2)] \mid \sigma(Y_i)\} = E[X \mid \sigma(Y_i)], \text{ a.s., } i = 1, 2. \qquad \text{(B.2)}$$

2. Let $\mathcal{A} = \{\Omega, \theta\}$, which is the trivial sigma-field. Then,
$$E[X \mid \mathcal{A}] = E[X].$$
3. Let $\mathcal{G} \subset \mathcal{F}$ be a (sub-) sigma-field and W be a $\mathcal{G}$-measurable random variable. Then,
$$E[WX \mid \mathcal{G}] = WE[X \mid \mathcal{G}] \text{ a.s.}$$
4. Let X and Y be two integrable random variables defined on a common probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathcal{G} \subset \mathcal{F}$ be a (sub-) sigma-field. Then, for any two scalars α and β,
$$E\{[\alpha X + \beta Y] \mid \mathcal{G}\} = \alpha E[X \mid \mathcal{G}] + \beta E[Y \mid \mathcal{G}].$$
5. *Orthogonality principle*. Let X and Z be two random variables defined on the common probability space $(\Omega, \mathcal{F}, \mathbf{P})$, where Z is $\mathcal{G}$-measurable, with $\mathcal{G} \subset \mathcal{F}$ (a sub-sigma-field on $\mathcal{F}$). Then the random variables $Y := X - E[X \mid \mathcal{G}]$ and Z are uncorrelated, i.e., $E[YZ] = E[Y]E[Z] = 0$, and this is 0 since $E[Y] = 0$.

The last property has important implications in *minimum mean-square estimation*: Given a random variable X as above, and a sub-sigma-field $\mathcal{G}$ (of $\mathcal{F}$), we wish to find a $\mathcal{G}$-measurable random variable W that yields the minimum value to $E[(X - W)^2]$ among all such $\mathcal{G}$-measurable random variables.

The unique solution to this "estimation" problem is $W = E[X \mid \mathcal{G}]$, as given below as a *fact*.

Fact. Let X be a random variable on $(\Omega, \mathcal{F}, \mathbf{P})$ and $\mathcal{G} \subset \mathcal{F}$ be a given sub-sigma-field. Then,

$$E[(X - E[X \mid \mathcal{G}])^2] \leq E[(X - Z)^2], \text{ for all } \mathcal{G}\text{-measurable random variables } Z,$$

and equality holds if, and only if, $Z = E[X \mid \mathcal{G}]$. Here $Y := X - E[X \mid \mathcal{G}]$ is the "error" in the estimation of X using $\mathcal{G}$, and $E[Y^2]$ is the minimum mean-square error (*mmse*), which can also be written as $E[Y^2] = E[X^2] - E[(E[X \mid \mathcal{G}])^2]$. ◇

B.4 Stochastic Processes

A stochastic process $X := \{x_t, t \in \mathbf{T}\}$ is a parameterized family of random variables defined on the same probability space, with the parameter t being the "time variable." If the *time index* set $\mathbf{T}$ is finite or countably infinite, then the stochastic process is called a *discrete-time process*, whereas if it is uncountably infinite, such as the closed interval $[0, T]$, then the process is called a *continuous-time process*. In addition to being a well-defined random variable for every $t \in T$, a stochastic process should also have the property that for every finite subset S of T, the finite family of random variables $\{x_t, t \in S\}$ is a well-defined random vector.

A stochastic process $X := \{x_t, t \in \mathbf{T}\}$ is said to be a *second-order process*, if $E[|x_t|^2] < \infty$ for all $t \in \mathbf{T}$. For a second-order stochastic process, three quantities of interest (and of importance) are:

Mean function: $\mu_X(t) := E[x_t]$
Autocorrelation function: $R_X(t, s) := E[x_t x_s]$
Autocovariance function: $C_X(t, s) := E[(x_t - \mu_X(t))(x_s - \mu_X(s))] \equiv R_X(t, s)$
$-\mu_X(t)\mu_X(s)$

Estimation

Let $X := \{x_t, t \in \mathbf{T}\}$ and $Z := \{z_s, s \in \mathbf{T}\}$ be two stochastic processes, where $\mathbf{T}$ is a discrete- or a continuous-time interval. Suppose we are interested in estimating the value of X at some point, say $t \in \mathbf{T}$, based on some observed values of Z. Denote this estimate by $\hat{x}_t = \hat{g}_t(y_t)$, where y_t is the collection of observed values of Z by time t, and the function $\hat{g}_t(\cdot)$ is known as an *estimator*. There are several possibilities for y_t:

(i) $\mathbf{T} = [0, T]$ and $y_t = \{z_{t_1}, \ldots, z_{t_n}\}$, where $z_{t_i}, t_i \in [0, T]$ is some sampled value of Z. If $t_n \leq t$, we say that $\hat{g}$ is a *causal estimator*.
(ii) $y_t = \{z_s, s \leq t\} \Rightarrow \hat{x}_t = \hat{g}_t(z_s, s \leq t)$. Here all past values of the measurement are used in the construction of the estimate $\hat{x}_t$. Here $\hat{g}$ is also a causal estimator.
(iii) $y_t = \{z_s, s \in \mathbf{T}\} \Rightarrow \hat{x}_t = \hat{g}_t(z_s, s \in \mathbf{T})$. This is a *noncausal* estimator, which is also called a *smoother*.

We say that an estimator $\hat{g}$ is linear (affine) if it is given by a linear (affine) map, e.g., for a causal estimator,

$$\hat{g}_t(y_t) = k(t) + \int_0^t h(t, s) z_s ds, \qquad T = [0, \infty)$$

or

$$\hat{g}_n(y_n) = k_n + \sum_{i=1}^{n} h(n, i) z_i, \qquad \text{for } n \in T = \{0, 1, \ldots\},$$

where we have let $t = n$ for the discrete-time process. In all cases, we call $\hat{g}_t$ a *mms estimator* if it minimizes the quantity $E\{[x_t - g(y_t)]^2\}$ over the allowable class of g's. For each fixed y_t, the lowest possible value is attained by

$$\hat{x}_t = \hat{g}_t(y_t) = E[x_t \mid \sigma(y_t)] \equiv E[x_t \mid y_t]$$

provided that the conditional mean is well defined, where $\sigma(y_t)$ is the smallest sigma-field generated by y_t. Here again the *orthogonality principle* applies

$$E\{(x_t - \hat{x}_t)g_t(y_t)\} = 0 \text{ for all functions } g_t.$$

Suppose that now we restrict our attention to affine estimators, where $g_t(\cdot)$ is *affine* in its arguments, as described above. Denote the class of such random variables (estimators) under any one of the information schemes above by H_y and let $\hat{x}_t$ denote the affine estimate for x_t based on Z for a given information scheme, under the mean-squared error criterion, i.e.,

$$E\{(x_t - \hat{x}_t)^2\} \leq E\{(x_t - w)^2\}, \text{ for all } \ w \in H_y.$$

Then, it again follows from the *orthogonality principle* that $\hat{x}_t \in H_y$ solves the estimation problem if, and only if,

$$E\{[x_t - \hat{x}_t]w\} = 0 \text{ for all } \ w \in H_y.$$

Appendix C
Markov Chains, Martingales, and Ergodic Processes

This appendix provides some preliminary background on Markov chains, martingales, and ergodic processes. More comprehensive treatment of Markov chains can be found in [271, 294], martingales in any standard book on stochastic processes, such as [402], and ergodic processes from an information theoretic angle in [170, 175, 338].

C.1 Markov Chains

A sequence of random variables, ordered temporally, is a Markov chain if the conditional probability of any element in the sequence given observed values of past elements in the sequence depends only on the most recent past element. That is, if $x_{[0,N]} = \{x_0, x_1, \ldots, x_N\}$ denotes the sequence defined over a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, with each element taking values in a measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, then for each $B \in \mathcal{B}(\mathbb{X})$ and each $t \in [0, N-1]$,

$$\mathbf{P}(\{\omega \in \Omega : x_{t+1}(\omega) \in B\} \mid x_t, x_{t-1}, \ldots, x_0) = \mathbf{P}(\{\omega \in \Omega : x_{t+1}(\omega) \in B\} \mid x_t).$$

If P is the probability measure induced by $x_{[0,N]}$ on $(\mathbb{X}^{N+1}, \mathcal{B}(\mathbb{X}^{N+1}))$, then the above can equivalently be written as

$$P(x_{t+1} \in D \mid x_t, x_{t-1}, \ldots, x_0) = P(x_{t+1} \in D \mid x_t)$$

for all $D \in \mathcal{B}(\mathbb{X})$ and each $t \in [0, N-1]$.

Thus a Markov chain is completely determined by the transition probability and the probability of the initial state, $P(dx_0)$, denoting $P(x_0 \in \cdot)$. Hence, the probability of the event $\{x_{t+1} \in D\}$ for any t can be computed recursively by

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,

starting at $t = 0$, with $P(x_1 \in D) = \int_x P(x_1 \in D|x_0)P(dx_0)$, and iterating in a similar manner for $t = 1, 2, \ldots$, that is, for an arbitrary t, $P(x_{t+1} \in D) = \int_x P(x_{t+1} \in D|x_t)P(dx_t)$.

In the remainder of this section, we discuss a number of properties of Markov chains.

Let $\{x_t, t \geq 0\}$ be a Markov chain taking values in a complete, separable, metric state space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ and defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. For each $D \in \mathcal{B}(\mathbb{X})$, let $P(x, D) := P(x_{t+1} \in D|x_t = x)$ denote the transition probability from x to D, that is, the probability of the event $\{x_{t+1} \in D\}$ given that $x_t = x$, which as we have seen completely determines the evolution of the Markov chain. We note that given the one-step transition kernel, $P(x, A)$, the n-step transitions can be obtained via composition, $P(x_{t+n} \in A \mid x_t = x) = P^n(x, A)$, for any $n \geq 1$. The transition law acts on measurable functions $f \colon \mathbb{X} \to \mathbb{R}$ and measures μ on $\mathcal{B}(\mathbb{X})$ via,

$$Pf\,(x) := \int_{\mathbb{X}} P(x, dy)f(y), \quad x \in \mathbb{X}, \quad \mu P\,(A) := \int_{\mathbb{X}} \mu(dx)P(x, A), \quad A \in \mathcal{B}(\mathbb{X}).$$

A probability measure π on $\mathcal{B}(\mathbb{X})$ is called *invariant* or *stationary* if $\pi P = \pi$. That is,

$$\int \pi(dx)P(x, A) = \pi(A), \qquad A \in \mathcal{B}(\mathbb{X}).$$

For any initial probability measure ν on $\mathcal{B}(\mathbb{X})$ we can construct a stochastic process with transition law P, satisfying $x_0 \sim \nu$. We let P_ν denote the resulting probability measure on the sample space, with the usual convention for $\nu = \delta_x$ when the initial state is $x \in \mathbb{X}$. When $\nu = \pi$, then the resulting process is stationary.

There is at most one stationary solution under the following irreducibility condition. For a set $A \in \mathcal{B}(\mathbb{X})$, let

$$\tau_A := \min(t \geq 1 : x_t \in A). \tag{C.1}$$

Definition C.1.1. Let φ denote a sigma-finite measure on $\mathcal{B}(\mathbb{X})$.

(i) The Markov chain is called *φ-irreducible* if for any $x \in \mathbb{X}$, and any $B \in \mathcal{B}(\mathbb{X})$ satisfying $\varphi(B) > 0$, we have

$$P_x\{\tau_B < \infty\} > 0\,.$$

(ii) A φ-irreducible Markov chain is *aperiodic* if for any $x \in \mathbb{X}$, and any $B \in \mathcal{B}(\mathbb{X})$ satisfying $\varphi(B) > 0$, there exists $n_0 = n_0(x, B)$ such that

$$P^n(x, B) > 0 \qquad \textit{for all} \;\; n \geq n_0.$$

(iii) A φ-irreducible Markov chain is *Harris recurrent* if $P_x(\tau_B < \infty) = 1$ for any $x \in \mathbb{X}$ and any $B \in \mathcal{B}(\mathbb{X})$ satisfying $\varphi(B) > 0$. It is *positive Harris recurrent* if in addition there is an invariant probability measure π. ◇

Definition C.1.2. A set α is called an atom if there exists a probability measure ν such that

$$P(x, A) = \nu(A), \quad \forall x \in \alpha, A \in \mathcal{B}(\mathbb{X}).$$

If the chain is μ-irreducible and $\mu(\alpha) > 0$, then α is called an *accessible atom.* ◇

In case there is an atom α, we have the following:

Theorem C.1.1. *For a Markov chain for which $E_\alpha[\tau_\alpha] < \infty$, the following is an invariant probability measure:*

$$\pi(A) = E_\alpha[\frac{\sum_{k=0}^{\tau_\alpha - 1} 1_{x_k \in A}}{E[\tau_\alpha]} | x_0 = \alpha], \quad \forall A \in \mathcal{B}(\mathbb{X}), \mu(A) > 0.$$

◇

In case there is no atom, one can construct atoms in an extended space through Nummelin's splitting technique [294] provided one of the following sets exists:

Definition C.1.3. A set $A \subset \mathbb{X}$ is μ-small on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if for some n and some positive measure μ,

$$P^n(x, B) \geq \mu(B), \quad \forall x \in A, \text{ and } B \in \mathcal{B}(\mathbb{X})\,.$$

◇

Definition C.1.4. A set $A \subset \mathbb{X}$ is μ-petite on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ if for some distribution $\mathcal{T}$ on $\mathbb{N}$ (set of natural numbers) and some positive measure μ,

$$\sum_{n=0}^{\infty} P^n(x, B)\mathcal{T}(n) \geq \mu(B), \quad \forall x \in A, \text{ and } B \in \mathcal{B}(\mathbb{X})\,.$$

◇

Note. By Theorem 5.5.7 of [271], under aperiodicity and irreducibility, every petite set is small.

The following is a key result.

Theorem C.1.2 (Theorem 4.1 of Meyn-Tweedie [270]). *Suppose that $\boldsymbol{X}$ is a φ-irreducible Markov chain and suppose that there is a set $A \in \mathcal{B}(\mathbb{X})$ satisfying the following:*

(i) *A is μ-petite for some μ.*
(ii) *The chain is recurrent in the sense that $P_x(\tau_A < \infty) = 1$ for any $x \in \mathbb{X}$.*
(iii) *A is finite mean recurrent: $\sup_{x \in A} E_x[\tau_A] < \infty$.*

Then $\boldsymbol{X}$ is positive Harris recurrent and there exists a unique invariant probability measure. ◇

An invariant measure satisfies the following:

Theorem C.1.3. *For a μ-irreducible Markov chain with the unique invariant probability measure π, the following is satisfied by the invariant probability measure:*

$$\pi(A) = \int_C \pi(dx) E_x[\sum_{k=0}^{\tau_C - 1} 1_{\{x_k \in A\}}], \quad A \in \mathcal{B}(\mathbb{X}),$$

where C satisfies $\sup_{x \in C} E_x[\tau_C] < \infty$. Furthermore, for measurable f,

$$\pi(f) = \int_C \pi(dx) E_x[\sum_{k=0}^{\tau_C - 1} f(x_k)].$$

◇

The existence of an invariant distribution is important primarily because of the following:

Theorem C.1.4 (Birkhoff's Sample Path Ergodic Theorem for Markov Chains). *Consider a positive Harris recurrent Markov chain $\{x_t\}$ taking values in $\mathbb{X}$, with invariant distribution $\pi(.)$. Let $f : \mathbb{X} \to \mathbb{R}$ be such that $\int f(x)\pi(dx) < \infty$. Then, the following holds almost surely:*

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} f(x_t) = \int f(x)\pi(dx).$$

◇

On Petite and Small Sets

Establishing petiteness may be difficult to directly verify. In the following, we present two conditions that may be used to establish the petiteness properties.

A kernel A from $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ to itself is called *sub-stochastic* if for every $x \in \mathbb{X}$, $A(x, \cdot)$ is nonnegative, $A(x, \mathbb{X}) \leq 1$ and for every $B \in \mathcal{B}(\mathbb{X})$, $A(\cdot, B)$ is a measurable function on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$.

For a Markov chain with transition kernel P on a Borel space $\mathbb{X}$, and for $\mathcal{K}$ a probability measure on $\mathbb{N}$, if there exists for every $B \in \mathcal{B}(\mathbb{X})$ a lower semicontinuous function $A(\cdot, B)$ such that

$$\sum_{n=0}^{\infty} P^n(x, B)\mathcal{K}(n) \geq A(x, B),$$

for a sub-stochastic kernel $A(\cdot, \cdot)$, the chain is called a T-chain. By [271], Theorem 6.0.1, every compact set in an irreducible T-chain is petite.

For a countable state space, under irreducibility, every finite set S is petite.

To establish the petite set property, Tweedie [370] considers the following test which only depends on the one-stage transition kernel of a Markov chain: If a set S is such that the following *uniform countable additivity* condition

$$\lim_{n\to\infty} \sup_{x\in S} P(x, B_n) = 0 \tag{C.2}$$

is satisfied for $B_n \downarrow \emptyset$, and if the Markov chain is irreducible, then S is petite. This condition may be easier to directly verify in a large class of applications.

C.2 Discrete-Time Martingales

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space. An increasing family $\{\mathcal{F}_n\}$ of sub σ-fields of $\mathcal{F}$ is called a *filtration*. A sequence of $\mathbb{X}$-valued random variables x_n defined on $(\Omega, \mathcal{F}, \mathbf{P})$ is said to be adapted to $\mathcal{F}_n$ if x_n is $\mathcal{F}_n$-measurable, that is, $x_n^{-1}(B) = \{\omega \in \Omega : x_n(\omega) \in B\} \in \mathcal{F}_n$ for all $B \in \mathcal{B}(\mathbb{X})$. This holds, for example, if $\mathcal{F}_n = \sigma(x_m, m \le n), n \ge 0$. Given a filtration $\mathcal{F}_n$ and a sequence of real random variables adapted to it, $(x_n, \mathcal{F}_n)$ is said to be a martingale if

$$E[|x_n|] < \infty$$

and

$$E[x_{n+1}|\mathcal{F}_n] = x_n.$$

We will occasionally take the sigma-fields to be $\mathcal{F}_n = \sigma(x_1, x_2, \ldots, x_n)$. Let $n > m \in \mathbb{Z}_+$. If $\{x_n\}$ is a martingale sequence, $E[x_n|\mathcal{F}_m] = x_m$. If we have that $E[x_n|\mathcal{F}_m] \ge x_m$, $\{x_n\}$ is called a submartingale. If $E[x_n|\mathcal{F}_m] \le x_m$, then $\{x_n\}$ is called a supermartingale.

A useful concept related to filtration is that of a stopping time. A stopping time is a random time, whose occurrence is measurable with respect to the filtration in the sense that for each $n \in \mathbb{N}$, $\{T \le n\} \in \mathcal{F}_n$.

The following is known as Doob's optional sampling theorem [120].

Theorem C.2.1. *Suppose* $(x_n, \mathcal{F}_n)$ *is a supermartingale and* ρ, τ *are bounded stopping times (say bounded by* $n \in \mathbb{Z}_+$*) with* $\rho \le \tau$ *almost surely. Then,* $E[x_\tau|\mathcal{F}_\rho] \le x_\rho$. ◇

Using Doob's upcrossing lemma [120], together with the optional sampling theorem above, one obtains the following:

Theorem C.2.2 (Submartingale Convergence Theorem). *Suppose* x_n *is a submartingale and* $\sup_{n\ge 0} E[|x_n|] < \infty$. *Then* $x := \lim_{n\to\infty} x_n$ *exists (almost surely) and* $E[|x|] < \infty$. ◇

C.3 Stochastic Stability of Dynamical Systems and Random Processes

C.3.1 Stationary, Ergodic, and Asymptotically Mean Stationary Processes

In this subsection, we review ergodic theory, in the context of information theory (i.e., with the transformations being specific to the shift operation). A comprehensive discussion is available in Shields [338] and Gray [170, 175].

Let $\mathbb{X}$ be a complete, separable, metric space. Let $\mathcal{B}(\mathbb{X})$ denote the Borel sigma-field of subsets of $\mathbb{X}$. Let $\Sigma = \mathbb{X}^{\infty}$ denote the sequence space of all one-sided or two-sided infinite sequences drawn from $\mathbb{X}$. Thus, for a two-sided sequence space, if $x \in \Sigma$ then $x = \{\dots, x_{-1}, x_0, x_1, \dots\}$ with $x_i \in \mathbb{X}$. Let $X_n : \Sigma \to \mathbb{X}$ denote the coordinate function such that $X_n(x) = x_n$. Let T denote the shift operation on Σ, that is, $X_n(Tx) = x_{n+1}$. That is, for a one-sided sequence space, $T(x_0, x_1, x_2, \dots) = (x_1, x_2, x_3, \dots)$.

Let $\mathcal{B}(\Sigma)$ denote the smallest sigma-field containing all cylinder sets of the form $\{x : x_i \in B_i, m \leq i \leq n\}$ where $B_i \in \mathcal{B}(\mathbb{X})$, for all integers m, n. Observe that $\cap_{n\geq 0} T^{-n}\mathcal{B}(\Sigma)$ is the tail σ-field $\cap_{n \geq 0}\sigma(x_n, x_{n+1}, \cdots)$, since $T^{-n}(A) = \{x : T^n x \in A\}$.

Let μ be a stationary measure on $(\Sigma, \mathcal{B}(\Sigma))$ in the sense that $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}(\Sigma)$. The sequence of random variables $\{x_n\}$ defined on the probability space $(\Sigma, \mathcal{B}(\Sigma), \mu)$ is a stationary process.

Definition C.3.1. Let μ be the measure on a process. This random process is ergodic if $A = T^{-1}A$ implies that $\mu(A) \in \{0, 1\}$. ◇

That is, the events that are unchanged with a shift operation are trivial events. Mixing is a sufficient condition for ergodicity. Thus, a source is ergodic if $\lim_{n\to\infty} P(A \cap T^{-n}B) = P(A)P(B)$, since the process forgets its initial condition. For the special case of Markov sources, we have the following: A positive Harris recurrent Markov chain is ergodic, since such a process is mixing and stationary.

Definition C.3.2. A random process with measure μ is N-stationary (cyclo-stationary or periodically stationary with period N) if $\mu(T^{-N}B) = \mu(B)$ for all $B \in \mathcal{B}(\Sigma)$ or equivalently for any $n \in \mathbb{N}$ samples $t_1, t_2, \dots, t_n$:

$$\mu(x_{t_1} \in A_1, \dots, x_{t_n} \in A_n) = \mu(x_{t_1+N} \in A_1, \dots, x_{t_n+N} \in A_n).$$

◇

Definition C.3.3. A random process is N-ergodic if $A = T^{-N}A$ implies that $\mu(A) \in \{0, 1\}$. ◇

Definition C.3.4. A set $A \in \mathcal{B}(\mathbb{X})$ is coordinate recurrent if for some $m \in \mathbb{Z}_+$

$$\sum_{m=0}^{\infty} 1_{\{X_m(x) \in A\}} = \infty, \quad a.s.$$

◇

Definition C.3.5. A process on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with process measure P is asymptotically mean stationary (AMS) if there exists a probability measure $\bar{P}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=0}^{N-1} P(T^{-k}F) = \bar{P}(F),$$

for all events $F \in \mathcal{B}(\Sigma)$. Here $\bar{P}$ is called the stationary mean of P and is a stationary measure. ◇

$\bar{P}$ is stationary since, by definition, $\bar{P}(F) = \bar{P}(T^{-1}F)$. A cyclo-stationary process is AMS; see, for example, [78, 175] or [170] (Theorem 7.3.1). Asymptotic mean stationarity is a very important property:

1. The Shannon-McMillan-Breiman theorem (the entropy ergodic theorem) applies to finite-alphabet AMS sources [175] (see an extension for a more general class [46]). In this case, the ergodic decomposition of the AMS process leads to almost sure convergence of the conditional entropies.
2. Birkhoff's ergodic theorem applies to bounded measurable functions f, if and only if the process is AMS [175].

Let

$$F = \{x : \lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} f(T^i x) \quad \text{exists.}\}$$

It follows that for an AMS process, $m(F) = 1$, with m being the stationary mean of the process. Birkhoff's Almost-Sure Ergodic Theorem states the following: If a dynamical system is AMS with stationary mean m, then all bounded measurable functions f have the ergodic property, and with probability 1,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=0}^{N-1} f(T^i x) = E_{m_x}[f], \quad x \in F,$$

where E_{m_x} denotes the expectation under measure m_x and m_x is the resulting ergodic measure with initial state x in the ergodic decomposition of the asymptotic mean (see Theorem 1.8.2 in [171]): $m(A) = \int m_x(A) m(dx)$.

Definition C.3.6. A random process x is second moment stable if the following holds:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=0}^{N-1} E[(X_m(x))^2] < \infty.$$

◇

Definition C.3.7. A random process x is quadratically stable (almost surely) if the following limit exists and is finite almost surely:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{m=0}^{N-1} X_m(x)^2 < \infty.$$

◇

We finally note that a positive Harris recurrent Markov chain (thus with a unique invariant distribution on the state space) is also ergodic in the more general sense of ergodic theory.

Appendix D
Markov Decision Theory and Optimality of Markov Policies

This appendix provides some preliminary background on optimization of controlled Markov chains. A comprehensive treatment can be found in [14, 55, 84, 194, 225, 269].

D.1 Controlled Markov Models

D.1.1 Fully Observed Markov Control Problem Model

Consider the following model:

$$x_{t+1} = f(x_t, u_t, w_t), \tag{D.1}$$

where x_t is a $\mathbb{X}$-valued state variable, u_t is a $\mathbb{U}$-valued control action variable, w_t a $\mathbb{W}$-valued i.i.d noise process, and f is a measurable function. We assume that $\mathbb{X}, \mathbb{U}, \mathbb{W}$ are subsets of Polish spaces. The model above in (D.1) contains (see [69]) the class of all stochastic processes which satisfy the following for all Borel sets $B \in \mathcal{B}(\mathbb{X})$, $t \geq 0$ and all realizations $x_{[0,t]}, u_{[0,t]}$:

$$P(x_{t+1} \in B | x_{[0,t]} = a_{[0,t]}, u_{[0,t]} = b_{[0,t]}) = \mathcal{T}(x_{t+1} \in B | a_t, b_t), \tag{D.2}$$

where $\mathcal{T}(\cdot|x,u)$ is a *stochastic kernel* from $\mathbb{X} \times \mathbb{U}$ to $\mathbb{X}$.

A stochastic process which satisfies (D.2) is called a controlled Markov chain.

A fully observed Markov control problem is a five tuple $(\mathbb{X}, \mathbb{U}, \mathbb{K}, \mathcal{T}, c)$, where

- $\mathbb{X}$ is the state space, a subset of a Polish (i.e., a complete, separable, metric) space.
- $\mathbb{U}$ is the action space, a subset of a Polish space.

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,

- $\mathbb{K} = \{(x, u) : u \in U(x), x \in \mathbb{X}\}$ is the set of state, control pairs that are feasible. There might be different states where different control actions are possible.
- $\mathcal{T}$ is a state transition kernel, that is, $\mathcal{T}(A|x_t, u_t) = P(x_{t+1} \in A|x_t, u_t)$.
- $c : \mathbb{K} \to \mathbb{R}$ is the cost function.

The objective function to be minimized for a finite-stage setup is given by

$$J(x_0, \Pi) := E_{x_0}^{\Pi}[\sum_{t=0}^{T-1} c(x_t, u_t) + c_T(x_T)], \tag{D.3}$$

where Π is a control policy, $E_{x_0}^{\Pi}[\cdot]$ denotes the expectation under policy Π and initial state x_0, and c_T is a terminal state cost. Given a class of admissible policies (to be defined below), the goal is to find an admissible policy Π^* such that

$$J(x_0, \Pi^*) \leq J(x_0, \Pi),$$

for all admissible policies Π. Such a Π^* is an *optimal policy*. Here Π can also be called a *strategy* or a *control law*.

D.1.2 Classes of Control Policies

Admissible Control Policies

Let $H_0 := \mathbb{X}, \quad H_t = H_{t-1} \times \mathbb{X} \times \mathbb{U}$ for $t = 1, 2, \ldots$. We let h_t denote an element of H_t, where $h_t = \{x_{[0,t]}, u_{[0,t-1]}\}$. A *deterministic admissible control policy* Π is a sequence of functions $\{\gamma_t\}$ from $H_t \to \mathbb{U}$; in this case $u_t = \gamma_t(h_t)$. A randomized control policy is a sequence $\Pi = \{\Pi_t, t \geq 0\}$ such that $\Pi : H_t \to \mathcal{P}(\mathbb{U})$ (with $\mathcal{P}(\mathbb{U})$ being the set of probability measures on $\mathbb{U}$) such that

$$\Pi_t(u_t \in U(x_t)|h_t) = 1, \quad h_t \in H_t.$$

Markov Control Policies

A policy is *randomized Markov* if

$$P_{x_0}^{\Pi}(u_t \in C|h_t) = \Pi_t(u_t \in C|x_t), \quad C \in \mathcal{B}(\mathbb{U}).$$

Hence, the control action only depends on the state and the time, and not the past history. If the control strategy is deterministic, that is, if

$$\Pi_t(u_t = f_t(x_t)|x_t) = 1$$

for some function f_t, the control policy is said to be *deterministic Markov*.

Stationary Control Policies

A policy is *randomized stationary* if

$$P^{\Pi}_{x_0}(u_t \in C|h_t) = \Pi(u_t \in C|x_t), \quad C \in \mathcal{B}(\mathbb{U}).$$

Hence, the control action only depends on the state, and not the past history or on time. If the control strategy is deterministic, that is, if $\Pi(u_t = f(x_t)|x_t) = 1$ for some function f, the control policy is said to be *deterministic stationary*.

D.1.3 Optimality of Markov Policies and Elimination of Irrelevant Information

The following is a useful result on the structure of optimal control policies.

Theorem D.1.1 (Blackwell's Irrelevant Information Theorem [62, 64]). *Let* $\mathbb{X}, \mathbb{Y}, \mathbb{U}$ *be Polish spaces and* P *be a probability measure on* $\mathcal{B}(\mathbb{X} \times \mathbb{Y})$ *and let* $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ *be a bounded Borel measurable cost function. Then, for any Borel measurable function* $\gamma : \mathbb{X} \times \mathbb{Y} \to \mathbb{U}$*, there exists another Borel measurable function* $\gamma^* : \mathbb{X} \to \mathbb{U}$ *such that*

$$\int_{\mathbb{X}} c(x, \gamma^*(x))P(dx) \leq \int_{\mathbb{X}\times\mathbb{Y}} c(x, \gamma(x, y))P(dx, dy).$$

Furthermore, policies based only on x almost surely are optimal. ◇

D.1.4 Markov Decision Processes (MDPs) and Optimality of Markov Policies

The following result is a well-known theorem in the control theory literature (e.g., see [193, 348]). A proof can be obtained through the dynamic programming algorithm and applying Theorem D.1.1 at every time stage.

Theorem D.1.2. *Given an MDP as above, consider the minimization of (D.3) over all admissible control policies. There is no loss in restricting policies to be Markov, that is, a policy which only uses the current state* x_t *and the time information* t*.* ◇

D.1.5 Dynamic Programming and Measurable Selection Criteria

The following dynamic programming arguments hold when there exist a minimizing control policy (sequence of selectors for each time stage) for (D.3).

Dynamic Programming

Let $\{J_t, t = T, T-1, \ldots, 0\}$, with $J_t : \mathbb{X} \to \mathbb{R}$, be a sequence of functions generated by the recursion

$$J_t(x) = \inf_{u \in \mathbb{U}_t(x)} (c(x,u) + \int_{\mathbb{X}} J_{t+1}(y) P(x_{t+1} \in dy | x_t = x, u_t = u)), \qquad (*)$$

for all $x \in \mathbb{X}$, for $t \in \{0, 1, 2 \ldots, T-1\}$, with

$$J_T(x_T) = c_T(x_T).$$

Let the infimum in (*) be achieved for all t. Then, there exists a sequence of measurable functions (selectors) $\{f_t\}$ such that

$$J_t(x_t) = c(x_t, f_t(x_t)) + \int_{\mathbb{X}} J_{t+1}(y) P(dy | x_t, f_t(x_t)),$$

and $\{f_t\}$ is optimal for (D.3).

Such a sequence of optimal functions exists under the following conditions:

Condition 1. The stage wise cost function to be minimized, $c(x_t, u_t)$, is continuous on both $\mathbb{U}$ and $\mathbb{X}$, $\mathbb{U}_t(x) = \mathbb{U}$ is compact, and $\int_{\mathbb{X}} Q(dy|x,u)v(y)$ is a continuous function on $\mathbb{X} \times \mathbb{U}$ for every continuous and bounded v on $\mathbb{X}$. ◇

Condition 2. For every $x \in \mathbb{X}$ the stage wise cost function to be minimized, $c(x_t, u_t)$, is continuous on $\mathbb{U}$; $\mathbb{U}_t(x)$ is compact; and $\int_{\mathbb{X}} Q(dy|x,u)v(y)$ is a continuous function on $\mathbb{U}$ for every bounded, measurable function v on $\mathbb{X}$. ◇

Theorem D.1.3. *Under Conditions 1 or 2, the measurable selection hypothesis applies, and there exists an optimal control policy (which is a Markov policy) obtained through the dynamic programming recursions in (*), that is, there exists a minimizing control policy given by* $f_t : \mathbb{X} \to \mathbb{A}_t(x_t)$. *Furthermore, under Condition 1, the function* $J_t(x)$ *is continuous, if* $c_N(x_N)$ *is continuous.* ◇

We can replace the compactness condition with an *inf-compactness* condition, and modify Condition 1 as below:

Condition 3. For every $x \in \mathbb{X}$, $c(x,u)$ is continuous on $\mathbb{X} \times \mathbb{U}$ and is nonnegative; $\{u : c(x,u) \leq \alpha\}$ is compact for all $\alpha > 0$ and all $x \in \mathbb{X}$, and $\int_{\mathbb{X}} Q(dy|x,u)v(y)$ is a continuous function on $\mathbb{X} \times \mathbb{U}$ for every continuous and bounded v. ◇

Theorem D.1.4. *Under Condition 3, the measurable selection hypothesis applies.* ◇

D.1.6 Partially Observable MDPs (POMDPs)

Consider the following dynamics:

$$x_{t+1} = f(x_t, u_t, w_t), \quad y_t = g(x_t, v_t).$$

Here, as before, x_t is the state, $u_t \in \mathbb{U}$ is the control, $(w_t, v_t) \in \mathbb{W} \times \mathbb{V}$ are zero-mean, i.i.d noise processes, and w_t is independent of v_t. In addition to the previous fully observed model, y_t denotes an observation variable taking values in $\mathbb{Y}$, taken here to be a subset of $\mathbb{R}^n$. The controller only has causal access to the process $\{y_t\}$. An admissible policy $\Pi = \{\Pi_t\}$ is measurable with respect to $\sigma(\{y_s, s \leq t\})$. We denote the observed history space as $H_0 := \mathbb{Y}, \quad H_t = H_{t-1} \times \mathbb{Y} \times \mathbb{U}$. Hence, the set of admissible control policies are such that $P(u(h_t) \in \mathbb{U}|h_t) = 1 \quad \forall h_t \in H_t$. One could transform a partially observable Markov Decision Problem to a fully observed Markov decision problem via an enlargement of the state space. In particular, one obtains via the properties of total probability the following dynamic nonlinear filter equation:

$$\begin{aligned}\pi_t(A) :&= P(x_t \in A | y_{[0,t]}, u_{[0,t-1]}) \\ &= \frac{\int_{\mathbb{X}} \pi_{t-1}(dx_{t-1}) r(y_t|x_t) P(dx_t|x_{t-1}, u_{t-1})}{\int_{\mathbb{X}} \int_{\mathbb{X}} \pi_{t-1}(dx_{t-1}) r(y_t|x_t) P(x_t|x_{t-1}, u_{t-1})},\end{aligned}$$

where we assume that $\int_B r(y|x)dy = P(y_t \in B | x_t = x)$ for any $B \in \mathcal{B}(\mathbb{Y})$ and r denotes the conditional density. The conditional measure process becomes a controlled Markov chain in $\mathcal{P}(\mathbb{X})$ under the weak convergence topology.

Theorem D.1.5. *The process $\{\pi_t, u_t\}$ is a controlled Markov chain. That is, under any admissible control policy, given the action at time $t \geq 0$ and π_t, π_{t+1} is conditionally independent of $\{\pi_s, u_s, s \leq t-1\}$.* ◇

As before, let the objective function to be minimized be $\sum_{t=0}^{T-1} E_{x_0}^{\Pi}[c(x_t, u_t)]$. We transform the system into a fully observed Markov model as follows. Define the new cost as $\tilde{c}(\pi, u) = \int_{\mathbb{X}} c(x,u)\pi(dx), \quad \pi \in \mathcal{P}(\mathbb{X})$. The stochastic transition kernel q is given by

$$q(dx, dy|\pi, u) = \int_{\mathbb{X}} P(dx, dy|x', u)\pi(dx'), \quad \pi \in \mathcal{P}(\mathbb{X}).$$

This kernel can be decomposed as $q(dx, dy|\pi, u) = P(dy|\pi, u)P(dx|\pi, u, y)$.

The second term here is the filtering equation, mapping $(\pi, u, y) \in (\mathcal{P}(\mathbb{X}) \times \mathbb{U} \times \mathbb{Y})$ to $\mathcal{P}(\mathbb{X})$. It follows that $(\mathcal{P}(\mathbb{X}), \mathbb{U}, \mathcal{K}, \tilde{c})$ defines a completely observable controlled Markov process. Here, we have

$$\mathcal{K}(B|\pi, u) = \int_{\mathbb{Y}} 1_{\{P(.|\pi,u,y)\in B\}} P(dy|\pi, u), \quad \forall B \in \mathcal{B}(\mathcal{P}(\mathbb{X}))$$

As such, one can obtain the optimal solution by using the filtering equation as a sufficient statistic in a centralized setting, as Markov policies (policies that use the Markov state as their sufficient statistics) are optimal for control of Markov chains, under sufficiency conditions for the existence of optimal selectors.

D.2 Kalman Filter and Linear-Quadratic-Gaussian Optimal Control Problem

As we observed above, for controlling partially observed controlled Markov sources, one can enlarge the state space and define a probability measure valued state which is fully observed. For a Gaussian source, however, the above can be done in a computationally very efficient manner, as a Gaussian measure is uniquely characterized by its mean and covariance matrix. We will discuss this further in this section.

Consider now the partially observed linear system:

$$x_{t+1} = Ax_t + Bu_t + w_t,$$

$$y_t = Cx_t + v_t,$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m$ and $w \in \mathbb{R}^n$, $y \in \mathbb{R}^p$, $v \in \mathbb{R}^p$. Suppose $\{w_t, v_t\}$ are i.i.d. zero-mean random Gaussian vectors with given covariance matrices $E[w_t w_t'] = W$ and $E[v_t v_t'] = V$ for all $t \geq 0$. Further, let x_0 be zero-mean Gaussian.

Let us now define

$$m_t = E[x_t|y_{[0,t-1]}, u_{[0,t-1]}]$$

and

$$\Sigma_{t|t-1} = E[(x_t - E[x_t|y_{[0,t-1]}, u_{[0,t-1]}])(x_t - E[x_t|y_{[0,t-1]}, u_{[0,t-1]}])' \\ |y_{[0,t-1]}, u_{[0,t-1]}].$$

Then, $\{m_t, \Sigma_{t|t-1}\}$ are generated by the following recursions, which constitute the Kalman filter equations:

$$m_{t+1} = Am_t + Bu_t + A\Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(y_t - Cm_t),$$

$$m_0 = E[x_0],$$

$$\Sigma_{t+1|t} = A\Sigma_{t|t-1}A' + W - (A\Sigma_{t|t-1}C')(C\Sigma_{t|t-1}C' + V)^{-1}(C\Sigma_{t|t-1}A'),$$

$$\Sigma_{0|-1} = E[x_0x_0'].$$

Let us now introduce

$$I_t = \{y_{[0,t]}, u_{[0,t-1]}\}$$

and define $\bar{m}_t = E[x_t|I_t]$. The following readily holds:

$$\bar{m}_t = m_t + \Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(y_t - Cm_t),$$

or

$$\begin{aligned}\bar{m}_t &= A\bar{m}_{t-1} + Bu_t \\ &\quad + \Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(CA(x_{t-1} - \bar{m}_{t-1}) + v_t).\end{aligned} \tag{D.4}$$

We note that the zero-mean variable $x_t - \bar{m}_t$ is orthogonal to $I_t = \{y_{[0,t]}, u_{[0,t-1]}\}$, in the sense that the error is independent of the information available at the controller, and since the information available is Gaussian, independence and orthogonality are equivalent. In view of this, the variable $\Sigma_{t|t-1}C'(C\Sigma_{t|t-1}C' + V)^{-1}(CA(x_{t-1} - \bar{m}_{t-1}) + v_t)$ is an additive Gaussian noise acting on the new state variable $\bar{m}_t$.

LQG Optimal Control Problem

Now, given the Gaussian POMDP setup considered, suppose that the goal is to obtain a solution to

$$\inf_{\Pi} J(\Pi, \mu_0),$$

where

$$J(\Pi, \mu_0) = E_{\mu_0}^{\Pi}[\sum_{t=0}^{T-1} x_t'Qx_t + u_t'Ru_t + x_T'Q_Tx_T],$$

with $R > 0$ and $Q, Q_T \geq 0$ (i.e., these matrices are positive definite and positive semi-definite), and μ_0 is the prior measure on the initial state, which is taken to be Gaussian.

The quadratic optimization problem can be reformulated as a function of $\bar{m}_t$, u_t and the error $(x_t - \bar{m}_t)$, which makes the optimal control problem equivalent to optimal control of fully observed state $\bar{m}_t$, with an additive time-varying independent Gaussian noise process. This thus features as the separation of estimation and control, and a more special version is known as the certainty equivalent principle. The absence of *dual effect* plays a key role in this analysis, in taking $E[(x_t - \bar{m})'Q(x_t - \bar{m})]$ out of the conditioning on I_t.

Thus, one obtains

$$J_t(I_t) = E[x_t' K_t x_t | I_t] + \sum_{k=t}^{T-1} \Bigg(E[(x_k - E[x_k|I_k])' Q (x_k - E[x_k|I_k])] + E[\tilde{w}_t' K_{t+1} \tilde{w}] \Bigg), \qquad \text{(D.5)}$$

where $\tilde{w}_t = \Sigma_{t|t-1} C' (C \Sigma_{t|t-1} C' + V)^{-1} (y_t - CE[x_t | I_{t-1}])$.

Solving a fully observed LQG optimal control problem, the optimal minimizing control is linear and has the form [225]:

$$u_t = -(BK_{t+1}B + R)^{-1} B' K_{t+1} A E[x_t | I_t]$$

where K_t is generated from the discrete-time Riccati equation:

$$K_t = Q + A' K_{t+1} A - A' K_{t+1} B ((BK_{t+1}B + R))^{-1} B' K_{t+1} A,$$

with terminal condition $K_T = Q_T$.

References

1. Abaya, E.F., Wise, G.L.: Convergence of vector quantizers with applications to optimal quantization. SIAM J. Appl. Math. **44**, 183–189 (1984)
2. Abaya, E.F., Wise, G.L.: Some remarks on the existence of optimal quantizers. Stat. Probab. Lett. **2**, 349–351 (1984)
3. Acemoglu, D., Ozdaglar, A.: Opinion dynamics and learning in social networks. Dyn. Games Appl. **1**, 3–49 (2010)
4. Aicardi, M., Davoli, F., Minciardi, R.: Decentralized optimal control of Markov chains with a common past information set. IEEE Trans. Autom. Control **32**, 1028–1031 (1987)
5. Akyol, E., Rose, K.: On linear transforms in zero-delay Gaussian source channel coding. In: Proceedings of the IEEE International Symposium on Information Theory, pp. 1548–1552, Boston, MA, 2012
6. Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis: A Hitchhikers Guide, 3rd edn. Springer, Berlin (2006)
7. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in Space of Probability Measures. Birkhäuser, Basel (2008)
8. Anastasopoulos, A.: A sequential transmission scheme for the multiple access channel with noiseless feedback. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, 2009
9. Andersland, M., Teneketzis, D.: Information structures, causality, and non-sequential stochastic control, I: design-independent properties. SIAM J. Control Optim. **30**, 1447–1475 (1992)
10. Andersland, M., Teneketzis, D.: Information structures, causality, and non-sequential stochastic control, II: design-dependent properties. SIAM J. Control Optim. **32**, 1726–1751 (1994)
11. Anderson, B.D.O., Moore, J.B.: Time-varying feedback decentralized control. IEEE Trans. Autom. Control **26**, 1133–1139 (1981)
12. Andrews, G.E.: Number Theory. Dover, New York (1995)
13. Aoki, M.: On decentralized linear stochastic control problems with quadratic cost. IEEE Trans. Autom. Control **18**, 243–250 (1973)
14. Arapostathis, A., Borkar, V.S., Fernandez-Gaucherand, E., Ghosh, M.K., Marcus, S.I.: Discrete-time controlled Markov processes with average cost criterion: a survey. SIAM J. Control Optim. **31**, 282–344 (1993)
15. Arrow, K.J.: The informational structure of the firm. Am. Econ. Rev. **75**, 303–307 (1985)
16. Ash, R.B.: Real Analysis and Probability. Academic, New York (1972)
17. Aström, K.J., Bernhardsson, B.: Comparison of Riemann and Lebesgue sampling for first order stochastic systems. In: Proceedings of the IEEE Conference on Decision and Control, pp. 2011–2016, Dec 2002

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control:
Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,

18. Athans, M.: Survey of decentralized control methods. In: 3rd NBER/FRB Workshop on Stochastic Control, Washington D.C, 1974
19. Aumann, R.J.: Agreeing to disagree. Ann. Stat. **4**, 1236–1239 (1976)
20. Ayanoğlu, E., Gray, R.M.: The design of joint source and channel trellis waveform coders. IEEE Trans. Inf. Theor. **33**(6), 855–865 (1987)
21. Başar, T.Ü.: Performance bounds and optimal linear coding for multichannel communication systems. Ph.D. Dissertation, Bogazici University (1978)
22. Başar, T.: Two-criteria LQG decision problems with one-step delay observation sharing pattern. Inf. Control **38**, 21–50 (1978)
23. Başar, T.: A trace minimization problem with applications in joint estimation and control under nonclassical information. J. Optim. Theor. Appl. **31**, 343–359 (1980)
24. Başar, T.: An equilibrium theory for multi-person decision making with multiple probabilistic models. IEEE Trans. Autom. Control **30**, 118–132 (1985)
25. Başar, T.: Stochastic incentive problems with partial dynamic information and multiple levels of hierarchy. Eur. J. Polit. Econ. (special issue on "Economic Design") **5**, 203–217 (1989)
26. Başar, T.: Variations on the theme of the witsenhausen counterexample. In: IEEE Conference on Decision and Control, pp. 1614–1619, Mexico, 2008
27. Başar, T., Bansal, R.: Optimum design of measurement channels and control policies for linear-quadratic stochastic systems. Eur. J. Oper. Res. **73**, 226–236 (1994)
28. Başar, T., Bernhard, P.: H-infinity Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach. Birkhäuser, Boston (1995)
29. Başar, T., Cruz, J.: Concepts and methods in multiperson coordination and control. In: Tzafestas, S.G. (ed.) Optimization and Control of Dynamic Operational Research Models, Chapter 11, pp. 351–387. North Holland, Amsterdam (1982)
30. Başar, T., Mintz, M.: Minimax estimation under generalized quadratic loss. In: IEEE Conference Decision Control, pp. 456–461, Miami, FL, Dec 1971
31. Başar, T., Mintz, M.: On the existence of linear saddle-point strategies for a two-person zero-sum stochastic game. In: IEEE Conference Decision Control, pp. 188–192, New Orleans, Louisiana, Dec 1972
32. Başar, T., Olsder, G.J.: Dynamic Noncooperative Game Theory. SIAM Classics in Applied Mathematics, Philadelphia (1999)
33. Başar, T., Salmon, M.: Credibility and the value of information transmission in a model of monetary policy and inflation. J. Econ. Dyn. Control **14**, 97–116 (1990)
34. Bagchi, A., Başar, T.: Team decision theory for linear continuous-time systems. IEEE Trans. Autom. Control **27**, 1154–1161 (1980)
35. Bahadur, R.R.: Sufficiency and statistical decision functions. Ann. Math. Stat. **25**, 423–462 (1954)
36. Bahadur, R.R.: Statistics and subfields. Ann. Math. Stat. **26**, 490–497 (1955)
37. Baillieul, J.: Feedback designs for controlling device arrays with communication channel bandwidth constraints. In: 4th ARO Workshop on Smart Structures, State College, PA, August 1999
38. Bamieh, B., Voulgaris, P.: A convex characterization of distributed control problems in spatially invariant systems with communication constraints. Syst. Control Lett. **54**, 575–583 (2005)
39. Bansal, R., Başar, T.: Solutions to a class of linear-quadratic-Gaussian (LQG) stochastic team problems with nonclassical information. Syst. Control Lett. **9**, 125–130 (1987)
40. Bansal, R., Başar, T.: Stochastic team problems with nonclassical information revisited: when is an affine law optimal? IEEE Trans. Autom. Control **32**, 554–559 (1987)
41. Bansal, R., Başar, T.: Simultaneous design of measurement and control strategies in stochastic systems with feedback. Automatica **45**, 679–694 (1989)
42. Bao, L., Skoglund, M., Johansson, K.H.: Iterative encoder-controller design for feedback control over noisy channels. IEEE Trans. Autom. Control **57**, 265–278 (2011)
43. Bar-Shalom, Y., Tse, E.: Dual effect, certainty equivalence and separation in stochastic control. IEEE Trans. Autom. Control **19**, 494–500 (1974)

44. Baras, J.S.: Noncommutative probability models in quantum communication and multi-agent stochastic control. Richerche di Automatica **55**, 217–264 (1979)
45. Baras, J.S.: Multi-agent stochastic control: Models inspired from quantum physics. In: International Conference, Physics and Control (PhysCon), pp. 747–758, Saint Petersburg, Russia, 2003
46. Barron, A.R.: The strong ergodic theorem for densities: generalized Shannon-McMillan-Breiman theorem. Ann. Probab. **13**, 1292–1303 (1985)
47. Barron, A.R.: Information theory and martingales. In: IEEE International Symposium on Information Theory, Recent Results Session, Budapest, Hungary, 1991
48. Barron, A.R.: Limits of information, Markov chains, and projections. In: Proceedings of the IEEE International Symposium on Information Theory, 2000
49. Barta, S.M.: On linear control of decentralized stochastic systems. Ph.D. Thesis, Department of EECS, MIT, Cambridge (1978)
50. Barvinok, A.: A course in convexity. In: Graduate Studies in Mathematics, vol. 54. American Mathematical Society, Providence (2002)
51. Berger, T.: Information rates of Wiener processes. IEEE Trans. Inf. Theor. **16**, 134–139 (1970)
52. Berger, T.: Rate Distortion Theory. Prentice-Hall, Englewood Cliffs (1971)
53. Berger, T.: Multiterminal source coding. In: Lecture Notes presented at the CISM Summer School, Udine, Italy, July 1977
54. Berlekamp, E.R.: Block coding with noiseless feedback. Ph.D. Dissertation, Massachusetts Institute of Technology, Department of Electrical Engineering (1964)
55. Bertsekas, D.: Nonlinear Programming. Athena Scientific, Belmont (1999)
56. Bertsekas, D.: Dynamic Programming and Optimal Control, vol. 1. Athena Scientific, Belmont (2000)
57. Bertsekas, D.P., Shreve, S.: Stochastic Optimal Control: The Discrete Time Case. Academic, New York (1978)
58. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
59. Billingsley, P.: Probability and Measure, 2nd edn. Wiley, New York (1986)
60. Bismut, J.-M.: An example of interaction between information and control: the transparency of a game. IEEE Trans. Autom. Control **18**(5), 518–522 (1972)
61. Blackwell, D.: Equivalent comparison of experiments. Ann. Math. Stat. **24**, 265–272 (1953)
62. Blackwell, D.: Memoryless strategies in finite-stage dynamic programming. Ann. Math. Stat. **35**, 863–865 (1964)
63. Blackwell, D., Dubins, L.: Merging of opinions with increasing information. Ann. Math. Stat. **33**, 882–887 (1962)
64. Blackwell, D., Ryll-Nadrzewski, C.: Non-existence of everywhere proper conditional distributions. Ann. Math. Stat. **34**, 223–225 (1963)
65. Blahut, R.E.: Principles of Information Theory. Addison-Wesley, Reading (1987)
66. Bogachev, V.I.: Measure Theory. Springer, Berlin (2007)
67. Boll, C.: Comparison of experiments in the infinite case. Ph.D. Dissertation, Stanford University (1955)
68. Borade, S., Nakibo glu, B., Zheng, L.: Unequal error protection: an information-theoretic perspective. IEEE Trans. Inf. Theor. **55**, 5511–5539 (2009)
69. Borkar, V.S.: White-noise representations in stochastic realization theory. SIAM J. Control Optim. **31**, 1093–1102 (1993)
70. Borkar, V.S.: Probability Theory: An Advanced Course. Springer, New York (1995)
71. Borkar, V.S.: Convex analytic methods in Markov decision processes. In: Feinberg, E.A., Shwartz, A. (eds.) Handbook of Markov Decision Processes, pp. 347–375. Kluwer, Boston (2001)
72. Borkar, V.S., Meyn, S.P.: The ODE method for convergence of stochastic approximation and reinforcement learning. SIAM J. Control Optim. **38**, 447–469 (2000)
73. Borkar, V.S., Mitter, S.K.: LQG control with communication constraints. In: Kailath Festschrift. Kluwer Academic Publishers, Boston (1997)

74. Borkar, V.S., Mitter, S.K., Tatikonda, S.: Optimal sequential vector quantization of Markov sources. SIAM J. Control Optim. **40**, 135–148 (2001)
75. Borkar, V.S., Varaiya, P.: Asymptotic agreement in distributed estimation. IEEE Trans. Autom. Control **27**, 650–655 (1982)
76. Borkar, V.S., Mitter, S.K., Sahai, A., Tatikonda, S.: Sequential source coding: An optimization viewpoint. In: Proceedings of the IEEE Conference on Decision and Control, pp. 1035–1042, Dec 2005
77. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
78. Boyles, R.A., Gardner, W.A.: Cycloergodic properties of discrete-parameter non-stationary stochastic processes. IEEE Trans. Inf. Theor. **29**, 105–114 (1983)
79. Brandenburger, A., Dekel, E.: Common knowledge with probability 1. J. Math. Econ. **16**, 237–245 (1987)
80. Braslavsky, J.H., Middleton, R.H., Freudenberg, J.S.: Feedback stabilization over signal-to-noise ratio constrained channels. IEEE Trans. Autom. Control **52**(8), 1391–1403 (2007)
81. Brockett, R., Liberzon, D.: Quantized feedback stabilization of linear systems. IEEE Trans. Autom. Control **45**, 1279–1289 (2000)
82. Bross, S., Lapidoth, A.: An improved achievable region for the discrete memoryless two-user multiple-access channel with noiseless feedback. IEEE Trans. Inf. Theor. **51**, 811–833 (2005)
83. Burnashev, M.V.: Data transmission over a discrete channel with feedback. Random transmission time. Prob. Inform. Transm. **12**, 10–30 (1976)
84. Caines, P.E.: Linear Stochastic Systems. Wiley, New York (1988)
85. Le Cam, L.: Sufficiency and approximate sufficiency. Ann. Math. Stat. **35**, 1419–1455 (1964)
86. Le Cam, L.: Comparison of experiments – a short review. In: Ferguson, T., Shapley, L. (eds.) Statistics, Probability and Game Theory Papers in Honor of David Blackwell. IMS Lecture Notes Monograph Series, Hayward (1996)
87. Casalino, G., Davoli, F., Minciardi, R., Puliafito, P.P., Zoppoli, R.: Partially nested information structures with a common past. IEEE Trans. Autom. Control **29**, 846–850 (1984)
88. Castanon, D.A., Teneketzis, D.: Distributed estimation algorithms for nonlinear systems. IEEE Trans. Autom. Control **30**, 418–425 (1985)
89. Castanon, D.A., Teneketzis, D.: Further results on the asymptotic agreement problem. IEEE Trans. Autom. Control **33**, 515–523 (1988)
90. Cesa-Bianchi, N., Lugosi, G.: Prediction, Learning, and Games. Cambridge University Press, New York (2006)
91. Charalambous, C.D., Farhadi, A., Denic, D., Rezaei, F.: Robust control over uncertain communication channels. In: Proceedings of the IEEE Med. Conference Control and Automation, pp. 737–742, June 2005
92. Charalambous, C.D., Farhadi, A., Denic, S.Z.: Control of continuous-time linear Gaussian systems over additive Gaussian wireless fading channels: a separation principle. IEEE Trans. Autom. Control **53**, 1013–1019 (2008)
93. Chen, C.T.: Linear Systems Theory and Design. Oxford University Press, Oxford (1999)
94. Chernoff, H.: Rational selection of decision functions. Econometrica **22**, 422–443 (1954)
95. Chong, C.Y., Athans, M.: On the periodic coordination of linear stochastic systems. Automatica **12**, 321–335 (1976)
96. Cloosterman, M.B.G., van de Wouw, N., Heemels, W.P.M.H., Nijmeijer, H.: Stability of networked control systems with uncertain time-varying delays. IEEE Trans. Autom. Control **54**, 1575–1580 (2009)
97. Como, G., Fagnani, F.: Scaling limits for continuous opinion dynamics systems. Ann. Appl. Probab. **21**, 1537–1567 (2011)
98. Como, G., Fagnani, F., Zampieri, S.: Anytime reliable transmission of real-valued information through digital noisy channels. SIAM J. Control Optim. **48**, 3903–3924 (2010)
99. Como, G., Yüksel, S.: On the capacity of memoryless finite state multiple access channels with asymmetric state information at the encoders. IEEE Trans. Inf. Theor. **56**, 1267–1273 (2011)

100. Connor, S.B., Fort, G.: State-dependent foster-lyapunov criteria for subgeometric convergence of Markov chains. Stoch. Process. Appl. **119**, 176–4193 (2009)
101. Corfmat, J.P., Morse, A.S.: Decentralized control of linear multivariable systems. Automatica **11**, 479–497 (1976)
102. Cover, T.M., Leung, C.S.K.: An achievable rate region for the multiple-access channel with feedback. IEEE Trans. Inf. Theor. **27**, 292–298 (1981)
103. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
104. Coviello, L., Minero, P., Franceschetti, M.: Stabilization over Markov feedback channels: the general case. In: Proceedings of the IEEE Conference on Decision and Control, pp. 3776–3782, Dec 2011
105. Crawford, V.P., Sobel, J.: Strategic information transmission. Econometrica **50**, 1431–1451 (1982)
106. Şen, N., Alajaji, F., Yüksel, S.: Feedback capacity of a class of symmetric finite-state Markov channels. IEEE Trans. Inf. Theor. **56**, 4110–4122 (2011)
107. Csiszar, I., Korner, J.: Information Theory: Coding Theorems for Discrete Memoryless Channels. Akademiai Kiado, Budapest (1981)
108. Curry, R.E.: Estimation and Control with Quantized Measurements. MIT Press, Cambridge (1969)
109. Curtain, R.F., Pritchard, A.J.: Functional Analysis in Modern Applied Mathematics. Academic, London (1977)
110. Hokayem, P., Chatterjee, D., Ramponi, F., Lygeros, J.: On mean square boundedness of stochastic linear systems with bounded controls. Syst. Control Lett. **61**, 375–380 (2012)
111. Dabora, R., Goldsmith, A.: On the capacity of indecomposable finite-state channels with feedback. In: Proceedings of the Allerton Conference on Communication, Control, and Computing, pp. 1045–1052, Sept 2008
112. Dai, J.G.: On positive harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. Ann. Appl. Probab. **5**, 49–77 (1995)
113. Dai, J.G., Meyn, S.P.: Stability and convergence of moments for multiclass queueing networks via fluid limit models. IEEE Trans. Autom. Control **40**, 1889–1904 (1995)
114. de Waal, P.R., van Schuppen, J.H.: A class of team problems with discrete action spaces: optimality conditions based on multimodularity. SIAM J. Control Optim. **38**, 875–892 (2000)
115. Devroye, L., Györfi, L.: Non-parametric Density Estimation: The L_1 View. Wiley, New York (1985)
116. Devroye, L., Györfi, L.: No empirical measure can converge in the total variation sense for all distributions. Ann. Stat. **18**, 1496–1499 (1990)
117. Diaconis, P., Freedman, D.: On the consistency of bayes estimates. Ann. Stat. **14**, 1–26 (1986)
118. Dobrushin, R.L.: An asymptotic bound for the probability error of information transmission through a channel without memory using the feedback. Problemy Kibemetiki **8**, 161–168 (1962)
119. Dobrushin, R.L., Tsybakov, B.S.: Information transmission with additional noise. IRE Trans. Inf. Theor. **18**, 293–304 (1962)
120. Doob, J.L.: Measure Theory. Springer, New York (1994)
121. Douc, R., Fort, G., Moulines, E., Soulier, P.: Practical drift conditions for subgeometric rates of convergence. Ann. Appl. Probab. **14**, 1353–1377 (2004)
122. Draper, S., Sahai, A.: Variable length channel coding with noisy feedback. Eur. Trans. Telecommun. **19**, 355–370 (2008)
123. Dubins, L., Freedman, D.: Measurable sets of measures. Pacific J. Math. **14**, 1211–1222 (1964)
124. Dudley, R.M.: Real Analysis and Probability, 2nd edn. Cambridge University Press, Cambridge (2002)
125. Dudley, R.M., Gine, E., Zinn, J.: Uniform and universal Glivenko-Cantelli classes. J. Theor. Probab. **4**, 485–510 (1991)
126. Duncan, T.E.: On the calculation of mutual information. SIAM J. Appl. Math. **19**, 215–220 (1970)

127. Dunford, N., Schwartz, J.: Linear Operators I. Interscience, New York (1958)
128. D'yachkov, A.G.: Upper bounds on the error probability for discrete memoryless channels with feedback. Problemy Peredachi Informatsii **II**, 13–28 (1975)
129. Ebeid, H., Coleman, T.P.: Source coding with feedforward using the posterior matching scheme. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT), Austin, Texas, USA, 2010
130. Elia, N.: When Bode meets Shannon: control-oriented feedback communication schemes. IEEE Trans. Autom. Control **49**(9), 1477–1488 (2004)
131. Elia, N.: Remote stabilization over fading channels. Syst. Control Lett. **54**, 237–249 (2005)
132. Elia, N., Mitter, S.K.: Stabilization of linear systems with limited information. IEEE Trans. Autom. Control **46**(9), 1384–1400 (2001)
133. Elias, P.: Channel capacity without coding. In: Quarterly Progress Report, pp. 90–93. MIT Research Laboratory of Electronics, Cambridge, 1956 (also in Lectures on Communication System Theory, Eds. E. Baghdady, New York:McGraw Hill, 1961)
134. Fardad, M., Lin, F., Jovanovic, M.R.: Optimal control of vehicular formations with nearest neighbor interactions. IEEE Trans. Autom. Control **57**, 2203–2218 (2012)
135. Fagnani, F., Zampieri, S.: Stability analysis and synthesis for scalar linear systems with a quantized feedback. IEEE Trans. Autom. Control **48**(9), 1569–1584 (2003)
136. Feldbaum, A.A.: Dual control theory. Autom. Remote Control **21**, 1033–1039 (1961)
137. Feller, W.: An Introduction to Probability Theory and Its Applications. Wiley, New York (1971)
138. Fine, T.: Optimum mean-square quantization of a noisy input. IEEE Trans. Inf. Theor. **11**, 293–294 (1965)
139. Fischer, T.R.: Optimal quantized control. IEEE Trans. Autom. Control **27**, 996–998 (1982)
140. Fleming, W.H.: Functions of Several Variables. Addison-Wesley Publishing Co., Reading, Mass.-London (1965)
141. Flynn, T.J., Gray, R.M.: Encoding of correlated observations. IEEE Trans. Inf. Theor. **33**, 773–787 (1987)
142. Fort, G., Meyn, S.P., Moulines, E., Priouret, P.: ODE methods for skip-free Markov chain stability with applications to MCMC. Ann. Appl. Probab. **18**, 664–707 (2008)
143. Foster, D.P., Young, H.P.: On the impossibility of predicting the behavior of rational agents. Proc. Nat.Acad. Sci. **22**, 12848–12853 (2001)
144. Fralix, B.H.: Foster-type criteria for Markov chains on general spaces. J. Appl. Probab. **43**, 1194–1200 (2006)
145. Freudenberg, J.S., Middleton, R.H., Solo, V.: Stabilization and disturbance attenuation over a Gaussian communication channel. IEEE Trans. Autom. Control **55**, 795–799 (2010)
146. Frey, B.J.: Graphical Models for Machine Learning and Digital Communication. MIT Press, Cambridge (1998)
147. Fu, M.: Lack of separation principle for quantized linear quadratic Gaussian control. IEEE Trans. Autom. Control **57**, 2385–2390 (2012)
148. Gaarder, N., Slepian, D.: On optimal finite-state digital transmission systems. IEEE Trans. Inf. Theor. **28**, 167–186 (1982)
149. Gabor, G., Szekeres, G., Györfi, Z.: On the Gaarder-Slepian tracking system conjecture. IEEE Trans. Inf. Theor. **37**, 1165–1168 (1991)
150. Gallager, R.G.: A simple derivation of the coding theorem and some applications. IEEE Trans. Inf. Theor. **11**, 3–18 (1965)
151. Gallager, R.G.: Information Theory and Reliable Communication. Wiley, New York (1968)
152. Gallager, R.G., Nakiboğlu, B.: Variations on a theme by schalkwijk and kailath. IEEE Trans. Inf. Theor. **12**, 6–17 (2010)
153. El Gamal, A., Kim, Y.H.: Network Information Theory. Cambridge University Press, Cambridge (2012)
154. Gastpar, M.: Cut-set arguments for source-channel networks. In: Proceedings of the IEEE International Symposium on Information Theory, vol. 34, 2004

155. Gastpar, M.: Uncoded transmission is exactly optimal for a simple Gaussian sensor network. IEEE Trans. Inf. Theor. **54**, 5247–5251 (2008)
156. Gastpar, M., Rimoldi, B., Vetterli, M.: To code, or not to code: Lossy source-channel communication revisited. IEEE Trans. Inf. Theor. **49**, 1147–1158 (2003)
157. Gattami, A., Bernhardsson, B.M., Rantzer, A.: Robust team decision theory. IEEE Trans. Autom. Control **57**, 794–798 (2012)
158. Geanakoplos, J.: Common knowledge. In: Aumann, R.J., Hart, S. (eds.) Handbook of Game Theory, Chapter 40, pp. 1437–1496. Elsevier, Amsterdam (1994)
159. Geanakoplos, J., Polemarchakis, H.M.: We can't disagree forever. J. Econ. Theor. **28**, 192–200 (1982)
160. Gharesifard, B., Cortés, J.: Distributed continuous-time convex optimization on weight-balanced digraphs, IEEE Transaction on Automatic Control, to appear (2013)
161. Ghosh, J.K., Ramamoorthi, R.V.: Bayesian Nonparametrics. Springer, New York (2003)
162. Giridhar, A., Kumar, P.R.: Toward a theory of in-network computation in wireless sensor networks. IEEE Commun. Mag. **44**, 98–107 (2006)
163. Gish, H., Pierce, J.N.: Asymptotically efficient quantization. IEEE Trans. Inf. Theor. **14**, 676–683 (1968)
164. Goldberg, A., Yüksel, S.: On the value of information in decentralized control. Queen's University Technical Report, August 2012
165. Gong, Z., Aldeen, M.: Stabilization of decentralized control systems. J. Math. Syst. Estimation Control **7**, 1–16 (1997)
166. Goodman, D.J., Gersho, A.: Theory of an adaptive quantizer. IEEE Trans. Commun. **22**, 1037–1045 (1974)
167. Graf, S., Luschgy, H.: Foundations of Quantization for Probability Distributions. Springer, Berlin (2000)
168. Graf, S., Luschgy, H.: Quantization for probability measures in the Prokhorov metric. Theor. Probab. Appl. **53**, 216–241 (2009)
169. Gray, R.M.: Information rates of autoregressive processes. IEEE Trans. Inf. Theor. **16**, 412–421 (1970)
170. Gray, R.M.: Probability, Random Processes, and Ergodic Properties. Springer, New York (1988)
171. Gray, R.M.: Entropy and Information Theory. Springer, New York (1990)
172. Gray, R.M., Boyd, S., Lookabaugh, T.: Low rate distributed quantization of noisy observations. In: Proceedings of the Allerton Conference Communication, Control and Computing, pp. 354–358, 1985
173. Gray, R.M., Davisson, L.D.: Quantizer mismatch. IEEE Trans. Commun. **23**, 439–443 (1975)
174. Gray, R.M., Hashimoto, T.: A note on rate-distortion functions for nonstationary Gaussian autoregressive processes. IEEE Trans. Inf. Theor. **54**, 1319–1322 (2008)
175. Gray, R.M., Kieffer, J.C.: Asymptotically mean stationary measures. Ann. Probab. **8**, 962–973 (1980)
176. Gray, R.M., Neuhoff, D.L.: Quantization. IEEE Trans. Inf. Theor. **44**, 2325–2383 (1998)
177. Grover, P., Park, S.Y., Sahai, A.: The finite-dimensional Witsenhausen counterexample. IEEE Trans. Autom. Control, to appear (2013)
178. Grover, P., Sahai, A.: Implicit and explicit communication in decentralized control. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, pp. 278–285, Monticello, IL, 2010
179. Gupta, V., Chung, T.H., Hassibi, B., Murray, R.M.: On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage. Automatica **42**, 251–260 (2006)
180. Gupta, V., Martins, N.C., Baras, J.S.: Optimal output feedback control using two remote sensors over erasure channels. IEEE Trans. Autom. Control **54**, 1463–1476 (2009)
181. Gupta, V., Martins, N.C., Baras, J.S.: On stability in the presence of analog erasure channels between controller and actuator. IEEE Trans. Autom. Control **545**, 175–179 (2010)
182. Gurt, A., Nair, G.N.: Internal stability of dynamic quantised control for stochastic linear plants. Automatica **45**, 1387–1396 (2009)

183. Györfi, L., Kohler, M.: Nonparametric estimation of conditional distributions. IEEE Trans. Inf. Theor. **53**, 1872–1879 (2007)
184. György, A., Linder, T.: Optimal entropy-constrained scalar quantization of a uniform source. IEEE Trans. Inf. Theor. **46**, 2704–2711 (2000)
185. György, A., Linder, T.: Codecell convexity in optimal entropy-constrained vector quantization. IEEE Trans. Inf. Theor. **49**, 1821–1828 (2003)
186. György, A., Linder, T., Chou, P.A., Betts, B.J.: Do optimal entropy constrained quantizers have a finite or infinite number of codewords? IEEE Trans. Inf. Theor. **49**, 3031–3037 (2003)
187. György, A., Linder, T., Lugosi, G.: Tracking the best quantizer. IEEE Trans. Inf. Theor. **54**, 1604–1625 (2008)
188. Hadjicostis, C.N., Touri, R.: Feedback control utilizing packet dropping network links. In: Proceedings of the IEEE Conference on Decision and Control, pp. 1205–1210, Dec 2002
189. Hairer, M.: Convergence of Markov Processes. Lecture Notes, University of Warwick, 2010
190. Haroutunian, E.A.: A lower bound of the probability of error for channels with feedback. Problemy Peredachi Informatsii **13**, 36–44 (1977)
191. Harremoës, P., Holst, K.K.: Convergence of Markov chains in information divergence. J. Theor. Probab. **22**, 186–202 (2009)
192. Hashimoto, T., Arimoto, S.: On the rate-distortion function for the nonstationary Gaussian autoregressive process. IEEE Trans. Inf. Theor. **26**, 478–480 (1980)
193. Hernandez-Lerma, O., Gonzales-Hernandez, J., Lopez-Martinez, R.R.: Constrained average cost Markov control processes in Borel spaces. SIAM J. Control Optim. **42**, 442–468 (2003)
194. Hernandez-Lerma, O., Lasserre, J.: Discrete-Time Markov Control Processes. Springer, New York (1996)
195. Hernandez-Lerma, O., Lasserre, J.B.: Markov Chains and Invariant Probabilities. Birkhäuser, Basel (2003)
196. Hespanha, J.P., Naghshtabrizi, P., Xu, Y.: A survey of recent results in networked control systems. Proc. IEEE **95**, 138–162 (2007)
197. Ho, Y.C.: Team decision theory and information structures. Proc. IEEE **68**, 644–654 (1980)
198. Ho, Y.C., Chu, K.C.: Team decision theory and information structures in optimal control problems – part I. IEEE Trans. Autom. Control **17**, 15–22 (1972)
199. Horstein, M.: Sequential decoding using noiseless feedback. IEEE Trans. Inf. Theor. **9**, 136–143 (1963)
200. Huang, M., Caines, P.E., Malhamé, R.P.: Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized ϵ-Nash equilibria. IEEE Trans. Autom. Control **52**, 1560–1571 (2007)
201. Huang, M., Dey, S.: Dynamic quantizer design for hidden Markov state estimation via multiple sensors with fusion center feedback. IEEE Trans. Signal Process. **54**, 2887–2896 (2006)
202. Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop McLean-Vlasov systems and the Nash certainty equivalence principle (Special issue in honour of the 65th birthday of Tyrone Duncan). Commun. Inf. Syst. **6**, 221–252 (2006)
203. Huber, P.J.: Robust estimation of a location parameter. Ann. Math. Stat. **35**, 73–101 (1964)
204. Imer, O.C., Başar, T.: Optimal control with limited controls. In: Proceedings of the American Control Conference, pp. 298–303, Minneapolis, MN, 2006
205. Imer, O.C., Yüksel, S., Başar, T.: Optimal control of LTI systems over unreliable communication links. Automatica **42**(9), 1429–1440 (2006)
206. Ishii, H., Başar, T.: Remote control of LTI systems over networks with state quantization. Syst. Control Lett. **54**(1), 15–32 (2005)
207. Ishii, H., Francis, B.A.: Limited data rate in control systems with networks. In: Lecture Notes in Control and Information Sciences, vol. 275. Springer, Berlin (2002)
208. Jadbabaie, A., Lin, J., Morse, A.S.: Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. Autom. Control **48**, 988–1001 (2003)

209. Johnston, A., Yüksel, S.: Stochastic stabilization of partially observed and multi-sensor systems driven by Gaussian noise under fixed-rate information constraints. arXiv:1209.4365
210. Jovanovic, M.R.: On the optimality of localized distributed controllers. Int. J. Syst. Control Commun. **2**, 82–99 (2010)
211. Goblick Jr, T.: Theoretical limitations on the transmission of data from analog sources. IEEE Trans. Inf. Theor. **11**, 558–567 (1965)
212. Kalai, E., Lehrer, E.: Weak and strong merging of opinions. J. Math. Econ. **23**, 73–100 (1994)
213. Khargonekar, P.P., Ozguler, A.B.: Decentralized control and periodic feedback. IEEE Trans. Autom. Control **39**, 877–882 (1994)
214. Kieffer, J.C.: Stochastic stability for feedback quantization schemes. IEEE Trans. Inf. Theor. **28**, 248–254 (1982)
215. Kieffer, J.C., Dunham, J.G.: On a type of stochastic stability for a class of encoding schemes. IEEE Trans. Inf. Theor. **29**, 793–797 (1983)
216. Kim, Y.H., Lapidoth, A., Weissman, T.: Error exponents for the Gaussian channel with active noisy feedback. IEEE Trans. Inf. Theor. **57**, 1223–1236 (2011)
217. Kobayashi, H., Hanafusa, H., Yoshikawa, T.: Controllability under decentralized information structure. IEEE Trans. Autom. Control **23**, 182–188 (1978)
218. Krainak, J.C., Marcus, S.I.: Static team problems – part I: sufficient conditions and the exponential cost criterion. IEEE Trans. Autom. Control **27**, 839–848 (1982)
219. Krainak, J.C., Speyer, J.S., Marcus, S.I.: Static team problems – part II: affine control laws, projections, algorithms, and the LEGT problem. IEEE Trans. Autom. Control **27**, 848–859 (1982)
220. Kumar, U., Gupta, V., Laneman, J.: Stability across a Gaussian product channel: Necessary and sufficient conditions. IEEE Trans. Autom. Control, under review (2012)
221. Kumar, U., Gupta, V., Laneman, J.: On stability across a gaussian product channel. In: IEEE Conference on Decision and Control, pp. 3142–3147, Florida, USA, 2011
222. Kumar, U., Gupta, V., Laneman, J.: Sufficient conditions for stabilizability over Gaussian relay and cascade channels. In: Proceedings of the IEEE Conf. Decision and Control, Atlanta, GA, USA, 2010
223. Kurtaran, B.: A concise derivation of the LQG one-step-delay sharing problem solution. IEEE Trans. Autom. Control **20**, 808–810 (1975)
224. Kurtaran, B.: Corrections and extensions to decentralized control with delayed sharing information pattern. IEEE Trans. Autom. Control **24**, 656–657 (1979)
225. Kushner, H.J.: Introduction to Stochastic Control Theory. Holt, Rinehart and Winston, New York (1972)
226. Lasry, J.M., Lions, P.L.: Mean field games. Jpn. J. Math. **2**, 229–260 (2007)
227. Lasserre, J.B.: Invariant probabilities for Markov chains on a metric space. Stat. Probab. Lett. **34**, 259–265 (1997)
228. Lee, K.H., Petersen, D.P.: Optimal linear coding for vector channels. IEEE Trans. Commun. **24**, 1283–1290 (1976)
229. Lehrer, E., Rosenberg, D., Shmaya, E.: Signaling and mediation in games with common interests. Games Econ. Behav. **68**, 670–682 (2010)
230. Lemmon, M.: Event-triggered feedback in control, estimation, and optimization. In: Bemporad, A., Heemels, M., Johansson, M. (eds.) Networked Control Systems, ser. Lecture Notes in Control and Information Sciences, pp. 293–358. Springer, New York, (2010)
231. Lessard, L., Lall, S.: A state-space solution to the two-player decentralized optimal control problem. In: Annual Allerton Conference on Communication, Control and Computing, Monticello, IL, 2011
232. Lewis, J.B., Tou, J.T.: Optimum sampled-data systems with quantized control signals. IEEE Trans. Appl. Ind. **82**, 229–233 (1965)
233. Li, S., Başar, T.: Asymptotic agreement and convergence of asynchronous stochastic algorithms. IEEE Trans. Autom. Control **32**, 612–618 (1987)

234. Liberzon, D., Nešić, D.: Input-to-state stabilization of linear systems with quantized state measurements. IEEE Trans. Autom. Control **52**(5), 767–781 (2007)
235. Linder, T.: On the training distortion of vector quantizers. IEEE Trans. Inf. Theor. **46**, 1617–1623 (2000)
236. Linder, T., Lugosi, G.: A zero-delay sequential scheme for lossy coding of individual sequences. IEEE Trans. Inf. Theor. **47**, 2533–2538 (2001)
237. Linder, T., Zamir, R.: Causal coding of stationary sources and individual sequences with high resolution. IEEE Trans. Inf. Theor. **52**, 662–680 (2006)
238. Lipsa, G.M., Martins, N.C.: Optimal memoryless control in Gaussian noise: a simple counterexample. Automatica **47**, 552–558 (2011)
239. Lipsa, G.M., Martins, N.C.: Remote state estimation with communication costs for first-order LTI systems. IEEE Trans. Autom. Control **56**, 2013–2025 (2011)
240. Lloyd, S.P.: Least squared quantization in pcm. IEEE Trans. Inform. Theor. **28**, 129–137 (1982)
241. Loève, M.: Probability Theory. Van Nostrand, New York (1963)
242. Luenberger, D.G.: Optimization by Vector Space Methods. Wiley, New York (1969)
243. Luenberger, D.G.: Linear and Nonlinear Programming. Addison-Wesley, Reading (1984)
244. Johansson, K.H., Rabi, M., Johansson, M.: Optimal stopping for event-triggered sensing and actuation. In: Proceedings of the IEEE Conference on Decision and Control, pp. 3607–3612, Dec 2008
245. Ma, N., Ishwar, P.: Some results on distributed source coding for interactive function computation. IEEE Trans. Inf. Theor. **57**, 6180–6195 (2011)
246. Mahajan, A.: Sequential decomposition of sequential teams: applications to real-time communication and networked control systems. Ph.D. Dissertation, University of Michigan, Ann Arbor (2008)
247. Mahajan, A., Martins, N.C., Rotkowitz, M., Yüksel, S.: Information structures in optimal decentralized control. In: IEEE Conference on Decision and Control, Hawaii, USA, 2012
248. Mahajan, A., Teneketzis, D.: On the design of globally optimal communication strategies for real-time noisy communication with noisy feedback. IEEE J. Sel. Areas Commun. **26**, 580–595 (2008)
249. Mahajan, A., Teneketzis, D.: Optimal design of sequential real-time communication systems. IEEE Trans. Inf. Theor. **55**, 5317–5338 (2009)
250. Mahajan, A., Teneketzis, D.: Optimal performance of networked control systems with non-classical information structures. SIAM J. Control Optim. **48**, 1377–1404 (2009)
251. Mahajan, A., Yüksel, S.: Measure and cost dependent properties of information structures. In: Proceedings of the American Control Conference, Baltimore, MD, USA, 2010
252. Mahmoud, M.S.: Decentralized Systems with Design Constraints. Springer, London (2011)
253. Malyšev, V.A., Menšikov, M.V.: Ergodicity, continuity and analyticity of countable Markov chains. Trans. Moscow Math. Soc. **39**, 1–48 (1981)
254. Marschak, J.: Elements for a theory of teams. Manag. Sci. **1**, 127–137 (1955)
255. Marshak, J., Radner, R.: Economic Theory of Teams. Yale University Press, New Haven (1972)
256. Martins, N.C., Dahleh, M.A.: Feedback control in the presence of noisy channels: Bode-like fundamental limitations of performance. IEEE Trans. Autom. Control **53**, 1604–1615 (2008)
257. Martins, N.C., Dahleh, M.A., Doyle, J.C.: Fundamental limitations of disturbance attenuation in the presence of side information. IEEE Trans. Autom. Control **52**, 56–66 (2007)
258. Martins, N.C., Dahleh, M.A., Elia, N.: Feedback stabilization of uncertain systems in the presence of a direct link. IEEE Trans. Autom. Control **51**(3), 438–447 (2006)
259. Massey, J.L.: Causality, feedback and directed information. In: IEEE International Symposium on Information Theory, pp. 303–305, Waikiki, Hawaii, 1990
260. Matveev, A.S.: State estimation via limited capacity noisy communication channels. Math. Control Signals Syst. **20**, 1–35 (2008)

261. Matveev, A.S., Savkin, A.: An analogue of Shannon information theory for detection and stabilization via noisy discrete communication channels. SIAM J. Control Optim. **46**, 1323–1367 (2007)
262. Matveev, A.S., Savkin, A.V.: The problem of LQG optimal control via a limited capacity communication channel. Syst. Control Lett. **53**, 51–64 (2004)
263. Matveev, A.S., Savkin, A.V.: Decentralized stabilization of linear systems via limited capacity communication networks, pp. 1155–1161. Seville, Spain (2005)
264. Matveev, A.S., Savkin, A.V.: Multirate stabilization of linear multiple sensor systems via limited capacity communication channels. SIAM J. Control Optim. **44**, 584–617 (2006)
265. Matveev, A.S., Savkin, A.V.: Decentralized stabilization of networked systems under data-rate constraints. In: Proceedings of the International Federation of Automatic Control World Congress, Seoul, Korea, 2008
266. Matveev, A.S., Savkin, A.V.: Estimation and Control over Communication Networks. Birkhäuser, Boston (2008)
267. Merhav, N.: Exponential error bounds on parameter modulation-estimation for discrete memoryless channels. IEEE Trans. Inf. Theor. under review (2012)
268. Merhav, N.: On optimum parameter modulation-estimation from a large deviations perspective. IEEE Trans. Inf. Theor. **58**, 172–182 (2012)
269. Meyn, S.P.: Control Techniques for Complex Networks. Cambridge University Press, Cambridge (2007)
270. Meyn, S.P., Tweedie, R.: Stability of Markovian processes I: criteria for discrete-time chains. Adv. Appl. Probab. **24**, 542–574 (1992)
271. Meyn, S.P., Tweedie, R.: Markov Chains and Stochastic Stability. Springer, London (1993)
272. Meyn, S.P., Tweedie, R.: State-dependent criteria for convergence of Markov chains. Ann. Appl. Probab. **4**, 149–168 (1994)
273. Minero, P., Franceschetti, M., Dey, S., Nair, G.N.: Data rate theorem for stabilization over time-varying feedback channels. IEEE Trans. Autom. Control **54**(2), 243–255 (2009)
274. Mitter, S., Sahai, A.: Information and control: Witsenhausen revisited. In Learning, Control, and Hybrid Systems, pp. 281–293. Springer, New York (1999)
275. Mooij, M., Kappen, H.J.: Sufficient conditions for convergence of the sum-product algorithm. IEEE Trans. Inf. Theor. **53**, 4422–4437 (2007)
276. Muraki, N., Ohya, M.: Note on continuity of information rate. Illinois J. Math. **36**, 529–550 (1992)
277. Nachbar, J.H.: Prediction, optimization, and learning in games. Econometrica **65**, 275–309 (1997)
278. Naiqi, W., Renhou, L., Baosheng, H.: Decentralized controllability and observability, and decentralized stabilization. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, pp. 598–601, Aug 1988
279. Nair, G.N.: A nonstochastic information theory for communication and state estimation. IEEE Trans. Autom. Control (2013, to appear)
280. Nair, G.N., Evans, R.J.: Stabilizability of stochastic linear systems with finite feedback data rates. SIAM J. Control Optim. **43**, 413–436 (2004)
281. Nair, G.N., Evans, R.J.: Cooperative networked stabilisability of linear systems with measurement noise. In: Proceedings of the 15th IEEE Mediterranean Conference Control and Automation, Athens, Greece, 2007
282. Nair, G.N., Evans, R.J., Caines, P.E.: Stabilising decentralised linear systems under data rate constraints. In: Proceedings of the IEEE Conference on Decision and Control, pp. 3992–3997, Dec 2004
283. Nair, G.N., Fagnani, F., Zampieri, S., Evans, J.R.: Feedback control under data constraints: an overview. Proc. IEEE **95**, 108–137 (2007)
284. Nakiboğlu, B., Gallager, R.G.: Error exponents for variable-length block codes with feedback and cost constraints. IEEE Trans. Inf. Theor. **54**, 945–963 (2008)

285. Nakiboğlu, B., Zheng, L.: Upper bounds to error probability with feedback. In: Proceedings of the IEEE International Symposium on Information Theory, South Korea, pp. 1515–1519, July 2009
286. Nayyar, A., Mahajan, A., Teneketzis, D.: Decentralized stochastic control with partial sharing information structures: a common information approach. IEEE Transactions on Automatic Control, to appear (2013)
287. Nayyar, A., Mahajan, A., Teneketzis, D.: Optimal control strategies in delayed sharing information structures. IEEE Trans. Autom. Control **56**, 1606–1620 (2011)
288. Nayyar, A., Mahajan, A., Teneketzis, D.: The common-information approach to decentralized stochastic control. In: Como, G., Bernhardsson, B., Rantzer, A. (eds.) Information and Control in Networks. Springer, New York (2013)
289. Nayyar, A., Teneketzis, D.: On the structure of real-time encoders and decoders in a multi-terminal communication system. IEEE Trans. Inf. Theor. **57**, 6196–6214 (2011)
290. Nedić, A., Özdaglar, A., Parrilo, P.A.: Constrained consensus and optimization in multi-agent networks. IEEE Trans. Autom. Control **55**, 922–938 (2010)
291. Nemirovsky, A.S., Yudin, D.B.: Problem Complexity and Method Efficiency in Optimization. Wiley-Interscience, New York (1983)
292. Neuhoff, D.L., Gilbert, R.K.: Causal source codes. IEEE Trans. Inf. Theor. **28**, 701–713 (1982)
293. Nielsen, L.: Common knowledge, communication and convergence of beliefs. Math. Soc. Sci. **8**, 1–14 (1984)
294. Nummelin, E.: General irreducible Markov chains and non-negative operators. Cambridge University Press, Cambridge (1984)
295. Oksendal, B.: Stochastic Differential Equations. Springer, Berlin (2003)
296. Olshevsky, A.: Efficient information aggregation methods in distributed control and signal processing. Ph.D. Thesis, Department of EECS, MIT, Cambridge, MA (2010)
297. Omura, J.K.: On general Gilbert bounds. IEEE Trans. Inf. Theor. **19**, 661–666 (1973)
298. Omura, J.K.: Expurgated bounds, bhattacharyya distance, and rate distortion function. Inform. Control **24**, 358–383 (1974)
299. Ooi, J., Verbout, S., Ludwig, J., Wornell, G.: A separation theorem for periodic sharing information patterns in decentralized control. IEEE Trans. Autom. Control **42**, 1546–1550 (1997)
300. Orlitsky, A., Roche, J.R.: Coding for computing. IEEE Trans. Inf. Theor. **47**, 903–917 (2001)
301. Ostrovsky, R., Rabani, Y., Schulman, L.: Error-correcting codes for automatic control. IEEE Trans. Inf. Theor. **55**(7), 2931–2941 (2009)
302. Özcelikkale, A., Yüksel, S., Ozaktas, H.M.: Unitary precoding and basis dependency of MMSE performance for Gaussian erasure channel. arXiv IEEE Trans. Inf. Theor., under review (2013)
303. Özguner, Ü., Davison, E.J.: Sampling and decentralized fixed modes. In: Proceedings of the in American Control Conference, pp. 257–262, 1985
304. Papoulis, A.: Probability, Random Variables and Stochastic Processes. McGraw-Hill, New York (1965)
305. Park, S.Y., Sahai, A.: Network coding meets decentralized control: capacity-stabilizabililty equivalence. In: IEEE Conference on Decision and Control, pp. 4817–4822, Florida, 2011
306. Permuter, H.H., Weissman, T., Goldsmith, A.J.: Finite state channels with time-invariant deterministic feedback. IEEE Trans. Inf. Theor. **55**(2), 644–662 (2009)
307. Phelps, R.P.: Lectures on Choquet's theorem. Van Nostrand, New York (1966)
308. Pilc, R.: The optimum linear modulator for a Gaussian source used with a Gaussian channel. IEEE Trans. Autom. Control **48**, 3075–3089 (1969)
309. Pollard, D.: Quantization and the method of k-means. IEEE Trans. Inf. Theor. **28**, 199–205 (1982)
310. Polyanskiy, Y., Poor, H.V., Verdú, S.: Variable-length coding with feedback in the non-asymptotic regime. In: Proceedings of the IEEE International Symposium on Information Theory, pp. 231–235, 2010

311. Poor, V.: An Introduction to Signal Detection and Estimation. Springer, New York (1994)
312. Pradhan, S.S., Ramchandran, K.: Distributed source coding using syndromes (discus): design and construction. IEEE Trans. Inf. Theor. **49**, 626–643 (2003)
313. Quevedo, D., Nešić, D.: Robust stability of packetized predictive control of nonlinear systems with disturbances and markovian packet losses. Automatica **48**(8), 1803–1811 (2012)
314. Quevedo, D., Østergaard, J., Nešić, D.: Packetized predictive control of stochastic systems over bit-rate limited channels with packet loss. IEEE Trans. Autom. Control **56**(12), 2854–2868 (2011)
315. Quevedo, D., Silva, E.I., Goodwin, G.C.: Control over unreliable networks affected by packet erasures and variable transmission delays. IEEE J. Sel. Areas Commun. **26**, 672–685 (2008)
316. Radner, R.: Team decision problems. Ann. Math. Stat. **33**, 857–881 (1962)
317. Raginsky, M.: Shannon meets Blackwell and Le Cam: channels, codes, and statistical experiments. In: Proceedings of the IEEE International Symposium on Information Theory, pp. 1220–1224, St. Petersburg, Russia, 2011
318. Raginsky, M., Rakhlin, A.: Information-based complexity, feedback and dynamics in convex programming. IEEE Trans. Inf. Theor. **57**, 7036– 7056 (2011)
319. Ramponi, F., Milias-Argeitis, A., Chatterjee, D., Hokayem, P., Lygeros, J.: Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input. IEEE Trans. Autom. Control **55**, 2414–2418 (2010)
320. Rantzer, A.: Linear quadratic team theory revisited. In: Proceedings of the American Control Conference, pp. 1637–1642, 2006
321. Rezaei, F., Charalambous, C.D., Ahmed, N.U.: Optimization of stochastic uncertain systems with variational norm constraints. In: Proceedings of the IEEE Conference on Decision and Control, pp. 2159–2163, Dec 2007
322. Roberts, G.O., Rosenthal, J.S.: General state space Markov chains and MCMC algorithms. Prob. Surv. **1**, 20–71 (2004)
323. Rotkowitz, M.: Linear controllers are uniformly optimal for the witsenhausen counterexample. In: Proceedings of the IEEE Conference on Decision and Control, pp. 553–558, San Diego, CA, Dec 2006
324. Rotkowitz, M., Lall, S.: A characterization of convex problems in decentralized control. IEEE Trans. Autom. Control **51**, 274–286 (2006)
325. Rudin, W.: Real and Complex Analysis, 3rd edn. McGraw-Hill, New York (1987)
326. Verdú, S., Shamai, S., Zamir, R.: Systematic lossy source/channel coding. IEEE Trans. Inf. Theor. **44**, 564–579 (1998)
327. Sabau, S., Martins, N.C.: Stabilizability and norm-optimal control design subject to sparsity constraints. IEEE Trans. Autom. Control, under review (2013)
328. Sahai, A.: Anytime information theory. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA (2001)
329. Sahai, A.: Coding unstable scalar Markov processes into two streams. In: Proceedings of the IEEE International Symposium on Information Theory, p. 462, 2004
330. Sahai, A.: Why block length and delay behave differently for channel coding with feedback. IEEE Trans. Inf. Theor. **54**, 1860–1886 (2008)
331. Sahai, A., Mitter, S.: The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link part I: Scalar systems. IEEE Trans. Inf. Theor. **52**(8), 3369–3395 (2006)
332. Sandell, N., Athans, M.: Solution of some nonclassical LQG stochastic decision problems. IEEE Trans. Autom. Control **19**, 108–116 (1974)
333. Savage, L.J.: The Foundations of Statistics. Wiley, New York (1954)
334. Schalkwijk, J., Kailath, T.: A coding scheme for additive noise channels with feedback I: No bandwidth constraint. IEEE Trans. Inf. Theor. **12**, 172–182 (1966)
335. Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., Sastry, S.S.: Foundations of control and estimation over lossy networks. Proc. IEEE **95**(1), 163–187 (2007)
336. Schutzenberger, M.P.: On the quantization of finite-dimensional messages. Inf. Control **1**, 153–158 (1958)

337. Shannon, C., Gallager, R.G., Berlekamp, E.R.: Lower bounds to error probability for coding on discrete memoryless channels. Inform. Control, **10**(65–103), 522–552 (1967)
338. Shields, P.C.: The Ergodic Theory of Discrete Sample Paths. American Mathematical Society, Providence (1996)
339. Shu, Z., Middleton, R.H.: Stabilization over power-constrained parallel Gaussian channels. IEEE Trans. Autom. Control **56**, 1718–1724 (2011)
340. Siljak, D.D.: Decentralized Control of Complex Systems. Academic, Boston (1991)
341. Silva, E.I., Derpich, M.S., Østergaard, J.: An achievable data-rate region subject to a stationary performance constraint for LTI plants. IEEE Trans. Autom. Control **56**, 1968–1973 (2011)
342. Singh, M.G.: Decentralised Control. North Holland, Amsterdam (1981)
343. Socratous, Y., Rezaei, F., Charalambous, C.D.: Nonlinear estimation for a class of systems. IEEE Trans. Inf. Theor. **55**, 1930–1938 (2009)
344. Srikant, R.: The Mathematics of Internet Congestion Control. Birkhäuser-Springer, USA (2004)
345. Srikant, R., Başar, T.: Asymptotic solutions to weakly coupled stochastic teams with nonclassical information. IEEE Trans. Autom. Control **37**, 163–173 (1992)
346. Stark, W.E., McEliece, R.J.: On the capacity of channels with block memory. IEEE Trans. Inf. Theor. **3**, 322–324 (1988)
347. Stettner, L.: Ergodic control of partially observed Markov control processes with equivalent transition probabilities. Applicationes Mathematicae **22**, 25–38 (1993)
348. Striebel, C.: Sufficient statistics in the optimum control of stochastic systems. J. Math. Anal. Appl. **12**, 576–592 (1965)
349. Sukhavasi, R.T., Hassibi, B.: Linear error correcting codes with anytime reliability. arXiv (2011)
350. Sundaram, S., Hadjicostis, C.N.: Distributed function calculation and consensus using linear iterative strategies. IEEE J. Sel. Areas Commun.: Issue Control Commun. **26**, 650–660 (2008)
351. Tabuada, P.: Event-triggered real-time scheduling of stabilizing control tasks. IEEE Trans. Autom. Control **52**, 1680–1685 (2007)
352. Tatikonda, S.: Control under communication constraints. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA (2000)
353. Tatikonda, S.: Some scaling properties of large distributed control systems. In: Proceedings of the 42nd IEEE CDC, pp. 3142–3147, Dec 2003
354. Tatikonda, S.: Cooperative control under communication constraints. In: Proceedings of the IEEE Information Theory Workshop, pp. 243–246, Porto, Portugal, May 2008
355. Tatikonda, S., Mitter, S.: Control under communication constraints. IEEE Trans. Autom. Control **49**(7), 1056–1068 (2004)
356. Tatikonda, S., Mitter, S.: Control under communication constraints. IEEE Trans. Autom. Control **49**(7), 1056–1068 (2004)
357. Tatikonda, S., Mitter, S.: The capacity of channels with feedback. IEEE Trans. Inf. Theor. **55**(1), 323–349 (2009)
358. Tatikonda, S., Sahai, A., Mitter, S.: Stochastic linear control over a communication channels. IEEE Trans. Autom. Control **49**, 1549–1561 (2004)
359. Teneketzis, D.: Communication in decentralized control. Ph.D. Dissertation, M.I.T. (1979)
360. Teneketzis, D.: On information structures and nonsequential stochastic control. CWI Q. **9**, 241–260 (1996)
361. Teneketzis, D.: On the structure of optimal real-time encoders and decoders in noisy communication. IEEE Trans. Inf. Theor. **52**, 4017–4035 (2006)
362. Teneketzis, D., Varaiya, P.: Consensus in distributed estimation. In: Poor, H.V. (ed.) Advances in Statistical Signal Processing, pp. 361–386. JAI Press, Greenwich (1988)
363. Topsœ, F.: Uniformity in convergence of measures. Z. Wahrscheinlichkeitsth **39**, 1–30 (1977)
364. Touri, B., Nedić, A.: On ergodicity, infinite flow and consensus in random models. IEEE Trans. Autom. Control **56**, 1593–1605 (2011)

365. Tsitsiklis, J.N.: Problems in decentralized decision making and computation. Ph.D. Thesis, Department of EECS, MIT, Cambridge, MA (1984)
366. Tsitsiklis, J.N.: Extremal properties of likelihood-ratio quantizers. IEEE Trans. Commun. **41**, 550–558 (1993)
367. Tsitsiklis, J.N., Athans, M.: Convergence and asymptotic agreement in distributed decision problems. IEEE Trans. Autom. Control **29**, 42–50 (1984)
368. Tsitsiklis, J.N., Athans, M.: On the complexity of decentralized decision making and detection problems. IEEE Trans. Autom. Control **30**, 440–446 (1985)
369. Tsitsiklis, J.N., Luo, Z.-Q.: Communication complexity of convex optimization. J. Complex. **3**, 231–243 (1987)
370. Tweedie, R.L.: Drift conditions and invariant measures for Markov chains. Stoch. Process. Appl. **92**, 345–354 (2001)
371. Ugrinovskii, V.A., Petersen, I.R.: Minimax LQG control of stochastic partially observed uncertain systems. SIAM J. Control Optim. **40**, 1189–1226 (2001)
372. van Zandt, T.: Decentralized Information Processing in the Theory of Organizations, in Contemporary Economic Development Reviewed, vol. 4: The Enterprise and its Environment. Murat Sertel, ed. London: MacMillan. (1997)
373. Vapnik, V.N.: The Nature of Statistical Learning Theory, 2nd edn. Springer, New York (2000)
374. Varaiya, P., Walrand, J.: On delayed-sharing patterns. IEEE Trans. Autom. Control **23**, 443–445 (1978)
375. Veeravalli, V.V., Başar, T., Poor, H.V.: Decentralized sequential detection with a fusion center performing the sequential test. IEEE Trans. Inf. Theor. **39**, 433–442 (1993)
376. Vembu, S., Verdú, S., Steinberg, Y.: The source-channel separation theorem revisited. IEEE Trans. Inf. Theor. **41**, 44–54 (1995)
377. Venkataramanan, R., Pradhan, S.S.: Source coding with feed-forward: Rate-distortion theorems and error exponents for a general source. IEEE Trans. Inf. Theor. **53**, 2154–2179 (2007)
378. Venkataramanan, R., Pradhan, S.S.: A new achievable rate region for the discrete memoryless multiple-access channel with feedback. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT), Seoul, South Korea, 2009
379. Verdú, S., Han, T.S.: A general formula for channel capacity. IEEE Trans. Inf. Theor. **40**, 1147–1157 (1994)
380. Viterbi, A.J., Omura, J.K.: Principles of Digital Communication and Coding. McGraw-Hill, Inc. New York (1979)
381. von Neumann, H.J.: A certain zero-sum two-person game equivalent to the optimal assignment problem. Contributions to the Theory of Games, vol. II (ed. by H. W. Kuhn and A. W. Tucker) Ann. of Math. Studies, no. 28, Princeton, pp. 5–12 (1953)
382. Voulgaris, P.G.: A convex characterization of classes of problems in control with specific interaction and communication structures. In: Proceedings of the American Control Conference, pp. 3128–3133, June 2001
383. Wald, A.: Statistical Decision Functions. Wiley, New York (1950)
384. Walrand, J.C., Varaiya, P.: Causal coding and control of markov chains. Syst. Control Lett. **3**, 189–192 (1983)
385. Walrand, J.C., Varaiya, P.: Optimal causal coding-decoding problems. IEEE Trans. Inf. Theor. **19**, 814–820 (1983)
386. Wang, S., Davison, E.J.: On the stabilization of decentralized control systems. IEEE Trans. Autom. Control **18**, 473–478 (1973)
387. Wang, S.H.: Stabilization of decentralized control systems via time-varying controllers. IEEE Trans. Autom. Control **27**, 741–744 (1982)
388. Weissman, T., Kim, Y.H., Permuter, H.H.: Directed information, causal estimation, and communication in continuous time. IEEE Trans. Inf. Theor. **59**, 1271–1287 (2013)
389. Weissman, T., Merhav, N.: On causal source codes with side information. IEEE Trans. Inf. Theor. **51**, 4003–4013 (2005)
390. Wernersson, N., Karlsson, J., Skoglund, M.: Distributed quantization over noisy channels. IEEE Trans. Commun. **57**, 1693–1700 (2009)

391. Wheeden, R.L., Zygmund, A.: Measure and Integral. Dekker, New York (1977)
392. Willems, J.L.: Time-varying feedback for the stabilization of fixed modes in decentralized control systems. Automatica **25**, 127–131 (1989)
393. Witsenhausen, H.S.: On information structures, feedback and causality. SIAM J. Control **9**, 149–160 (1971)
394. Witsenhausen, H.S.: A standard form for sequential stochastic control. Math. Syst. Theor. **7**, 5–11 (1973)
395. Witsenhausen, H.S.: The zero-error side information problem and chromatic numbers. IEEE Trans. Inf. Theor. **22**, 592–593 (1976)
396. Witsenhausen, H.S.: On the structure of real-time source coders. Bell Syst. Tech. J. **58**, 1437–1451 (1979)
397. Witsenhausen, H.S.: Indirect rate-distortion problems. IEEE Trans. Inf. Theor. **26**, 518–521 (1980)
398. Witsenhausen, H.S.: A counterexample in stochastic optimal control. SIAM J. Control **6**, 131–147 (1968)
399. Witsenhausen, H.S.: Separation of estimation and control for discrete time systems. Proc. IEEE **59**, 1557–1566 (1971)
400. Witsenhausen, H.S.: The intrinsic model for discrete stochastic control: some open problems. In: Lecture Notes in Mathematical Economics and System Theory, vol. 107, pp. 322–335. Springer, New York (1975)
401. Witsenhausen, H.S.: Equivalent stochastic control problems. Math. Control Signals Syst. **1**, 3–11 (1988)
402. Wong, E., Hajek, B.E.: Stochastic Processes in Engineering Systems. Springer, New York (1985)
403. Wong, P.W., Gray, R.M.: Sigma-delta modulation with i.i.d. Gaussian inputs. IEEE Trans. Inf. Theor. **36**, 784–778 (1990)
404. Wong, W.S.: Control communication complexity of distributed control systems. SIAM J. Control Optim. **48**, 1722–1742 (2009)
405. Wong, W.S., Baillieul, J.: Control communication complexity of distributed actions. IEEE Trans. Autom. Control. **57**, 2731–2345 (2012)
406. Wong, W.S., Brockett, R.W.: Systems with finite communication bandwidth constraints - part ii: Stabilization with limited information feedback. IEEE Trans. Autom. Control **42**, 1294–1299 (1997)
407. Wu, W., Arapostathis, A.: Optimal sensor querying: General markovian and lqg models with controlled observations. IEEE Trans. Autom. Control **53**, 1392–1405 (2008)
408. Wu, Y., Verdú, S.: Functional properties of MMSE. In: Proceedings of the IEEE International Symposium on Information Theory, pp. 13–18, 2010
409. Wu, Y., Verdú, S.: Witsenhausen's counterexample: a view from optimal transport theory. In: Proceedings of the IEEE Conference on Decision and Control, Florida, USA, pp. 5732–5737, Dec 2011
410. Wyner, A.D.: Another look at the coding theorem of information theory: a tutorial. Proc. IEEE **58**, 894–913 (1970)
411. Yamamoto, H., Itoh, K.: Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback. IEEE Trans. Inf. Theor. **25**(6), 729–733 (1979)
412. Yao, A.C.C.: Some complexity questions related to distributive computing. In: Proceedings of the of the 11th Annual ACM Symposium on Theory of Computing, 1979
413. Yoshikawa, T.: Dynamic programming approach to decentralized control problems. IEEE Trans. Autom. Control **20**, 796–797 (1975)
414. Yoshikawa, T.: Decomposition of dynamic team decision problems. IEEE Trans. Autom. Control **23**, 627–632 (1978)
415. You, K., Xie, L.: Minimum data rate for mean square stabilization of discrete LTI systems over lossy channels. IEEE Trans. Autom. Control. **55**, 2373–2378, Oct 2010

416. Yüksel, S.: A random time stochastic drift result and application to stochastic stabilization over noisy channels. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, 2009
417. Yüksel, S.: Stochastic nestedness and the belief sharing information pattern. IEEE Trans. Autom. Control **54**, 2773–2786 (2009)
418. Yüksel, S.: On optimal causal coding of partially observed Markov sources under classical and non-classical information structures. In: Proceedings of the IEEE International Symposium on Information Theory, pp. 81–85, Austin, TX, 2010
419. Yüksel, S.: Stochastic stabilization of noisy linear systems with fixed-rate limited feedback. IEEE Trans. Autom. Control **55**, 2847–2853 (2010)
420. Yüksel, S.: Stochastic stability and ergodicity of linear Gaussian systems controlled over discrete channels. In: Proceedings of the IEEE Conference on Decision and Control, pp. 4680–4685, Florida, USA, 2011
421. Yüksel, S.: A tutorial on quantizer design for networked control systems: Stabilization and optimization. Appl. Comput. Math. **11**, 365–403 (2011)
422. Yüksel, S.: Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems. IEEE Trans. Inf. Theor. **58**, 6332–6354 (2012)
423. Yüksel, S.: Jointly optimal LQG quantization and control policies for multi-dimensional linear Gaussian sources. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, 2012
424. Yüksel, S.: Design of information channels for optimization and stabilization in networked control. In: Como, G., Bernhardsson, B., Rantzer, A. (eds.) Information and Control in Networks. Springer, New York (2013)
425. Yüksel, S.: On optimal causal coding of partially observed Markov sources in single and multi-terminal settings. IEEE Trans. Inf. Theor. **59**, 424–437 (2013)
426. Yüksel, S., Başar, T.: Coding and control over discrete noisy forward and feedback channels. In: Proceedings of the IEEE Conference on Decision and Control, pp. 2517–2522, Seville, Spain, Dec 2005
427. Yüksel, S., Başar, T.: Minimum rate coding for LTI systems over noiseless channels. IEEE Trans. Autom. Control **51**(12), 1878–1887 (2006)
428. Yüksel, S., Başar, T.: On the absence of rate loss in decentralized sensor and controller structure for asymptotic stability. In: Proceedings of the American Control Conference, Minneapolis, MN, 2006
429. Yüksel, S., Başar, T.: Communication constraints for decentralized stabilizability with time-invariant policies. IEEE Trans. Autom. Control **52**, 1060–1066 (2007)
430. Yüksel, S., Başar, T.: Information theoretic study of the signaling problem in decentralized stabilization of noisy linear systems. In: Proceedings of the IEEE Conference on Decision and Control, New Orleans, 2007
431. Yüksel, S., Başar, T.: Optimal signaling policies for decentralized multi-controller stabilizability over communication channels. IEEE Trans. Autom. Control **52**, 1969–1974 (2007)
432. Yüksel, S., Başar, T.: Control over noisy forward and reverse channels. IEEE Trans. Autom. Control **56**, 1014–1029 (2011)
433. Yüksel, S., Başar, T., Meyn, S.P.: Optimal causal quantization of Markov sources with distortion constraints. In: Proceedings of the Information Theory and Applications (ITA) Workshop, UCSD, 2008
434. Yüksel, S., Linder, T.: On optimal zero-delay quantization of vector Markov sources. Technical Report, Queen's University
435. Yüksel, S., Linder, T.: On optimization and convergence of observation channels and quantizers in stochastic control. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, pp. 286–293, Monticello, IL, 2010
436. S. Yüksel and T. Linder. Optimization and convergence of observation channels in stochastic control. In: American Control Conference, pp. 637–642, San Francisco, CA, 2011

437. Yüksel, S., Linder, T. On optimal zero-delay quantization of vector Markov sources. Queen's University Technical Report, 2013
438. Yüksel, S., Linder, T.: Optimization and convergence of observation channels in stochastic control. SIAM J. Control Optim. **50**, 864–887 (2012)
439. Yüksel, S., Meyn, S.P.: Random-time, state-dependent stochastic drift for Markov chains and application to stochastic stabilization over erasure channels. IEEE Trans. Autom. Control **58**, 47–59 (2013)
440. Yüksel, S., Tatikonda, S.: Distributed sensing and control over communication channels and a counterexample. In: Proceedings of the Annual Allerton Conference on Communications, Control and Computing, Monticello, IL, Sept 2007
441. Yüksel, S., Tatikonda, S.: A counterexample in distributed optimal sensing and control. IEEE Trans. Autom. Control **54**, 841–844 (2009)
442. Zador, P.: Asymptotic quantization error of continuous signals and the quantization dimension. IEEE Trans. Inf. Theor. **28**, 139–148 (1982)
443. Zaidi, S.A., Oechtering, T., Yüksel, S., Skoglund, M.: Stabilization over Gaussian networks. IEEE Trans. Autom. Control, under review (2013)
444. Zaidi, S.A., Yüksel, S., Oechtering, T., Skoglund, M.: On optimal policies for control and estimation over a Gaussian relay channel. In: Proceedings of the IEEE Conference on Decision and Control, Dec 2011
445. Zaidi, S.A., Oechtering, T., Yüksel S., Skoglund, M.: Stabilization and Control over Gaussian Networks, In: Como, G., Bernhardsson, B., Rantzer, A. (eds.) Information and Control in Networks. Springer, New York (2013)
446. Zigangirov, K.S.: Upper bounds for the error probability for channels with feedback. Problemy Peredachi Informatsii **6**, 87–92 (1970)
447. Ziv, J.: The behavior of analog communication systems. IEEE Trans. Inf. Theor. **16**, 587–594 (1970)

Index

S. Yüksel and T. Başar, *Stochastic Networked Control Systems*, Systems & Control: Foundations & Applications, DOI 10.1007/978-1-4614-7085-4,