Lecture Notes in Control and Information Sciences 450

Giacomo Como
Bo Bernhardsson
Anders Rantzer *Editors*

Information and Control in Networks

LNCIS

Springer

Lecture Notes in Control and Information Sciences 450

Giacomo Como • Bo Bernhardsson •
Anders Rantzer
Editors

Information and Control in Networks

Springer

Editors
Giacomo Como
Department of Automatic Control
Lund University
Lund, Sweden

Anders Rantzer
Department of Automatic Control
Lund University
Lund, Sweden

Bo Bernhardsson
Department of Automatic Control
Lund University
Lund, Sweden

ISSN 0170-8643 ISSN 1610-7411 (electronic)
Lecture Notes in Control and Information Sciences
ISBN 978-3-319-02149-2 ISBN 978-3-319-02150-8 (eBook)
DOI 10.1007/978-3-319-02150-8
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013953916

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In the control of complex networked systems, a central role is played by information. System dynamics and information flows evolve in an intertwined way. While Control Theory and Information Theory have traditionally developed independently from each other, the need for a convergence of the two has strongly emerged in the last 15 years, and is now a very active research field. In addition to efforts in Control and Information Theory, strong research is witnessed in such diverse fields as Computer Science, Mathematics, and Statistics. Potential applications of the theory include remote robot control, automated highway navigation using wireless sensor systems, automatic control for unmanned aerial vehicles, and design of critical energy, transportation, and health care systems, and more.

This book collects contributions from some of the world-leading researchers in the area who have gathered in Lund on October 17–19, 2012 for the LCCC Workshop in Information and Control in Networks. The workshop was held in the middle of a five-weeks-long focus period on the same theme which has created exciting cross-fertilization and new ideas. (Please, refer to the webpage http://www.lccc.lth.se/index.php?page=workshop2012-10 for more detailed information.)

The book is organized in three parts, consisting of three chapters each. The first part collects contributions dealing with the problem of stabilizability of dynamical systems with communication constraints in the feedback loop. In particular, Franceschetti and Minero provide a survey of several results that have appeared in the last 15 years on this topic; Zaidi et al. study the stabilization problem of a linear system driven by Gaussian noise with a Gaussian relay channel in the feedback loop; Cardoso de Castro et al. design energy-efficient radio-mode switching controllers for stabilization over noisy channels.

The second part includes chapters exploring the fundamental relationships between Control and Information from three different angles: Nayyar et al. discuss a novel ‘common information approach’ solving decentralized stochastic control problems; Asnani et al. present a unified framework for deriving the relations between information and estimation in the presence of feedback; Yüksel studies the design problem of information channels for stabilization and optimization in networked control.

The third and final part of the book includes contributions dealing with different information dynamics over networks: Nair investigates the problem of characterizing those noiseless network topologies for which the information transmission problem is equivalent to the existence of an admissible multi-commodity flow; Elia et al. adopt a dynamical system viewpoint in order to tackle the problem of distributed computation over unreliable networks whereby transmission links are subject to both additive noise and packet drops; Sedghi and Jonckheere apply techniques from Markov Random Fields to deal with the problem of detection of faults or malicious attacks in the smart power grid.

Any attempt to sketch a brief summary of the historical development of the subject here would necessary fail to do justice to its breadth and complexity. We rather address the reader to the reference lists of the different chapters, and in particular in the ones by Franceschetti and Minero, Nayyar et al., for an overview of the relevant literature. A complete list of the titles of the LCCC workshop's seminars is also included in the next two pages, with the aim of giving the reader an idea of the full range of themes discussed there, including those who have not resulted in a contribution to this book.

Lund, Sweden
June 2013

Giacomo Como
Bo Bernhardsson
Anders Rantzer

LCCC Workshop on Information and Control in Networks

Lund, October 17–19, 2012

Program

(Slides available at http://www.lccc.lth.se/index.php?page=workshop1210program)

Massimo Franceschetti, UC San Diego, USA
The Complex Braid of Communication and Control

Girish Nair, University of Melbourne, Australia
Elements of a Nonstochastic Information Theory

Anant Sahai, UC Berkeley, USA
Implicit Communication in Decentralized Control Systems

Lars Rasmussen, KTH Royal Institute of Technology, Sweden
Asymptotic Anytime Reliability of LDPC Convolutional Codes

Tsachy Weissman, Stanford University, USA
Relations Between Information and Estimation in the Presence of Feedback

Edmond Jonckheere, USC, USA
On the Conditional Mutual Information in Gauss–Markov Structured Grids

Sandro Zampieri, Università di Padova, Italy
An Algorithm for Cameras' Network's Calibration

Carlos Canudas-de-Wit, CNRS Grenoble, France
Optimal Radio Mode Control for Intelligent Sensor Nodes in NCS

Gerhard Kramer, Technische Universität München, Germany
Feedback for Channels with In-Block Memory

Amos Lapidoth, ETH Zurich
Some Problems Are Easier with Feedback

Demosthenis Teneketzis, University of Michigan, USA
A Survey of the Common-Information Approach to Decentralized Stochastic Control

Nuno Martins, University of Maryland, USA
Necessary and Sufficient Conditions for Stabilizability in Decentralized Control and Estimation

Mikael Skoglund, KTH Royal Institute of Technology, Sweden
Stabilization and Control over Gaussian Networks

Serdar Yüksel, Queen's University, Canada
Optimal Design of Information Channels in Networked Control

Michelle Effros, CalTech, USA
On Networks, Capacities, and Controls

Sekhar Tatikonda, Yale University, USA
Sparse Regression Codes

Nicola Elia, Iowa State University, USA
Computing over Unreliable Communication Networks

Dragan Obradovic, Siemens AG, Munich, Germany
Precise Clock Synchronization for Industrial Automation and Other Networked Applications

Murat Arcak, UC Berkeley, USA
Pattern Formation by Lateral Inhibition: A Case Study in Networked Dynamical Systems

Fabio Fagnani, Politecnico di Torino, Italy
Democracy and the Role of Minorities in Markov Chain Models

Peter Caines, Mc Gill University, Canada
A Mean-Field Games' Formulation of Network-Based Auction Dynamics

Laurent Lessard, Lund University, Sweden
Optimal Collaborative Control in the Absence of Communication

Giacomo Como, Lund University, Sweden
Resilient Distributed Routing in Dynamical Flow Networks

Sanjoy Mitter (chair), Massachusetts Institute of Technology, USA
Panel Discussion: *Open Problems and Future Challenges in Information and Control in Networks*

Acknowledgements

The contributions were presented at the LCCC Workshop on Information and Control held in Lund on October 17–19, 2012. LCCC is a Linnaeus center at Lund University funded by the Swedish Research Council: Ref. VR 2007-8646. The ten principal investigators of LCCC are from the Department of Automatic Control and the Department of Electrical and Information Technology.

The research vision of LCCC is to make fundamental contributions to a general theory and methodology for control of complex engineering systems. This includes scalable methods and tools for modeling, analysis and control synthesis, as well as reliable implementations using networked embedded systems. LCCC strives to maintain a leading role in a worldwide effort involving partners of many kinds.

The editors would like to thank the participants of the LCCC workshop and focus period for rewarding contributions to the workshop and to this book. Special thanks go to Prof. Sanjoy Mitter for his continuous and enthusiastic encouragement throughout the project, and to the other members of the scientific committee, Prof. Michelle Effros, Prof. Gerhard Kramer, and Prof. Nuno Martins, for helping organize the event and choosing the invited speakers. Finally, thanks are extended to Dr. Eva Westin and Mr. Leif Andersson and the Springer editors for valuable editorial help and advice.

Contents

Contributors

Himanshu Asnani Electrical Engineering Department, Stanford University, Stanford, CA, USA

Carlos Canudas de Wit Department of Automatic Control, GIPSA-Lab, CNRS, NeCS Team, Grenoble, France

Nicolas Cardoso de Castro INRIA Rhône-Alpes, NeCS Team, Grenoble, France

Nicola Elia Dept. of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA

Massimo Franceschetti Department of Electrical and Computer Engineering, University of California, San Diego, CA, USA

Federica Garin INRIA Rhône-Alpes, NeCS Team, Grenoble, France

Edmond Jonckheere Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

Aditya Mahajan Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada

Paolo Minero Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN, USA

Girish N. Nair Dept. Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC, Australia

Ashutosh Nayyar Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA

Tobias J. Oechtering KTH Royal Institute of Technology, Stockholm, Sweden

Hanie Sedghi Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

Mikael Skoglund KTH Royal Institute of Technology, Stockholm, Sweden

Demosthenis Teneketzis Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA

Andalam Satya Mohan Vamsi Dept. of Electrical and Computer Engineering, Iowa State University, Ames, IA, USA

Kartik Venkat Electrical Engineering Department, Stanford University, Stanford, CA, USA

Jing Wang Cummins Inc., Columbus, IN, USA

Tsachy Weissman Electrical Engineering Department, Stanford University, Stanford, CA, USA

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

Ali A. Zaidi KTH Royal Institute of Technology, Stockholm, Sweden

Part I
Control with Information Constraints

Next generation cyber-physical systems challenge the standard assumption of classical control theory that communication can be performed instantaneously, reliably, and with infinite precision. This has motivated the birth of a new chapter of control theory relating the controllability dynamical systems to the characteristics of the communication channels used in the feedback loop. The first part of the book collects three contributions in this field.

In Chap. 1, Franceschetti and Minero review a series of contributions at the intersection of information and control theories, briefly describing applications, sketching mathematical arguments, and illustrating in a tutorial style the main information-theoretic and control-theoretic tools used to derive these results. The authors also draw a connection between control over networks and some recent advancements in feedback communication, and mention some open problems related to error correcting codes for interactive communications.

In Chap. 2, Zaidi et al. provide an overview and some recent results on real-time communication and control over Gaussian channels. Their contribution is focused on the problem of remote stabilization of linear systems driven by Gaussian noise over Gaussian relay channels, and necessary and sufficient conditions are presented for mean-square stabilization which are tight in a certain class of settings. The authors construct optimal linear policies and investigate global optimality and sub-optimality of such policies in a variety of settings. They also consider the design of low-delay sensing and transmit schemes for real-time communication.

Chapter 3 by Cardoso de Castro et al. deals with energy efficiency of the radio chip of sensor nodes. The authors exploit the use of different radio-modes: various transmitting modes having a different transmission power, where increased power results in better transmission quality (fewer errors) but higher energy cost; and various non-transmitting modes which correspond to switching off only some components of the chip, and result in having higher energy cost than Sleep but faster/less costly transition to transmission. An event-based radio-mode switching policy is proposed to perform a trade-off between energy saving and performance of the control application.

Chapter 1
Elements of Information Theory for Networked Control Systems

Massimo Franceschetti and Paolo Minero

1.1 Introduction

Next generation cyber-physical systems [35] will integrate computing, communication, and control technologies, to respond to the increased societal need to build large-scale systems using digital technology and interacting with the physical world. These include energy systems where the generation, transmission, and distribution of energy is made more efficient through the integration of information technologies; transportation systems that integrate intelligent vehicles and intelligent infrastructures; and health care systems where medical devices have high degree of intelligence and interoperability, integrating wireless networking and sensing capabilities.

One of the fundamental characteristics of cyber-physical systems is that communication among computing and physical entities occurs over communication channels of limited bandwidth and is subject to interference and noise. This challenges the standard assumption of classical control theory that communication can be performed instantaneously, reliably, and with infinite precision, and leads to the development of a new theory of networked control systems (NCS) [7, 8, 24, 30].

This chapter complements the surveys [2, 50] that focus on the communication constraints imposed by the network on the ability to estimate and control dynamical systems. We describe in a tutorial style the main ideas and techniques that contributed shaping the field, with particular attention to the connections with Shannon's information theory. A compendium of additional related results can be found in the recent monograph [44], relating results to Kolmogorov's approach to information theory via the concept of topological entropy [1].

M. Franceschetti (✉)
Department of Electrical and Computer Engineering, University of California, San Diego, CA 92093, USA
e-mail: massimo@ece.ucsd.edu

P. Minero
Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA
e-mail: pminero@nd.edu

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_1,

We shall not repeat proofs here that are readily available in the literature, but rather concentrate on providing specific illustrative examples and on bridging between different results, with the objective of outlining the leitmotiv and the central theoretical issues underlying this research area. We also present some new results that were not mentioned in the above works, and draw attention to a recent approach, based on the theory of Markov jump linear systems (MJLS) [15], that can be used to derive in a unified way many earlier results obtained using different techniques. Finally, we give a perspective on the open problems that are the natural candidates for future research in the field.

The rest of the chapter is organized as follows. In the next section, we describe a standard model of NCS. In Sect. 1.3, we present a basic result on the data-rate required in the feedback loop to guarantee system's stabilization. This is an important point of contact between communication and control theories and can be written in various forms. These are illustrated in Sect. 1.4, along with their connections with different notions of information capacity and their associated reliability constraints. Section 1.5 focuses on challenges in the design of suitable error correcting codes to satisfy these constraints. Section 1.6 looks more closely at a specific communication channel, illustrating how the theory of MJLS can be used to recover in a unified way many of the results on system stabilization that appeared in the literature. Finally, Sect. 1.7 discusses some of the problems and challenges that lay ahead.

1.2 Networked Control Systems

The block diagram of a typical NCS is depicted in Fig. 1.1. The state of a dynamical system evolves over time according to deterministic plant dynamics, possibly affected by stochastic disturbances. Sensors feed back the plant's output to a controller over a digital communication channel. The control action is then sent back to the plant over another digital communication channel for actuation. Communication is affected by noise, and the channel has limited bandwidth as it may be shared among different components in a network setting. This limits the amount of information that can be transferred in the feedback loop at each time step of the evolution of the system.

A natural mathematical abstraction of the above scenario considers the plant to be a discrete-time, linear, dynamical system, affected by additive disturbances

$$\begin{cases} x_{k+1} = Ax_k + Bu_k + v_k, & \text{(1.1a)} \\ y_k = Cx_k + w_k, & \text{(1.1b)} \end{cases}$$

where $k = 0, 1, \ldots$ is time, $x_k \in \mathbb{R}^d$ represents the state variable of the system, $u_k \in \mathbb{R}^m$ is the control input, $v_k \in \mathbb{R}^d$ is an additive disturbance, $y_k \in \mathbb{R}^p$ is the sensor measurement, $w_k \in \mathbb{R}^p$ is the measurement disturbance, and A, B, C are constant real matrices of matching dimensions. Standard conditions on (A, B) to

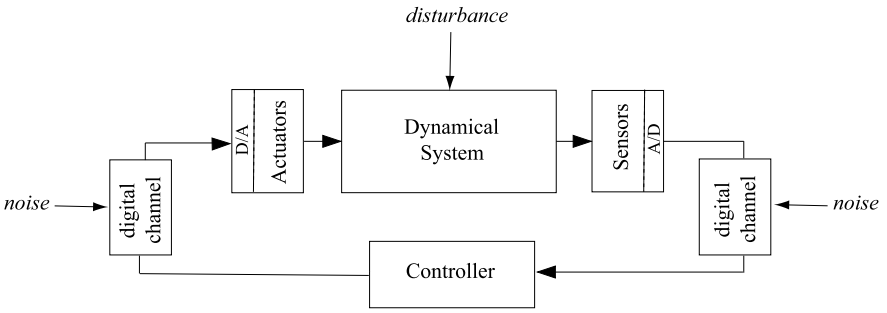


Fig. 1.1 Feedback loop model of a networked control system

be reachable, (C, A) observable, are added to make the problems considered well-posed.

In a first approximation, noise and bandwidth limitations in the communication channels can be captured by modeling the channels as "bit pipes" capable of transmitting only a fixed number r of bits in each time slot of the system's evolution. In this way, each channel can represent a network connection with a limited available bit-rate. This approach was originally proposed in [10] in the context of linear quadratic Gaussian (LQG) control of *stable* dynamical systems. In this case, by sending to the controller a quantized version of the innovation step of the minimum variance estimator, it was shown that the separation principle between estimation and control holds, and the optimal controller is a linear function of the state. Hence, the estimation problem is formally equivalent to the control one. Extensions of this result to LQG control of unstable systems and to other kind of channel models are highly dependent on the information pattern available to the sender and receiver and are explored in [26, 27, 56, 66]. In particular, when channel errors make the en-decoder uncertain of what the decoder received, the optimal controller is in general nonlinear [56], a result reminiscent of Witsenhausen's famous counter example [67].

1.3 The Data-Rate Theorem

For unstable systems under the bit-pipe communication model, when the control objective is to keep the state of the system bounded, or asymptotically drive it to zero, the control law is always a linear function of the state, and the central issue is to characterize the ability to perform a reliable estimate of the state at the receiving end of the communication channel. The central result in this case is the *data-rate theorem*. Loosely speaking, this states that the information rate r supported by the channel to keep the system stable must be large enough compared to the unstable modes of the system, so that it can compensate for the expansion of the state during the communication process. Namely,

$$r > \sum_{i \in \mathscr{U}} \log |\lambda_i|, \tag{1.2}$$

where the $\mathscr{U}$ is the set of indexes of the unstable eigenvalues of the open loop system and the logarithm is base 2. In the simple setting considered, the result is oblivious to the presence of two communication channels between the sensor and the controller and between the controller and the actuator. From the perspective of the system, the location of the controller is purely nominal. Since the key issue is communication of a reliable state estimate, the "bottleneck" link determines the effective data rate of the feedback loop. This intuitive reasoning can easily be made rigorous [49, Proposition 2.2]. The situation is, of course, different in the presence of channel uncertainties that, as already mentioned, make the problem highly dependent on the available information pattern at different points in the feedback loop. In this case, (1.2) should be modified using an appropriate notion of information capacity available in the feedback loop that depends, as we shall see, on the particular notion of stability employed, and on the characteristics of the disturbance.

The intuition behind the data-rate theorem is evident by considering the scalar case of (1.2)

$$2^r > |\lambda|, \tag{1.3}$$

and noticing that while the squared volume of the state of the open loop system increases by $|\lambda|^2$ at each time step, in closed loop this expansion is compensated by a factor 2^{-2r} due to the partitioning induced by the coder providing r bits of information through the communication channel. By imposing the product to be less than one, the result follows. Another interpretation arises if one identifies the logarithm of the right-hand side of (1.2) as a measure of the rate at which information is generated by the unstable plant, then the theorem essentially states that to achieve stability the channel must transport information as fast as it is produced.

Early incarnations of this fundamental result appeared in [5, 6, 68, 69] where it was shown that the state of an *undisturbed*, scalar, unstable plant with mode λ can be kept bounded if and only if the data rate in the feedback loop is at least $\log|\lambda|$ bits per unit time. While an improvement of the result from maintaining a bounded state to obtaining a state that asymptotically approaches zero cannot be achieved using a fixed quantizer [18], the works [12, 22, 37] showed that this can be obtained letting the encoder to have memory and using of an adaptive "zoom-in, zoom-out" strategy that adjusts the range of the quantizer so that it increases as the plant's state approaches the target and decreases if the state diverges from the target. This follows the intuition that in order to drive the state to zero, the quantizer's resolution should become higher close to the target.

In the presence of system disturbances, asymptotic stability can only be guaranteed within the range of the disturbances. Disturbances of unbounded support can drive the state arbitrarily far from zero. In this case, it is possible to guarantee stability only in a weaker, probabilistic sense. The work [65] proved the data-rate theorem for vector systems affected by unknown, but bounded disturbances, while the work [49] proved the data-rate theorem under the weaker condition of stochastic disturbances having unbounded support but a uniformly bounded higher moment, and using the probabilistic notion of mean-square stability. The work in [72] provides a related result by characterizing the limit for the second moment of the state in the infinite time horizon.

Since η-moment stability requires

$$\sup_{k\in\mathbb{N}} \mathbb{E}\big(\|X_k\|^\eta\big) < \infty, \tag{1.4}$$

the index η gives an estimate of the quality of the stability attainable: large stabilization errors occur more rarely as η increases and in this sense the system is better stabilized. One interpretation of the results in [49, 65] is that in order to achieve stability in a strong, almost deterministic sense ($\eta \to \infty$), one needs to assume almost surely bounded disturbances and bounded initial condition; on the other hand, relaxing the condition on stability to the weaker mean-square sense ($\eta = 2$), one can use the weaker assumption of bounded higher moments

$$\exists \varepsilon > 0: \quad \mathbb{E}\big(\|X_0\|^{2+\varepsilon}\big) < \infty, \qquad \sup_{k\in\mathbb{N}} \mathbb{E}\big(\|V_k\|^{2+\varepsilon}\big) < \infty, \qquad \sup_{k\in\mathbb{N}} \mathbb{E}\big(\|W_k\|^{2+\varepsilon}\big) < \infty. \tag{1.5}$$

In short, better stability is guaranteed with better behaved disturbances, while "wild disturbances" can only guarantee second moment stability.

The strict necessity of the data-rate theorem is proven in the deterministic setting of bounded disturbances by a recursive argument using the Brunn–Minkowski inequality, which states that the effective radius of the union of two sets is greater than the sum of their effective radii. In the stochastic setting, it is proven using the stochastic counterpart of the inequality, namely the entropy power inequality of information theory which states that the effective variance (entropy power) of the sum of two independent random variables is greater than the sum of their effective variances. The similarity between these two tools is well documented in [14]. In the stochastic case, it is required that the disturbances and the initial state have finite differential entropy.

The difficulty in proving the sufficiency of the data-rate theorem in the unbounded support case is due to the uncertainty about the state that cannot be confined in any bounded interval. This is overcome by using an adaptive quantizer depicted in Fig. 1.2 whose number of levels N depends on the rate process and whose resolution exponentially increases near the origin and diverges far from it, so that it can avoid saturation. The constant ξ depends on the statistics of the disturbance and it is used to recursively split the open semi-infinite intervals on the real axis into two, while every other finite interval is recursively divided in half. The main idea is then to divide time into cycles of length τ and at the beginning of each cycle quantize the estimated state using $N = 2^{R\tau}$ levels. Using this strategy, it can be shown that the estimated state satisfies a recurrence of the type

$$\mathbb{E}\big(\|X_{k\tau}\|^2\big) \le c_1 \left(\frac{|\lambda|^2}{2^{2R}}\right)^\tau \mathbb{E}\big(\|X_{(k-1)\tau}\|^2\big) + c_2, \tag{1.6}$$

where c_1 and c_2 are constants. This converges in virtue of (1.2) and by choosing τ large enough. In practice, the strategy allows the system to evolve in open loop for τ time steps and then applies a sufficiently refined control input that makes the state decrease at an exponential rate higher than the exponential divergence rate of the system.

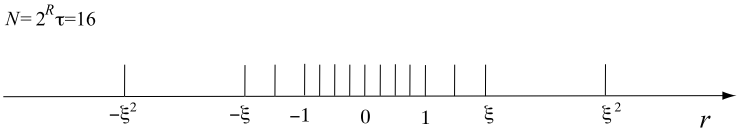


Fig. 1.2 Adaptive quantizer used to avoid saturation due to unbounded disturbances

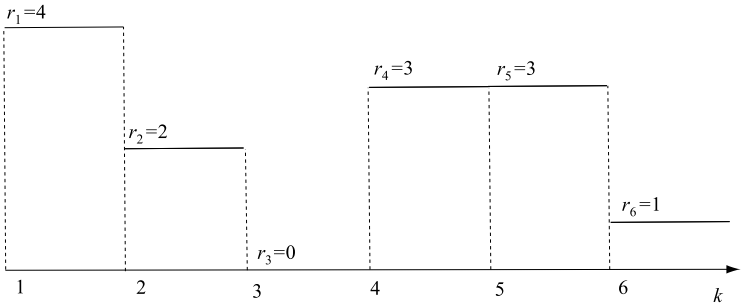


Fig. 1.3 Example of a realization of a stochastic rate channel R_k

1.4 Stochastic Time-Varying Channels

1.4.1 Stochastic Rate Channel

A different set of extensions concern the stochastic variability of the channel depicted in Fig. 1.3. This can be a first-order approximation of a wireless communication channel where the rate varies randomly in a slotted fashion. When the channel rate varies randomly with time in an independent, identically distributed (i.i.d.) fashion $R_k \sim R$ and there is causal knowledge of the rate process at both ends of the communication channel, the data-rate theorem for second moment stability in the scalar case becomes

$$|\lambda|^2 \mathbb{E}\left(2^{-2R}\right) < 1. \tag{1.7}$$

The work [39] proves the result for scalar systems with bounded disturbances and also provides the extension to η-moment stability

$$|\lambda|^\eta \mathbb{E}\left(2^{-\eta R}\right) < 1. \tag{1.8}$$

The intuition that to keep the state bounded it is required to balance the expansion of the state variable of the unstable system with the contraction provided by the received information bits still holds. The contraction rate is now a random variable, whose η-moment trades off the η-power of the unstable mode.

The work [46] proves the result for unbounded disturbances and second moment stability, and also provides necessary and sufficient conditions for vector systems that are tight in some special cases. The tools required to prove these results are the

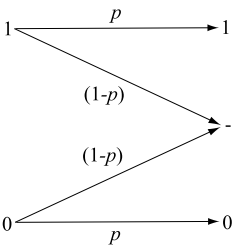


Fig. 1.4 The binary erasure channel

same as the ones described in the previous section. The additional complication due to the time-varying nature of the channel in the unbounded support case is solved using the idea of *successive refinements*. Namely, at the beginning of each cycle of duration τ the quantizer sends an initial estimate of the state using the quantizer depicted in Fig. 1.2, with a resolution dictated by the current value of the rate. In the remaining part of the cycle, the initial estimate is refined using the appropriate quantizer resolution allowed by the channel at each step. The refined state is then used for control at the end of the cycle. Notice that in this case the number of bits per cycle is a random variable dependent on the rate process and the mean square of the state is with respect to both the channel variations and the system disturbances.

The difficulties associated with the vector extension amount to the design of a bit allocation algorithm that dynamically allocates the available rate to the different unstable modes of the plant. The work [46] solves the problem using time-sharing techniques reminiscent of the ones developed in the context of network information theory for the multiple access channel [19]. Some extensions showing the tightness of the construction for some specific class of vector systems are provided in [70].

The stochastic rate channel includes the erasure channel as a special case that corresponds to the rate distribution

$$\begin{cases} \mathbb{P}(R = r) = p, \\ \mathbb{P}(R = 0) = 1 - p. \end{cases} \tag{1.9}$$

This reduces, for $r = 1$, to the binary erasure channel depicted in Fig. 1.4 and, for $r \to \infty$, to the continuous intermittent channel. We explore these reductions in more detail in the next section.

In real networks, many channels exhibit correlations over time. When the rate process follows a two-state Markov chain that corresponds to an erasure channel with two-state memory called the Gilbert–Elliott channel and depicted in Fig. 1.5, the data-rate theorem for mean-square stability in the scalar case with unbounded disturbances becomes [71]

$$r > \frac{1}{2} \log \mathbb{E}\big(|\lambda|^{2T}\big), \tag{1.10}$$

where T is the excursion time of state r. A more general result is provided in [17] that models the time-varying rate of the channel as an arbitrary time-invariant, positive-recurrent Markov chain of n states. This allows arbitrary temporal correlations of the channel variations and includes all previous models mentioned above,

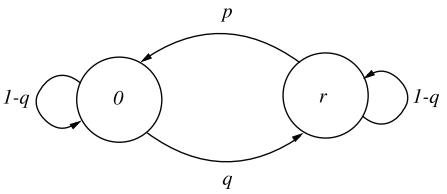


Fig. 1.5 The r-bit erasure channel with two-state memory (Gilbert–Elliott channel)

including extensions to the vector case. The technique used to provide this extension is based on the theory of MJLS.

In the scalar case, it is shown that stabilizing the system is equivalent to stabilizing

$$z_{k+1} = \frac{\lambda}{2^{r_k}} z_k + c, \tag{1.11}$$

where $z_k \in \mathbb{R}$ with $z_0 < \infty$, $c > 0$, $\{R_k\}_{k\geq 0}$ is the Markov rate process whose evolution through one time step is described by the transition probabilities

$$p_{ij} = \mathbb{P}\{R_{k+1} = r_j | R_k = r_i\}, \tag{1.12}$$

for all $k \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$. This equivalent MJLS describes the stochastic evolution of the estimation error $\|x_k - \hat{x}_k\|$ at the decoder, which at every time step k increases by λ because of the system dynamics, and is reduced by 2^{R_k} because of the information sent across the channel. A tight condition for second-moment stability is then expressed in terms of the spectral radius of an augmented matrix describing the dynamics of the second moment of this MJLS.

Letting H be the $n \times n$ matrix with elements

$$h_{ij} = \frac{p_{ij}}{2^{2r_j}}, \tag{1.13}$$

with spectral radius $\rho(H)$, the data-rate theorem becomes

$$|\lambda|^2 < \frac{1}{\rho(H)}. \tag{1.14}$$

A similar approach provides stability conditions for the case of vector systems. Necessary conditions use the idea of a "genie"-aided proof. First, it is assumed that a genie helps the channel decoder by stabilizing a subset of the unstable states. Then, the stability of the reduced vector system is related to the one of a scalar MJLS whose evolution depends on the remaining unstable modes. By considering all possible subsets of unstable modes, a family of conditions is obtained that relate the degree of instability of the system to the parameters governing the rate process. On the other hand, a sufficient condition for mean-square stability is given using a control scheme in which each unstable component of the system is quantized using a separate scalar quantizer. A bit allocation function determines how the bits available for communication over the Markov feedback channel are distributed among

the various unstable sub-systems. Given a bit allocation function, the sufficient condition is then given as the intersection of the stability conditions for the scalar jump linear systems that describe the evolution of the estimation error for each unstable mode.

The data-rate theorem for general Markovian rates presented in [17] recovers all results in [39, 46, 49, 65, 71] for constant, i.i.d., or two-state Markov data rates, with bounded or unbounded disturbances, in the scalar or vector cases. In addition, it also recovers results for the intermittent continuous channel and for the erasure channel, as discussed next. We discuss the techniques used to derive the results using the theory of MJLS in more detail in Sect. 1.6.

1.4.2 Intermittent Channel

The study of the intermittent continuous channel for estimation of the state of a dynamical system first initiated in [48]. The study of this channel was boosted in more recent times by the paper [61] in the context of Kalman filtering with intermittent observations. This work was inspired by computer networks in which packets can be dropped randomly and are sufficiently long that can be thought as representing real, continuous values. The analysis does not involve quantization, but only erasures occurring at each time step of the evolution of the system. Hence, the system in Fig. 1.1 is observed "intermittently", through an analog, rather than digital channel, and y_k in (1.1a), (1.1b) can be lost, with some probability, at each time step k. Similar to the data-rate theorem, it is of interest to characterize the *critical packet loss probability*, defined in [61], above which the mean-square estimation error remains bounded and below which it grows unbounded. This threshold value depends, once again, on the unstable modes of the system. Extensions providing large deviation bounds on the error covariance and conditions on its weak convergence to a stationary distribution are given in [47, 59, 62].

The model is easily extended to stabilization and control by considering an intermittent continuous channel also between the controller and the actuator. The work [56] considers LQG control over i.i.d. packet dropping links and shows that in the presence of acknowledgement of received packets the separation between estimation and control holds and the optimal controller is a linear function of the state. On the other hand, when there is uncertainty regarding the delivery of the packet, the optimal control is in general nonlinear. Similar results in the slightly more restrictive setting of the system being fully observable and the disturbance affecting only the system and not the observation, also appear in [32]. The critical role of the available information pattern on the optimal control is well known [67] and is further explored for stochastic rate channel models in [66].

The critical packet loss probability for mean-square stabilization is characterized in [26], under the assumption of i.i.d. erasures, and in [28] in the case of Markov erasures. The work [21] shows that such critical packet loss probabilities can be obtained as a solution of a robust control synthesis problem. These results can also

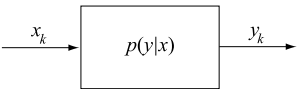


Fig. 1.6 The discrete memoryless channel

be obtained from the stochastic rate channel model, considering the erasure channel in (1.9) and letting $r \to \infty$. An easy derivation of the critical packet loss probability for stabilization is obtained in the scalar case by evaluating the expectation in (1.7), immediately yielding the result in [26]

$$p < \frac{1}{\lambda^2}. \tag{1.15}$$

Similarly, evaluating the condition in [71] for the Gilbert–Elliott channel as $r \to \infty$, one recovers the critical probability for the two-state Markov model of [28]. The works [17, 46] give matching reductions for the vector case as well. The latter of these works considers the most general channel model described so far, being an arbitrary Markov chain of n states, where r can be as low as zero (erasure) and as high as ∞ (continuous channel).

1.4.3 Discrete Memoryless Channels

Information theory treats the communication channel as a stochastic system described by the conditional probability distribution of the channel output under the given input. Figure 1.6 gives a visual representation of this information-theoretic model for the discrete memoryless channel (DMC).

In this context, the Shannon capacity of the channel is the supremum of the achievable rates of transmissions with an arbitrarily small error probability. It follows that the erasure channel of bit-rate r described previously is a special case of the DMC and has Shannon capacity [16]

$$C = (1 - p)r. \tag{1.16}$$

In the presence of system disturbances, for the erasure channel it follows from (1.7) that to ensure second moment stability a necessary and sufficient condition is

$$|\lambda|^2\left(2^{-2r}(1 - p) + p\right) < 1. \tag{1.17}$$

Comparing (1.16) with (1.17), it is evident that the Shannon capacity does not capture the ability to stabilize the system: not only the left-hand side of (1.17) is different from (1.16), but as $r \to \infty$ the Shannon capacity of the channel grows unboundedly, while the data-rate condition for stabilization reduces to (1.15) and critically depends on the erasure probability. Despite the infinite channel capacity, the system may be unstable when the erasure probability is high.

The reason for the insufficiency of the of Shannon capacity to characterize the trade-off between communication and information rate production of a dynamical

system lies in its operational definition. Roughly speaking, the notion of Shannon capacity implies that the message is encoded into a finite length codeword that is then transmitted over the channel. The message is communicated reliably only *asymptotically*, as the length of the codeword transmitted over the channel increases. The probability of decoding the wrong codeword is never zero, but it approaches zero as the length of the code increases. This asymptotic notion clashes with the dynamic nature of the system. A very large Shannon capacity can be useless from the system's perspective if it cannot be used in time for control. As argued at the end of Sect. 1.3, the system requires to receive without error a *sufficiently refined* control signal *every time* τ that makes the state decrease by a factor exponential in τ. The ability to receive a control input without error in a given time interval can be characterized in a classical information-theoretic setting using the notion of error exponent. However, for the control signal to be effective it must also be appropriate to the current state of the system. The state depends on the history of whether previous codewords were decoded correctly or not, since decoding the wrong codeword implies applying a wrong signal and driving the system away from the stability. In essence, this problem is an example of *interactive communication*, where two-way communication occurs through the feedback loop between the plant and the controller to stabilize the system. Error correcting codes developed in this context have a natural tree structure representing past history [51, 57] and are natural candidates to be used for control over channels with errors. They satisfy more stringent reliability constraints than the ones required to achieve Shannon capacity and can be used, as we shall see in Sect. 1.5, to obtain moment stabilization over the DMC.

Alternative notions of capacity have been proposed to capture the hard reliability constraints dictated by the control problem. The zero-error capacity C_0 was also introduced by Shannon [58] and considers the maximum data rate that can be communicated over the channel with no error. Assuming that the encoder knows the channel output perfectly, this notion of capacity can be used to obtain a data-rate theorem for systems with bounded disturbances with probability one in the form [43]

$$C_0 \gtrsim \sum_{i \in \mathscr{U}} \log |\lambda_i|, \tag{1.18}$$

where we have used the symbol $\gtrsim$ to indicate that the inequality is strict for the sufficient but not for the necessary condition. It was noted in [43] that even if a feedback channel from decoder to encoder is not available, in the absence of bounded external disturbances "virtual feedback" from decoder to encoder can always be established because the controller affects the plant's motion in a deterministic way and the sensor observes such motion. The controller can then encode its message depending on the observed state motion. For this reason, it is customary in the literature to assume the presence of communication feedback. This assumption is particularly important in the case of (1.18) because, unlike in the classical Shannon capacity, the zero-error capacity of the DMC increases in the presence of feedback.

The insufficiency of classical Shannon capacity to describe stabilization with probability one in the presence of disturbances over erasure channels was first

pointed out in [41], which led to the zero-error capacity framework of [43]. Unfortunately, the zero-error capacity (with or without feedback) of most practical channels (including the erasure channel) is zero [36], which implies that unstable systems cannot keep a bounded state with probability one when controlled over such channels. In practice, a long sequence of decoding errors always arises with probability one, and the small unknown disturbances that accumulate in this long time interval can always drive the system state without bound.

The situation drastically changes for *undisturbed* systems. In this case, the classical Shannon capacity C can be used to derive a data-rate theorem with probability one in the form [42]

$$C \gtrsim \sum_{i \in \mathscr{U}} \log |\lambda_i|. \tag{1.19}$$

This result was proven for the special case of the erasure channel in [64] and in the more general form for the DMC in [42].

Zero-error capacity and Shannon capacity provide data-rate theorems for plants with and without disturbances, respectively, over the DMC. They both require the strong notion of keeping the state bounded with probability one. Another notion of capacity arises by relaxing the constraint on stabilization with probability one to the weaker constraint of moment stability (1.4) that we used to describe stabilization over stochastic rate channels with unbounded system disturbances. In this case, the data-rate theorem can be written in terms of a parametric notion of channel capacity called *anytime capacity* [52]. Consider a system for information transmission that allows the time for processing the received codeword at the decoder to be infinite, and improves the reliability as time progresses. More precisely, at each step k in the evolution of the plant a new message m_k of r bits is generated that must be sent over the channel. The coder sends a bit over the channel at each k and the decoder upon reception of the new bit updates the estimates for all messages up to time k. It follows that at time k messages

$$m_0, m_1, \ldots, m_k$$

are considered for estimation, while estimates

$$\hat{m}_{0|k}, \hat{m}_{1|k}, \ldots, \hat{m}_{k|k}$$

are constructed, given all the bits received up to time k. Hence, the processing operation for any message m_i continues indefinitely for all $k \geq i$. A reliability level α is achieved in the given transmission system if for all k the probability that there exists at least one message in the past whose estimate is incorrect decreases α-exponentially with the number of bits received, namely

$$\mathbb{P}\big((\hat{M}_{0|k}, \ldots, \hat{M}_{d|k}) \neq (M_0, \ldots, M_d)\big) = O\big(2^{-\alpha d}\big) \quad \text{for all } d \leq k. \tag{1.20}$$

The described communication system is then characterized by a rate–reliability pair (r, α). It turns out that the ability to stabilize a dynamical system depends on the

ability to construct such a communication system, in terms of achievable coding and decoding schemes, with a given rate–reliability constraints.

Let the supremum of the rate r that can be achieved with reliability α be the α-anytime capacity $C_A(\alpha)$ of a given DMC with channel feedback. The necessary and sufficient condition of the data-rate theorem for η-moment stabilization of a scalar system with bounded disturbances and in the presence of channel output feedback is [53]

$$C_A\big(\eta \log|\lambda| + \varepsilon\big) \gtrsim \log|\lambda|. \tag{1.21}$$

Extensions to vector systems appear in preprint form in [54].

The anytime capacity has been introduced as an intermediate quantity between the hard notion of zero-error capacity and the soft notion of Shannon capacity. Not surprisingly, we have

$$C_0 \le C_A(\alpha) \le C, \tag{1.22}$$

and in the limiting cases

$$C_A\big(0^+\big) = C, \qquad C_A(\infty) = C_0. \tag{1.23}$$

Zero-error capacity requires transmission without error. Shannon capacity requires the decoding error go to zero with the length of the code. In the presence of disturbances, only the zero-error capacity can guarantee the almost sure stability of the system. The anytime capacity requires transmission with codeword reliability increasing exponentially in the delay of the single received bit. For scalar systems in presence of bounded disturbances, it is able to characterize the ability to stabilize the system in the weaker η-moment sense [53].

Unfortunately, the anytime capacity can be computed only for the special cases of the erasure channel and the additive white Gaussian noise channel with input power constraint, and in both of these cases it provides data-rate theorems that can also be derived directly in a more classical setting. For the r-bit erasure channel with feedback, we have

$$C_A(\alpha) = \frac{r\alpha}{\alpha + \log[(1-p)(1-2^{\alpha}p)^{-1}]}. \tag{1.24}$$

Substituting (1.24) into (1.21), we obtain after some algebra

$$|\lambda|^{\eta}\big(2^{-\eta r}(1-p) + p\big) \lesssim 1. \tag{1.25}$$

Comparing (1.25) with (1.17), it follows that (1.25) is consistent with the result for the stochastic rate channel in [17], which, in fact, gives a stronger version of the anytime capacity data-rate theorem for the case of the erasure channel with feedback, providing a single (necessary and sufficient) strict inequality condition for second moment stability. Furthermore, it also extends the result for this particular channel to disturbances with unbounded support.

For the additive white Gaussian noise channel with input power constraint, the anytime capacity is independent of the reliability level α and it coincides with the

Shannon capacity. In this case, the data-rate theorem can be given in terms of signal-to-noise ratio and available bandwidth [11, 25, 66].

The anytime capacity of more general channel models remains unknown. In addition, there may be cases in which the output of the noisy channel may not be available at the encoder and is impracticable to use the plant to signal from the decoder to the encoder. In this case, it is only known that the anytime capacity of a DMC without feedback is lower bounded by the exponent β of the error probability of block codes; namely, for any rate $r < C$ we have

$$C_A\big(\beta(r)\log_2 e\big) \geq r \log_2 e. \tag{1.26}$$

The work [53] proposes an ingenious control scheme to achieve (1.26) based on the idea of random binning: the observer maps to state using a time-varying randomly labeled lattice quantizer and outputs a random label for the bin index; the controller, on the other hand, makes use of the common randomness used to select the random bin labels to decode the quantized state value. This proof technique, however, only applies to plants with bounded disturbances.

Despite these shortcomings, the anytime capacity has been influential in the definition of the reliability constraints for the coding–decoding schemes that can achieve moment stabilization of linear systems in the presence of bounded disturbances, thus providing inspiration for further research in coding [13, 51, 60, 63].

1.4.4 Additive Gaussian channels

The additive white Gaussian noise communication channel with power constraint P is defined as the system

$$y_k = x_k + z_k, \tag{1.27}$$

where z_k is the realization of an i.i.d. Gaussian process with zero mean and variance σ^2, and the input is constrained by

$$\mathbb{E}\big(X_k^2\big) \leq P, \quad \forall k. \tag{1.28}$$

The Shannon capacity of this channel is perhaps the most notorious formula in information theory

$$C = \frac{1}{2}\log\big(1 + P/\sigma^2\big). \tag{1.29}$$

In this case, the data-rate theorem for second moment stabilization becomes [11, 25]

$$\frac{P}{\sigma^2} > \prod_{i \in \mathscr{U}} |\lambda_i|^2 - 1, \tag{1.30}$$

that is equivalent to

$$C > \sum_{i \in \mathscr{U}} \log |\lambda_i|. \tag{1.31}$$

The work in [11] also shows that stabilization can be achieved, provided (1.31) holds, using a linear controller with constant gain, if the system's output sent to the controller consists of the entire state vector. If the output consists only of a linear combination of state elements, then the required signal-to-noise ratio for stabilization using linear constant feedback exceeds the bound in (1.30), unless the plant is minimum phase. The work in [25] also shows that (1.31) is also required for second moment stability using nonlinear, time-varying control and provides an explicit lower bound on the second moment of the state that diverges as one approaches the data-rate capacity threshold. Earlier incarnation of these results go back to [66], with slightly stronger assumptions on the available information pattern, and to [20] that connected the recursive capacity-achieving scheme in [55] for the AWGN with feedback to the stabilization problem of scalar systems over AWGN channels.

Extensions to additive *colored* Gaussian channels (ACGC) provide additional connections between the ability to stabilize dynamical systems and the *feedback capacity* C_F of the channel. This is defined as the capacity, in Shannon's sense, in the presence of an additional noiseless feedback link between the output and the input of the channel. While for the AWGN channel feedback does not improve capacity, for ACGC it does improve it. The feedback capacity of the first order moving average (MA1) additive Gaussian channel has been determined in [33] and for the general case of stationary ACGC in [34]. The work in [45] exploits the result in [33] to show that mean-square stabilization of an undisturbed minimum phase plant with a single unstable pole over a MA1 additive Gaussian channel is possible if and only if

$$C_F > \log |\lambda|. \tag{1.32}$$

The work in [3] exploits the result in [34] to show that the feedback capacity of the general stationary ACGC with power constraint P is

$$C_F = \sup_{\mathcal{L}} U, \tag{1.33}$$

where

$$U = \sum_{i \in \mathscr{U}} \log |\lambda_i| \tag{1.34}$$

and $\mathcal{L}$ is the set of all undisturbed (vector) linear systems that can be stabilized using a linear controller over the same additive Gaussian channel, with power constraint

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \left| T\left(e^{j\omega}\right) \right|^2 S_Z(\omega)\, d\omega \leq P, \tag{1.35}$$

Table 1.1 Summary of data-rate theorems for stabilization over noisy channels

Condition	Channel	Stabilization	Disturbance
$C \gtrsim U$	DMC	a.s.	0
$C_0 \gtrsim U$	DMC	a.s.	bounded
$C_A(\eta \log \lvert\lambda\rvert) \gtrsim \eta \log \lvert\lambda\rvert$	DMC	η-moment	bounded
$\lvert\lambda\rvert^2(2^{-2r}(1-p)+p) < 1$	Erasure	2nd moment	unbounded
$C > U$	AWGN	η-moment	unbounded
$C_F = \sup U$	ACGN	2nd moment	0

where $S_z(\omega)$ is the power spectral density of the noise, and T is the complementary sensitivity function of the system. This result shows that the maximum "tolerable instability" U of an LTI system with a given power constraint P, controlled by a linear controller over a general stationary Gaussian channel, corresponds to the feedback capacity of that channel subject to the same power constraint P. Hence, there is a natural duality between feedback stabilization and communication over the Gaussian channel. This duality can also be exploited to construct efficient communication schemes over the Gaussian channel with feedback in the context of network information theory, using control tools. This theme was first explored in [20] and later expanded in [4].

We provide a summary of the results for different noisy channels Table 1.1.

1.5 Error Correcting Codes for Control

Independent of research in stabilization and control, error correcting codes with exponential reliability constraints in the form of (1.20) were introduced in the context of interactive communication [57]. These codes possess a natural tree structure that can be used to maintain synchronization between the controller and system when communication occurs over noisy channels. Although it is not known whether tree codes are anytime capacity achieving, they can be used for stabilization of networked control systems when their rate-reliability parameters fall within a region needed for stabilization of the given system. We motivate them with the following example.

Consider the problem of tracking a scalar unstable process with dynamics

$$x_{k+1} = \lambda x_k + v_k, \tag{1.36}$$

with $\lambda > 1$. The initial condition and the additive disturbance are supposed to be random but bounded, i.e., $|X_0| \le \alpha$ and $|V_k| \le \beta$ for some $\alpha, \beta < \infty$. We consider the setup where a coder having access to the state communicates over a binary noisy channel to a decoder that wishes to track the state of the system. The objective is to

design a coder–decoder pair such that

$$\sup_k \mathbb{E}\big(|X_k - \hat{X}_k|^2\big) < \infty. \tag{1.37}$$

If the communication channel is noiseless and allows transmission without errors of r bits per unit of time, then we obtain the usual data-rate theorem in the form (1.3). The strategy used for estimation follows the one described in [65]. Let $\mathcal{U}_0 = [-\alpha, +\alpha]$ denote the set containing the initial condition x_0. At time $k = 0$, the coder and the decoder partition $\mathcal{U}_0$ into 2^r intervals $\mathcal{U}_0(1), \ldots, \mathcal{U}_0(2^r)$ of equal size. The coder communicates to the decoder the index m_0 of the interval $\mathcal{U}_0(m_0)$ containing the state, so the decoder can form a state estimate $\bar{x}_0$ as the midpoint of $\mathcal{U}_0(m_0)$. This construction implies

$$|x_0 - \bar{x}_0| \leq \alpha 2^{-r}$$

for any $x_0 \in \mathcal{U}_0$. Also, notice that x_1 is contained inside the set $\mathcal{U}_1 := \lambda \mathcal{U}_0(m_0) + [-\beta, +\beta]$, where the sum denotes the Minkowski sum of sets. This means that the same scheme can be used at time $k = 1$ to estimate the state x_1. Specifically, the coder and the decoder partition the set $\mathcal{U}_1$ into 2^r intervals $\mathcal{U}_1(1), \ldots, \mathcal{U}_1(2^r)$ of equal size, the coder transmits the index m_1 of the subinterval containing the state, and the decoder sets $\bar{x}_1$ equal to the midpoint of $\mathcal{U}_1(m_1)$, so that

$$|x_1 - \bar{x}_1| \leq \alpha \lambda 2^{-2r} + \beta 2^{-r}.$$

By iterating the same procedure k times, at time k the coder and the decoder agree that x_k belongs to a set $\mathcal{U}_k := \lambda \mathcal{U}_{k-1}(m_{k-1}) + [-\beta, +\beta]$. Then, the coder sends over the channel the index m_k of the subinterval $\mathcal{U}_k(m_k) \subseteq \mathcal{U}_k$ containing x_k and the decoder forms an estimate $\bar{x}_k$ as the midpoint of the uncertainty interval $\mathcal{U}_k(m_k)$. It can be shown by induction that

$$|x_k - \bar{x}_k| \leq \big(\lambda 2^{-r}\big)^k \alpha 2^{-r} + \beta 2^{-r} \sum_{j=0}^{k-1} \big(\lambda 2^{-r}\big)^{k-1-j}.$$

It follows that a sufficient condition for the estimation error at the decoder to remain bounded for all k coincides with (1.3).

Consider now the case of a noisy channel in which synchronism between coder and decoder can be lost in the event that the sequence $m_0, \ldots, m_k$ is not correctly decoded at the estimator. To prevent this, at every time k a channel encoder maps the sequence $m_0, \ldots, m_k$ into an r-bit channel input sequence $f_k(m_0, \ldots, m_k)$ that is transmitted over the channel. A channel decoder maps the received channel bits up to time k into an estimate $\hat{m}_{0|k}, \ldots, \hat{m}_{k|k}$ for the input sequence, which, in turn, is used to form the state estimate $\hat{x}_k$ as the midpoint of the interval $\mathcal{U}_k(\hat{m}_{k|k})$ which is formed by recursively partitioning $\lambda \mathcal{U}_j(\hat{m}_{j|k}) + [-\beta, \beta]$, $j = 0, \ldots, k-1$, into 2^r intervals.

If the index of the first wrong estimate at the decoder is $k - d$, that is, if $\hat{m}_{0|k} = m_0, \ldots, \hat{m}_{k-d-1|k} = m_{k-d-1}$ and $\hat{m}_{m-d|k} \neq m_{m-d}$, then the error between

the estimators at coder and decoder is

$$|\bar{x}_k - \hat{x}_k| = O\big(\lambda^d\big), \tag{1.38}$$

because the difference between the two estimates at time $k-d$ is amplified by λ at each iteration due to the expansion of the state process. It follows that the mean-square estimation error can be upper bounded as

$$\begin{aligned}\mathbb{E}\big(|X_k - \hat{X}_k|^2\big) &\le 2\mathbb{E}\big(|X_k - \bar{X}_k|^2\big) + 2\mathbb{E}\big(|\bar{X}_k - \hat{X}_k|^2\big) \\ &= O\left(\frac{\lambda^{2k}}{2^{2kr}} + \sum_{d=0}^{k-1} P_{d,k}\lambda^{2d}\right),\end{aligned} \tag{1.39}$$

where

$$P_{d,k} = P\{\hat{M}_{0|k} = M_0, \dots, \hat{M}_{k-d-1|k} = M_{k-d-1}, \hat{M}_{k-d|k} \neq M_{k-d}\},$$

denotes the probability that the index of the first wrong estimate at time k is $k-d$, $d = 0, 1, \dots, k$. Observe that (1.39) is obtained by separately bounding two terms, the first of which represents the mean-square estimation error under the assumption that the channel is noise free, that goes to zero if (1.3) is satisfied, while the second denotes the mean-square error between the estimator $\bar{x}_k$ available at the encoder and the estimator $\hat{x}_k$ available at the decoder, and is bounded provided $P_{d,k}$ decays fast enough as d grows. It follows that a sufficient condition for second moment stabilization is given by

$$\begin{cases} r \ge \log|\lambda|, & \text{(1.40a)} \\ P_{d,k} = O(2^{-2(\log|\lambda|+\varepsilon)d}) \quad \text{for all } d \le k, & \text{(1.40b)} \end{cases}$$

that corresponds to the sufficient condition given in (1.21) in terms of anytime capacity.

1.5.1 Tree Codes

The reliability condition imposed by (1.40a), (1.40b) is amenable to the following visual interpretation. First, notice that the coding–decoding scheme can be visualized on a tree of depth k, as depicted in Fig. 1.7, where the nodes at level i denote the uncertainty intervals $\mathcal{U}_j(1), \dots, \mathcal{U}_j(2^r)$, while the label on each branch denotes the r-bit sequence transmitted over the channel at each time instant. The codeword associated to a given path in the tree is given by the concatenation of the branch symbols along that path. The sequence $m_0, \dots, m_k$ determines the path in the tree followed up to time k by the encoder, while $\hat{m}_0, \dots, \hat{m}_k$ determines the path followed by the decoder. Then, (1.40a), (1.40b) implies that the uncertainty at the controller about

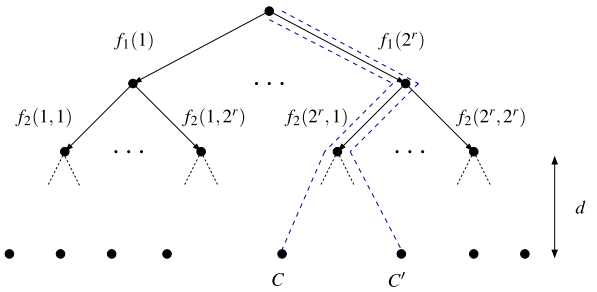


Fig. 1.7 Binary tree visualizing the evolution of the uncertainty set containing the initial condition. The coding–decoding scheme described in Sect. 1.5 can be visualized on this tree by labeling each branch with the symbols sent over the channel. The codeword associated to a given path is given by the concatenation of the branch symbols along that path

the path followed in the binary tree must decrease exponentially at rate $2(\log|\lambda|+\varepsilon)$ with the distance d from the bottom of the tree.

Tree codes and their maximum likelihood analysis were first introduced in [23], but finding explicit deterministic constructions of codes achieving a given rate-reliability pair (r,α) is still an important open problem. The work [57] applied the random coding argument in [23] to prove the existence of codes within a specific (r,α) region. The codes introduced in [57] are defined by the property that the Hamming distance between any two codewords associated with distinct paths of equal depth in the binary tree is proportional to the height from the bottom of the tree of the least common ancestor between the two paths. For example, the Hamming distance between the codewords C and C' illustrated in Fig. 1.7 should be proportional to h. This property on the minimum distance translates into different guarantees on the reliability of the code depending on the communication channel. The preprint [63] proves the existence with high probability of *linear* (r,α) tree codes, i.e., codes where the channel input sequence $f_k(m_0,\dots,m_k)$ transmitted over the channel at time k is a linear function of $m_0,\dots,m_k$. The (r,α) region of existence obtained in [63] is currently the largest known region of existence. An important open problem is to show the existence of (possibly nonlinear) $(2\log|\lambda|)$-reliable codes for any rate r greater than $\log|\lambda|$. This result would show that tree codes are anytime-capacity achieving and therefore they are both necessary and sufficient for moment stabilization of unstable scalar systems over noisy channels.

The argument in [57] relies on the probabilistic method and only ensures the existence of tree codes, not their explicit construction. A new class of codes with explicit constructions that are computationally efficient have been presented in [51], but they exhibit weaker reliability constraints that are only useful for stabilization of plants whose state space grows polynomially with time. The preprint [63] offers an explicit construction for the binary erasure channel that does not require causal knowledge of the erasure process, as it was assumed to derive the data-rate theorem in [17].

It is important to emphasize that explicit constructions require coding and decoding operations to be computationally efficient. One could, in principle, consider

using traditional convolutional codes developed in the context of wireless communication to stabilize dynamical systems [38]. These codes perform "on-line" encoding and decoding in which the estimate of the received message is refined as more bits are received within the *constraint length* window of the code. The constraint length is analogous to the block length of traditional block codes, but it allows incremental, on-line refinement of the received message estimate at the decoder. The error probability decreases exponentially with the constraint length of the code, thus providing the required reliability constraint. Unfortunately, the complexity of the construction increases with the constraint length and computationally efficient convolutional codes only exist for small constraint lengths. Convolutional codes are heavily used in mobile phones, where occasional errors translate in call drops or audio disturbances. In control applications, however, the accumulation of errors over long time periods resulting from finite constraint lengths would make them unsuitable for practical implementations as they would drive the system to instability.

1.6 Stochastic Time-Varying Rate: An In-Depth Look

We now provide a more rigorous treatment of the data-rate theorem for stochastic time-varying rate channels, with the objective of illustrating recently developed techniques based on the theory of MJLS that can be used to derive many of the results available in the literature. We follow the approach developed in [17]; however, we consider here the special case of a scalar system in which there are only system disturbances and no observation disturbances. This allows presenting simplified proofs that are considerably shorter, more easily accessible, and better suited to grasp the main ideas behind them.

Consider the special case of a scalar system with state feedback

$$\begin{cases} x_{k+1} = \lambda x_k + u_k + v_k, & \text{(1.41a)} \\ y_k = x_k, & \text{(1.41b)} \end{cases}$$

where $k = 0, 1, \ldots$ and $|\lambda| \geq 1$, and suppose that the following assumptions hold:

Assumption 1.1 The initial condition X_0 and the plant disturbance V_k, $k \geq 0$, are zero mean and have continuous probability density functions of finite differential entropy, so there exists a constant $\beta > 0$ such that $e^{2h(V_k)} \geq \beta$ for all k.

Assumption 1.2 The initial condition X_0 and the plant disturbance V_k, $k \geq 0$, have uniformly bounded $(2+\varepsilon)$th moments so there exists a constant $\alpha < \infty$ such that $\mathbb{E}(|V_k|^{2+\varepsilon}) \leq \alpha$ for all k.

We also assume that the sensor measurements y_k are transmitted from the state observer to the actuator over a noiseless digital communication link that at each time k allows transmission without errors of r_k bits. The rate sequence $r_0, r_1, \ldots$ is the

realization of a stochastic process $R_1, R_2, \ldots$, that is modeled as a homogeneous positive-recurrent Markov chain taking values in a finite subset of the nonnegative integers

$$\mathcal{R} = \{\bar{r}_1, \ldots, \bar{r}_n\},$$

and whose evolution through one time step is described by the transition probabilities (1.12), i.e.,

$$p_{ij} = \mathbb{P}\{R_{k+1} = \bar{r}_j | R_k = \bar{r}_i\}$$

for all $k \in \mathbb{N}$ and $i, j \in \{1, \ldots, n\}$. The rate process is independent of the other quantities describing the system and is causally known at observer and controller.

At each time k, a coding function (coder) $s_k = s_k(y_0, \ldots, y_k)$ maps all past and present measurements into the set $\{1, \ldots, 2^{r_k}\}$. The digital link is mathematically modeled as the identity function on the set $\{1, \ldots, 2^{r_k}\}$, so the symbols s_k are reliably transmitted without distortion. The received channel outputs are transformed by a decoding function (decoder) $u_k = \hat{x}_k(s_0, \ldots, s_k)$ that maps all past and present symbols sent over the digital link into a control input u_k that is sent to the plant.

The problem is to find conditions on the rate process and the system parameters to ensure stability of the closed loop system. We adopt the probabilistic notion of mean-square stability and require that

$$\sup_k \mathbb{E}\big[|X_k|^2\big] < \infty, \tag{1.42}$$

where the expectation is taken with respect to the rate process, the initial condition, and the plant disturbance.

We now proceed to establish necessary and sufficient conditions for mean-square stability of the scalar linear system (1.41a), (1.41a).

Theorem 1.1 *Let H be the $n \times n$ matrix with nonnegative real elements*

$$h_{ij} = \frac{1}{2^{2\bar{r}_j}} p_{ji} \tag{1.43}$$

for all $1 \le i, j \le n$. If Assumption 1.1 holds, then (1.41a), (1.41b) is mean-square stable only if

$$|\lambda|^2 < \frac{1}{\rho(H)}. \tag{1.44}$$

Conversely, if Assumption 1.2 holds, then there exists a coder–decoder pair that stabilizes (1.41a), (1.41b) is mean-square sense if (1.44) is satisfied.

If both Assumptions 1.1 and 1.2 hold, then Theorem (1.1) asserts that condition (1.44) is both necessary and sufficient to ensure mean-square stability. Application of Theorem 1.1 yields the following results as special cases.

(a) *Constant rate.* When the channel supports a constant rate, i.e., the rate process is identically equal to $\bar{r}$ at all times, the matrix H is equal to $1/2^{2\bar{r}}$ and thus (1.44) reduces to

$$\bar{r} > \log|\lambda|, \tag{1.45}$$

which is the condition given by the data-rate theorem in its basic formulation. It should be remarked that here $\bar{r}$ is restricted to be an integer, but this assumption can be relaxed by taking the approach followed in [49, 65], where the rate process is allowed to vary deterministically and $\bar{r}$ is defined as the infinite horizon time-average of the process.

(b) *Independent rate process*. Consider the special case of an independent rate process where each random variable R_k in the rate process is identically distributed as a random variable R with probability mass function $p_i = \mathbb{P}\{R = \bar{r}_i\}$, $\bar{r}_i \in \mathcal{R}$. It can be easily seen that in this case H reduces to a rank-one matrix with only one nonzero eigenvalue equal to $\sum_{i=1}^{n} p_i |\lambda|^2 2^{-2\bar{r}_i}$. Therefore, (1.44) specializes to

$$\begin{aligned} |\lambda|^2 \rho(H) &= \sum_{i=1}^{n} p_i |\lambda|^2 2^{-2\bar{r}_i} \\ &= \mathbb{E}\big(|\lambda|^2 2^{-2R}\big) < 1. \end{aligned} \tag{1.46}$$

The necessity and sufficiency of (1.46) for mean-square stability in this setting was established in [46]. This condition is also a special case of a result in [39], where it is established under the assumption of bounded disturbances that necessary and sufficient condition for ηth moment stability, i.e., boundedness of the ηth moment of the plant, is $\mathbb{E}(|\lambda|^\eta 2^{-\eta R}) < 1$.

(c) *Two-state Markov process.* Consider the special case of a rate process that randomly switches between two different states, state $\bar{r}_1$ and $\bar{r}_2$, and where the transition probabilities from $\bar{r}_1$ to $\bar{r}_2$ and from $\bar{r}_2$ to $\bar{r}_1$ are denoted by p and q, respectively. In this case, it is possible to relate the spectral radius of H to its determinant $\det(H)$ and its trace $\mathrm{tr}(H)$. Specifically, the condition in Theorem 1.1 reduces to

$$\frac{|\lambda|^2}{2}\Big(\mathrm{tr}(H) + \sqrt{\mathrm{tr}(H)^2 - 4\det(H)}\Big) < 1. \tag{1.47}$$

(d) *Erasure Channel.* Another special case that has been studied in the literature is the case of an erasure channel, which is further specialization of the two-state Markov process described above in the case where $\bar{r}_1 = 0$, $\bar{r}_2 = \bar{r}$. Necessary and sufficient conditions for mean-square stability under this channel model were established in [71], for the Markovian case, and in [46, 52] in the special case of independent rate process. If we further specialize to the case where $\bar{r} \to \infty$, then (1.47) recovers a result that was first established in [26].

1.6.1 Necessity

The following lemma states that if Assumption 1.1 is satisfied, then the second moment of the state in (1.41a), (1.41b) is lower bounded by the first moment of a MJLS whose dynamics depends on the Markov rate process $\{R_k\}$ and on the constant β defined in Assumption 1.1.

Lemma 1.1 *Let Assumption 1.1 hold. Then, for every $k = 0, 1, \ldots$ the second moment of X_k satisfies*

$$\mathbb{E}\big(|X_k|^2\big) > \frac{1}{2\pi e}\mathbb{E}(Z_k),$$

where $\{Z_k\}$ is a non-homogeneous MJLS with dynamics $z_0 = e^{2h(X_0)}$ and

$$z_{k+1} = \frac{|\lambda|^2}{2^{2R_k}} z_k + \beta, \quad k = 0, 1, \ldots. \tag{1.48}$$

Proof Let $S^k = \{S_0, \ldots, S_k\}$ denote the symbols transmitted over the digital link up to time k. By the law of total expectation and the maximum entropy theorem [19], we have

$$\begin{aligned}
\mathbb{E}\big(|X_{k+1}|^2\big) &= \sum_{s^k} P\big\{S^k = s^k\big\}\mathbb{E}\big(|X_{k+1}|^2 | S^k = s^k\big) \\
&= \frac{1}{2\pi e}\sum_{s^k} P\big\{S^k = s^k\big\} e^{\ln 2\pi e \mathbb{E}(|X_{k+1}|^2 | S^k = s^k)} \\
&\geq \frac{1}{2\pi e}\sum_{s^k} P\big\{S^k = s^k\big\} e^{\ln 2\pi e h(X_{k+1} | S^k = s^k)} \\
&=: \frac{1}{2\pi e}\,\mathbb{E}_{S^k}\big(e^{2h(X_{k+1}|S^k = s^k)}\big),
\end{aligned} \tag{1.49}$$

where the summation is over $s_i \in \mathcal{S} := \cup_{r\in\mathcal{R}}\{1, \ldots, 2^{2r}\}$, $0 \leq i \leq k$. It follows that the second moment of the state is lower bounded by the average entropy power of X_k conditional on S^k. From the translation invariance property of the differential entropy, the conditional version of entropy power inequality [19], and Assumption 1.1, it follows that

$$\begin{aligned}
\mathbb{E}_{S^k}\big(e^{2h(X_{k+1}|S^k = s^k)}\big) &= \mathbb{E}_{S^k}\big(e^{2h(\lambda X_k + \hat{x}(s^k) + V_k | S^k = s^k)}\big) \\
&\geq \mathbb{E}_{S^k}\big(e^{2h(\lambda X_k | S^k = s^k)}\big) + e^{2h(v_k)} \\
&\geq |\lambda|^2 \mathbb{E}_{S^k}\big(e^{2h(X_k | S^k = s^k)}\big) + \beta.
\end{aligned} \tag{1.50}$$

We can further lower bound (1.50) making use of a result proved in [46, 49], which states that for every time $k \geq 0$, $s^{k-1} \in \mathcal{S}^{k-1}$, and $r \in \mathcal{R}$

$$\sum_{s_k} P\left\{S^k = s_k \mid S^{k-1} = s^{k-1}, R_k = r\right\} e^{2h(X_k|S^k=s^k)} \geq \frac{1}{2^{2r}} e^{2h(X_k|S^{k-1}=s^{k-1})}, \tag{1.51}$$

where $S_{-1} := \emptyset$. By the tower rule of conditional expectation, it then follows that

$$\mathbb{E}_{S^k}\left(e^{2h(X_k|S^k=s^k)}\right) \geq \mathbb{E}_{S^{k-1},R_k}\left(\frac{1}{2^{2R_k}} e^{2h(X_k|S^{k-1}=s^{k-1})}\right). \tag{1.52}$$

Combining (1.52) and (1.50) gives

$$\begin{aligned} &\mathbb{E}_{S^k}\left(e^{2h(X_{k+1}|S^k=s^k)}\right) \\ &\quad \geq \mathbb{E}_{R_k}\left(\frac{|\lambda|^2}{2^{2R_k}} \mathbb{E}_{S^{k-1}|R_k}\left(e^{2h(X_k|S^{k-1}=s^{k-1})}\right)\right) + \beta. \end{aligned} \tag{1.53}$$

Following similar steps and using the Markov chain $S^{k-1} \to (S^{k-2}, R_{k-1}) \to R_k$, we obtain

$$\begin{aligned} &\mathbb{E}_{S^{k-1}|R_k}\left(e^{2h(X_k|S^{k-1}=s^{k-1})}\right) \\ &\quad \geq |\lambda|^2 \mathbb{E}_{S^{k-1}|R_k}\left(e^{2h(X_{k-1}|S^{k-1}=s^{k-1})}\right) + \beta \\ &\quad \geq \mathbb{E}_{S^{k-2},R_{k-1}|R_k}\left(\frac{|\lambda|^2}{2^{2R_{k-1}}} e^{2h(X_{k-1}|S^{k-2}=s^{k-2})}\right) + \beta \\ &\quad = \mathbb{E}_{R_{k-1}|R_k}\left(\frac{|\lambda|^2}{2^{2R_{k-1}}} \mathbb{E}_{S^{k-2}|R_{k-1},R_k}\left(e^{2h(X_{k-1}|S^{k-2}=s^{k-2})}\right)\right) + \beta. \end{aligned} \tag{1.54}$$

Substituting (1.54) into (1.53) and re-iterating k times, it follows that

$$\begin{aligned} &\mathbb{E}_{S^k}\left(e^{2h(X_{k+1}|S^k=s^k)}\right) \\ &\quad \geq \mathbb{E}_{R_{k-1},R_k}\left(\frac{|\lambda|^4}{2^{2(R_{k-1}+R_k)}} \mathbb{E}_{S^{k-2}|R_{k-1},R_k}\left(e^{2h(X_{k-1}|S^{k-2}=s^{k-2})}\right)\right) \\ &\qquad + \beta\left(1 + \mathbb{E}_{R_k}\left(\frac{|\lambda|^4}{2^{2R_k}}\right)\right) \\ &\quad \geq \mathbb{E}_{R_1,\ldots,R_k}\left(\frac{|\lambda|^{2k}}{2^{2(R_1+\cdots+R_k)}} \mathbb{E}_{S_1|R_1,\ldots,R_k}\left(e^{2h(X_1|S_0=s_0)}\right)\right) \\ &\qquad + \beta\left(1 + \sum_{j=2}^{k} \mathbb{E}_{R_1,\ldots,R_k}\left(\frac{|\lambda|^{2(k-j+1)}}{2^{2(R_j+\cdots+R_k)}}\right)\right) \end{aligned} \tag{1.55}$$

$$= \mathbb{E}\left(\frac{|\lambda|^{2(k+1)}}{2^{2(R_1+\cdots+R_k)}} \right) e^{2h(X_0)} + \beta \left(1 + \sum_{j=1}^{k} \mathbb{E}\left(\frac{|\lambda|^{2(k-j+1)}}{2^{2(R_j+\cdots+R_k)}} \right) \right), \tag{1.56}$$

where (1.55) uses the fact that the initial condition of the state X_0 is independent of the rate process R_k. By taking the expectation on both sides of (1.48) and iterating k times, it is easy to see that the right hand side of (1.56) is the first moment of the non-homogeneous MJLS z_{k+1} with dynamics given in (1.48). Hence, combining (1.53)–(1.56), we conclude that $\mathbb{E}(|X_k|^2) > \frac{1}{2\pi e}\mathbb{E}(Z_k)$, which is the claim. □

Lemma 1.1 shows that the state cannot be mean-square stable if the average of the $\{Z_k\}$ process is unbounded. Next, we establish that (1.44) is a necessary condition for the first-moment stability of $\{Z_k\}$. For every $k \geq 0$, let $\mu_{k,i} = \mathbb{E}[Z_k 1_{\{R_k=\bar{r}_i\}}]$ denote the expectation of Z_k in the event that the rate at time k is $\bar{r}_i$. Since $Z_{k+1} \to R_k \to R_{k+1}$ form a Markov chain, the following recursion holds for every $1 \leq i, j \leq n$:

$$\mu_{k+1,j} = \sum_{i=1}^{n} \frac{|\lambda|^2}{2^{2\bar{r}_i}} p_{ij} \mu_{k,i} + \beta \sum_{i=1}^{n} p_{ij} P\{R_k = \bar{r}_i\}, \quad k = 0, 1, \ldots.$$

It follows that the vector $\mu_k = (\mu_{k,1}, \ldots, \mu_{k,n})^T \in \mathbb{R}^n$ evolves over time according to the linear system

$$\mu_{k+1} = |\lambda|^2 H \, \mu_k + b_k, \quad k = 0, 1, \ldots, \tag{1.57}$$

where H is the transition probability matrix defined in (1.43) and $b_k \in \mathbb{R}^n$ is a vector with jth element equal to $\beta \sum_{i=1}^{n} p_{ij} P\{R_k = \bar{r}_i\}$. Notice that $\rho(|\lambda|^2 H) < 1$ is a necessary condition to ensure that the linear system (1.57) is stable, i.e., $\sup_k \|\mu_k\|_1 < \infty$. On the other hand, by the law of total probability, $\mathbb{E}(Z_k) = \sum_{i=1}^{n} \mu_{k,i} = \|\mu_k\|_1$ and so the plant is mean-square stable only if $\sup_k \|\mu_k\|_1 < \infty$. This establishes that (1.44) is a necessary condition for the second moment stability of the plant.

1.6.2 Sufficiency

Consider now the system (1.41a), (1.41b) and suppose that Assumption 1.2 is satisfied. In this section, we build a coder–decoder pair that stabilizes the system under the assumption that (1.44) holds. We first describe the adaptive quantizer that is at the base of the constructive scheme. This is based on the construction given in [49].

Adaptive Quantizer For any $r \geq 2$, the quantizer q_r proposed in [49] induces the following partition of the real line:

- The set $[-1, 1]$ is divided into 2^{r-1} intervals of the same length;
- The sets $(\xi^{i-2}, \xi^{i-1}]$ and $(-\xi^{i-1}, -\xi^{i-2}]$ are divided into 2^{r-1-i} intervals of the same length, for each $i \in \{2, \ldots, r-1\}$;

- The leftmost and rightmost intervals are the semi-open sets $(-\infty, -\xi^{r-2}]$ and (ξ^{r-2}, ∞).

A sketch of the quantizer for $r = 4$ is depicted in Fig. 1.2. Here $\xi > 1$ is a parameter that determines the concentration of intervals around the origin. We can see that the width of the quantization regions increases with ξ, so the partition becomes more spread out as ξ increases. Given a real number x, the output value of the quantizer $q_r(x)$ is the midpoint of the interval in the partition containing x. In the sequel, we will also make use of the function $\kappa_r(x)$, which instead returns the half-length of such interval, such that the quantization error is bounded by $\kappa_r(x)$. If x is in one of the two semi-open sets at the two extremes of the partition, then we set $q_r(x) = \text{sign}(x)\xi^r$ and $\kappa_r(x) = \xi^r - \xi^{r-1}$.

A fundamental property of this construction is that, loosely speaking, the estimation error produced by the mapping q_r decays exponentially fast r. The precise statement of this property involves a functional that was first introduced in [49]. For any pair of random variables (X, L), where $L \geq 0$, let

$$\|X, L\| := \sqrt{\mathbb{E}\big[L^2 + |X|^{2+\varepsilon} L^{-\varepsilon}\big]}. \tag{1.58}$$

In [29], it is shown that the non-negative functional $\|X, L\|$ is a *pseudo-norm* in the space of random vectors $(X, L) \in \mathbb{R} \times \mathbb{R}_+$ and satisfies the following properties:

(i) Second moment bound:

$$\mathbb{E}\big(|X|^2\big) \leq \|dX, dL\|^2. \tag{1.59}$$

(ii) Positive homogeneity: For any $d \geq 0$

$$\|dX, dL\| = d\|X, L\|. \tag{1.60}$$

(iii) Triangle inequality: For any $X_1, X_2 \in \mathbb{R}$ and $L_1, L_2 \geq 0$,

$$\|X_1 + X_2, L_1 + L_2\| \leq \|X_1, L_1\| + \|X_2, L_2\|. \tag{1.61}$$

Lemma 5.2 in [49] proves that if $\xi > 2^{2/\varepsilon}$, then the average quantization error produced by q_r satisfies

$$\left\| X - Lq_r\left(\frac{X}{L}\right), L\kappa_r\left(\frac{X}{L}\right)\right\|^2 \leq \frac{\zeta}{2^{2r}}\|X, L\|^2, \tag{1.62}$$

for some constant $\zeta > 0$ only determined by ε and ξ.

Another important property of this quantizer is that it is *successively refinable*. Observe in fact that the partition of the r-bit quantizer can be obtained recursively from the one of the $(r-1)$-bit quantizer by dividing each bounded interval into two intervals of the same length and the two semi-open intervals into two intervals each. In particular, the interval (ξ^{r-2}, ∞) is divided into the bounded interval $(\xi^{r-2}, \xi^{r-1}]$ and the semi-open interval (ξ^{r-1}, ∞), and similarly for the interval

$(-\infty, -\xi^{r-2}]$. Thus, $q_{r+r'}(x)$ can be computed recursively starting from $q_r(x)$ by repeating the above procedure r' times. We will make use of this property in our control scheme, where we use the fact that if coder and decoder know $q_{r_k}(x)$ at time k, then the coder can communicate to the decoder $q_{r_k+r_{k+1}}(x)$ by sending r_{k+1} bits at time $k+1$.

The stabilizing scheme can be described as follows. Coder and decoder share at each time k a state estimator $\hat{x}_k$ that is recursively updated using the symbols sent over the digital link. Time is divided into cycles of fixed duration τ. At the beginning of each cycle, the coder sends a scaled version of the estimation error that is quantized at a resolution dictated by the current value of the rate. In the remaining part of the cycle, the coder sends refinements of the original transmission at a resolution determined by the rate process at each step. At the end of each cycle, the decoder updates the state estimator and sends a control signal to the plant. The scaling factor that is applied to the error prior to quantization is updated at the end of each cycle. The basic idea is to adjust the range of the quantizer as in the zoom-in zoom-out strategy proposed in [37, 69]: the range is increased (zoom-out phase) when atypically large disturbances affect the system, and decreased as the state reduces its size (zoom-in phase). Next, the coder and decoder are described in detail.

Coder At the beginning of the jth cycle, i.e., at time $j\tau$, the coder computes

$$q_{r_{j\tau}}\big((x_{j\tau} - \hat{x}_{j\tau})/l_j\big), \tag{1.63}$$

where l_j is the scaling factor updated at the beginning of each cycle, and communicates to the decoder the index $s_{j\tau} \in \{1, \dots, 2^{r_{j\tau}}\}$ of the quantization interval containing the scaled estimation error. At time $j\tau + 1$, coder and decoder divide the quantization interval into $2^{r_{j\tau+1}}$ subintervals according to the recursive procedure described above. The coder sets $s_{j\tau+1} \in \{1, \dots, 2^{r_{j\tau+1}}\}$ equal to the subinterval containing $(x_{j\tau} - \hat{x}_{j\tau})/l_j$, so the decoder can compute

$$q_{r_{j\tau}+r_{j\tau+1}}\big((x_{j\tau} - \hat{x}_{j\tau})/l_j\big).$$

By repeating the same procedure for the rest of the cycle, at time $(j+1)\tau - 1$ the decoder knows $(x_{j\tau} - \hat{x}_{j\tau})/l_j$ at the resolution provided by a quantizer with

$$r(j) = r_{j\tau} + \cdots + r_{(j+1)\tau-1}$$

bits. Before the beginning of the next cycle, coder and decoder compute

$$\hat{x}_{(j+1)\tau} = \lambda^\tau \left(\hat{x}_{j\tau} + l_j q_{r(j)} \left(\frac{x_{j\tau} - \hat{x}_{j\tau}}{l_j} \right) \right), \tag{1.64}$$

and

$$l_{j+1} = \max \left\{ \varphi, |\lambda|^\tau l_j \kappa_{r(j)} \left(\frac{x_{j\tau} - \hat{x}_{j\tau}}{l_j} \right) \right\}, \tag{1.65}$$

with $\hat{x}_0 = 0$, $l_0 = \varphi$, where φ is any constant that only depends on ε.

Decoder At every time k the decoder sends to the plant the control signal

$$u_k = \begin{cases} -\lambda \hat{x}_k & \text{if } k = \tau, 2\tau, \dots, \\ 0 & \text{otherwise,} \end{cases} \tag{1.66}$$

where $\hat{x}_{j\tau}$ is updated as in (1.64) at the beginning of each cycle.

Analysis First, we prove that if (1.44) holds, then the second moment of the mean-squared estimation error at the beginning of each cycle is bounded. The following lemma shows that $\mathbb{E}(|X_{j\tau} - \hat{X}_{j\tau}|^2)$ is lower bounded by the second moment of a MJLS whose dynamics depends on the Markov rate process $\{R_k\}$ and on the constants α and ε defined in Assumption 1.2.

Lemma 1.2 *Let Assumption 1.2 hold. Then, for every $k = 0, 1, \dots$, the estimation error $X_{j\tau} - \hat{X}_{j\tau}$ satisfies*

$$\mathbb{E}\big(|X_{j\tau} - \hat{X}_{j\tau}|^2\big) \le \mathbb{E}\big(Z_{j\tau}^2\big),$$

where $\{Z_{j\tau}\}$ is a non-homogeneous MJLS with dynamics

$$z_{(j+1)\tau} = \phi \frac{|\lambda|^\tau}{2^{R_{j\tau} + \cdots + R_{(j+1)\tau - 1}}} z_{j\tau} + \varsigma, \quad j = 0, 1, \dots, \tag{1.67}$$

for some constants $z_0 > 0$, $\phi > 1$, and $\varsigma > 0$ that are only determined by ε, τ, and α.

Proof Let $e_{j\tau} = x_{j\tau} - \hat{x}_{j\tau}$ denote the estimation error at the beginning each cycle. By (1.59) and the fact that scaling factor L_j updated by coder and controller at the end of each cycle is nonnegative,

$$\mathbb{E}\big(|E_{(j+1)\tau}|^2\big) \le \|E_{(j+1)\tau}, L_{j+1}\|^2. \tag{1.68}$$

Notice from (1.65) that

$$l_{j+1} \le |\lambda|^\tau l_j \kappa_{R(j)}\left(\frac{x_{j\tau} - \hat{x}_{j\tau}}{l_j}\right) + \varphi,$$

and that by iteration of (1.41a), (1.41b) and (1.64) for τ time steps

$$e_{(j+1)\tau} = |\lambda|^\tau \left(e_{j\tau} - l_j q_{R(j)}\left(\frac{e_{j\tau}}{l_j}\right)\right) + \eta_j,$$

where $\eta_j := \sum_{i=0}^{\tau-1} \lambda^{\tau-1-i} v_{j\tau+i}$. Thus, properties (1.60) and (1.61) yield

$$\begin{aligned} \|E_{(j+1)\tau}, L_{j+1}\|^2 \le 2|\lambda|^{2\tau} \left\| E_{j\tau} - L_j q_{R(j)}\left(\frac{E_{j\tau}}{L_j}\right), L_j \kappa_{R(j)}\left(\frac{X_{j\tau} - \hat{X}_{j\tau}}{L_j}\right) \right\|^2 \\ + 2\|H_j, \varphi\|^2. \end{aligned} \tag{1.69}$$

Notice that $\|H_j, \varphi\|^2$ is upper bounded by a constant ς^2 that only depends on ε, τ, and α.

Let

$$\theta_{j,i} = \mathbb{E}\big(\big(L_j^2 + |E_{j\tau}|^{2+\varepsilon} L_j^{-\varepsilon}\big) 1_{\{R_{j\tau}=r_i\}}\big), \quad i \in \mathcal{R}.$$

Combining (1.62) and (1.69) and making use of the law of total probability,

$$\begin{aligned}\theta_{j+1,i_\tau} \le 2\zeta \sum_{i_0}\Bigg(\sum_{i_1,\ldots,i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0}+\cdots+r_{i_{\tau-1}})}} p_{i_0,i_1}\cdots p_{i_{\tau-1},i_\tau}\Bigg)\theta_{j,i_0} \\ + \varsigma^2 P\{R_{(j+1)\tau} = r_{i_\tau}\},\end{aligned} \tag{1.70}$$

which provides a recursive formula for the $\theta_{j,i}$ subsequences.

Next, we claim that, for every $j \ge 0$,

$$\theta_{j+1,i} \le \mathbb{E}\big[Z_{(j+1)\tau}^2 1_{\{R_{(j+1)\tau}=r_i\}}\big], \quad r_i \in \mathcal{R}, \tag{1.71}$$

where the process $\{Z_{j\tau}\}$ is formed recursively from $z_0 = \theta_0$ as

$$z_{(j+1)\tau} = \phi \frac{|\lambda|^\tau}{2^{r_{j\tau}+\cdots+r_{(j+1)\tau-1}}} z_{j\tau} + \varsigma, \quad j \ge 1, \tag{1.72}$$

where $\phi = \sqrt{2\zeta} > 1$. To see this, consider the following inductive argument. By construction $z_0 = \theta_0$, hence the claim holds for $k = 0$. Now, suppose that the claim is true up to time j. Then, for any $r_{i_\tau} \in \mathcal{R}$,

$$\begin{aligned}&\mathbb{E}\big[Z_{(j+1)\tau}^2 1_{\{R_{(j+1)\tau}=r_{i_\tau}\}}\big] \\ &= \mathbb{E}\Bigg(\Bigg(\sqrt{2\zeta}\frac{|\lambda|^\tau}{2^{R_{j\tau}+\cdots+R_{(j+1)\tau-1}}} Z_{j\tau} + \varsigma\Bigg)^2 1_{\{R_{(j+1)\tau}=r_{i_\tau}\}}\Bigg) \\ &\ge \mathbb{E}\Bigg(\Bigg(\sqrt{2\zeta}\frac{|\lambda|^\tau}{2^{R_{j\tau}+\cdots+R_{(j+1)\tau-1}}} Z_{j\tau}\Bigg)^2 1_{\{R_{(j+1)\tau}=r_{i_\tau}\}}\Bigg) + \varsigma^2 P\{R_{(j+1)\tau} = r_{i_\tau}\} \\ &= 2\zeta \sum_{i_0,\ldots,i_{\tau-1}} \frac{|\lambda|^{2\tau} p_{i_0,i_1}\cdots p_{i_{\tau-1},i_\tau}}{2^{2(r_{i_0}+\cdots+r_{i_{\tau-1}})}} \mathbb{E}\big[Z_{j\tau}^2 1_{\{R_{j\tau}=r_{i_0}\}}\big] + \varsigma^2 P\{R_{(j+1)\tau} = r_{i_\tau}\} \\ &\ge 2\zeta \sum_{i_0,\ldots,i_{\tau-1}} \frac{|\lambda|^{2\tau}}{2^{2(r_{i_0}+\cdots+r_{i_{\tau-1}})}} p_{i_0,i_1}\cdots p_{i_{\tau-1},i_\tau}\theta_{j,i_0} + \varsigma^2 P\{R_{(j+1)\tau} = r_{i_\tau}\} \\ &\ge \theta_{j+1,i_\tau}\end{aligned}$$

where the first inequality follows from the fact that $(a+b)^2 \ge a^2 + b^2$ for all nonnegative numbers a and b, the second inequality uses the induction hypothesis, while the last inequality uses (1.70). Hence, the claim holds at time $k+1$ as well.

Summing both sides of (1.71) over $r_i \in \mathcal{R}$ and making use of (1.68), it follows that $\mathbb{E}(E_{j\tau}^2) \le \mathbb{E}(Z_{j\tau}^2)$, as claimed. □

Lemma 1.2 shows that the mean-squared estimation error at the beginning of each cycle if finite if the process $\{Z_k\}$ is mean-square stable. Next, we establish that (1.44) is a sufficient condition for the second-moment stability $\{Z_k\}$. Let $\sigma_{k,i}^2 = \mathbb{E}[Z_k^2 1_{\{R_k=\bar{r}_i\}}]$ denote the second moment of Z_k in the event that the rate at time k takes value $\bar{r}_i$. Making use of the fact that $(a+b)^2 \le 2(a^2+b^2)$, it can be verified that the vector $\sigma_k^2 = (\sigma_{k,1}^2, \dots, \sigma_{k,n}^2)^T \in \mathbb{R}^n$ satisfies

$$\sigma_{k+1}^2 \le 2\phi^2 |\lambda|^{2\tau} H^\tau \, \sigma_k^2 + 2\varsigma_k^2, \quad k = 0, 1, \dots, \tag{1.73}$$

where H is the transition probability matrix defined in (1.43) and $\varsigma_k \in \mathbb{R}^n$ is a vector with the ith component equal to $\varsigma P\{R_k = \bar{r}_i\}$. A sufficient condition for the recursion in (1.73) to be bounded is

$$2\phi^2 \big(|\lambda|^2 \rho(H)\big)^\tau < 1. \tag{1.74}$$

Since by the law of total probability $\mathbb{E}(|Z|_k^2) \le \sum_{i=1}^n \sigma_{k,i}^2 = \|\sigma_k^2\|_1$, it follows that (1.74) is a sufficient condition for Z_k to be mean-square stable. On the other hand, if the condition of Theorem 1.1 is satisfied, that is, if $|\lambda|^2 \rho(H) < 1$, then we can choose the duration of a cycle τ large enough to ensure that (1.74) holds and, as a consequence, the second moment of the estimation error at the beginning of each cycle is bounded. Notice that the choice of a larger τ translates into larger oscillations of the system state because, according to our quantization scheme, the system evolves in open loop during a cycle.

Finally, for any $i = 1, \dots, \tau - 1$, the triangle inequality implies that $|x_{j\tau+i}| \le |\lambda|^i ||x_{j\tau} - \hat{x}_{j\tau}| + \sum_{k=0}^{i-1} |\lambda^{i-1-k}||v_{j\tau+i}|$, so the state remain bounded at all times. This establishes that (1.44) is a sufficient condition for the second moment stability of the plant.

1.7 Conclusion

Understanding the operational mechanism of feedback loops over limited data-rate communication channels will be of outmost importance in the near future, as cyber-physical systems (CPS) continue to impact our society more broadly. This requires the development of a rigorous theory of information transmission for control systems. This theory must identify the trade-offs between the amount of information that can be communicated through the control loop and the ability of achieving the required control objectives.

In the past decade, a number of results appeared in the literature, but much remains to be done. Obtained results show that the control objective is fundamentally limited by *both* the channel noise and the intrinsic system noise that affects the plant

in the form of external disturbances. For channels that allow transmission of a given number of bits without error, the "quality" of the achievable stabilization in terms of moment constraints depends on the corresponding constraints on the noise process disturbances. Loosely speaking, better stability can only be guaranteed with better behaved disturbances, while "wild disturbances" can only guarantee lower moment stability. In all cases, the region where the system can be stabilized is clearly demarcated by a data-rate theorem relating the amount of instability of the system to the available communication rate.

For noisy channels, the quality of the stabilization depends on the notion of channel capacity employed. Zero-error capacity, guaranteeing reliable transmission without error, allows for almost sure stabilization. Shannon capacity, guaranteeing reliable transmission with error that decays to zero asymptotically, allows for almost sure stabilization only for systems without disturbances. The parametric notion of anytime capacity, with communication reliability stronger than Shannon's capacity, but weaker than zero-error capacity, can be used to characterize stabilization of disturbed systems in a moment sense. Again, the region where the system can be stabilized is determined by a data-rate theorem written using the appropriate notion of capacity.

For limited rate channels, the theory of MJLS provides a general framework that can be used to develop data-rate theorems characterizing necessary and sufficient conditions for stabilization that hold in a variety of cases, including for the erasure channel, and for the continuous intermittent channel, with or without memory. On the other hand, the study of the DMC with memory in the context of control remains an important open problem.

Beside the formulation of data-rate theorems for different channels and noise models, a field open for further research is error correcting codes for automatic control over noisy channels. For the Gaussian channel, uncoded transmission is sufficient to achieve stabilization when the Shannon capacity is above the threshold dictated by the data-rate theorem, but for the DMC stabilization requires development of error correcting codes with specific rate-reliability constraints dictated by the corresponding data-rate theorem. These constructions are, at present, largely unknown, although recent advancements in tree codes for the erasure channel appear promising.

We conclude this chapter by mentioning some open problems. As remarked in Sect. 1.4, tight conditions for moment stability of a vector system over a time-varying bit pipe link are not known, in general. Even in the simple setting where the process on the feedback link is an i.i.d. process, only partial results are available. All existing works on stability of linear systems under stochastic disturbance of unbounded support focus on the restrictive notion of second-moment stability [17, 46, 49, 70, 71]. The generalization to η-moment stability, which is currently known only in the case where the disturbance is bounded [39, 53], is an open problem. Similarly, most of the existing works assume a perfect channel from the controller to the actuator. The case where both the sensor–controller and the controller–actuator channels are noisy was studied in [73], which provides conditions for second moment stability using Markov stability theory. In general, however, it is not known when the criteria summarized in this chapter continue to hold

after replacing the relevant notion of capacity with the capacity of the bottleneck channel. Our previous work [46] has revealed a connection between stabilization over the intermittent continuous channel and the rate-limited channel. It would be of interest to establish a similar connection in the case of optimal control over finite–capacity channels. Previous works [26, 56] have considered the LQG problem under the network-theoretic approach where packets can be lost, while [9, 31, 40] studied the same problem under the assumption that the feedback channel is a bit pipe with constant rate R. In order to create a connection between these two lines of work, one would have to formulate an LQG problem over a time-varying bit pipe channel whose rate oscillates independently over time between 0 and R. As a final remark, notice that the proof techniques used in [53] only apply to plants with bounded disturbances. A question that requires further investigation is to extend the result in [53] to the case of noise with infinite support. A possible approach based on variable rate coding is outlined in [52, 73].

As control systems gradually evolve towards usage of wireless platforms, the developed theory will have a direct applicability in a practical setting. The move towards wireless is dictated by both technological advancements and economic factors, as the cost of "wiring" large CPS can easily dominate development costs. The theory developed so far has shown that existing error correcting codes for wireless communication are not immediately applicable in the context of control, due to their soft reliability constraints that are not sufficient to ensure even low-moment stability for safety critical applications. In the next decades, we will witness a refinement of the theory to gain additional understanding of fundamental limitations, as well as the development of new communication schemes needed to address the growing industrial need for control over noisy channels.

Acknowledgement This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Adler, R.L., Konheim, A.G., McAndrew, M.H.: Topological entropy. Trans. Am. Math. Soc. **114**, 309–319 (1965)
2. Andrievsky, B., Matveev, A., Fradkov, A.: Control and estimation under information constraints: toward a unified theory of control, computation and communications. Autom. Remote Control **71**, 572–633 (2010). Original Russian text in Autom. Telemekh. **4**, 34–99 (2010)
3. Ardestanizadeh, E., Franceschetti, M.: Control-theoretic approach to communication with feedback. IEEE Trans. Autom. Control (2012)
4. Ardestanizadeh, E., Minero, P., Franceschetti, M.: LQG control approach to Gaussian broadcast channels with feedback. IEEE Trans. Inf. Theory **58**(8), 5267–5278 (2012)
5. Baillieul, J.: Feedback designs for controlling device arrays with communication channel bandwidth constraints. In: ARO Workshop on Smart Structures, Penn. State U., USA (1999)
6. Baillieul, J.: Feedback designs in information-based control. In: Pasik-Duncan, B. (ed.) Proceedings of the Workshop on Stochastic Theory and Control. Springer, Lawrence (2001)
7. Baillieul, J., Antsaklis, P.J. (eds.): Special issue on Networked Control Systems. IEEE Trans. Autom. Control **49**(9) (2004)

8. Baillieul, J., Antsaklis, P.J.: Control and communication challenges in networked real-time systems. Proc. IEEE **95**(1), 9–28 (2007)
9. Borkar, V., Mitter, S.: LQG Control with Communication Constraints. LIDS-P. Massachusetts Institute of Technology, Laboratory for Information and Decision Systems (1995)
10. Borkar, V., Mitter, S.: LQG Control with Communication Constraints. Communications, Computation, Control and Signal Processing: A Tribute to Thomas Kailath. Kluwer Academic, Dordrecht (1997)
11. Braslavsky, J., Middleton, R., Freudenberg, J.: Feedback stabilization over signal-to-noise ratio constrained channels. IEEE Trans. Autom. Control **52**(8), 1391–1403 (2007)
12. Brockett, R., Liberzon, D.: Quantized feedback stabilization of linear systems. IEEE Trans. Autom. Control **45**(7), 1279–1289 (2000)
13. Como, G., Fagnani, F., Zampieri, S.: Anytime reliable transmission of real-valued information through digital noisy channels. SIAM J. Control Optim. **48**(6), 3903–3924 (2010)
14. Costa, M., Cover, T.: On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. IEEE Trans. Inf. Theory **30**(6), 837–839 (1984)
15. Costa, O., Fragoso, D., Marques, R.: Discrete-Time Markov Jump Linear Systems. Probability and Its Applications. Springer, Berlin (2004)
16. Cover, T., Thomas, J.: Elements of Information Theory. Wiley, New York (2006)
17. Coviello, L., Minero, P., Franceschetti, M.: Stabilization over Markov feedback channels: the general case. IEEE Trans. Autom. Control **58**(2), 349–362 (2013)
18. Delchamps, D.: Stabilizing a linear system with quantized state feedback. IEEE Trans. Autom. Control **35**(8), 916–924 (1990)
19. El Gamal, A., Kim, Y.H.: Network Information Theory. Cambridge University Press, Cambridge (2011)
20. Elia, N.: When Bode meets Shannon: control-oriented feedback communication schemes. IEEE Trans. Autom. Control **49**(9), 1477–1488 (2004)
21. Elia, N.: Remote stabilization over fading channels. Syst. Control Lett. **54**(3), 237–249 (2005)
22. Elia, N., Mitter, S.K.: Stabilization of linear systems with limited information. IEEE Trans. Autom. Control **46**(9), 1384–1400 (2001)
23. Forney, G.D.: Convolutional codes II. Maximum-likelihood decoding. Inf. Control **25**(3), 222–266 (1974)
24. Franceschetti, M., Javidi, T., Kumar, P.R., Mitter, S.K., Teneketzis, D. (eds.): Special issue on Control and Communications. IEEE J. Sel. Areas Commun. **26**(4) (2008)
25. Freudenberg, J., Middleton, R.H., Solo, V.: Stabilization and disturbance attenuation over a Gaussian communication channel. IEEE Trans. Autom. Control **55**(3), 795–799 (2010)
26. Gupta, V., Spanos, D., Hassibi, B., Murray, R.M.: Optimal LQG control across packet-dropping links. Syst. Control Lett. **56**(6), 439–446 (2007)
27. Gupta, V., Dana, A., Hespanha, J., Murray, R., Hassibi, B.: Data transmission over networks for estimation and control. IEEE Trans. Autom. Control **54**(8), 1807–1819 (2009)
28. Gupta, V., Martins, N., Baras, J.: Optimal output feedback control using two remote sensors over erasure channels. IEEE Trans. Autom. Control **54**(7), 1463–1476 (2009)
29. Gurt, A., Nair, G.N.: Internal stability of dynamic quantised control for stochastic linear plants. Automatica **45**(6), 1387–1396 (2009)
30. Hespanha, J., Naghshtabrizi, P., Xu, Y.: A survey of recent results in networked control systems. Proc. IEEE **95**(1), 138–162 (2007)
31. Huang, M., Nair, G., Evans, R.: Finite horizon LQ optimal control and computation with data rate constraints. In: 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC'05, pp. 179–184 (2005)
32. Imer, O.C., Yüksel, S., Başar, T.: Optimal control of LTI systems over unreliable communication links. Automatica **42**(9), 1429–1439 (2006)
33. Kim, Y.H.: Feedback capacity of the first-order moving average Gaussian channel. IEEE Trans. Inf. Theory **52**(7), 3063–3079 (2006)
34. Kim, Y.H.: Feedback capacity of stationary Gaussian channels. IEEE Trans. Inf. Theory **56**(1), 57–85 (2010)

35. Kim, K.D., Kumar, P.R.: Cyber-physical systems: a perspective at the centennial. Proc. IEEE **100**(13), 1287–1308 (2012)
36. Korner, J., Orlitsky, A.: Zero-error information theory. IEEE Trans. Inf. Theory **44**(6), 2207–2229 (1998)
37. Liberzon, D.: On stabilization of linear systems with limited information. IEEE Trans. Autom. Control **48**(2), 304–307 (2003)
38. Lin, S., Costello, D.J. Jr.: Error Control Coding: Fundamentals and Applications. Prentice Hall, New York (1983). TUB-HH 2413-469 3
39. Martins, N., Dahleh, M., Elia, N.: Feedback stabilization of uncertain systems in the presence of a direct link. IEEE Trans. Autom. Control **51**(3), 438–447 (2006)
40. Matveev, A.S., Savkin, A.V.: The problem of LQG optimal control via a limited capacity communication channel. Syst. Control Lett. **53**(1), 51–64 (2004)
41. Matveev, A., Savkin, A.: Comments on "control over noisy channels" and relevant negative results. IEEE Trans. Autom. Control **50**(12), 2105–2110 (2005)
42. Matveev, A.S., Savkin, A.V.: An analogue of Shannon information theory for detection and stabilization via noisy discrete communication channels. SIAM J. Control Optim. **46**(4), 1323–1367 (2007)
43. Matveev, A.S., Savkin, A.V.: Shannon zero error capacity in the problems of state estimation and stabilization via noisy communication channels. Int. J. Control **80**(2), 241–255 (2007)
44. Matveev, A., Savkin, A.: Estimation and Control over Communication Networks. Birkhäuser, Basel (2009). Control Engineering
45. Middleton, R., Rojas, A., Freudenberg, J., Braslavsky, J.: Feedback stabilization over a first order moving average Gaussian noise channel. IEEE Trans. Autom. Control **54**(1), 163–167 (2009)
46. Minero, P., Franceschetti, M., Dey, S., Nair, G.: Data rate theorem for stabilization over time-varying feedback channels. IEEE Trans. Autom. Control **54**(2), 243–255 (2009)
47. Mo, Y., Sinopoli, B.: Kalman filtering with intermittent observations: tail distribution and critical value. IEEE Trans. Autom. Control **57**(3), 677–689 (2012)
48. Nahi, N.: Optimal recursive estimation with uncertain observation. Automatica **15**(4), 457–462 (1969)
49. Nair, G.N., Evans, R.J.: Stabilizability of stochastic linear systems with finite feedback data rates. SIAM J. Control Optim. **43**(2), 413–436 (2004)
50. Nair, G., Fagnani, F., Zampieri, S., Evans, R.: Feedback control under data rate constraints: an overview. Proc. IEEE **95**(1), 108–137 (2007)
51. Ostrovsky, R., Rabani, Y., Schulman, L.: Error-correcting codes for automatic control. IEEE Trans. Inf. Theory **55**(7), 2931–2941 (2009)
52. Sahai, A.: Anytime information theory. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA (2001)
53. Sahai, A., Mitter, S.K.: 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. Theory **52**(8), 3369–3395 (2006)
54. Sahai, A., Mitter, S.K.: The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link—Part II: vector systems (2006). Available on-line at arXiv:cs/0610146v2 [cs.IT]
55. Schalkwijk, J.P.M., Kailath, T.: A coding scheme for additive noise channels with feedback—I: no bandwidth constraint. IEEE Trans. Inf. Theory **12**, 172–182 (1966)
56. Schenato, L., Sinopoli, B., Franceschetti, M., Poolla, K., Sastry, S.: Foundations of control and estimation over lossy networks. Proc. IEEE **95**(1), 163–187 (2007)
57. Schulman, L.: Coding for interactive communication. IEEE Trans. Inf. Theory **42**(6), 1745–1756 (1996)
58. Shannon, C.E.: The zero-error capacity of a noisy channel. IRE Trans. Inf. Theory **2**, 8–19 (1956)
59. Shi, L., Epstein, M., Murray, R.: Kalman filtering over a packet-dropping network: a probabilistic perspective. IEEE Trans. Autom. Control **55**(3), 594–604 (2010)

60. Simşek, T., Jain, R., Varaiya, P.: Scalar estimation and control with noisy binary observations. IEEE Trans. Autom. Control **49**(9), 1598–1603 (2004)
61. Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M., Sastry, S.: Kalman filtering with intermittent observations. IEEE Trans. Autom. Control **49**(9), 1453–1464 (2004)
62. Soummya, K., Sinopoli, B., Moura, J.M.F.: Kalman filtering with intermittent observations: weak convergence to a stationary distribution. IEEE Trans. Autom. Control **57**(2), 405–420 (2012)
63. Sukhavasi, R., Hassibi, B.: Error correcting codes for distributed control (2011). Available on-line at arXiv:1112.4236v2 [cs.IT]
64. Tatikonda, S., Mitter, S.K.: Control over noisy channels. IEEE Trans. Autom. Control **49**(7), 1196–1201 (2004)
65. Tatikonda, S., Mitter, S.K.: Control under communication constraints. IEEE Trans. Autom. Control **49**(7), 1056–1068 (2004)
66. Tatikonda, S., Sahai, A., Mitter, S.K.: Stochastic linear control over a communication channel. IEEE Trans. Autom. Control **49**(9), 1549–1561 (2004)
67. Witsenhausen, H.S.: A counterexample in stochastic optimum control. SIAM J. Control **6**(1), 131–147 (1968)
68. Wong, W.S., Brockett, R.: Systems with finite communication bandwidth constraints. I. State estimation problems. IEEE Trans. Autom. Control **42**(9), 1294–1299 (1997)
69. Wong, W.S., Brockett, R.: Systems with finite communication bandwidth constraints. II. Stabilization with limited information feedback. IEEE Trans. Autom. Control **44**(5), 1049–1053 (1999)
70. You, K., Xie, L.: Minimum data rate for mean square stabilization of discrete LTI systems over lossy channels. IEEE Trans. Autom. Control **55**(10), 2373–2378 (2010)
71. You, K., Xie, L.: Minimum data rate for mean square stabilizability of linear systems with Markovian packet losses. IEEE Trans. Autom. Control **56**(4), 772–785 (2011)
72. Yüksel, S.: Stochastic stabilization of noisy linear systems with fixed-rate limited feedback. IEEE Trans. Autom. Control **55**(12), 2847–2853 (2010)
73. Yüksel, S., Başar, T.: Control over noisy forward and reverse channels. IEEE Trans. Autom. Control **56**(5), 1014–1029 (2011)

Chapter 2
Stabilization and Control over Gaussian Networks

Ali A. Zaidi, Tobias J. Oechtering, Serdar Yüksel, and Mikael Skoglund

2.1 Introduction

In this chapter, we consider a setup where a linear time-invariant system (plant) with a random initial state and driven by Gaussian noise has to be remotely stabilized. A group of sensor nodes monitor the plant and communicate their observations (measurements) to a remotely situated control unit over wireless links that are modeled as additive white Gaussian channels. The common goal of the sensors and the controller is to stabilize the plant in closed-loop. Usually, in remote control applications, sensing and transmission under strict delay and power constraints is required. Therefore, we focus on delay-free and power efficient sensing and transmit schemes throughout the chapter. Our objective is to provide an overview and some recent results on real-time communication and stabilization over Gaussian channels. In order to grasp the fundamental principles, we consider setups with one or two sensor nodes under some basic topologies, however, useful references to more general setups are provided throughout the chapter.

The main focus of this chapter is on the mean-square stabilization of a linear dynamical system over some basic Gaussian network settings. Some real-time sensing and transmission schemes are proposed and stabilizability of the plant under those schemes is studied. The chapter is organized as follows. In Sect. 2.2, the problem

A.A. Zaidi (✉) · T.J. Oechtering · M. Skoglund
KTH Royal Institute of Technology, Stockholm, Sweden
e-mail: zaidi@kth.se

T.J. Oechtering
e-mail: oech@kth.se

M. Skoglund
e-mail: skoglund@kth.se

S. Yüksel
Queen's University, Kingston, Canada
e-mail: yuksel@mast.queensu.ca

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_2,

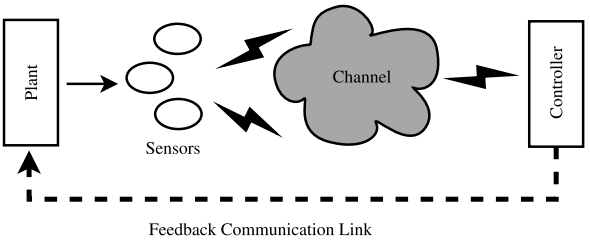


Fig. 2.1 Feedback control over sensor network

of remote stabilization of a discrete-time LTI plant over Gaussian sensor network is formulated. In Sect. 2.3, a single sensor setup is considered, i.e., stabilization over a point-to-point Gaussian channel. Section 2.4 and Sect. 2.6 consider two sensor setups where one sensor node merely acts as a relay for communicating state information to the remote controller. Section 2.4 focuses on the mean-square stabilization of an LTI plant in various relaying topologies. Section 2.5 and 2.4 addresses the problem of real-time transmission of a Gaussian source over a Gaussian relay channel for delay-sensitive and energy limited applications such as closed-control over wireless sensor networks. Finally, in Sect. 2.7, we discuss distributed sensing schemes for control over Gaussian channels. The chapter ends with an overview of the existing literature on the problem of control over Gaussian channels, highlighting the important relevant contributions.

2.2 Remote Stabilization of a Linear System

Consider the following linear time invariant system:

$$X_{t+1} = AX_t + U_t + W_t, \quad t \in \mathbb{N}, \tag{2.1}$$

where $X_t := [x_{1,t}, x_{2,t}, \ldots, x_{n,t}]^T$ is an $\mathbb{R}^n$-valued state process with an initial Gaussian distribution, $U_t := [u_{1,t}, u_{2,t}, \ldots, u_{n,t}]^T$ is an $\mathbb{R}^n$-valued control process, $W_t := [w_{1,t}, w_{2,t}, \ldots, w_{n,t}]^T$ is an $\mathbb{R}^n$-valued independent and identically distributed sequence of Gaussian random variables with zero mean and covariance K_W, and A is the system matrix of appropriate dimensions. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ denote the eigenvalues of the system matrix A. Without loss of generality, we assume that all the eigenvalues of the system matrix are outside the unit disc ($|\lambda_i| \geq 1$ for all i), i.e., all modes are unstable. Otherwise unstable modes can be decoupled from the stable modes by a similarity transformation. If the system in (2.1) is one-dimensional then A is scalar and we use the notation $A = \lambda$, where $|\lambda| > 1$. The initial state of the system X_0 is assumed to be a random variable with zero mean and covariance Λ_0.

We consider a remote control setup as shown in Fig. 2.1, where a sensor or a group of sensor nodes observe the state process and communicate their observations directly to a remotely situated controller over a wireless channel. In a sensor

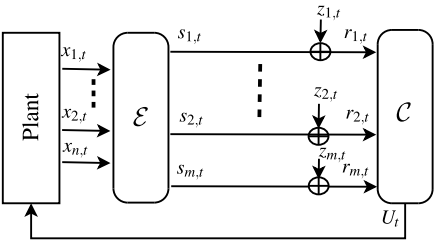


Fig. 2.2 Control over a point–point channel

network, some nodes can act as relays to support the communication with the remote controller. Upon receiving the signals from the sensors, the remotely located controller aims at stabilizing the system in the mean-square sense, which is defined as follows.

Definition 2.1 A system is said to be *mean-square stable* if there exists a constant $M < \infty$ such that $\mathbb{E}[\|X_t\|^2] < M$ for all t.

In practice, sensor nodes have limited power to spend. Therefore, we assume an average transmit power constraint at each sensor node. The communication links between all agents (sensors and controller) are modeled as independent Gaussian channels. Since control applications are usually quite sensitive to delays, the sensing and transmission schemes are restricted to be delay-free. In order to make the implementation simple, we assume that the controller has a separation structure based on the minimum mean-square estimator state estimator. This separation structure is not optimal in general but it makes the design and implementation much simpler, by employing Kalman filter as state estimator.

2.3 Stabilization over a Point–Point Channel

Consider the scenario shown in Fig. 2.2 where a sensor node $\mathcal{E}$ observes an n-dimensional state process and transmits it to a remote controller $\mathcal{C}$ over an m-dimensional parallel Gaussian channel. We assume that the initial state is zero mean Gaussian distributed. At any time instant t, $S_t := [s_{1,t}, s_{2,t}, \ldots, s_{m,t}]$ and $R_t := [r_{1,t}, r_{2,t}, \ldots, r_{m,t}]$ are the input and output of the channel, where $r_{i,t} = s_{i,t} + z_{i,t}$ and $z_{i,t} \sim \mathcal{N}(0, N_i)$ are zero mean white Gaussian noise components. Let $f_t : \mathbb{R}^{n(t+1)} \to \mathbb{R}^m$ denote the sensing policy such that $S_t = f_t(X_{[0,t]})$, where $X_{[0,t]} := \{X_0, X_1, \ldots, X_t\}$. The sensor is assumed to have an average transmit power constraint $\mathbb{E}[\|S_t\|^2] = \sum_{i=1}^m P_i \leq P_S$, where $P_i := \mathbb{E}[(s_{i,t})^2]$. Further, let $\pi_t : \mathbb{R}^{m(t+1)} \to \mathbb{R}^n$ be the controller policy, then we have $U_t = \pi_t(R_{[0,t]})$. The common goal of the sensor and the controller is to stabilize the LTI system (2.1) in the mean-square sense.

We first present a necessary condition for the *mean-square* stabilization over the point–point Gaussian channel depicted in Fig. 2.2.

Theorem 2.1 *The linear system in* (2.1) *can be mean-square stabilized over the given parallel Gaussian channel only if*

$$\log\big(|A|\big) < \frac{1}{2}\sum_{i=1}^{m}\log\bigg(1+\frac{P_i}{N_i}\bigg), \tag{2.2}$$

where $P_i = \max\{\gamma - N_i, 0\}$ *and* γ *is chosen such that* $\sum_{i=1}^{m} P_i = P_S$.

Proof In order to prove Theorem 2.1, we make use of the following lemma.

Lemma 2.1 [67, Theorem 2.1] *The linear system in* (2.1) *can be mean-square stabilized over a channel only if*

$$\log\big(|A|\big) \leq \liminf_{T\to\infty}\frac{1}{T} I\big(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\big), \tag{2.3}$$

where $\{\bar{X}_t\}$ *is the control free state process given by substituting* $U_t = 0$ *in* (2.1), $R_{[0,T-1]}$ *is the sequence of variables received by the controller over the given channel and* $I\big(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\big) = \sum_{t=0}^{T-1} I\big(\bar{X}_{[0,t]}; R_t | R_{[0,t\text{-}1]}\big)$ *denotes the directed information.*

Proof The proof can be found in [67]. This proof essentially follows from the same steps as in Theorem 4.1 of [54], however, with some differences due to the network structure. Similar considerations have appeared in different contexts in [28, 41]. □

We can bound the directed information $I\big(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\big)$ as

$$\begin{aligned}
&I\big(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\big)\\
&\overset{(a)}{\leq} I\big(\bar{X}_{[0,T-1]}; R_{[0,T-1]}\big) \overset{(b)}{\leq} I\big(S_{[0,T-1]}; R_{[0,T-1]}\big)\\
&\overset{(c)}{\leq} \sum_{t=0}^{T-1} I\big(s_{1,t}, s_{2,t}, \ldots, s_{m,t}; r_{1,t}, r_{2,t}, \ldots, r_{m,t}\big)\\
&= \sum_{t=0}^{T-1}\Big[h(r_{1,t}, r_{2,t}, \ldots, r_{m,t}) - h(r_{1,t}, r_{2,t}, \ldots, r_{m,t}|s_{1,t}, s_{2,t}, \ldots, s_{m,t})\Big]\\
&= \sum_{t=0}^{T-1}\Big[h(r_{1,t}, r_{2,t}, \ldots, r_{m,t}) - h(z_{1,t}, z_{2,t}, \ldots, z_{m,t}|s_{1,t}, s_{2,t}, \ldots, s_{m,t})\Big]\\
&= \sum_{t=0}^{T-1}\Big[h(r_{1,t}, r_{2,t}, \ldots, r_{m,t}) - h(z_{1,t}, z_{2,t}, \ldots, z_{m,t})\Big]\\
&\overset{(d)}{\leq} \sum_{t=0}^{T-1}\bigg[\sum_{i=1}^{m} h(r_{i,t}) - \sum_{i=1}^{m} h(z_{i,t})\bigg] \overset{(e)}{\leq} \sum_{t=0}^{T-1}\bigg[\sum_{i=1}^{m}\log\bigg(1+\frac{P_i}{N_i}\bigg)\bigg]
\end{aligned}$$

$$= \frac{T}{2} \sum_{i=1}^{m} \log\left(1 + \frac{P_i}{N_i}\right), \tag{2.4}$$

where (a) follows from [29, Theorem 1]; (b) following data processing inequality with Markov chain $\bar{X}_{[0,T-1]} - S_{[0,T-1]} - R_{[0,T-1]}$; (c) follows from the fact that the channels are memoryless and conditioning reduces entropy; (d) follows from conditioning reduces entropy and mutual independence of the noise sequence $\{z_{1,t}, z_{2,t}, \ldots, z_{m,t}\}$; and (e) follows from the fact that the Gaussian distribution maximizes differential entropy for a fixed variance. Now using (2.4) in Lemma 2.1, we get the necessary condition given in (2.2). The function $\sum_{i=1}^{L} \log(1 + \frac{P_i}{N_i})$ is jointly concave in $\{P_i\}_{i=1}^{m}$; therefore, we can solve this optimization problem by the Lagrangian method. The optimal power allocation using the Lagrangian method is given by $P_i = \max\{\gamma - N_i, 0\}$, where γ is chosen such that $\sum_{i=1}^{m} P_i = P_S$, which is the well-known water-filling solution. □

We now discuss some sensing and control schemes for stabilization over the given point–point Gaussian channel. By employing these schemes, we obtain sufficient conditions for stabilization, which are also presented in the following sections. The schemes for scalar and vector channels are discussed in Sect. 2.3.1 and Sect. 2.3.2, respectively.

2.3.1 Schemes for Scalar Channels

In this section, we consider the mean-square stability of the system in (2.1) over a scalar Gaussian channel, i.e., we assume that $m = 1$ in the system model shown in Fig. 2.2. The state encoder $\mathcal{E}$ observes the n-dimensional state process and transmits it over a one-dimensional Gaussian channel. We restrict our study to the class of encoders that are linear in the observed state with an average transmit power constraint P_S. Therefore, at any time t, the signal transmitted by the state encoder is given by $S_t = E_t X_t$, where E_t is an $1 \times n$ row vector. The power constraint at the encoder is given by

$$\mathbb{E}\left[S_t^2\right] = E_t \Lambda_t E_t^T \leq P,$$

where $\Lambda_t := \mathbb{E}[X_t X_t^T]$. The remotely located controller receives the following signal,

$$R_t = S_t + Z_t,$$

where Z_t is an i.i.d. Gaussian variable with zero mean and variance N. The information set available to the controller is $I_t^{\mathcal{C}} = \{R_{[0,t]}, U_{[0,t-1]}\}$. The controller applies an action which is linear in the information set, that is, $U_t = m_t I_t^{\mathcal{C}}$. In the following, we study the mean-square stability under the above linear sensing and control scheme.

We have restricted ourselves to linear schemes because they are easy to design and implement. At this point, we highlight some interesting questions that may arise

in the reader's mind: (i) Is there any loss in restricting sensing and control policies to be linear? (ii) Should the policies be time-invariant or time-variant? (iii) What is an optimal linear scheme? We try to address these questions in the Sects. 2.3.1.1–2.3.1.3.

2.3.1.1 Linear Time Invariant Scheme

Consider the linear scheme presented above to be time invariant, i.e., at any time t, the encoder output is given by $S_t = EX_t$. The controller receives $R_t = EX_t + Z_t$, then it runs a Kalman filter to estimate the state and applies the following action $U_t = -A\mathbb{E}[X_t|I_t^{\mathcal{C}}]$, which is optimal for stabilization under the given sensing scheme. Thus the closed loop system is given by

$$\begin{aligned} X_{t+1} &= A\big(X_t - \mathbb{E}\big[X_t\big|I_t^{\mathcal{C}}\big]\big) + W_t \\ &\overset{(a)}{=} A\big(X_t - \Lambda_t E^T\big[E\Lambda_t E^T + \sigma_z^2\big]^{-1}R_t\big) + W_t \\ &\overset{(b)}{=} \big(A_t - \Lambda_t E^T\big[E\Lambda_t E^T + \sigma_z^2\big]^{-1}E\big)X_t + \tilde{Z}_t, \end{aligned} \tag{2.5}$$

where (a) follows from the fact that the control actions whiten the state process and the Gaussian distribution of state process is preserved via linear actions of the encoder and the controller, which results in $\mathbb{E}[X_t|I_t^{\mathcal{C}}] = \mathbb{E}[X_t|R_t] = \mathbb{E}[X_t R_t^T]\mathbb{E}[R_t R_t^T]^{-1}R_t$; and (b) follows by substituting $R_t = EX_t + Z_t$ and summing up all the white Gaussian noise terms and denoting the sum by $\tilde{Z}_t$. The state covariance matrix Λ_t satisfies the following recursion

$$\Lambda_{t+1} = A\Lambda_t A^T - A\Lambda_t E^T\big[E\Lambda_t E^T + \sigma_z^2\big]^{-1}E\Lambda_t A^T + K_W, \tag{2.6}$$

which is the well-known Riccati equation. In [8], the authors studied such a scheme. According to [8], a noiseless plant can be mean-square stabilized by any time-invariant encoding matrix E over a Gaussian channel capacity C as long as the following two conditions are fulfilled: (i) $\log\{|A|\} < C$, (ii) the pair (A, E) is observable.

We now give a simple example where the LTI scheme fails to stabilize the system. Consider a diagonal system matrix $A = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$ with two equal eigenvalues and let $E = (e_1\ e_2\ e_3)$. The observability matrix $\mathcal{O}$ is then given by

$$\mathcal{O} \triangleq \begin{pmatrix} E \\ EA \\ EA^2 \end{pmatrix} = \begin{pmatrix} e_1 & e_2 & e_3 \\ e_1\lambda_1 & e_2\lambda_2 & e_3\lambda_3 \\ e_1\lambda_1^2 & e_3\lambda_2^2 & e_3\lambda_3^2 \end{pmatrix}.$$

For the pair (A, E) to observable, the observability matrix $\mathcal{O}$ is required to have full rank. In the above example, if any two eigenvalues of A are equal, then there can be at most two linearly independent columns in $\mathcal{O}$ and thus rank of $\mathcal{O}$ can never be made full by any choice of E. (One can also use the Hautus–Rosenbrock test for

observability.) Therefore, an LTI scheme can never stabilize if two or more eigenvalues of a diagonal system matrix are equal, no matter how large power the encoder is allowed to spend. In the following section, we present a linear time varying scheme and show that this scheme can always stabilize the system.

2.3.1.2 Linear Time Variant Scheme

Consider that the linear system (2.1) has to be stabilized over a Gaussian channel having information capacity $C := \frac{1}{2}\log(1+\frac{P}{N})$, which means that the sensor transmits with an average power P and the channel is disturbed by a zero mean Gaussian noise with variance N. In the following, we state a sufficient condition for the mean-square stability under a linear time varying scheme.

Theorem 2.2 *The linear system* (2.1) *can be mean-square stabilized by a linear time-variant scheme over a scalar Gaussian channel of capacity* C *if* $\log(|A|) < C$.

Proof Without loss of generality, we assume that the system matrix is in real Jordan form. Depending on the nature of eigenvalues, a real Jordan matrix J has the following structure:

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix}, \quad \text{where}$$

$$J_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix} \quad \text{for } \lambda_i \in \mathbb{R},$$

$$J_i = \begin{pmatrix} C_i & I & & \\ & C_i & \ddots & \\ & & \ddots & I \\ & & & C_I \end{pmatrix} \quad \text{for } \lambda_i = \sigma_i \pm j\omega_i \in \mathbb{C},$$

$$\text{with } C_i = \begin{pmatrix} \sigma_i & \omega_i \\ -\omega_i & \sigma_i \end{pmatrix}.$$

Consider a scheme in which the sensor transmits only one component of the state vector at each time step. Since the system matrix is in Jordan form, it can transmit the state components corresponding to more unstable modes more often. In the following, we show that with such a time varying mode-by-mode transmission scheme,

the plant can be stabilized if $\log(|A|) < C$. We justify this with the help of two simple examples. In the first example, we consider the system matrix with repeated real and complex eigenvalues having equal magnitude, whereas in the second example, we consider eigenvalues with unequal magnitude. These two examples capture the general principle. In the end, we outline a general transmission scheme.

Example 1 (Repeated eigenvalues with equal magnitude) Consider an LTI system with the following state equation:

$$\bar{X}_{t+1} = A\bar{X}_t + \bar{U}_t + \bar{W}_t, \tag{2.7}$$

where A is an $n \times n$ matrix with eigenvalues λ_i such that $|\lambda_i| = |\lambda_j|$ for all $1 \leq i, j \leq n$. Assume that the control actions U_t are taken periodically after every n time steps, i.e., at $t = l(n-1)$ for $l = 1, 2, 3, \ldots$. Under this control scheme, the state at times steps $t = ln$ is given by

$$\bar{X}_{t+n} = A^n \bar{X}_t + \bar{U}_{t+n-1} + \sum_{i=0}^{n-1} A^{n-i-1} \bar{W}_{t+i}, \quad t = ln,\ l \in \mathbb{N}.$$

Let T be a linear transformation such that $T^{-1}A^nT$ is in real Jordan form. It is known that such a transformation always exists [17]. Now apply the linear transformation $X_t = T^{-1}\bar{X}_t$, which gives

$$\begin{aligned} X_{t+n} &= T^{-1}A^nTX_t + T^{-1}\bar{U}_{t+n-1} + T^{-1}\sum_{i=0}^{n-1} A^{n-i-1}\bar{W}_{t+i}, \\ &= \tilde{A}X_t + U_{t+n-1} + V_t, \quad \text{for } t = ln,\ l \in \mathbb{N}, \end{aligned} \tag{2.8}$$

where we have defined $\tilde{A} := T^{-1}A^nT$, $U_t := T^{-1}\bar{U}_t$, and $V_t := T^{-1}\sum_{i=0}^{n-1} A^{n-i-1} \times \bar{W}_{t+i}$. The matrix $\tilde{A}$ is in real Jordan form with eigenvalues $\tilde{\lambda}_i = \lambda_i^n$, where λ_i are the eigenvalues of A for $i = 1, 2, \ldots, n$. Now consider the following sensing scheme for stabilization. The sensor periodically observes the state vector $X_t \in \mathbb{R}^n$ after every n time steps, i.e., at $t, t+n, t+2n, \ldots$. The sensor linearly amplifies each component of the state vector under an average transmit power constraint P and sequentially transmits n state components over the Gaussian channel. The state vector is thus transmitted to the controller by using the Gaussian channel n times. The controller computes the MMSE estimate of the state vector $\hat{X}_t$ based on the received signals and periodically takes actions after every n time steps, i.e., $U_{t+n-1} = -\tilde{A}\hat{X}_t$. Under the above scheme, we can write (2.8) as

$$X_{t+n} = \tilde{A}(X_t - \hat{X}_t) + V_t, \quad t = ln,\ l \in \mathbb{N}. \tag{2.9}$$

In the following, we consider an example with $n = 6$ and show that above scheme is sufficient for stabilization.

Consider a six-dimensional plant ($n = 6$) with state vector $X_t \in \mathbb{R}^6$, that is, a plant with six poles (eigenvalues). Assume that the matrix $\tilde{A}$ has a real eigenvalue

and a complex conjugate pair, each with algebraic multiplicity two. That is we have $\tilde{\lambda}_1 = \tilde{\lambda}_3 = \tilde{\sigma} + j\tilde{\omega}$, $\tilde{\lambda}_2 = \tilde{\lambda}_4 = \tilde{\sigma} - j\tilde{\omega}$, and $\tilde{\lambda}_5 = \tilde{\lambda}_6 = \tilde{\lambda}$. Since $\tilde{A}$ is in real Jordan form, we have

$$\tilde{A} = \begin{pmatrix} \tilde{\sigma} & \tilde{\omega} & 1 & 0 & 0 & 0 \\ -\tilde{\omega} & \tilde{\sigma} & 0 & 1 & 0 & 0 \\ 0 & 0 & \tilde{\sigma} & \tilde{\omega} & 0 & 0 \\ 0 & 0 & -\tilde{\omega} & \tilde{\sigma} & 0 & 0 \\ 0 & 0 & 0 & 0 & \tilde{\lambda} & 1 \\ 0 & 0 & 0 & 0 & 0 & \tilde{\lambda} \end{pmatrix}. \tag{2.10}$$

By substituting $\tilde{A}$ from (2.10) in (2.9), each component of the state vector is given by

$$\begin{aligned} x_{1,t+n} &= \tilde{\sigma}(x_{1,t} - \hat{x}_{1,t}) + \tilde{\omega}(x_{2,t} - \hat{x}_{2,t}) + (x_{3,t} - \hat{x}_{3,t}) + v_{1,t}, \\ x_{2,t+n} &= -\tilde{\omega}(x_{1,t} - \hat{x}_{1,t}) + \tilde{\sigma}(x_{2,t} - \hat{x}_{2,t}) + (x_{4,t} - \hat{x}_{4,t}) + v_{2,t}, \\ x_{3,t+n} &= \tilde{\sigma}(x_{3,t} - \hat{x}_{3,t}) + \tilde{\omega}(x_{4,t} - \hat{x}_{4,t}) + v_{3,t}, \\ x_{4,t+n} &= -\tilde{\omega}(x_{3,t} - \hat{x}_{3,t}) + \tilde{\sigma}(x_{4,t} - \hat{x}_{4,t}) + v_{4,t}, \\ x_{5,t+n} &= \tilde{\lambda}(x_{5,t} - \hat{x}_{5,t}) + (x_{6,t} - \hat{x}_{6,t}) + v_{5,t}, \\ x_{6,t+n} &= \tilde{\lambda}(x_{6,t} - \hat{x}_{6,t}) + v_{6,t}. \end{aligned} \tag{2.11}$$

We now find conditions for all modes to be stable. We start with the lowest mode. The second moment of $x_{6,t}$ is given by

$$\begin{aligned} \mathbb{E}\big[x_{6,t+n}^2\big] &= \tilde{\lambda}^2 \mathbb{E}\big[(x_{6,t} - \hat{x}_{6,t})\big] + \mathbb{E}\big[v_{6,t}^2\big] \\ &\overset{(a)}{=} \tilde{\lambda}^2 2^{-2C} \mathbb{E}\big[x_{6,t}^2\big] + n_{v,6}, \end{aligned} \tag{2.12}$$

where (a) follows from the linear mean-square estimation of a Gaussian variable over a scalar Gaussian channel of capacity C and $n_{v,k} := \mathbb{E}[v_{k,t}^2]$ for $k = 1, 2, \ldots, 6$. We observe that $\mathbb{E}[x_{6,t}^2]$ is bounded if $\tilde{\lambda}^2 2^{-2C} < 1$. Since $|\tilde{\lambda}_6| = \tilde{\lambda}$, the state component $x_{6,t}$ is stable if

$$|\tilde{\lambda}_6|^2 2^{-2C} < 1 \quad \Rightarrow \quad \log\big(|\tilde{\lambda}_6|\big) < C. \tag{2.13}$$

Now consider $x_{5,t}$, whose second moment can be bounded as

$$\begin{aligned} \mathbb{E}\big[x_{5,t+n}^2\big] &\overset{(a)}{=} \tilde{\lambda}^2 \mathbb{E}\big[(x_{5,t} - \hat{x}_{5,t})^2\big] + 2\tilde{\lambda}\mathbb{E}\big[(x_{5,t} - \hat{x}_{5,t})(x_{6,t} - \hat{x}_{6,t})\big] \\ &\quad + \mathbb{E}\big[(x_{6,t} - \hat{x}_{6,t})^2\big] + n_{v,5} \\ &\overset{(b)}{=} \tilde{\lambda}^2 2^{-2C} \mathbb{E}\big[x_{5,t}^2\big] + 2\tilde{\lambda}\mathbb{E}\big[(x_{5,t} - \hat{x}_{5,t})(x_{6,t} - \hat{x}_{6,t})\big] + 2^{-2C}\mathbb{E}\big[x_{6,t}^2\big] + n_{v,5} \\ &\overset{(c)}{\leq} \tilde{\lambda}^2 2^{-2C} \mathbb{E}\big[x_{6,t}^2\big] + 2\tilde{\lambda}\sqrt{\mathbb{E}\big[(x_{5,t} - \hat{x}_{5,t})^2\big]\mathbb{E}\big[(x_{6,t} - \hat{x}_{6,t})^2\big]} \end{aligned}$$

$$
\begin{aligned}
&+2^{-2C}\mathbb{E}\big[x_{6,t}^2\big]+n_{v,5}\\
&=\tilde{\lambda}^2 2^{-2C}\mathbb{E}\big[x_{5,t}^2\big]+2\tilde{\lambda}\sqrt{2^{-2C}\mathbb{E}\big[x_{5,t}^2\big]}\sqrt{2^{-2C}\mathbb{E}\big[x_{6,t}^2\big]}\\
&\quad+2^{-2C}\mathbb{E}\big[x_{6,t}^2\big]+n_{v,5}\\
&\overset{(d)}{\leq} k_1\mathbb{E}\big[x_{5,t}^2\big]+k_2\sqrt{\mathbb{E}\big[x_{5,t}^2\big]}+k_3, \qquad (2.14)
\end{aligned}
$$

where (a) follows from (2.11); (b) follows from the linear mean-square estimation of a Gaussian variable over a scalar Gaussian channel of capacity C; (c) follows Cauchy–Schwarz inequality; (d) follows from the fact $\mathbb{E}[x_{6,t}^2] < M$ (assuming that (2.13) is satisfied) and by defining $k_1 := \tilde{\lambda}^2 2^{-2C}$, $k_2 := 2\tilde{\lambda}2^{-2C}\sqrt{M}$, and $k_3 := 2^{-2C}M + n_{v,5}$. We now want to a find condition which ensures convergence of the following sequence:

$$\alpha_{t+1} = k_1\alpha_t + k_2\sqrt{\alpha_t} + k_3. \qquad (2.15)$$

In order to show convergence, we make use of the following lemma.

Lemma 2.2 [67, Lemma 6.1] *Let $T : \mathbb{R} \mapsto \mathbb{R}$ be a non-decreasing continuous mapping with a unique fixed point $x^\star \in \mathbb{R}$. If there exists $u \leq x^\star \leq v$ such that $T(u) \geq u$ and $T(v) \leq v$, then the sequence generated by $x_{t+1} = T(x_t)$, $t \in \mathbb{N}$ converges starting from any initial value $x_0 \in \mathbb{R}$.*

Proof The proof is can be found in [67]. □

We observe that the mapping $T(\alpha) = k_1\alpha + k_2\sqrt{\alpha} + k_3$ with $\alpha \geq 0$ is monotonically increasing since $k_1, k_2 > 0$. It will have a unique fixed point $\alpha^\star$ if and only if $k_1 < 1$, since $k_2, k_3 > 0$. Assuming that $k_1 < 1$, there exists $u < \alpha^\star < v$ such that $T(u) \geq u$ and $T(v) \leq v$. Therefore, by Lemma 2.2, the sequence $\{\alpha_t\}$ is convergent if $k_1 = \tilde{\lambda}^2 2^{-2C} < 1 \Rightarrow \log(\tilde{\lambda}) < C$. Since $|\tilde{\lambda}_5| = \tilde{\lambda}$, the state $x_{5,t}$ is stable if

$$\log\big(|\tilde{\lambda}_5|\big) < C. \qquad (2.16)$$

The second moments of $x_{3,t}$ and $x_{4,t}$ are given by

$$
\begin{aligned}
\mathbb{E}\big[x_{3,t+n}^2\big] &= \tilde{\sigma}^2 2^{-2C}\mathbb{E}\big[x_{3,t}^2\big] + \tilde{\omega}^2 2^{-2C}\mathbb{E}\big[x_{4,t}^2\big]\\
&\quad + 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\big[(x_{3,t}-\hat{x}_{3,t})(x_{4,t}-\hat{x}_{4,t})\big] + n_{v,3},\\
\mathbb{E}\big[x_{4,t+n}^2\big] &= \tilde{\omega}^2 2^{-2C}\mathbb{E}\big[x_{3,t}^2\big] + \tilde{\sigma}^2 2^{-2C}\mathbb{E}\big[x_{4,t}^2\big]\\
&\quad - 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\big[(x_{3,t}-\hat{x}_{3,t})(x_{4,t}-\hat{x}_{4,t})\big] + n_{v,4}.
\end{aligned} \qquad (2.17)
$$

By using the above equations, we can write

$$\mathbb{E}\big[x_{3,t+n}^2\big] + \mathbb{E}\big[x_{4,t+n}^2\big] = \big(\tilde{\sigma}^2+\tilde{\omega}^2\big)2^{-2C}\big(\mathbb{E}\big[x_{3,t}^2\big] + \mathbb{E}\big[x_{4,t}^2\big]\big) + n_{v,3} + n_{v,4}.$$

We observe that the sum $\mathbb{E}[x_{3,t+n}^2] + \mathbb{E}[x_{4,t+n}^2]$ is bounded if $(\tilde{\sigma}^2 + \tilde{\omega}^2)2^{-2C} < 1$. Since $|\tilde{\lambda}_3|^2 = |\tilde{\lambda}_4|^2 = (\tilde{\sigma}^2 + \tilde{\omega}^2)^2$, the state components $x_{3,t}$ and $x_{4,t}$ are stable if

$$\log\big(|\tilde{\lambda}_3|\big) < C, \qquad \log\big(|\tilde{\lambda}_4|\big) < C. \tag{2.18}$$

Finally, the second moments of $x_{1,t}$ and $x_{2,t}$ are given by

$$\begin{aligned}
\mathbb{E}\big[x_{1,t+n}^2\big] &= \tilde{\sigma}^2 2^{-2C}\mathbb{E}\big[x_{1,t}^2\big] + \tilde{\omega}^2 2^{-2C}\mathbb{E}\big[x_{2,t}^2\big] + 2^{-2C}\mathbb{E}\big[x_{3,t}^2\big] \\
&\quad + 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{2,t} - \hat{x}_{2,t})\big] + 2\tilde{\sigma}\mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{3,t} - \hat{x}_{3,t})\big] \\
&\quad + 2\tilde{\omega}\mathbb{E}\big[(x_{2,t} - \hat{x}_{2,t})(x_{3,t} - \hat{x}_{3,t})\big] + n_{v,1}, \\
\mathbb{E}\big[x_{1,t+n}^2\big] &= \tilde{\omega}^2 2^{-2C}\mathbb{E}\big[x_{1,t}^2\big] + \tilde{\sigma}^2 2^{-2C}\mathbb{E}\big[x_{2,t}^2\big] + 2^{-2C}\mathbb{E}\big[x_{4,t}^2\big] \\
&\quad - 2\tilde{\sigma}\tilde{\omega}\mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{2,t} - \hat{x}_{2,t})\big] - 2\tilde{\omega}\mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{4,t} - \hat{x}_{4,t})\big] \\
&\quad + 2\tilde{\sigma}\mathbb{E}\big[(x_{2,t} - \hat{x}_{2,t})(x_{4,t} - \hat{x}_{4,t})\big] + n_{v,2}.
\end{aligned} \tag{2.19}$$

By using the above equations, we can write

$$\begin{aligned}
&\mathbb{E}\big[x_{1,t+n}^2\big] + \mathbb{E}\big[x_{2,t+n}^2\big] \\
&\quad = \big(\tilde{\sigma}^2 + \tilde{\omega}^2\big)2^{-2C}\big(\mathbb{E}\big[x_{1,t}^2\big] + \mathbb{E}\big[x_{2,t}^2\big]\big) \\
&\qquad + 2\tilde{\sigma}\big(\mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{3,t} - \hat{x}_{3,t})\big] + \mathbb{E}\big[(x_{2,t} - \hat{x}_{2,t})(x_{4,t} - \hat{x}_{4,t})\big]\big) \\
&\qquad + 2\tilde{\omega}\big(\mathbb{E}\big[(x_{2,t} - \hat{x}_{2,t})(x_{3,t} - \hat{x}_{3,t})\big] - \mathbb{E}\big[(x_{1,t} - \hat{x}_{1,t})(x_{4,t} - \hat{x}_{4,t})\big]\big) \\
&\qquad + n_{v,1} + n_{v,2}.
\end{aligned} \tag{2.20}$$

We can now bound $\mathbb{E}[x_{1,t+n}^2] + \mathbb{E}[x_{2,t+n}^2]$ as

$$\begin{aligned}
\mathbb{E}\big[x_{1,t+n}^2\big] + \mathbb{E}\big[x_{2,t+n}^2\big] &\overset{(a)}{\leq} \big(\tilde{\sigma}^2 + \tilde{\omega}^2\big)2^{-2C}\big(\mathbb{E}\big[x_{1,t}^2\big] + \mathbb{E}\big[x_{2,t}^2\big]\big) \\
&\quad + 4\tilde{\sigma}^2 2^{-2C}\Big(\sqrt{\mathbb{E}\big[x_{1,t}^2\big]\mathbb{E}\big[x_{3,t}^2\big]} + \sqrt{\mathbb{E}\big[x_{2,t}^2\big]\mathbb{E}\big[x_{4,t}^2\big]}\Big) \\
&\quad + 4\tilde{\omega}^2 2^{-2C}\Big(\sqrt{\mathbb{E}\big[x_{2,t}^2\big]\mathbb{E}\big[x_{3,t}^2\big]} + \sqrt{\mathbb{E}\big[x_{1,t}^2\big]\mathbb{E}\big[x_{4,t}^2\big]}\Big) \\
&\quad + n_{v,1} + n_{v,2} \\
&\overset{(b)}{\leq} k_1\big(\mathbb{E}\big[x_{1,t}^2\big] + \mathbb{E}\big[x_{2,t}^2\big]\big) + k_2\sqrt{\mathbb{E}\big[x_{1,t}^2\big] + \mathbb{E}\big[x_{2,t}^2\big]} + k_3,
\end{aligned} \tag{2.21}$$

where (a) follows from the Cauchy–Schwarz inequality; and (b) follows from $\mathbb{E}[x_{3,t}^2] < M$ and $\mathbb{E}[x_{4,t}^2] < M$ (assuming that the condition in (2.18) is satisfied) and by defining $k_1 := (\tilde{\sigma}^2 + \tilde{\omega}^2)2^{-2C}$, $k_2 := 16(\tilde{\sigma}^2 + \tilde{\omega}^2)^2 2^{-2C} M$, and $k_3 :=$

$n_{v,1} + n_{v,2}$. By using Lemma 2.2, we can show that $x_{1,t}$ and $x_{2,t}$ are stable if $k_1 = (\tilde{\sigma}^2 + \tilde{\omega}^2)^2 2^{-2C} < 1$. Since $|\tilde{\lambda}_1|^2 = |\tilde{\lambda}_2|^2 = (\tilde{\sigma}^2 + \tilde{\omega}^2)^2$, we get

$$\log\big(|\tilde{\lambda}_1|\big) < C, \qquad \log\big(|\tilde{\lambda}_2|\big) < C. \tag{2.22}$$

It follows from (2.22), (2.22), and (2.22) that the system is stable if

$$\sum_{i=1}^{n} \log\big(|\tilde{\lambda}_i|\big) = \log\big(|\tilde{A}|\big) < nC.$$

Since $|\tilde{A}| = |T^1 A^n T| = |A|^n$, we have

$$\log\big(|A|\big) < C.$$

Having shown sufficiency of the linear time variant scheme for the system matrix having equal magnitude eigenvalues with algebraic multiplicity, we next consider an example of a system matrix having eigenvalues with unequal magnitude.

Example 2 (Eigenvalues with unequal magnitude) Consider a system matrix A with three eigenvalues, $\lambda_1 \in \mathbb{R}$ and $\lambda_2, \lambda_3 \in \mathbb{C}$ with $|\lambda_1| = |\lambda_2|^2 = |\lambda_3|^2$. For this system, consider the following scheme. The transmission from the sensor to the controller happens periodically, where each transmission period consists for four time slots. In the first two time slots, the state corresponding to λ_1 is transmitted and in the last two slots the states corresponding to λ_2 and λ_3 are transmitted. Note that the sensor is serving more unstable modes more often. In the following, we show again that under this transmit scheme, the system is mean-square stable if $\log(|A|) < C$.

Let us assume that the transmission period starts at t and ends at $t + 4$. At time t, the sensor transmits $x_{1,t}$ and the controller takes action $U_t = [-\lambda_1 \hat{x}_{1,t}, 0, 0]$. At time $t + 1$, the sensor transmits $x_{1,t+1}$ and the controller takes action $U_t = [-\lambda_1 \hat{x}_{1,t+1}, 0, 0]$. At time $t + 2$, the sensor transmits $x_{2,t}$ and the does not take any action. Like in the previous example, we are using such a scheme to make the analysis simpler, although it is better to transmit the most recent state and apply control action as early as possible. At time $t + 3$, the sensor transmits $x_{3,t}$ and the controller takes the following action: $U_t = [0, -\lambda_1 \hat{x}_{2,t}, -\lambda_1 \hat{x}_{2,t}]$. Under this transmit and control scheme, the second moments of $x_{1,t}$ are given by

$$\begin{aligned}
&\mathbb{E}\big[x_{1,t+1}^2\big] = \lambda_1^2 2^{-2C} \mathbb{E}\big[x_{1,t}^2\big] + n_1, \qquad \mathbb{E}\big[x_{1,t+2}^2\big] = \lambda_1^2 2^{-2C} \mathbb{E}\big[x_{1,t+1}^2\big] + n_1, \\
&\mathbb{E}\big[x_{1,t+3}^2\big] = \lambda_1^2 \mathbb{E}\big[x_{1,t+2}^2\big] + n_1, \qquad \mathbb{E}\big[x_{1,t+4}^2\big] = \lambda_1^2 \mathbb{E}\big[x_{1,t+4}^2\big] + n_1,
\end{aligned} \tag{2.23}$$

where n_1 is the variance of the process noise. Using the above equations, the second moment of the state $x_{1,t}$ at the start of each transmission period is given by

$$\mathbb{E}\big[x_{1,t+4}^2\big] = \lambda_1^8 2^{-4C} \mathbb{E}\big[x_{1,t}^2\big] + \tilde{n}_1, \quad t = 4l,\ l \in N, \tag{2.24}$$

where $\tilde{n}_1 = n_1(1+\lambda_1^2 2^{-2C}+\lambda_1^2 2^{-6C}+\lambda_1^2 2^{-8C})$. Similarly, using the approach that was used in the previous example, we can show that

$$\mathbb{E}\big[x_{2,t+4}^2\big]+\mathbb{E}\big[x_{2,t+4}^2\big] = |\lambda_2|^8 2^{-2C}\big(\mathbb{E}\big[x_{2,t+4}^2\big]+\mathbb{E}\big[x_{2,t+4}^2\big]\big)+\tilde{n}_2, \quad t=4l,\ l\in N, \tag{2.25}$$

where $\tilde{n}_2$ is the term due to the process noise. From (2.24) and (2.25) we observe that all modes will be stable if

$$\begin{aligned} &\lambda_1^8 2^{-4C}<1, \qquad \lambda_2^8 2^{-2C}<1, \qquad \lambda_3^8 2^{-2C}<1 \\ \Rightarrow\quad &\log\big(|\lambda_1|\big)<\frac{1}{2}C, \qquad \log\big(|\lambda_2|\big)=\log\big(|\lambda_3|\big)<\frac{1}{4}C \\ \Rightarrow\quad &\sum_{i=1}^{n}\log\big(|\lambda_i|\big)=\log\big(|A|\big)<C. \end{aligned} \tag{2.26}$$

For a general n-dimensional system, the transmit scheme can be generalized as follows: Choose k, k_m, such that $\frac{k_m}{k}=\frac{\log(|\lambda_m|)}{\sum_{i=1}^{n}\log(|\lambda_i|)}$ for $m=1,2,\ldots,n$. The sensor transmits periodically with a period equal to k time slots, in which the state x_m corresponding to λ_m is transmitted k_m times. Note that $\sum_{m=1}^{n}k_m=k$. The system will be stable if $\log(|\lambda_m|)<\frac{\log(|\lambda_m|)}{\sum_{i=1}^{n}\log(|\lambda_i|)}C$ for all $m\in\{1,2,\ldots,n\}$, which is equivalent to $\sum_{i=1}^{n}\log(|\lambda_i|)=\log(|A|)<C$. □

Remark 2.1 Although we have proved Theorem 2.2 for Gaussian distributed initial states, it is also valid for other distributions with finite variance. For any non-Gaussian distributed initial state with finite variance, we can use the approach in [64, Sect. IV] to make the state process Gaussian distributed and then the schemes discussed earlier in this section can be applied.

According to Theorem 2.2, there is no loss in the mean-square stabilizability by restricting the sensing and control scheme to be linear. This makes linear policies a good choice for stabilization over scalar Gaussian channels. Notice, while deriving an achievable stability region, the objective was to keep the second moment of the state process bounded and we did not aim at minimizing the second moment. One might be interested in minimizing the second moment of the state process over a finite or an infinite time horizon. In the following, we consider a finite horizon stabilization problem and present an optimal linear scheme.

2.3.1.3 An Optimal Linear Scheme for Stabilization

Consider a linear system with diagonalizable system matrix A and diagonal K_W. For this linear system, we derive an optimal linear time varying sensing policy $E_t^\star$ which minimizes the following cost: $\sum_{i=1}^{t_f}\mathbb{E}[\|X_t\|^2]$. We have restricted the matrices A and K_W to be diagonal for the ease of analysis. The optimal time varying sensing scheme is presented in the following theorem.

Theorem 2.3 *Let* $\tilde{G} := [\sqrt{P}, 0, 0, \ldots, 0]$, $K_t = A^T(I + K_{t+1})A(I - \tilde{G}^T\tilde{G}(\frac{1}{N+P}))$ *with* $K_{t_f} = 0$, *and* π_t *be a unitary matrix such that* $\pi_t^T(\Lambda_t^{\frac{1}{2}}A^T(I + K_{t+1})A\Lambda_t^{\frac{1}{2}})\pi_t = \mathrm{diag}(\upsilon_{1,t}, \ldots, \upsilon_{N,t})$ *with* $\upsilon_{1,t} \geq \upsilon_{2,t} \geq \cdots > 0$. *The optimal linear time varying sensing is given by* $E_t^\star = \tilde{G}\pi_t\Lambda_t^{-\frac{1}{2}}$.

Proof We rewrite the Riccati equation (2.6) as

$$\begin{aligned}\Lambda_{t+1} &= A\Lambda_t^{\frac{1}{2}}\left(I - \frac{\Lambda_t^{\frac{1}{2}}E_t^T}{\sqrt{N}}\left[\frac{E_t\Lambda_t E_t^T}{\sqrt{N}} + 1\right]^{-1}\frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}\right)\Lambda_t^{\frac{1}{2}}A^T + K_W \\ &\overset{(a)}{=} A\Lambda_t^{\frac{1}{2}}\left(I - C_t^T\left[C_tC_t^T + 1\right]^{-1}C_t\right)\Lambda_t^{\frac{1}{2}}A^T + K_W \\ &\overset{(b)}{=} A\Lambda_t^{\frac{1}{2}}\left[I + C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W, \end{aligned} \tag{2.27}$$

where (a) follows from $C_t := \frac{E_t\Lambda_t^{\frac{1}{2}}}{\sqrt{N}}$; and (b) follows from the matrix inversion lemma $[I + UWV]^{-1} = I - U[W^{-1} + VU]^{-1}V$, by choosing $U = C_t^T$, $W = 1$, $V = C_t$.

The finite horizon optimal stabilization problem can be stated as

$$\left\{C_i^\star\right\}_{i=0}^{t_f-1} = \underset{\{C_i\}_{i=0}^{t_f-1}: C_tC_t^T \leq \frac{P}{\sigma^2}}{\arg\min} \sum_{t=0}^{t_f-1} \mathrm{tr}[\Lambda_{t+1}],$$

subject to

$$\Lambda_{t+1} = A\Lambda_t^{\frac{1}{2}}\left[I + C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W. \tag{2.28}$$

This is a nonlinear dynamic optimization problem. In order to solve this problem, we follow a dynamic programming approach. Such an approach has also been considered for continuous time systems in [7]. At any time t, let the value function be $V_t(\Lambda_t) = \mathrm{tr}[K_t\Lambda_t + L_t]$. We have to find C_t such that

$$\begin{aligned}&\mathrm{tr}[K_t\Lambda_t + L_t] \\ &= \min_{C_t: C_tC_t^T \leq \frac{P}{N}} \left\{\mathrm{tr}[\Lambda_{t+1}] + \mathrm{tr}[K_{t+1}\Lambda_{t+1} + L_{t+1}]\right\} \\ &= \min_{C_t: C_tC_t^T \leq \frac{P}{N}} \mathrm{tr}\left[(I + K_{t+1})\Lambda_{t+1} + L_{t+1}\right] \\ &\overset{(a)}{=} \min_{C_t: C_tC_t^T \leq \frac{P}{N}} \mathrm{tr}\left[(I + K_{t+1}) \times \left(A\Lambda_t^{\frac{1}{2}}\left[I + C_t^TC_t\right]^{-1}\Lambda_t^{\frac{1}{2}}A^T + K_W\right) + L_{t+1}\right] \\ &\overset{(b)}{=} \mathrm{tr}\left[(I + K_{t+1})K_W + L_{t+1}\right]\end{aligned}$$

$$
\begin{aligned}
&+ \min_{C_t: C_t C_t^T \le \frac{P}{N}} \operatorname{tr}\big[(I + K_{t+1}) A \Lambda_t^{\frac{1}{2}} \big[I + C_t^T C_t\big]^{-1} \Lambda_t^{\frac{1}{2}} A^T\big] \\
&= \operatorname{tr}\big[(I + K_{t+1}) K_W + L_{t+1}\big] \\
&\quad + \min_{C_t: C_t C_t^T \le \frac{P}{N}} \operatorname{tr}\big[\Lambda_t^{\frac{1}{2}} A^T (I + K_{t+1}) A \Lambda_t^{\frac{1}{2}} \big[I + C_t^T C_t\big]^{-1}\big] \\
&\overset{(c)}{=} \operatorname{tr}\big[(I + K_{t+1}) K_W + L_{t+1}\big] + \operatorname{tr}\big[\Lambda_t^{\frac{1}{2}} A^T (I + K_{t+1}) A \Lambda_t^{\frac{1}{2}} \big[I + \pi_t G^T G \pi_t^T\big]^{-1}\big] \\
&\overset{(d)}{=} \operatorname{tr}\big[(I + K_{t+1}) K_W + L_{t+1}\big] \\
&\quad + \operatorname{tr}\big[\Lambda_t^{\frac{1}{2}} A^T (I + K_{t+1}) A \Lambda_t^{\frac{1}{2}} \big(I - \pi_t G^T \big(1 + G \pi_t^T \pi_t G^T\big)^{-1} G \pi_t^T\big)\big] \\
&\overset{(e)}{=} \operatorname{tr}\big[(I + K_{t+1}) K_W + L_{t+1}\big] \\
&\quad + \operatorname{tr}\left[A^T (I + K_{t+1}) A \left(\Lambda_t - \Lambda_t^{\frac{1}{2}} \pi_t G^T G \pi_t^T \Lambda_t^{\frac{1}{2}} \left(1 + \frac{P}{N}\right)^{-1}\right)\right] \\
&\overset{(f)}{=} \operatorname{tr}\big[(I + K_{t+1}) K_W + L_{t+1}\big] \\
&\quad + \operatorname{tr}\left[A^T (I + K_{t+1}) A \left(I - G^T G \left(\frac{N}{N+P}\right)\right) \Lambda_t\right],
\end{aligned}
\tag{2.29}
$$

where (a) follows by substituting Λ_{t+1} using (2.28); (b) follows from the fact that K_{t+1} and L_{t+1} do not depend on C_t; (c) follows from the fact that according to [6] the unique solution to the trace minimization problem unique minimizer is given by $C_t^\star = G \pi_t^T$, where $G := [\sqrt{\frac{P}{N}}, 0, 0, \ldots, 0]$, and π_t is a unitary matrix which diagonalizes $(\Lambda_t^{\frac{1}{2}} A^T (I + K_{t+1}) A \Lambda_t^{\frac{1}{2}})$ such that $\pi_t^T (\Lambda_t^{\frac{1}{2}} A^T (I + K_{t+1}) A \Lambda_t^{\frac{1}{2}}) \pi_t = \operatorname{diag}(\upsilon_{1,t}, \ldots, \upsilon_{N,t})$ with $\upsilon_{1,t} \ge \upsilon_{2,t} \ge \cdots > 0$; (d) follows from the matrix inversion lemma, $[I + UWV]^{-1} = I - U[W^{-1} + VU]^{-1} V$, by choosing by choosing $V = G \pi_t^T$, $W = 1$, $U = \pi_t G^T$; (e) follows from $\pi_t \pi_t^T = I$ and $G G^T = \frac{P}{N}$; and (f) follows from the assumption that A and Λ_t are diagonal, which implies that K_{t+1} and π_t are also diagonal. (Diagonality of K_{t+1} will become clear shortly.) Therefore, we have $\Lambda_t^{\frac{1}{2}} \pi_t G^T G \pi_t^T \Lambda_t^{\frac{1}{2}} = \Lambda_t^{\frac{1}{2}} G^T G \Lambda_t^{\frac{1}{2}} = G^T G \Lambda_t$, since $G^T G$ is diagonal. In order to satisfy the above equality (2.29), we choose $K_{t_f} = L_{t_f} = 0$ and $\{K_{t+1}, L_{t+1}\}$ according to

$$
\begin{aligned}
K_t &= A^T (I + K_{t+1}) A \left(I - G^T G \left(\frac{N}{N+P}\right)\right), \\
L_t &= (I + K_{t+1}) K_W + L_{t+1}.
\end{aligned}
\tag{2.30}
$$

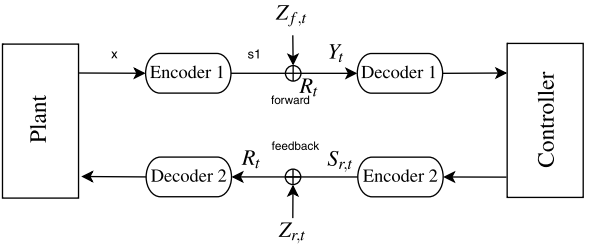


Fig. 2.3 Stabilization over noisy forward and reverse Gaussian channels

We can observe that K_t is also diagonal if A and K_W are diagonal, since $G^T G$ is diagonal. We have found the optimal $C_t^\star = G\pi_t$ and we know that $C_t = \frac{E_t \Lambda_t^{\frac{1}{2}}}{\sqrt{N}}$; therefore, $E_t^\star = \tilde{G}\pi_t \Lambda_t^{-\frac{1}{2}}$ where $\tilde{G} := \sqrt{N}G = [\sqrt{P}, 0, 0, \ldots, 0]$. □

2.3.1.4 Noisy Feedback Link

So far we have considered the communication link from the controller to the plant to be noiseless. Noiseless communication link from the controller to the plant can be a good assumption for certain scenarios in which the controller is either connected to the plant via a cable or the controller has very large power to spend for transmission of signals over the air. However, in some practical situations, it may not be reasonable to model the communication link from the controller to the plant as noiseless. In these situations, the remote controller can be equipped with an encoder to encode the control actions before transmitting them over a noisy channel and the remotely located actuator can be equipped with a decoder, to decode the control actions using the signal received over the noisy channel. In Fig. 2.3, such a setup is shown, where there are two encoders to encode the plant's state and the control actions with average transmit powers $\mathbb{E}[(S_{f,t})^2] = P_f$ and $\mathbb{E}[(S_{r,t})^2] = P_r$, respectively. The forward and the reverse channels are disturbed by white Gaussian noises $Z_{f,t} \sim \mathcal{N}(0, N_f)$ and $Z_{r,t} \sim \mathcal{N}(0, N_r)$, respectively. Thus the capacities of the forward channel and the reverse channel are given by $C_f := \frac{1}{2}\log(1 + \frac{P_S}{N_f})$ and $C_r := \frac{1}{2}\log(1 + \frac{P_r}{N_r})$. In the following, present necessary and sufficient conditions for the mean-square stability over the given channel.

Theorem 2.4 *The linear system* (2.1) *can be mean-square stabilized over a noisy forward channel with capacity* C_f *and a noisy reverse channel with capacity* C_r *only if*

$$\log\big(|A|\big) \leq \min\{C_f, C_r\}. \tag{2.31}$$

Proof The system in Fig. 2.3 can be viewed as the system depicted in Fig. 2.4, where the encoders, decoders, and controllers are viewed as decision makers. The encoder in the forward link (Encoder 1) is the first decision maker DM1. The decoder in

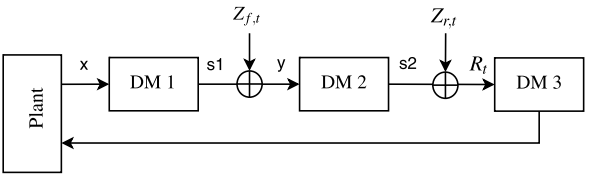


Fig. 2.4 Another representation of the system model in Fig. 2.3

the forward link (Decoder 1), the controller, and the encoder in the reverse link (Encoder 2) are altogether can be viewed as second decision maker DM2. Finally, the DM3 represents decoder of the reverse link (Decoder 2). For the system shown in Fig. 2.4, we know from Lemma 2.1 that the following condition is necessary for stabilization:

$$\log\big(|A|\big) \leq \liminf_{T\to\infty} \frac{1}{T} I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right). \tag{2.32}$$

According to the proof of Theorem 3.1 in [67], the directed information can be bounded as

$$\begin{aligned} I\left(\bar{X}_{[0,T-1]} \to R_{[0,T-1]}\right) &\leq \min\left\{\sum_{t=0}^{T-1} I\left(S_{f,t}; Y_t\right), \sum_{t=0}^{T-1} I\left(S_{r,t}; R_t\right)\right\} \\ &\leq \frac{1}{2}\min\left\{\log\left(1+\frac{P_f}{N_f}\right), \log\left(1+\frac{P_r}{N_r}\right)\right\} \\ &= \min\{C_f, C_r\}, \end{aligned} \tag{2.33}$$

where the last equality follows from the definition of channel capacities. By using (2.33) in (2.32), we get (2.31). □

Theorem 2.4 shows that the reliability of the reverse channel is as important as the forward channel. The necessary condition (2.31) was first obtained for memoryless sensors and controller in [56, Theorem 8.1]. In [56], the authors have also obtained a sufficient condition for stabilization by restricting the encoders and the decoders to be linear and memoryless, which is stated in the following theorem.

Theorem 2.5 [56, Theorem 8.1] *The linear system* (2.1) *can be mean-square stabilized over a Gaussian forward channel with capacity C_f and a Gaussian reverse channel with capacity C_r using a linear memoryless sensing and control scheme if*

$$\log\big(|A|\big) \leq \frac{1}{2}\log\left(\frac{1}{2^{-2C_f}+2^{-2C_r}-2^{-2(C_f+C_r)}}\right). \tag{2.34}$$

In Fig. 2.5, we have fixed forward channel capacity $C_f = 5$ bits/channel use, and have plotted stability region achievable with linear memoryless scheme as a function of reverse channel capacity C_r using (2.34). The figure shows that an LTI plant with system matrix A is stabilizable with linear scheme if $\log(|A|)$ is below the stability curve drawn in the figure. For the sake of comparison, the outer bound on the

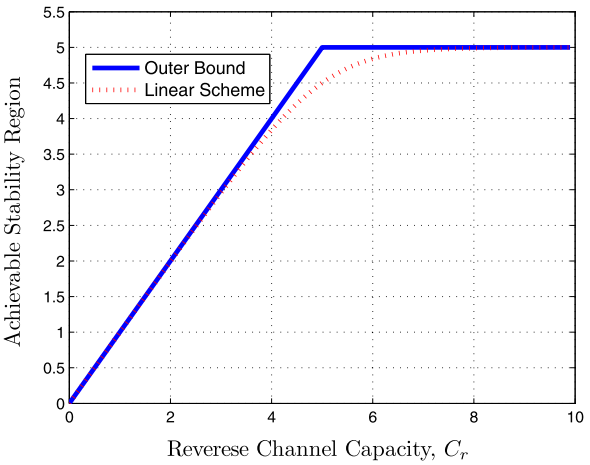


Fig. 2.5 Stability Region

achievable stability region is also plotted according to (2.31). One can observe that the linear memoryless scheme is quite efficient because its performance gets close to optimal as one of the two links becomes relatively more reliable. In Sect. 2.5, we discuss sub-optimality of linear policies for estimation over multi-hop relay networks. It is shown that a simple three-level quantizer policies can outperform the best linear policy even over a two-hop network. The setup of noisy forward and noisy reverse channel can be viewed as a two-hop network, as illustrated in Fig. 2.4. Use of such nonlinear schemes in forward and reverse channels can potentially improve stability of the closed-loop system; however, this is yet to explored. Control over noisy forward and reverse channels have been considered also for more general channels in [56].

In the remainder of the chapter, we keep the assumption of noiseless link from the controller to the plant, in order to simplify the analysis and to avoid tedious computations. As long as the encoders and decoders are linear and the channels are modeled as Gaussian, the nature of the problem does not change and one can obtain stability results for noisy reverse channels with some straightforward analysis.

2.3.2 Schemes for Vector Channel

In the previous section, we showed that linear schemes can achieve the minimum signal-to-noise required for stabilization of multi-dimensional LTI system over a scalar Gaussian channel. That is there is no rate loss in restricting the scheme to be linear. In fact, for the stabilization of scalar plant over a scalar Gaussian channel, linear policies are optimal. However, for transmission over vector channels, linear schemes may not be good enough. It is known from the information theory literature [38] that a distributed joint source–channel code is optimal in the MMSE sense, if the following conditions hold: (i) The information transmitted on all available channels is independent, (ii) Capacity is utilized by all channels (source-channel needs

to be matched). By using any linear scheme, it is not possible to make the transmitted signals on parallel channels independent. However, independent signals can be transmitted on parallel channels by employing nonlinear schemes. Some nonlinear sensing schemes for stabilization and control of a scalar system over parallel Gaussian channels are given in [24, 58]. In [58], the authors considered two parallel channels and proposed to send the magnitude of the observed state process on one channel and the phase value (plus or minus) on the second channel. The phase and magnitude of a signal are shown to be independent, thus satisfying the first condition of optimality. Although the second condition of optimality is not met, the proposed nonlinear schemes outperforms the best linear scheme. In [24], the authors proposed to use a hybrid digital–analog scheme, in which the state process is quantized and the quantized signal is transmitted on one channel and quantization error is transmitted on the other channel. This scheme can be extended to arbitrary number of parallel channels and it achieves the minimum signal-to-noise ratio requirement for the mean-square stabilization of a scalar noiseless plant.

In order to demonstrate inefficiency of linear schemes over vector channels, let us consider the following example.

Example Consider a scalar LTI system that has to be stabilized over M parallel white Gaussian channels. Assume that the sensor has an average transmit power constraint P_S and all channels are disturbed by noises of equal power, i.e., $z_{i,t} \sim \mathcal{N}(0, N)$ for all $i \in \{1, 2, \ldots, M\}$. It can be easily shown that the system can be stabilized over the given channel by a linear sensing and control scheme if $\log(\lambda) < \frac{1}{2}(1 + \frac{P}{N})$. However, with the nonlinear scheme proposed in [24], the achievable stability region is given by $\log(\lambda) < \frac{M}{2}(1 + \frac{P}{MN})$. The stability regions achieved by linear and nonlinear schemes have been plotted in Fig. 2.6 for $P = 1$ and $M = 2, 10,$ and 100. Note that in the given example, the achievable stability region of linear scheme is independent of the number of available parallel channels M. But with the nonlinear scheme, the stability region significantly enlarges as the number of parallel channels increases. This example shows that linear schemes can be very inefficient in some parallel channel settings.

The nonlinear scheme proposed in [24] works for scalar plants. For stabilization of a multi-dimensional plant, the non-linear scheme of [24, 58] can be used together with the time varying (mode-by-mode transmission) scheme discussed in Sect. 2.3.1.2. Such a nonlinear time varying scheme can achieve the minimum rate required for stabilization. An interesting open problem is to determine tight conditions for optimality of linear scheme for stabilization of multi-dimensional systems over vector Gaussian channel. In the following, we state a sufficient condition for optimality of a linear time varying scheme proposed in [68].

Theorem 2.6 [68, Theorem 3.2] *A linear time varying scheme is optimal for mean-square stabilizing an n-dimensional plant over m parallel Gaussian channels if*

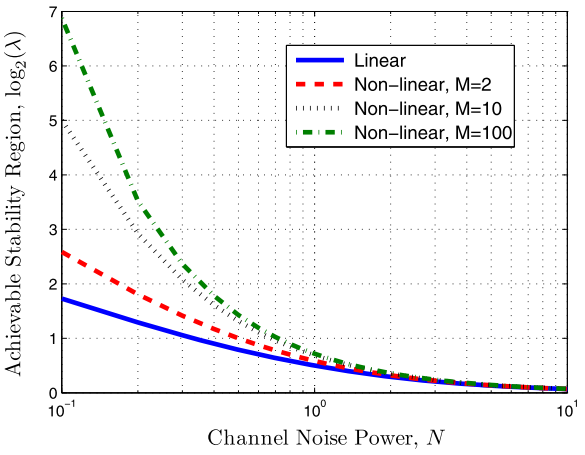


Fig. 2.6 Comparison of linear and nonlinear schemes

there exist $f_{ij} \in \mathbb{Q}$ *such that* $f_{ij} \geq 0$, $\sum_{j=1}^{m^\star} f_{ij} \leq 1$, $\sum_{i=1}^{n} f_{ij} = 1$ *and*

$$\log\bigl(|\lambda_i|\bigr) < \sum_{j=1}^{m^\star} \frac{f_{ij}}{2} \log\left(1 + \frac{P_j^\star}{N_j}\right),$$

for all $i \in \{1, 2, \ldots, n\}$ *and* $j \in \{1, 2, \ldots, m^\star\}$, *where* $P_j^\star$ *is the optimal power allocation given by the water-filling solution* [46, *pp.* 204–205] *and* $m^\star \leq m$ *is the number of active channels for which optimal transmit power is nonzero.*

Some relevant works on the source–channel matching and optimality of linear estimation can be found in [1, 15, 25, 34, 45, 47, 51].

2.4 Stabilization over Relay Channels

In this section, we study stabilization of linear systems over Gaussian relay channels. The basic relay channel consists of one sender (source), one receiver (destination), and an intermediate node (relay) whose sole purpose is to help the communication between the source and the destination [10]. The basic three node relay channel is a basic block of a large sensor network where a group of sensor nodes cooperate to communicate information from a source to a destination. In order to understand the problem of stabilization over a general relay network, we study some basic relay network topologies such as non-orthogonal relay channel and orthogonal relay channel. By the orthogonality of the relay channel we mean that the signal spaces of the encoder and the relay are orthogonal. For example, if the source node and the relay node transmit in disjoint frequency bands or non-overlapping time slots, then the relay channel is considered to be orthogonal. These topologies serve

as the basic building blocks of a large network. In practice, the relay node can be either half-duplex or full-duplex. A node which is capable of transmitting and receiving signals simultaneously using the same frequency band is known as full-duplex while a half-duplex node cannot simultaneously receive and transmit signals. It is expensive and hard to a build a communication device which can transmit and receive signals at the same time using the same frequency, due to the self-interference created by the transmitted signal to the received signal. Therefore, half-duplex systems are mostly used in practice.

The problem of control over a basic three node Gaussian relay channel was first introduced in [59, 61], some sufficient conditions for the mean-square stability were derived. Further related work on control over noisy relay channels can be found in [23, 64].

We know from [36] that the concept of Shannon capacity is not sufficient to characterize moment stability. Moreover, even for the general three node Gaussian relay channel, a single-letter expression for Shannon capacity is still not known. In [14], Gastpar and Vetterli determined capacity of a particular large Gaussian relay network in the limit as the number of relays tends to infinity. The achievable information rate over the relay channel depends on the processing strategy of the relay. The most well known relaying strategies are amplify-and-forward (AF), compress-and-forward, and decode-and-forward [22]. The AF strategy is well suited for delay sensitive control applications and is therefore addressed here.

In this section, we discuss the mean-square stabilization of the system in (2.1) over some fundamental relay channels such as non-orthogonal half-duplex relay channels, non-orthogonal full-duplex relay channel, and orthogonal relay channel in Sects. 2.4.1, 2.4.2, and 2.4.3, respectively. For each relay channel, we present necessary conditions and sufficient conditions for stabilization. In Sect. 2.4.4, we briefly compare achievable stability regions using linear schemes over these basic relay channels.

2.4.1 Non-orthogonal Half-duplex Relay

Consider a non-orthogonal half-duplex Gaussian relay channel shown in Fig. 2.7. A sensor node (state encoder) $\mathcal{E}$ senses the state of the plant and transmits it to the remote controller, while another sensor node $\mathcal{R}$ acts as a relay to support communication from $\mathcal{E}$ to the controller. The state encoder and the relay transmit in the same frequency band and the relay node is assumed to be half-duplex, i.e., it cannot transmit and receive signals simultaneously. In [64], the authors proposed a transmission protocol having two transmission phases, as shown in Fig. 2.7. The signals transmitted by $\mathcal{E}$ and $\mathcal{R}$ are denoted as $S_{e,t}$ and $S_{r,t}$, respectively. The variables $Z_{r,t}$ and Z_t denote two mutually independent white Gaussian noise components with zero mean and variances N_r and N, respectively. In the first transmission phase (odd time steps), $\mathcal{E}$ transmits signal with an average power $2\beta P_S$, where $0 < \beta \leq 1$ is a parameter that distributes power between the two transmission phases. In this transmission

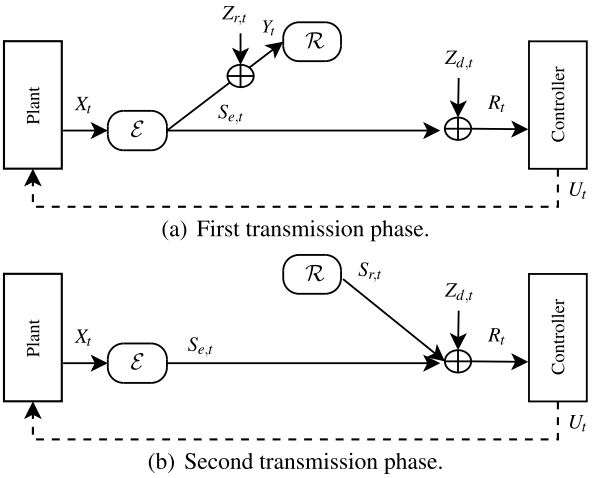


Fig. 2.7 Half-duplex AWGN relay channel

phase, the relay $\mathcal{R}$ receives a noisy signal Y_t from the encoder but it does not transmit any signal. In the second transmission phase (even time steps), both the encoder $\mathcal{E}$ and the relay $\mathcal{R}$ transmit with average powers $2(1-\beta)P_S$ and P_r, respectively. The relay node employs amplify-and-forward (linear) strategy, where amplification at the relay is done under an average power constraint P_R. The multiplicative gain of the $\mathcal{E}$–$\mathcal{D}$ link is assumed to be a constant $h \in \mathbb{R}$ and the gain of the $\mathcal{R}$–$\mathcal{D}$ link is assumed to be one, without loss of generality. The presence of relay node can be more useful in scenarios where the direct link is weaker, i.e., $|h|$ is small. The controller in the first transmission phase receives $R_t = hS_{e,t} + Z_t$ and in the second phase receives $R_t = hS_{e,t} + S_{r,t} + Z_t$. At any time, the controller estimates the present state of the plant using all the signals it has received so far, and then takes an action to stabilize the plant using its state estimate. In the following, we discuss a linear control and communication scheme based on the above transmission protocol and give necessary and sufficient conditions for stabilization. The communication and control scheme is presented in Sect. 2.4.1.1 and the mean-square stability of the system under the given scheme is analyzed in Sect. 2.4.1.2.

2.4.1.1 Sensing and Control Scheme

The control and communication scheme has an initialization step, which is done to make the state distribution Gaussian. This initialization step works as follows. The encoder $\mathcal{E}$ observes X_0 and transmits $S_{e,0} = \sqrt{\frac{P_S}{\alpha_0}} X_0$. The controller $\mathcal{D}$ receives $R_0 = hS_{e,0} + Z_0$ and estimates the initial state as

$$\hat{X}_0 = \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}} R_0 = X_0 + \frac{1}{h}\sqrt{\frac{\alpha_0}{P_S}} Z_0.$$

The controller $\mathcal{C}$ then takes an action $U_0 = -\lambda \hat{X}_0$ which results in

$$X_1 = \lambda X_0 + U_0 + W_0 = \lambda(X_0 - \hat{X}_0) + W_0 = -\frac{\lambda}{h}\sqrt{\frac{\alpha_0}{P_S}}Z_0 + W_0. \tag{2.35}$$

The new plant state X_1 is Gaussian distributed with zero mean and variance $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0 + n_w$. This initialization step is not required if the initial state is already Gaussian distributed. After the initialization, further transmissions are divided into two separate transmission phases as discussed earlier. In the first transmission phase, i.e., for $t = 1, 3, 5, \ldots$, the encoder $\mathcal{E}$ transmits $S_{e,t} = \sqrt{\frac{2\beta P_S}{\alpha_t}}X_t$ to the controller. The relay $\mathcal{R}$ operates in the receiving mode, i.e., it receives the signal transmitted by $\mathcal{E}$. The controller $\mathcal{C}$ observes $R_t = hS_{e,t} + Z_t$ and computes the MMSE estimate of X_t based on $R_{[0,t]}$. It can be shown that $\mathbb{E}[X_t R_{t-j}] = 0$ for $j \geq 1$; therefore, the optimal MMSE estimator uses only the latest received signal R_t to estimate the state. Further, the optimal estimator is linear in the received signal due to the Gaussian distribution. The optimal MMSE state estimate is computed as $\hat{X}_t = (\frac{h\sqrt{2\beta P_S \alpha_t}}{2h^2\beta P_S + N})R_t$. Based on the estimate $\hat{X}_t$, the controller $\mathcal{C}$ takes an action $U_t = -\lambda \hat{X}_t$ which results in $X_{t+1} = \lambda(X_t - \hat{X}_t) + W_t$. The new plant state X_{t+1} is a linear combination of zero mean Gaussian variables $\{X_t, \hat{X}_t, W_t\}$; therefore, it is also zero mean Gaussian. The variance of X_{t+1} can be computed as

$$\alpha_{t+1} := \mathbb{E}\big[X_{t+1}^2\big] = \lambda^2 \mathbb{E}\big[(X_t - \hat{X}_t)^2\big] + \mathbb{E}\big[W_t^2\big] = \lambda^2\left(\frac{N}{2h^2\beta P_S + N}\right)\alpha_t + n_w. \tag{2.36}$$

In the second transmission phase, i.e., for $t = 2, 4, 6, \ldots$, the encoder $\mathcal{E}$ transmits $S_{e,t} = \sqrt{\frac{2(1-\beta)P_S}{\alpha_t}}X_t$ to the controller. The relay now operates in the transmitting mode. It amplifies the previously received signal under the average transmit power constraint and transmits the following signal to the controller,

$$S_{r,t} = \sqrt{\frac{P_r}{(2\beta P_S + N_r)}}(S_{e,t-1} + Z_{r,t-1}).$$

The controller $\mathcal{C}$ thus receives a linear combination of the signal transmitted from the relay and the state encoder. The signal received at the controller is given by

$$R_t = hS_{e,t} + S_{r,t} + Z_t = L_1 X_t + L_2 X_{t-1} + \tilde{Z}_t, \tag{2.37}$$

where $L_1 = \sqrt{\frac{2(1-\beta)h^2 P_S}{\alpha_t}}$, $L_2 = \sqrt{\frac{2\beta P_S P_r}{(2\beta P_S + N_r)\alpha_{t-1}}}$, and $\tilde{Z}_t = Z_t + \sqrt{\frac{P_r}{2\beta P_S + N_r}}Z_{r,t-1}$ with $\tilde{Z}_t \sim \mathcal{N}(0, \tilde{N}(\beta, P_r))$. Next, the controller computes the MMSE estimate of X_t given all previous channel outputs $R_{[0,t]}$ in the following three steps: (i) Compute the MMSE prediction of R_t from $R_{[0,t]}$ as $\hat{R}_t = L_2 \hat{X}_{t-1}$, where $\hat{X}_{t-1}$ is the MMSE estimate of X_{t-1}, (ii) Compute the innovation as $I_t = R_t - \hat{R}_t$, and (iii) Estimate

the state using only the innovation I_t as $\hat{X}_t = \mathbb{E}[X_t|I_t]$. Note that this is the optimum MMSE estimate since X_t is independent of $\{R_1, R_2, \ldots, R_{t-1}\}$ due to the orthogonality property of MMSE estimation. The optimal MMSE state estimate is computed as

$$\hat{X}_t = \mathbb{E}[X_t|I_t] = \frac{\lambda(\lambda L_1 + L_2)\alpha_t}{(\lambda L_1 + L_2)^2\alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)} I_t.$$

Based on the state estimate, the controller $\mathcal{C}$ takes an action $U_t = -\lambda \hat{X}_t$ that results in $X_{t+1} = \lambda(X_t - \hat{X}_t) + W_t$. The new plant state X_{t+1} is a linear combination of zero mean Gaussian variables $\{X_t, \hat{X}_t, W_t\}$; therefore, it is also zero mean Gaussian distributed. The variance of the new plant state X_{t+1} follows from simple computations as

$$\begin{aligned}
\alpha_{t+1} &= \lambda^2 \mathbb{E}\big[(X_t - \hat{X}_t)^2\big] + \mathbb{E}\big[W_t^2\big] \\
&= \lambda^2 \alpha_t \left(\frac{L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)}{(\lambda L_1 + L_2)^2 \alpha_t + L_2^2 n_w + \lambda^2 \tilde{N}(\beta, P_r)} \right) + n_w \\
&= \lambda^2 \big(\lambda^2 k \alpha_{t-1} + n_w\big) \\
&\quad \times \left(\frac{(\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)}{(k_2 + \sqrt{k_1 k + \frac{n_w k_1}{\lambda^2}\frac{1}{\alpha_{t-1}}})^2 + (\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-1}} + \tilde{N}(\beta, P_r)} \right) + n_w, \qquad (2.38)
\end{aligned}$$

where the last equality follows by substituting α_t from (2.36) and by defining $k := \frac{N}{2h^2\beta P_S + N}$, $k_1 := \frac{2\beta P_S P_r}{2\beta P_S + N_r}$, $k_2 := q\sqrt{2h^2(1-\beta)P_S}$. Having presented the sensing and control scheme, we now discuss stability of the plant under the given scheme.

2.4.1.2 Stability Analysis

We wish to find the values of the system parameter λ for which the second moment of the state remains bounded, i.e., the sequence $\{\alpha_t\}$ has to be bounded. Rewriting (2.36) and (2.38), the variance of the state at any time t is given by

$$\alpha_t = \lambda^2 \left(\frac{N}{2h^2\beta P_S + N} \right) \alpha_{t-1} + n_w, \quad t = 2, 4, 6, \ldots, \qquad (2.39)$$

$$\alpha_t = \lambda^2 \big(\lambda^2 k \alpha_{t-2} + n_w\big) f(\alpha_{t-2}) + n_w, \quad t = 3, 5, 7, \ldots, \qquad (2.40)$$

where

$$\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0 + n_w \quad \text{and}$$

$$f(\alpha_{t-2}) \triangleq \left(\frac{(\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-2}} + \tilde{N}(\beta, P_r)}{(k_2 + \sqrt{k_1 k + \frac{n_w k_1}{\lambda^2}\frac{1}{\alpha_{t-2}}})^2 + (\frac{n_w k_1}{\lambda^2})\frac{1}{\alpha_{t-2}} + \tilde{N}(\beta, P_r)} \right).$$

If the odd indexed sub-sequence $\{\alpha_{2t+1}\}$ in (2.40) is bounded, then the even indexed sub-sequence $\{\alpha_{2t}\}$ in (2.39) is also bounded. Therefore, it is sufficient to consider the odd indexed sub-sequence $\{\alpha_{2t+1}\}$. A complicated structure of $f(\alpha_t)$ in (2.40) makes it difficult to find a condition on λ for which this sequence is bounded. Therefore, in [64], we use the following approach. We construct a sequence $\{\alpha'_t\}$ which upper bounds the sub-sequence $\{\alpha_{2t+1}\}$ and is easier to analyze. Then we derive a condition on the system parameter λ for which the sequence $\{\alpha'_t\}$ converges to a limit point as $t \to \infty$, and consequently the boundedness of $\{\alpha_{2t+1}\}$ is guaranteed. We will show later that there is no loss in considering the majorizing sequence instead of the original sequence. The detailed analysis is given in [64], and here we merely give the condition under which the system is stable:

$$\lambda^4 < \left(\frac{(k_2 + \sqrt{k_1 k})^2 + \tilde{N}(\beta, P_r)}{k\tilde{N}(\beta, P_r)}\right) \tag{2.41}$$

$$\Rightarrow \quad \log(\lambda) \leq \frac{1}{4}\left(\log\left(1 + \frac{2h^2\beta P_S}{N}\right) + \log\left(1 + \frac{\tilde{M}(\beta, P_r)}{\tilde{N}(\beta, P_r)}\right)\right), \tag{2.42}$$

where in the last equality we substituted $k = \frac{N}{2h^2\beta P_S + N}$ and $M(\beta, P_r) = (k_2 + \sqrt{k_1 k})^2$ in order to show the dependencies on the average relay power P_r and the power allocation parameter β at the encoder. Since the relay node amplifies the desired signal as well as the noise which is then superimposed at the decoder to the signal coming directly from the encoder, an optimal choice of the relay transmit power $0 \leq P_r \leq P_R$ depends on the relay channel parameters $\{P_S, N_r, N, h, \beta\}$. Moreover, an optimal choice of the power allocation factor β at the encoder also depends on the relay channel parameters $\{P_S, P_r, N_r, N, h\}$. Therefore, we can rewrite (2.42) as

$$\log(\lambda) < \frac{1}{4} \max_{\substack{0<\beta\leq 1 \\ 0\leq P_r \leq P_R}} \left(\log\left(1 + \frac{2h^2\beta P_S}{N}\right) + \log\left(1 + \frac{\tilde{M}(\beta, P_r)}{\tilde{N}(\beta, P_r)}\right)\right), \tag{2.43}$$

which is a sufficient condition for the mean-square stability of a scalar plant.

It is interesting to see that the sufficient condition for the mean-square stability does not depend on the process noise. This provides motivation to study stabilizability of the system in (2.1) without process noise, i.e., $W_t = 0$. In the absence of the process noise in (2.1), the state variance of the noiseless system at any time step t is then given by substituting $n_w = 0$ in (2.38), that is,

$$\alpha_t = \left(\frac{\lambda^2 N}{2h^2\beta P_S + N}\right)\alpha_{t-1}, \quad t = 2, 4, 6, \ldots,$$

$$\alpha_t = \left(\frac{\lambda^4 k\tilde{N}(\beta, P_r)}{(k_2 + \sqrt{k_1 k})^2 + \tilde{N}(\beta, P_r)}\right)\alpha_{t-2}, \quad t = 3, 5, 7, \ldots.$$

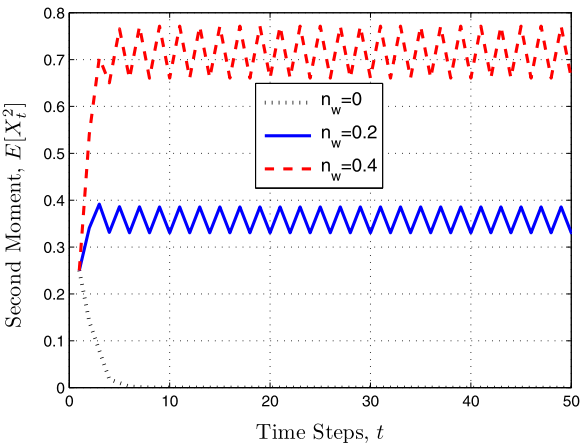


Fig. 2.8 Comparison of second moments of the plant state process at three different levels of process noise

Since $\alpha_1 = \frac{\lambda^2 N}{h^2 P_S}\alpha_0 + n_w$, the state variance $\alpha_t \to 0$ as $t \to \infty$ if

$$\left(\frac{\lambda^4 k \tilde{N}(\beta, P_r)}{(k_2 + \sqrt{k_1 k})^2 + \tilde{N}(\beta, P_r)}\right) < 1.$$

This is the same condition as in (2.41). Thus by using the proposed linear coding and control scheme, we obtain identical sufficient conditions for the mean-square stability of noisy and noiseless first LTI system over *half-duplex* relay channel. Although the sufficient conditions are identical, the state variance in the noisy plant scenario cannot converge to zero like in the noiseless scenario.

A comparison of the second moments of the plant's state process at three different power levels of the process noise is illustrated in Fig. 2.8. In this figure, we have fixed the relay channel parameters $\{P_S = 2, P_r = 2, h = 1, \beta = 0.5, N = 0.5, N_r = 0.1\}$, the plant parameters $\{\alpha_0 = 0.25, \lambda = 1.5\}$, and have plotted the second moment $\mathbb{E}[X_t^2]$ of the state process as a function of time t for three power levels of the process noise, i.e., $n_w = 0$, 0.2, and 0.4. For the given set of channel parameters, the *mean-square stability* of the system requires $\lambda < 1.975$, according to Theorem 2.7. In Fig. 2.8, we have fixed $\lambda = 1.5$ (i.e., less than 1.975); therefore, starting from an arbitrary initial value the second moment of the state process stays bounded for all levels of the process noise. For $n_w = 0$ the second moment converges to zero, starting from an initial value equal to 0.25 as shown in Fig. 2.8. For nonzero values of the process, the second moment keeps alternating between two different values. This happens due to the first and the second transmission phases. As shown in Fig. 2.8, for $n_w = 0.2$ and $n_w = 0.4$ the second moment converges to a unique nonzero value for each transmission phase, and thus it keeps alternating between these two unique limit points. In Fig. 2.8, we can also observe that the rate of convergence is similar in the three examples, and seems to be unaffected by the power level of the process noise.

The sufficient condition for a multi-dimensional plant can be obtained by using the time varying (mode-by-mode transmission) scheme proposed in Sect. 2.3.1.2 to-

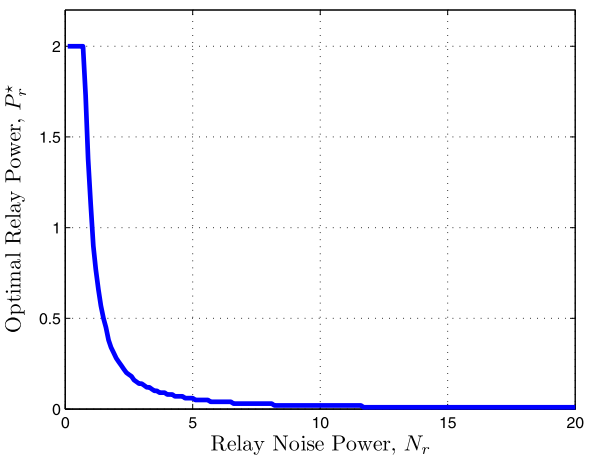


Fig. 2.9 Optimal relay power P_r for $P_R = 2$, $P_S = 10$, $N = 1$

gether with the linear scheme used for the scalar plant above. With a similar analysis as above, we can prove the following theorem.

Theorem 2.7 [64, Theorem 3.1] *The linear time invariant system in* (2.1) *can be mean-square stabilized over the* half-duplex *AWGN relay channel if*

$$\log\big(|A|\big) < \frac{1}{4} \max_{\substack{0<\beta\leq 1 \\ 0\leq P_r \leq P_R}} \left(\log\left(1+\frac{2h^2\beta P_S}{N}\right) + \log\left(1+\frac{\tilde{M}(\beta, P_r)}{\tilde{N}(\beta, P_r)}\right)\right), \tag{2.44}$$

where $\tilde{N}(\beta, P_r) = \frac{P_r N_r}{2\beta P_S + N_r} + N$, $\beta \imath n [0, 1]$, *and*

$$\tilde{M}(\beta, P_r) = \left(\sqrt{2h^2(1-\beta)P_S} + \sqrt{\frac{2\beta P_S P_r N}{(2\beta P_S + N_r)(2h^2\beta P_S + N)}}\right)^2.$$

Remark 2.2 It has been shown in [64, Appendix I] that the term on the right-hand side of (2.44) is the information rate over the half-duplex AWGN relay channel with noiseless feedback.

Optimal choices of the power allocation parameter β at the encoder and the relay transmit power P_r which maximize the term on the right hand side of (2.44) depend on the quality (i.e., SNR) of $\mathcal{E}$–$\mathcal{D}$, $\mathcal{E}$–$\mathcal{R}$, and $\mathcal{R}$–$\mathcal{D}$ links. To illustrate this, we have plotted optimal relay power $P_r^\star$ as a function of the relay noise power N_r for fixed values of $P_R = 2$, $P_S = 10$, $N = 1$ in Fig. 2.9, with the help of numerical computations. We observe that for low values of N_r, the relay uses all available power. As the N_r increases, $P_r^\star$ decreases because the relay is using an amplify-and-forward strategy in which noise also gets amplified along with the signal of interest (state information). Eventually, $P_r^\star$ goes to zero for very high values of N_r, indicating the fact that the relay is not useful anymore if linear strategy is employed

at the relay. However, nonlinear strategies might be useful for high values of N_r. Some nonlinear relaying protocols have been proposed in [62] for the given half-duplex non-orthogonal relay channel, which give significantly higher transmission rates than the linear relaying. Such nonlinear schemes [62] can potentially enlarge the achievable stability region if one uses them for remote stabilization. However, a careful analysis is yet to be carried out to quantify the gains one can obtain in terms of stability with those non-linear relaying strategies.

The optimal choices of power allocation parameter β has not been plotted; however, we have observed via numerical experiments that $\beta = 0.5$ is usually a good choice which corresponds to an equal power allocation to the two transmission phases. For very low values of N_r (i.e., very reliable $\mathcal{E}$–$\mathcal{R}$ link), an optimal β can be slightly greater than 0.5, which is due to the reason that the communication via the relay can be more helpful.

We now consider a special case, where there is no direct communication link from the encoder to the decoder and the information can be communicated only via the relay. We call this setup as a *two-hop* relay channel, since the communication from the sensor to the controller takes places in two hops: the first hop is from the $\mathcal{E}$ to $\mathcal{R}$ and the second hop is from $\mathcal{R}$ to $\mathcal{C}$. The half-duplex relay channel discussed earlier becomes *two-hop* if $h = 0$. Naturally, for this case, we choose $\beta = 1$ and $P_r = P_R$ and obtain the following sufficient condition for stabilization.

Corollary 2.1 [64, Corollary 3.2] *The linear time invariant system in* (2.1) *can be mean-square stabilized over a* two-hop *half-duplex AWGN relay channel if*

$$\log\big(|A|\big) < \frac{1}{4}\log\left(1 + \frac{2P_S P_R}{P_R N_r + N(2P_S + N_r)}\right). \tag{2.45}$$

For a setup which is equivalent to the *two-hop* relay channel, we find a necessary condition in [56, Theorem 4.1] which reads as

$$\log\big(|A|\big) < \frac{1}{4}\min\left\{\log\left(1 + \frac{2P_S}{N_r}\right), \log\left(1 + \frac{P_R}{N}\right)\right\}.$$

The condition in (2.45) becomes both necessary and sufficient if either the $\mathcal{E}$–$\mathcal{R}$ link is noiseless ($N_r = 0$) or the $\mathcal{R}$–$\mathcal{D}$ link is noiseless ($N = 0$).

Consider a *two-hop* relay channel with a causal noiseless feedback link from the controller to the relay. For this setup, the condition in (2.45) becomes necessary and sufficient if we restrict the encoder to be linear in the state. This result is an application of a result in [53, 55]. It follows from the following arguments:

For the *two-hop* relaying scenario with a noiseless causal feedback link from the controller to the relay, we have a partially nested type of information pattern. It is known that the separation of estimation and control holds for such an information pattern and there is no dual effect of control [5]. The optimal control strategy using dynamic programming is $U_t = -\lambda \mathbb{E}[X_t | R_0^t]$, where $R_0^t = \{R_i, 0 \le i \le t\}$. By applying the optimal control action, the plant's state at any time t is given by $X_{t+1} = \lambda(X_t - E[X_t | R_0^t]) + W_t$. If we restrict the state encoder policy to be

linear, then an innovation (memoryless) encoder is optimal since the control actions whiten the state process. Given a linear and memoryless policy at the encoder, let us now find an optimal relaying policy which minimizes $\mathbb{E}[X_{t+1}^2] = \lambda^2\mathbb{E}[(X_t - E[X_t|R_0^t])^2] + \mathbb{E}[W_t^2]$. The cost to be minimized is

$$\begin{aligned}\mathbb{E}\big[\big(X_t - E\big[X_t|R_0^t\big]\big)^2\big] &\overset{(a)}{=} \mathbb{E}\big[\big(X_t - E\big[X_t|Y_0^t\big]\big)^2\big] + \mathbb{E}\big[\big(E\big[X_t|Y_0^t\big] - E\big[X_t|R_0^t\big]\big)^2\big] \\ &\overset{(b)}{=} \mathbb{E}\big[\big(X_t - E\big[X_t|Y_0^t\big]\big)^2\big] + \mathbb{E}\big[\big(cY_t - E\big[X_t|R_0^t\big]\big)^2\big],\end{aligned} \tag{2.46}$$

where (a) follows from $\mathbb{E}[(X_t - E[X_t|Y_0^t])(E[X_t|Y_0^t] - E[X_t|R_0^t])] = 0$ (by the orthogonality principle of MMSE estimation); and (b) follows from the fact that the encoder transmits only innovation at each time step and the MMSE estimation of a Gaussian variable is linear, i.e., $E[X_t|Y_0^t] = cY_t$, where c is a scalar. An optimal relaying policy is the one which minimizes $E[(cY_t - E[X_t|R_0^t])^2]$, since the remaining term of the cost function in (2.46) is independent of the relaying policy. A similar problem was studied in [3], from which it follows that an optimal relaying policy is linear and memoryless. We have earlier obtained the sufficient condition in (2.45) by using optimal linear (memoryless) communication and control policies; therefore, this condition is also necessary provided that the encoder is constrained to be linear in the state. Moreover, if we restrict the relay to be linear, then the two-hop relay channel becomes equivalent to a scalar Gaussian channel. For this channel, it has been shown earlier that a linear scheme is optimal. Therefore, if the relay is restricted to be linear in the received signal, then the condition in (2.45) becomes necessary and sufficient.

2.4.2 Non-orthogonal Full-duplex Relay Channel

Although a half-duplex relay node is easier to build compared to a full-duplex node, there is some loss in the performance due to its inability to communicate simultaneously with state encoder and the controller. A full-duplex system can be realized by placing transmit and receive antennas far enough to ensure sufficient isolation and/or by incorporating some interference cancellation schemes in analog and/or digital domain. In the following, we consider remote stabilization of linear plant over non-orthogonal full-duplex Gaussian relay channel depicted in Fig. 2.10. The variables $\{Z_{r,t}, Z_{d,t}\}$ denote mutually independent white noise components with $Z_{r,t} \sim \mathcal{N}(0, N_r)$ and $Z_{d,t} \sim \mathcal{N}(0, N)$. The gain of $\mathcal{R}$–$\mathcal{C}$ link is denoted by h. At time step t, the encoder $\mathcal{E}$ inputs $S_{e,t}$ to the relay channel with an average power P_S. The relay simultaneously listens to $S_{e,t}$ and transmits $S_{r,t}$ which is the amplified version of the noisy signal received in the time step $t-1$. The amplification at the relay is done under an average power constraint P_r, where $0 \le P_r \le P_R$. That is, the relay transmits, $S_{r,t} = \sqrt{\frac{P_r}{P_S+N_R}}(S_{e,t-1} + Z_{r,t-1})$. The controller receives $R_t = S_{e,t} + hS_{r,t} + Z_t$, computes the MMSE estimate of the state

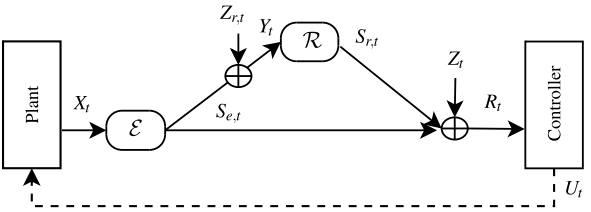


Fig. 2.10 Non-orthogonal full-duplex AWGN relay channel

as, $\hat{X}_t = \mathbb{E}[X_t|R_{[0,t]}]$, and then applies an action to stabilize the system, as done in the half-duplex case. The stabilization of linear plant over full-duplex Gaussian relay channel under a linear scheme has been in studied in [59]. In the following, we present sufficient condition for stabilization under the best linear scheme.

Theorem 2.8 [59, Theorem 6] *The linear system in* (2.1) *can be mean-square stabilized over the* non-orthogonal full-duplex *AWGN relay channel if*

$$\log\big(|A|\big) < \frac{1}{2}\max_{0\leq P_r\leq P_R} \log\left(1+\frac{(\sqrt{P_S(P_S+N_R)}+\eta^\star h\sqrt{P_S P_R})^2}{h^2 P_R N_R + N(P_S+N_R)}\right), \tag{2.47}$$

where $\eta^\star$ *is the unique root in the interval* $[0,1]$ *of the following fourth order polynomial*

$$\left(\frac{h^2 P_S P_R}{P_S+N_R}\right)\eta^4 + \left(2h P_S\sqrt{\frac{P_R}{P_S+N_R}}\right)\eta^3 + \left(P_S+N+\frac{h^2 P_R N_R}{P_S+N_R}\right)\eta^2$$
$$= \left(N+\frac{h^2 P_R N_R}{P_S+N_R}\right).$$

Proof This theorem can be proved by employing linear sensing and control policies, and by following the same analysis as in the proof of Theorem 2.7. The detailed proof for a scalar plant can be found in [59]. The result can be extended to multi-dimensional systems using a similar time-sharing (mode-by-mode) transmission scheme as discussed in Sect. 2.3.1.2. □

2.4.3 Orthogonal Relay Channel

An orthogonal AWGN relay channel is depicted in Fig. 2.11. The variables $\{Z_{r,t}, Z_{1,t}, Z_{2,t}\}$ denote mutually independent white noise components with $Z_{r,t} \sim \mathcal{N}(0, N_r)$ and $Z^i_{d,t} \sim \mathcal{N}(0, N_i)$ for $i \in \{1,2\}$. At any discrete time step t the encoder $\mathcal{E}$ inputs $S_{e,t}$ to the relay channel with an average power P_S. The relay observes $S_{e,t}$ in noise, amplifies it under an average power constraint P_R and forwards it to the

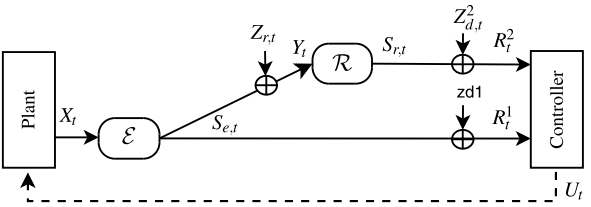


Fig. 2.11 Orthogonal half-duplex AWGN relay channel

controller. Accordingly, the relay transmits

$$S_{r,t} = \alpha(S_{e,t} + Z_{r,t}) = \sqrt{\frac{P_R}{P_S + N_R}}(S_{e,t} + Z_{r,t}),$$

where the amplification factor α is chosen equal to $\sqrt{\frac{P_R}{P_S+N_R}}$ in order to satisfy the average power constraint, i.e., $\mathbb{E}[S^2_{r,t}] \leq P_R$. The output of the relay channel at the decoder $\mathcal{D}$ is $\{R^1_t, R^2_t\}$, which is given by

$$\begin{aligned} R^1_t &= S_{e,t} + Z^1_{d,t}, \\ R^2_t &= S_{r,t} + Z^2_{d,t} = \alpha S_{e,t} + \tilde{Z}_t, \end{aligned} \tag{2.48}$$

where $\tilde{Z}_t \sim \mathcal{N}(0, \alpha^2 N_r + N_2)$.

By using a linear sensing and control scheme over the given channel (as done in the previous sections), we obtain the following sufficient condition for stabilization.

Theorem 2.9 [59, Theorem 2] *The linear time invariant system in* (2.1) *can be mean-square stabilized over the* orthogonal half-duplex *AWGN relay channel if*

$$\log(|A|) < \frac{1}{2}\log\left(1 + \frac{P_S(P_S N_2 + P_R N_R + N_2 N_R + P_R N_1)}{N_1(P_R N_R + P_S N_2 + N_R N_2)}\right). \tag{2.49}$$

Proof This theorem can be proved by employing linear sensing and control policies, and by following the same analysis as in the proof of Theorem 2.7. The detailed proof for a scalar plant can be found in [59]. □

By using information theoretic arguments, we can obtain the following necessary condition for stabilization.

Theorem 2.10 *The linear time invariant system in* (2.1) *can be mean-square stabilized over the* orthogonal half-duplex *AWGN relay channel only if*

$$\log(|A|) < \frac{1}{2}\min\left\{\log\left(1 + \frac{P_S}{N_1}\right) + \log\left(1 + \frac{P_R}{N_2}\right), \log\left(1 + \frac{P_S}{N_1} + \frac{P_S}{N_r}\right)\right\}. \tag{2.50}$$

In order to see the performance of the proposed linear scheme over the orthogonal relay channel, we plot two achievable stability region in Figs. 2.12 and 2.13 as

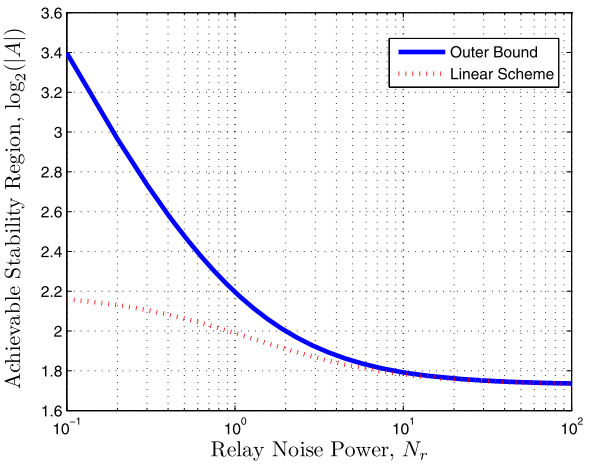


Fig. 2.12 Achievable Stability Region for $P_S = P_R = 10$, $N_1 = N_2 = 1$

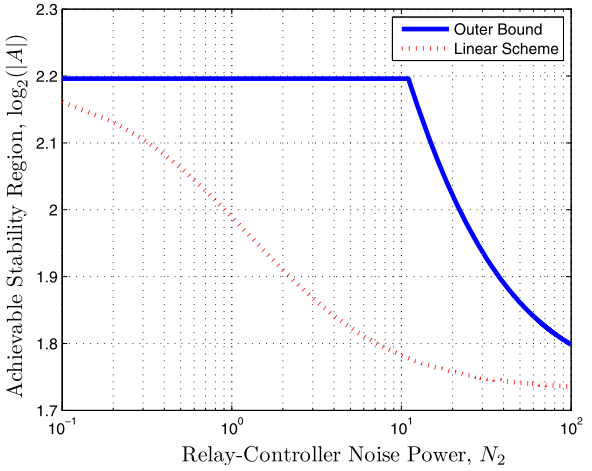


Fig. 2.13 Achievable Stability Region for $P_S = P_R = 10$, $N_1 = N_r = 1$

functions of N_r and N_2 according to (2.49). For comparison we also show the outer bound on stability region using (2.50). We can observe that the linear scheme usually performs good when either N_r is much greater than N_2 or when N_2 is much greater than N_r, i.e., when one of the two links (either $\mathcal{E}$–$\mathcal{R}$ or $\mathcal{R}$–$\mathcal{C}$) is much stronger than the other. However, in some regimes, there is a large gap between stability region achieved by the linear scheme and the outer bound, indicating that the linear schemes can be highly suboptimal in general for orthogonal Gaussian relay networks. In Sect. 2.6, we will present some nonlinear relaying schemes for real-time transmission of a Gaussian source over an orthogonal Gaussian relay channel. Those nonlinear schemes significantly outperform linear schemes, which makes them potential candidates to be used in remote control or stabilization over Gaussian relay networks.

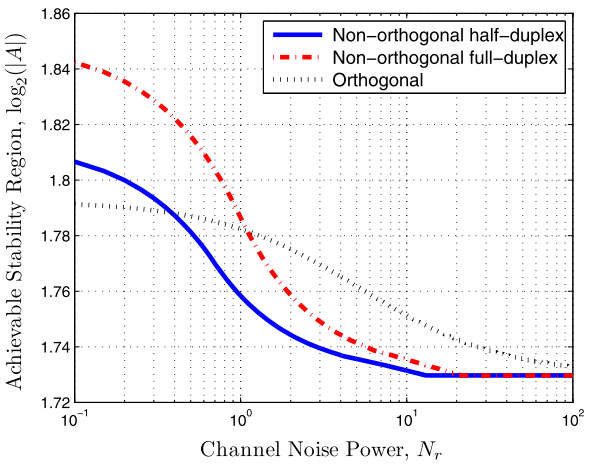


Fig. 2.14 Comparison of linear and nonlinear schemes

2.4.4 Comparison of Relaying Topologies

We have so far studied following three relaying topologies: (i) non-orthogonal half-duplex relay channel, (ii) non-orthogonal full-duplex relay channel, (iii) orthogonal relay channel. In Fig. 2.14, we make a comparison of achievable stability regions over these network topologies. We fix $P_S = 10, P_R = 1, N = N_1 = N_2 = 1$ and show the stability regions that are achieved with linear schemes as functions of increasing relay noise power N_r, according to Theorems 2.7, 2.8, and 2.9. Since in the half-duplex setting the relay transmits in alternate time steps, it can transmit with double power compared to the full-duplex case. Therefore, for the full-duplex relay and the orthogonal relay channels we have used $P_R = 1$, whereas for the half-duplex relay channel we have used $P_R = 2$ while plotting achievable stability regions in Fig. 2.14. The figure shows that the full-duplex relaying is superior to the half-duplex relaying. The reader should keep in mind that the full-duplex sensor nodes are usually more expensive due to the implementation issues discussed earlier. Moreover, we can observe that the orthogonal relaying outperforms non-orthogonal relaying for higher values of N_r; however, this gain is obtained at the cost of using more channel resources (for example, using extra bandwidth).

2.5 Sub-optimality of Linear Policies for Multi-hop Networks

It is known from [27, 65] that linear schemes are not optimal in general for estimation and control over Gaussian multi-hop relay networks. The paper [65] considers the problem of transmission of an i.i.d. Gaussian source over most basic two-hop relay channel illustrated in Fig. 2.15. The problem formulation is as follows: Consider a sequence of independent and identically distributed real-valued Gaussian random variables $\{X_t\}_{t\in\mathbb{N}}$ having zero mean and variance σ_x^2, where t denotes a

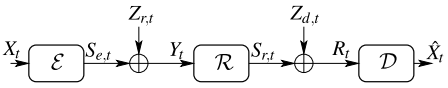


Fig. 2.15 Real-time transmission of a memoryless Gaussian source over a two-hop relay channel

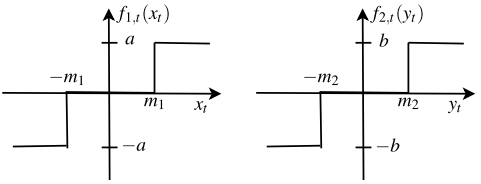


Fig. 2.16 Source and Relay Policies

discrete time index. According to the figure, at a discrete time $t \in \mathbb{N}$ the source encoder $\mathcal{E}$ observes X_t and produces $S_{e,t} = f_{1,t}(X_{[0,t]})$ suitable for transmission, where $f_{1,t} : \mathbb{R}^t \mapsto \mathbb{R}$ is a causal measurable mapping. The mapping $f_{1,t}$ has to satisfy the following average power constraint

$$\mathbb{E}\big[S_{e,t}^2\big] \leq P_S.$$

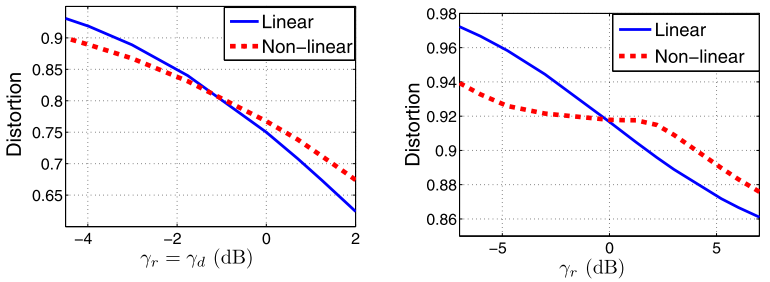
The signal $S_{e,t}$ is then observed in noise by the relay node $\mathcal{R}$ as $Y_t = S_{e,t} + Z_{r,t}$, where $\{Z_{r,t}\}_{t\in\mathbb{N}}$ is a zero mean white Gaussian noise sequence of variance N_r. Since there is no direct link from the source encoder to the destination, we neglect transmission and processing delays at the relay, i.e., the relay node applies a causal mapping on the received signal $f_{2,t} : \mathbb{R}^t \mapsto \mathbb{R}$ to produce $S_{r,t} = f_{2,t}(Y_{[0,t]})$ under the power constraint

$$\mathbb{E}\big[S_{r,t}^2\big] \leq P_r. \tag{2.51}$$

The signal $S_{r,t}$ is then transmitted over a Gaussian channel. Accordingly, the destination node $\mathcal{D}$ receives $R_t = S_{r,t} + Z_{d,t}$, where $\{Z_{d,t}\}_{t\in\mathbb{N}}$ is a zero mean white Gaussian noise sequence of variance N. Upon receiving R_t, the decoder wishes to reconstruct the transmitted variable X_t by applying a mapping $g_t : \mathbb{R}^t \mapsto \mathbb{R}$ to produce $\hat{X}_t = g_t(R_{[0,t]})$. Let us define the signal-to-noise ratios of the $\mathcal{E}$–$\mathcal{R}$ and $\mathcal{R}$–$\mathcal{D}$ links as $\gamma_r := P_S/N_r$ and $\gamma_d := P_r/N$, respectively. The encoder, the relay, and the decoder are all causal and delay-free (zero delay). The objective is to choose the encoder, relay, and decoder mappings such that following distortion

$$D = \limsup_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \mathbb{E}\big[(X_t - \hat{X}_t)^2\big] \tag{2.52}$$

is minimized subject to the power constraints.



(a) $\sigma_x^2 = P_S = P_r = 1$ (b) $\sigma_x^2 = P_S = P_r = 1, N = 5$

Fig. 2.17 Comparison of the linear and the nonlinear schemes

It has been shown in [65] that the following simple time invariant nonlinear source and relay policies can beat the best linear scheme in some cases:

$$f_{2,t}(y_t) = \begin{cases} b, & \text{for } y_t > m_2, \\ 0, & \text{for } |y_t| \le m_2, \\ -b, & \text{for } y_t < -m_2, \end{cases}$$

$$f_{1,t}(x_t) = \begin{cases} a, & \text{for } x_t > m_1, \\ 0, & \text{for } |x_t| \le m_1, \\ -a, & \text{for } x_t < -m_1. \end{cases}$$

The functions $f_{1,t}(\cdot)$ and $f_{2,t}(\cdot)$ are illustrated in Fig. 2.16.

In Fig. 2.17, we have plotted the distortion achieved with the nonlinear and the optimal linear schemes as functions of signal-to-noise ratios for some fixed parameters. These figures demonstrate that the simple three-level quantizer policies can outperform the best linear policies. The proposed nonlinear scheme is not always better than the optimal linear scheme as demonstrated in Fig. 2.17, where we have plotted distortion achieved with the nonlinear and the optimal linear schemes as functions of signal-to-noise ratios for some fixed parameters. The nonlinear scheme outperforms the linear scheme in low SNR regions; however, there might exist better nonlinear strategies which may outperform the linear strategy also in high SNR regions. When the channels are very noisy, the proposed nonlinear strategy is superior because it does not amplify the large values of channel noise at its input unlike the linear (amplify-and-forward) strategy. When linear schemes are employed in multi-hop relay networks, noise is accumulated in every hop, whereas nonlinear schemes can suppress noise. In Sect. 2.6, we discuss an algorithm to numerically optimize the source and relay mappings for an orthogonal relay channel. One can use a similar approach to numerically optimize the source and relay mappings for the given two-hop relay channel as well.

The motivation for choosing three-level quantizer policies comes from the multi-stage decision problem studied in [27] where a binary quantizer was shown to beat the best linear policy for five or more stages. The given two-hop relaying setup corresponds to three stages. It is, however, not known whether binary quantizers

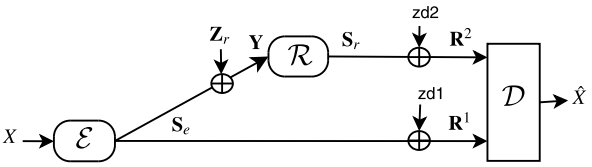


Fig. 2.18 Real-time transmission of a Gaussian source over an orthogonal half-duplex Gaussian relay channel

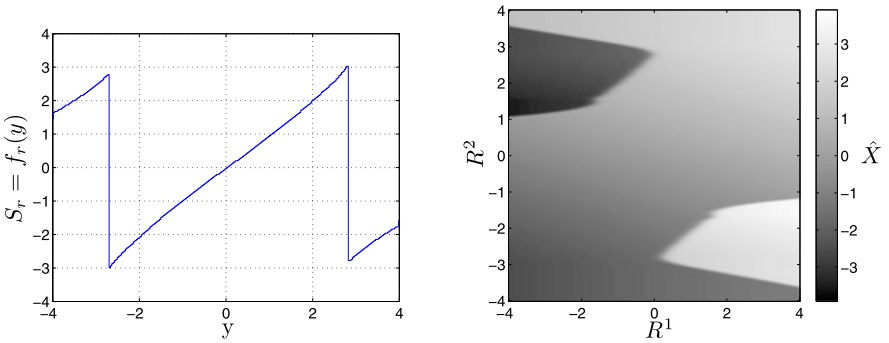
are always worse than the best linear scheme for the given three-stage problem. The intuition for choosing symmetric quantizer comes from the fact that symmetry in distribution is preserved when symmetric functions are applied to sources with symmetric distributions. Moreover, with centering the quantizer at zero, the encoders can utilize the available transmit power in an efficient way by transmitting signals with power equal to zero more often.

2.6 Real-Time Transmission over an Orthogonal Relay Channel

In this section, we consider real-time transmission of a memoryless source over an orthogonal half-duplex Gaussian relay channel. The system model is depicted in Fig. 2.18, where a sensor node $\mathcal{E}$ observes a Gaussian variable $X \sim \mathcal{N}(0, \sigma_x^2)$ and transmits it to the destination $\mathcal{D}$ over a Gaussian channel. An intermediate sensor node $\mathcal{R}$ called relay, overhears the signal transmitted from sensor $\mathcal{E}$ and relays its received information to the destination $\mathcal{D}$ over an orthogonal channel. We assume that the relay is half-duplex, i.e., it cannot simultaneously receive and transmit signals. For real-time coding of i.i.d. sources, memoryless coding is optimal [55]. Therefore, we consider memoryless encoders. For each source sample X_i, the source encoder $\mathcal{E}$ uses channel K_1 times and the relay encoder $\mathcal{R}$ uses channel K_2, where K_1, K_2 are positive integers. Thus the transmission of source each sample takes $K = K_1 + K_2$ channel uses. For each source sample, $\mathcal{E}$ transmits $\mathbf{S}_e = f_e(X)$, where $f_e : \mathbb{R} \mapsto \mathbb{R}^{K_1}$, subject to an average power constraint $\mathbb{E}[f_e^2(X)] \leq P_S$. The relay and the destination accordingly receive:

$$\begin{aligned} \mathbf{Y} &= \mathbf{S}_e + \mathbf{Z}_r, \\ \mathbf{R}^1 &= \mathbf{S}_e + \mathbf{Z}_d^1, \end{aligned} \tag{2.53}$$

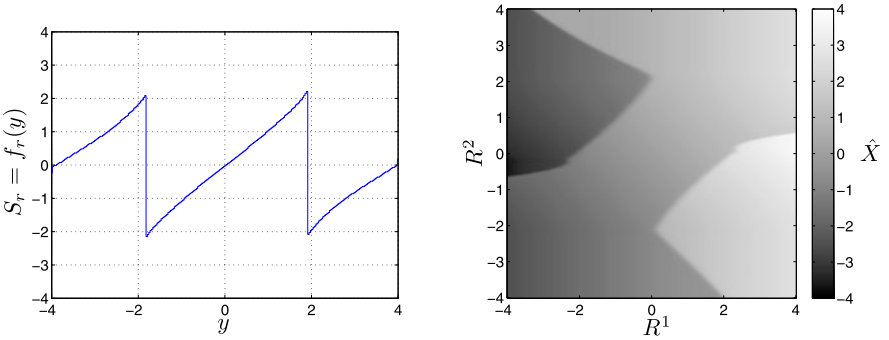
where $\mathbf{Z}_r, \mathbf{Z}_d^1 \in \mathbb{R}^{K_1}$ are mutually independent zero mean white Gaussian noise vectors with $\mathbb{E}[\mathbf{Z}_d^1(\mathbf{Z}_d^1)^T] = N_d^1 I$ and $\mathbb{E}[\mathbf{Z}_r \mathbf{Z}_r^T] = N_r I$. Upon receiving $\mathbf{Y}$, $\mathcal{R}$ transmits $\mathbf{S}_r = f_r(\mathbf{Y})$, where $f_r : \mathbb{R}^{K_1} \mapsto \mathbb{R}^{K_2}$ is subject to an average power constraint $\mathbb{E}[f_r^2(\mathbf{Y})] \leq P_r$. Accordingly, the destination $\mathcal{D}$ receives $\mathbf{R}^2 = \mathbf{S}_r + \mathbf{Z}_d^2$, where $\mathbf{Z}_d^2 \in \mathbb{R}^{K_2}$ is a white Gaussian noise vector with $\mathbb{E}[\mathbf{Z}_d^2(\mathbf{Z}_d^2)^T] = N_d^2 I$. After receiving signals from both $\mathcal{E}$ and $\mathcal{R}$, the decoder $\mathcal{D}$ reconstructs X as $\hat{X} = f_d(\mathbf{R}^1, \mathbf{R}^2)$ where the mapping $f_d : \mathbb{R}^K \mapsto \mathbb{R}$ is chosen such that the mean-squared error $\mathbb{E}[(X - \hat{X})^2]$ is minimized.



(a) Relay mapping (b) Decoder mapping

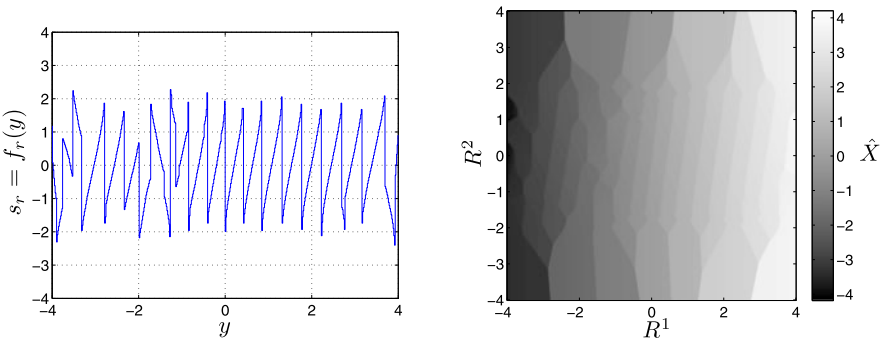
Fig. 2.19 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 5$ dB, $\gamma_{er} = 10$ dB, and $\gamma_{rd} = 25$ dB

The problem of real-time transmission of a Gaussian source over the given three node orthogonal Gaussian relay channel has been studied in [20], where the authors propose an algorithm to numerically optimize the source, relay, and destination mappings. Since there are three mappings to be optimized for a given set of channel parameters, the design algorithm in [20] uses a common strategy of optimizing one mapping at a time while keeping the other two fixed. Moreover, each dimension of the channel space is discretized into equally spaced points in the design algorithm. The optimized mappings obtained in [20] are in general nonlinear and are shown to provide significant gains over linear mappings in terms lower achievable distortion. In the following, we use the design algorithm of [20] to optimize mappings for some fixed channel parameters. We give examples of optimized mappings for only two cases: (i) $K_1 = K_2 = 1$, and (ii) $K_1 = 1, K_2 = 2$. Let us define the SNRs of the $\mathcal{E}$–$\mathcal{D}$, $\mathcal{E}$–$\mathcal{R}$, and $\mathcal{R}$–$\mathcal{D}$ links as $\gamma_{ed} := P_S/N_d^1$, $\gamma_{er} := P_S/N_r$, and $\gamma_{rd} := P_r/N_d^2$, respectively. For the case $K_1 = K_2 = 1$, we provide some examples of optimized mappings in Figs. 2.19, 2.20, 2.21. In these examples, we have fixed the source mapping to be linear and optimized the other two mappings (relay and decoder mappings) for different values of signal-to-noise ratios as given in the captions of the respective figures. We observe that these optimized relay mappings are non-invertible and have an almost periodic like behavior. Several input values are mapped to the same output value; this way of reusing output values can be seen as Wyner–Ziv type compression. This reuse of output values (an almost periodic behavior) makes the relay mappings more power efficient. Such non-invertible mappings have become possible due to the availability of the side information via the $\mathcal{E}$–$\mathcal{D}$ link. In Figs. 2.19–2.21, we have also plotted decoder mappings, which basically estimate the source X using the two received signals R^1 and R^2. The decision regions along with the reconstructions $\hat{X}$ are also shown in the figures. From these examples of optimized mappings, we observe that the number of periods in the relay mappings increase as the reliability of side information increases, thus making the relay more power efficient.



(a) Relay mapping (b) Decoder mapping

Fig. 2.20 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 10$ dB, $\gamma_{er} = 10$ dB, and $\gamma_{rd} = 25$ dB



(a) Relay mapping (b) Decoder mapping

Fig. 2.21 Relay and decoder mappings ($K_1 = K_2 = 1$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 20$ dB, $\gamma_{er} = 30$ dB, and $\gamma_{rd} = 25$ dB

In Figs. 2.22 and 2.23, we give two examples of optimized relay mappings for $K_1 = 1$, $K_2 = 2$. That is the case where the relay performs an expansion—from its one-dimensional input to its two-dimensional output. Once again, there is a reuse of the output symbols which is only possible due to the side information from the direct link. As reliability of the side information increases, the reuse of the same output values also increases. The mappings have a spiral like shape. As reliability of direct link increases, the reuse of same output values also increase. Looking at the spiral from above, a similarity to the polynomial based source–channel codes proposed in [19, 42] can be seen.

In order to see the gains of nonlinear optimized mappings over linear mappings in terms of achievable distortion, we refer the reader to [20]. In [20], the authors have analyzed the performance in detail for various values of channel dimensions (K_1 and K_2) and signal-to-noise ratios. It is observed that with these optimized mappings significantly lower distortion can be achieved, which makes them very useful for

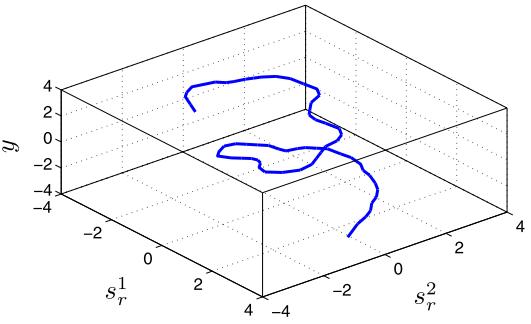


Fig. 2.22 Relay mapping ($K_1 = 1$, $K_2 = 2$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 5$ dB, $\gamma_{er} = 15$ dB, and $\gamma_{rd} = 10$ dB

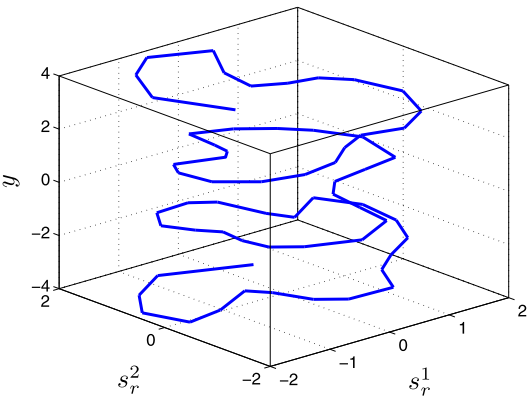


Fig. 2.23 Relay mapping ($K_1 = 1$, $K_2 = 2$) optimized for $\sigma_x^2 = P_s = P_r = 1$ and $\gamma_{ed} = 15$ dB, $\gamma_{er} = 15$ dB, and $\gamma_{rd} = 10$ dB

large sensor networks and remote control scenarios where delay is a critical factor and transmit powers are limited.

2.7 Distributed Sensing for Control

In this section, we consider a multi-sensor setup, where multiple sensors in parallel observe noisy versions of the state process and communicate their observations to a remotely situated controller over orthogonal (parallel) channels. This scenario is different from the one studied earlier since each sensor has access to a different observation due to the addition of the measurement noise. The sensors then transmit their local observations to the controller. The schemes of [24, 58] can also be used for distributed sensing. However, it is not known how useful they are in the presence of measurement noise.

For the sake of simplicity, consider a two sensor setup shown in Fig. 2.24 with a scalar plant whose state equation is given by (2.1) with $A = \lambda$. The state X_t is observed in noise by the sensors $\mathcal{E}_1$ and $\mathcal{E}_2$ as

$$Y_t^i = X_t + Z_{e,t}^i, \quad i = 1, 2,$$

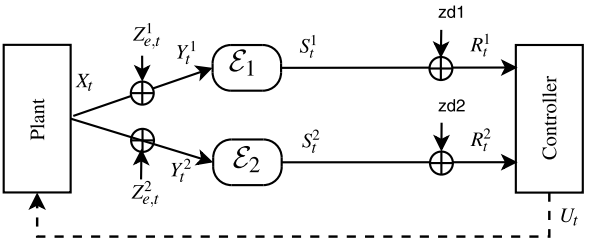


Fig. 2.24 A closed-loop control system with state measurements transmitted over wireless channels

where $Z^1_{e,t}$ and $Z^1_{e,t}$ are two i.i.d. mutually independent measurement noise components, which are Gaussian distributed with zero means and variances N^1_e and N^2_e, respectively. Based on their noisy observations, the two sensors transmit the following signals:

$$S^i_t = f_{i,t}\big(Y^i_t\big), \quad i = 1, 2,$$

subject to the following power constraints:

$$E\big[\big(S^i_t\big)^2\big] \le P_i, \quad i = 1, 2. \tag{2.54}$$

Accordingly, the remote controller receives

$$R^i_t = S^i_t + Z^i_{d,t}, \quad i = 1, 2, \tag{2.55}$$

where $Z^i_{d,t}$, $i = 1, 2$, are independent and i.i.d. zero-mean Gaussian with power N^i_d. We have assumed orthogonal channels from the sensors to the controller; therefore, there is no interference between the two received signals (i.e., we have two parallel Gaussian channels from the sensors to the sink node). Based on the received signals, the controller takes an action $U_t = \pi_t(R^1_{[0,t]}, R^2_{[0,t]})$. The objective is to minimize the following finite horizon quadratic cost function

$$J_T = \mathbb{E}\left[\sum_{t=1}^{T} X^2_t\right], \tag{2.56}$$

where the expectation is taken over the initial state X_0, the process noise W_t, the measurement noise $Z^i_{e,t}$, and the channel noise $Z^i_{d,t}$.

In the following, we present a nonlinear distributed sensing scheme which outperforms the best linear scheme. This scheme was first introduced in [47] for the transmission of a Gaussian source over orthogonal Gaussian channels and was later used in control context in [2].

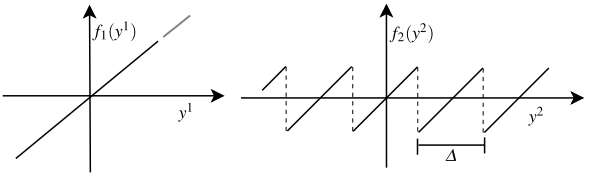


Fig. 2.25 Nonlinear distributed sensing

2.7.1 Sensing Scheme

The nonlinear distributed sending and control scheme works as follows. The signals transmitted by the two sensors are given by

$$S_t^1 = \eta_t Y_t^1, \tag{2.57}$$

$$S_t^2 = \eta_t \left(Y_t^2 - \Delta_t \left\lfloor \frac{Y_t^2}{\Delta_t} \right\rceil \right), \tag{2.58}$$

where $\lfloor \cdot \rceil$ denotes rounding to the nearest integer. A pictorial illustration of this nonlinear scheme is given in Fig. 2.25. The parameter Δ_t controls the length of each period in the periodic sawtooth function. The values $\Delta_{[0,t]}$ are chosen such that the cost function J_T in (2.56) is minimized. The procedure of choosing $\Delta_{[0,t]}$ can be found in [2]. The parameters $\{\eta_t, \Delta_{[0,t]}\}$ are chosen such that the average transmit power constraints (2.54) are met.

2.7.2 Control Scheme

The controller is assumed to have a separation structure where it first computes an estimation of the state and then take action using the state estimate. Since the computation of optimal MMSE estimate based on all previously received signals $\{R^1_{[0,t]}, R^2_{[0,t]}\}$ is not practical, the following sub-optimal algorithm is proposed in [2]:

1. Compute estimates $\tilde{X}_{0|t}, \ldots, \tilde{X}_{t|t}$ of $X_0, \ldots, X_t$ based on the previous estimate $\hat{X}_{t-1}$ and R_t^1 using a Kalman filter (cf. Kalman Filter 1 in the Fig. 2.26).
2. Assume that $|(\tilde{X}_{s|t} - Y_s^2 - Z_{d,s}^2)/\eta_s| \leq \Delta_s/2 \ \forall s$ and compute the Maximum Likelihood estimates $\hat{Y}_s^2$ as (cf. ML decoder in Fig. 2.26):

$$\hat{Y}_s^2 = \text{argmin}_{Y_s \in \mathcal{Y}} \left(\left(S^2(Y_s) - R_s^2 \right)^2 \right), \tag{2.59}$$

   where $\mathcal{Y} = \{Y_s : |\tilde{X}_{s|t} - Y_s| \leq \eta_s \Delta_s/2\}$.
3. Finally, assume that the estimates $\hat{Y}_s^2$ had been linearly encoded (multiplied by η_0^t) and find the estimate $\hat{X}_t$ from a Kalman filter using $\{R^1_{[0,t]}, \eta_{[0,t]}, U_{[0,t-1]}\}$ and $U_{[0,t-1]}$ as input (cf. Kalman Filter 2 in Fig. 2.26).

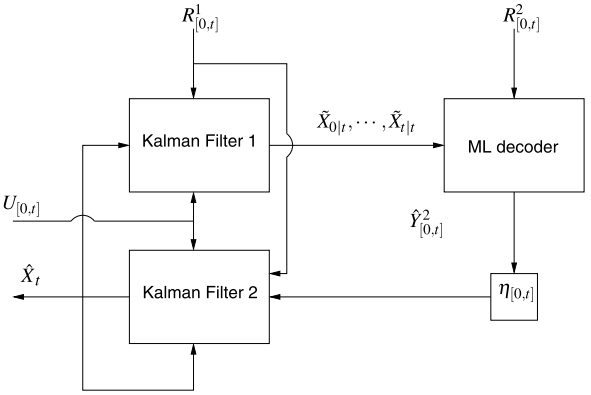
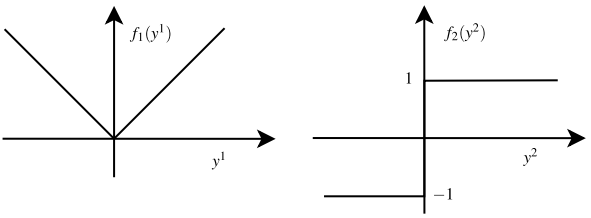


Fig. 2.26 State estimator for the nonlinear distributed sensing scheme



1st Sensor Mapping. 2nd Sensor Mapping.

Fig. 2.27 Nonlinear distributed sensing

The above nonlinear sensing and control scheme is delay-free and can be implemented with reasonable complexity. This nonlinear scheme has been shown to outperform the best linear strategy in [2]. Furthermore, it is robust to the knowledge of noise statistics at the sensors as demonstrated in [2]. Intuitively, this scheme can be easily extended to an arbitrary number of sensors by employing a linear mapping at the first sensor node and sawtooth mappings at the remaining sensor nodes with successively decreasing time periods Δ_t. How the number of sensor nodes will affect the system performance compared to the best linear scheme is yet to be studied.

Another nonlinear distributed sensing scheme has been proposed in [58] for the two-sensor setup, where one sensor transmits magnitude of the received signal and the other sensor transmits phase value of the received signal. The mappings employed by the two distributed sensors are shown in Fig. 2.27. It has been shown in [58] that the outputs of these two sensor mappings are mutually independent and thus enable us to send independent information over the two parallel channels. This nonlinear scheme has been shown to outperform linear scheme in absence of measurement noise. A careful comparison of these two nonlinear sensing schemes discussed in this section is yet to be made. There might be certain regimes where one scheme may perform better than the other.

2.8 Bibliographic Notes

There exists a diverse literature on the problem of control and real-time communication over Gaussian channels, focusing on different models, objectives, and design constraints. For instance, the plant and the channel models can be either discrete-time or continuous with different network topologies and different assumption on Gaussian noise. There can be different design constraints such as transmission delay-constraints, sum and individual power constraints, average and peak power constraints, bandwidth constraint, etc. And the commonly studied objectives are minimizing a quadratic cost function of the state and the control variables, achieving moment stability or invariant state distribution on the state of the plant. In this chapter, the discussion was mostly limited to the problem of the mean-square stabilization of an LTI discrete-time plant and real-time communication over some specific discrete-time white Gaussian channels with average transmit power constraints. In the following, we highlight some of the important and related research contributions on the problem of control over Gaussian channels.

Some of the earliest papers addressing the control of linear systems over Gaussian channels include [43, 50]. These papers show that for linear systems subject to Gaussian noise with linear sensing policies having perfect memory (recall), the optimal control policies are linear and there exists a separation property between estimation and control. However, in [48], Witsenhausen showed via a simple counter example that linear policies may not be optimal when there are more two or more decision makers (sensors/controllers) without perfect memory (recall). At this point, we emphasize the importance of information structures and recommend some fundamental papers on stochastic team decision problems [16, 35, 49, 57]. The information structure can be classical, quasiclassical, and non-classical. The problem studied in [43, 50] falls in the class of classical information structure, for which linear control policies were shown to be optimal. In the quasiclassical information structure, decision maker A effects the information of decision maker B, and the decision maker B knows what is known by decision maker A. For LQG systems having a quasiclassical information structure, linear policies have been shown to be optimal, for example, see [16, 35]. The Witsenhausen problem [48] has a non-classical information structure in which decision maker A effects the information of decision maker B, but the decision maker B does not have access to what is known by decision maker A. The Witsenhuasen problem is unsolved till today, which indicates the hardness of such problems. However, for some LQG systems with non-classical information structures, linear policies have been shown to be optimal, for example, see [3–5, 52]. These papers have used tools from information theory. The paper [3] studies the problem of a causal memoryless transmission of a noisy Gaussian source over a Gaussian channel and shows that linear coding and decoding policies are optimal. The optimality of linear sensing and control policies for a first order scalar LTI system with an objective of minimizing a quadratic cost function of state and control variables was established in [5], where some concepts from rate distortion theory were used. Another paper [44] used tools from source coding and channel coding to establish necessary condition for stabilization over a large class of communication channel including a memoryless Gaussian channel. The authors of [45]

found the conditions under which separation property between estimation and control holds for LQG problems where there is a communication link (for example, a memoryless Gaussian channel) between the sensor and the controller. Moreover, they introduced a framework of sequential rate distortion theory for designing the encoders and the decoders. In [8, 13, 31, 39], the reader can find relevant studies on signal-to-noise ratio requirements for stabilization over some Gaussian channel models. The papers [18, 40] have proposed some techniques for designing linear controllers for Gaussian channels. Some recent results on control over Gaussian fading channels can be found in [9, 26].

The problem of control over communication channels is closely related to the problem of communication over channels with feedback. In [11], a general equivalence was shown between feedback stabilization over an analog communication channel and a communication scheme for channels with noiseless feedback. This communication scheme is a generalization of Schalkwijk–Kailath coding scheme [37] for a single user channel. And for multi-user channels such as broadcast, multiple-access and interference channels, this scheme is a generalization of coding scheme given in [21, 32, 33]. Using the communication schemes proposed in [21, 32, 33] for multi-user Gaussian channels with noiseless feedback, necessary and sufficient conditions for stabilization of multiple plants over multi-user Gaussian channels are obtained in [60, 63, 66]. A decentralized design of linear sensors and controllers over white Gaussian channels with the objective of the mean-square stability is studied in [12]. A comprehensive study of stabilization and optimization of networked control, and information structures is present in [57] and [30].

Acknowledgements We would like to thank the editors, B. Bernhardsson, G. Como, and A. Rantzer, for giving us an opportunity to write this chapter. We are also very grateful to Johanness Kron (formerly Johanness Karlsson) for performing numerical simulations to generate the figures included in Sect. 2.6 of this chapter. Some of these results are part of his PhD thesis.

This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Akyol, E., Viswanatha, K., Rose, K.: On conditions for linearity of optimal estimation. IEEE Trans. Inf. Theory **58**(6), 3497–3508 (2012)
2. Andersson, M., Zaidi, A.A., Wernersson, N., Skoglund, M.: Nonlinear distributed sensing for closed-loop control over Gaussian channels. In: IEEE Swe–CTW, pp. 19–23 (2011)
3. Bansal, R., Başar, T.: Solutions to a class of linear-quadratic-Gaussian LQG stochastic team problems with nonclassical information. Syst. Control Lett. **9**(2), 125–130 (1987)
4. Bansal, R., Başar, T.: Stochastic teams with nonclassical information revisited: when is an affine law optimal? IEEE Trans. Autom. Control **32**(6), 554–559 (1987)
5. Bansal, R., Başar, T.: Simultaneous design of measurement and control strategies for stochastic systems with feedback. Automatica **25**(5), 679–694 (1989)
6. Başar, T.: A trace minimization problem with applications in joint estimation and control under nonclassical information. J. Optim. Theory Appl. **31**(3), 343–359 (1980)
7. Başar, T., Bansal, R.: Optimum design of measurement channels and control policies for linear-quadratic stochastic systems. Eur. J. Oper. Res. **73**(2), 226–236 (1994)

8. 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)
9. 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**(4), 1013–1019 (2008)
10. Cover, T., Thomas, J.: Elements of Information Theory. Wiley, New York (2006)
11. Elia, N.: When Bode meets Shannon: control-oriented feedback communication schemes. IEEE Trans. Autom. Control **49**(9), 1477–1488 (2004)
12. Farhadi, A., Ahmed, N.U.: Suboptimal decentralized control over noisy communication channels. Syst. Control Lett. **60**(4), 285–293 (2011)
13. Freudenberg, J.S., Middleton, R.H., Solo, V.: Stabilization and disturbance attenuation over a Gaussian communication channel. IEEE Trans. Autom. Control **55**(3), 795–799 (2010)
14. Gastpar, M., Vetterli, M.: On the capacity of large Gaussian relay networks. IEEE Trans. Inf. Theory **51**(3), 765–779 (2005)
15. Gastpar, M., Rimoldi, B., Vetterli, M.: To code, or not to code: lossy source-channel communication revisited. IEEE Trans. Inf. Theory **49**(5), 1147–1158 (2003)
16. Ho, Y.C.: Team decision theory and information structures. Proc. IEEE **68**(6), 644–654 (1980)
17. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, Cambridge (1990)
18. Johannesson, E., Rantzer, A., Bernhardsson, B.: Optimal linear control for channels with signal-to-noise ratio constraints. In: ACC, pp. 521–526 (2011)
19. Karlsson, J., Skoglund, M.: Analog distributed source-channel coding using sinusoids. In: IEEE ISWCS, pp. 279–282 (2009)
20. Karlsson, J., Skoglund, M.: Optimized low-delay source-channel-relay mapping. IEEE Trans. Commun. **58**(5), 1397–1404 (2010)
21. Kramer, G.: Feedback strategies for white Gaussian interference networks. IEEE Trans. Inf. Theory **48**(6), 1423–1438 (2002)
22. Kramer, G., Maric, I., Yates, R.: Cooperative communications. Found. Trends Netw. **1**(3–4), 271–425 (2006)
23. Kumar, U., Gupta, V., Laneman, J.N.: Sufficient conditions for stabilizability over Gaussian relay channel and cascade channels. In: IEEE CDC, pp. 4765–4770 (2010)
24. Kumar, U., Gupta, V., Laneman, J.N.: On stability across a Gaussian product channel. In: IEEE CDC, pp. 3142–3147 (2011)
25. Lee, K.H., Petersen, D.P.: Optimal linear coding for vector channels. IEEE Trans. Commun. **24**(12), 1283–1290 (1976)
26. Leong, A.S., Dey, S., Anand, J.: Optimal LQG control over continuous fading channels. In: 18th IFAC World Congress (2011)
27. Lipsa, G.M., Martins, N.C.: Optimal memoryless control in Gaussian noise: a simple counterexample. Automatica **47**, 552–558 (2011)
28. 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**(7), 1604–1615 (2008)
29. Massey, J.L.: Causality, feedback and directed information. In: IEEE ISITA (1990)
30. Matveev, A.S., Savkin, A.V.: Estimation and Control over Communication Networks. Birkhäuser, Boston (2008)
31. Middleton, R.H., Rojas, A.J., Freudenberg, J.S., Braslavsky, J.H.: Feedback stabilization over a first order moving average Gaussian noise channel. IEEE Trans. Autom. Control **54**(1), 163–167 (2009)
32. Ozarow, L.H.: The capacity of the white Gaussian multiple-access channel with feed-back. IEEE Trans. Inf. Theory **30**(4), 623–629 (1984)
33. Ozarow, L., Leung-Yan-Cheong, S.: An achievable region and outer bound for the Gaussian broadcast channel with feedback. IEEE Trans. Inf. Theory **30**(4), 667–671 (1984)
34. Pilc, R.J.: The optimum linear modulator for a Gaussian source used with a Gaussian channel. Bell Syst. Tech. J. 3075–3089 (1969)
35. Radner, R.: Team decision problems. Ann. Math. Stat. **33**, 857–881 (1962)

36. Sahai, A., Mitter, S.: The necessity and sufficiency of anytime capacity for stabilization of a linear system over noisy communication links—part I: scalar systems. IEEE Trans. Inf. Theory **52**(8), 3369–3395 (2006)
37. Schalkwijk Kailath, T.: A coding scheme for additive noise channels with feedback—I: no bandwidth constraint. IEEE Trans. Inf. Theory **12**(2), 172–182 (1966)
38. Shamai, S., Verdu, S., Zamir, R.: Systematic lossy source/channel coding. IEEE Trans. Inf. Theory **44**(2), 564–579 (1998)
39. Shu, Z., Middleton, R.H.: Stabilization over power-constrained parallel Gaussian channels. IEEE Trans. Autom. Control **56**(7), 1718–1724 (2011)
40. Silva, E.I., Goodwin, G.C., Quevedo, D.E.: Control system design subject to SNR constraints. Automatica **46**(2), 428–436 (2010)
41. Silva, E.I., Derpich, M.S., Ostergaard, J.: A framework for control system design subject to average data-rate constraints. IEEE Trans. Autom. Control **56**(8), 1886–1899 (2011)
42. Skoglund, N.W.M., Ramstad, T.: Polynomial based analog source-channel codes. IEEE Trans. Commun. **57**(9), 2600–2606 (2009)
43. Striebel, C.: Sufficient statistics in the optimum control of stochastic systems. J. Math. Anal. Appl. **12**, 576–592 (1962)
44. Tatikonda, S., Mitter, S.: Control over noisy channels. IEEE Trans. Autom. Control **49**(7), 1196–1201 (2004)
45. Tatikonda, S., Sahai, A., Mitter, S.: Stochastic linear control over a communication channel. IEEE Trans. Autom. Control **49**(9), 1549–1561 (2004)
46. Tse, D., Viswanath, P.: Fundamentals of Wireless Communication. Cambridge University Press, Cambridge (2005)
47. Wernersson, N., Skoglund, M.: Nonlinear coding and estimation for correlated data in wireless sensor networks. IEEE Trans. Commun. **57**(10), 2932–2939 (2009)
48. Witsenhausen, H.S.: A counterexample in stochastic optimum control. SIAM J. Control **6**, 131–147 (1968)
49. Witsenhausen, H.S.: Separation of estimation and control for discrete time systems. Proc. IEEE **59**(11), 1557–1566 (1971)
50. Wonham, W.M.: On the separation theorem of stochastic control. SIAM J. Control **6**, 312–326 (1968)
51. Wu, Y., Verdú, S.: Functional properties of MMSE. In: IEEE ISIT, pp. 1453–1457 (2010)
52. Yüksel, S.: Stochastic nestedness and the belief sharing information pattern. IEEE Trans. Autom. Control **54**(12), 2773–2786 (2009)
53. Yüksel, S.: On optimal causal coding of partially observed Markov sources under classical and non-classical information structures. In: IEEE ISIT, pp. 81–85 (2010)
54. Yüksel, S.: Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems. IEEE Trans. Inf. Theory **58**(10), 6332–6354 (2012)
55. Yüksel, S.: On optimal causal coding of partially observed Markov sources in single and multi-terminal settings. IEEE Trans. Inf. Theory **59**, 424–437 (2013)
56. Yüksel, S., Başar, T.: Control over noisy forward and reverse channels. IEEE Trans. Autom. Control **56**, 1014–1029 (2011)
57. Yüksel, S., Başar, T.: Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints. Birkhäuser, Boston (2013)
58. Yüksel, S., Tatikonda, S.: A counterexample in distributed optimal sensing and control. IEEE Trans. Autom. Control **54**(4) (2009)
59. Zaidi, A.A., Oechtering, T.J., Skoglund, M.: Rate sufficient conditions for closed–loop control over AWGN relay channels. In: IEEE ICCA, pp. 602–607 (2010)
60. Zaidi, A.A., Oechtering, T.J., Skoglund, M.: Sufficient conditions for closed-loop control over multiple-access and broadcast channels. In: IEEE CDC (2010)
61. Zaidi, A.A., Oechtering, T.J., Yüksel, S., Skoglund, M.: Closed-loop control over half-duplex AWGN relay channels. In: Reglermöte (2010)

62. Zaidi, A.A., Khormuji, M.N., Skoglund, M.: Nonlinear transmission strategies for a general half-duplex AWGN relay channel. In: IEEE Swe–CTW, pp. 58–61 (2011)
63. Zaidi, A.A., Oechtering, T.J., Skoglund, M.: Closed-loop stabilization over Gaussian interference channel. In: 18th IFAC World Congress (2011)
64. Zaidi, A.A., Oechtering, T.J., Yüksel, S., Skoglund, M.: Sufficient conditions for closed-loop control over a Gaussian relay channel. In: IEEE ACC, pp. 2240–2245 (2011)
65. Zaidi, A.A., Yüksel, S., Oechtering, T.J., Skoglund, M.: On optimal policies for control and estimation over a Gaussian relay channel. In: IEEE CDC (2011)
66. Zaidi, A.A., Oechtering, T.J., Skoglund, M.: On stabilization over a Gaussian interference channel. In: IEEE ECC (2013)
67. Zaidi, A.A., Oechtering, T.J., Yüksel, S., Skoglund, M.: Stabilization of linear systems over Gaussian networks. IEEE Trans. Autom. Control (2012, accepted)
68. Zaidi, A.A., Yüksel, S., Oechtering, T.J., Skoglund, M.: On the tightness of linear policies for stabilization of linear systems over Gaussian networks. Syst. Control Lett. (2013, under review)

Chapter 3
Optimal Radio-Mode Switching for Wireless Networked Control

Nicolas Cardoso de Castro, Federica Garin, and Carlos Canudas de Wit

3.1 Introduction

Networked Control Systems (NCS) are systems where the communication between the sensors, the controller, and the actuators occurs through a network, see Fig. 3.1(a) and [18]. Energy is a key resource in those systems, in particular in applications concerning wireless networks. Energy-efficiency in Wireless Sensor Networks (WSN) has given rise to a rich literature in the last ten years, see, e.g., [1, 8, 11, 13, 20, 32].

However, an important number of these contributions are focused on the transmission techniques regardless of the nature of the data or the final application. In a closed-loop system, the control performance is a crucial point that is omitted in the works cited in the last paragraph. On the other hand, the control community has been interested in saving energy in wireless NCS, see [3, 12, 15, 24, 28, 29] among others.

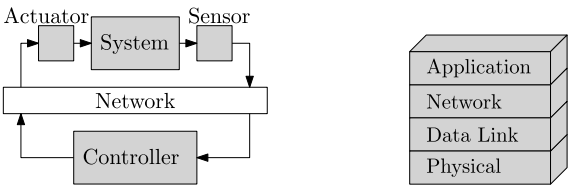
The authors of [10, 22] state that additional energy can be saved in control applications by a multi-layer design. The notion of a layer is taken from communications theory, and illustrated in Fig. 3.1(b). Contributions considering this approach are few. Liu and Goldsmith in [22] optimize the control performance taking into account some network parameters (throughput, packet delay, and packet loss probabilities). The authors of [23, 25] state that there does not exist a communication protocol dedicated to NCS. Then they derive communication protocols suited to NCS because

N. Cardoso de Castro (✉) · F. Garin
INRIA Rhône-Alpes, NeCS Team, Grenoble, France
e-mail: nicolas.cardoso@karrus-its.com

F. Garin
e-mail: federica.garin@inria.fr

C. Canudas de Wit
Department of Automatic Control, GIPSA-Lab, CNRS, NeCS Team, Grenoble, France
e-mail: carlos.canudas-de-wit@gipsa-lab.grenoble-inp.fr

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_3,



(a) Block diagram of a typica l Networked Control System.

(b) Networked Control System stack.

Fig. 3.1 (**a**) Block diagram of a typical Networked Control System: the sensor sends its measurement to the controller over the network, which sends the control input to the actuator over the same network. (**b**) Different steps involved in a networked communication are commonly abstracted with a layer approach, grouped in a stack, where each layer has a dedicated function. The Physical layer is in charge of the radio modulation of the digital data. The Data Link (MAC) layer manages and shares the transmission medium. The Network layer routes the data to the destination in an efficient manner. Finally, the Application layer is related to the control scheme and possibly includes source encoding and decoding

they expose various protocol parameters to the application layer, which can provide the desired trade-off between reliability/latency and control performance. Quevedo et al. in [28] derive a state estimator that accounts for packet loss probabilities. Indeed, these probabilities depend upon time-varying channel gains, packet lengths, and transmission power levels of the sensors. By adapting the source coding scheme and the transmission level at the sensor side, they are able to find a trade-off between energy and estimation performance.

Our main goal in this chapter is to design control laws and switching power control policies to save energy in an NCS by considering communication and control co-design. An often-quoted rule of thumb is that executing 3 million instructions is equivalent to transmitting 1000 bits at a distance of 100 meters in terms of expended energy [26]. From such an observation that the radio is an important energy consumer in a wireless node, we decide to focus on the management of the radio chip. Our strategy gathers two main components. The first one is Event-Based Control (EBC) which consists in relaxing the time-triggering paradigm. Indeed, instead of closing the loop periodically at each sampling interval, the control input applied to the system is only updated when a given event occurs. The second component is radio-mode management. Indeed, when a node is not transmitting data, its radio can be switched to one of the several non-transmitting radio-modes. Each mode is characterized by its consumption and the transition costs, and the time needed to switch back to a transmitting mode. On the other hand, when the radio emits a message, it can use one of the available power levels. Increasing the power level permits improving the quality of the transmission (modeled as a lower erasure probability in a memoryless packet-erasure channel) at the price of consuming more energy. Our strategy is a joint design of the radio-mode switching policy and of the feedback control law. It is derived in the framework of optimal control with the use of a suitable cost function. The optimization problem is solved using Dynamic Programming with the Value Iteration method.

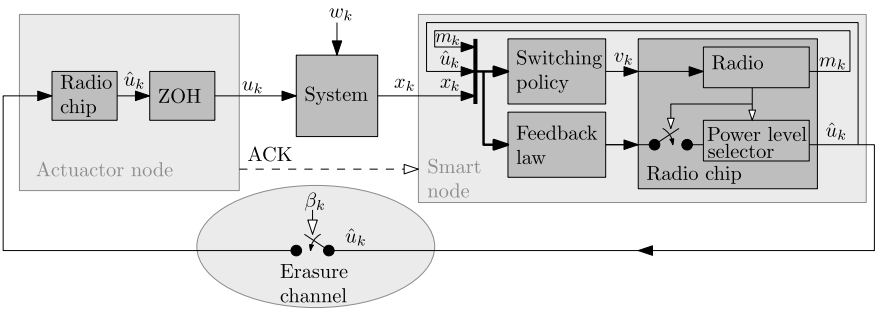


Fig. 3.2 Block diagram of the problem setup. The smart node measures the noisy state x_k from the system, it computes the feedback law $\hat{u}_k$, and decides whether to send it or not to the actuator node. $\hat{u}_k$ can actually be $\varnothing$ when no transmission is scheduled or when the transmitted message is dropped. Then, the receiver is able to determine if it has received an update or not. β_k equals 1 when the transmission is successful or 0 when there is a dropout

The rest of this chapter is organized as follows: Sect. 3.2 gives the mathematical model of the NCS that we consider, which is turned into a switched model formulation in Sect. 3.2.6. The optimization problem that describes the trade-off between energy saving and control performance is derived and solved in Sect. 3.3 with the use of Dynamic Programming. Section 3.4 provides simulation results and comparison to standard periodic approaches. Finally, the chapter is concluded in Sect. 3.5.

3.2 Control and Communication Joint Modeling

We restrict our focus to a setup composed of two nodes, as depicted in Fig. 3.2 and described hereafter. A two-nodes setup captures the challenges of energy efficiency without introducing the difficulties appearing in a multi-nodes setup (such as medium access control or routing).

The first node, called the smart sensor node, has sensing and computing capabilities. It is in charge of sensing the system output, computing the feedback law, and deciding whether or not to send the control input to the second node, in charge of applying the control law to the actuator. We assume that the receiver node is co-located with the actuator. This configuration, also called *one-channel feedback NCS*, is commonly considered in the literature, as it is described in [18]. We focus our attention on the problem of energy-saving at the sensor side only because the assumption that the receiver node is co-located with the actuator usually implies that it has access to a large energy supply, since actuators often require more energy than radio-receivers and computers.

The communications from the smart sensor to the receiver node are event-based. This means that the smart sensor can decide to let the system run open

loop when the control performance is good enough. We define in Sect. 3.3 a cost function that is used by the smart sensor to determine how good the control performance is with respect to the energy needed for a transmission. When the receiver node does not receive any new update from the smart sensor, it holds the memory of the last received control input. The memory is implemented in a Zero Order Holder (ZOH) device in Fig. 3.2. The control update computed at the smart sensor side, denoted $\hat{u}_k$, can be different from the control input actually applied to the system, denoted u_k. Both $\hat{u}_k$ and u_k take values in $\mathbb{R}^{n_u}$.

This event-based behavior is considered on the base of a discrete-time monitoring. This means that the smart sensor only monitors the system at a given sampling period. Then, depending on the state of the system, it decides if whether or not a transmission occurs at the current sampling time. This last point is important, as highlighted by the authors in [17], since it relaxes a widely used assumption in Event-Based control that the events are monitored in continuous time, which is barely implementable in practice.

In addition to the event-based approach, we consider a radio-mode management scheme. Indeed, not only the smart node decides whether or not to use the radio, but also it decides upon the radio-modes. When the radio is active, a transmission power level is chosen; when it is not active, it is switched to one of the several non-transmitting radio-modes. The radio can use different power levels to change the transmission success probability in the case of unchanged channel conditions. In the case of a varying channel, the transmission power can be increased to face bad channel conditions and keep the transmission success probability constant, but the study of varying channel conditions is beyond the scope of this work. The non-transmitting modes (e.g., `Idle`, `Sleep`), explained in details in Sect. 3.2.2, allow saving energy by turning off some components in the radio chip when there is no transmission. If no transmission happens in a long time span, it is clear that the best choice is the lowest-consuming radio-mode. However, a certain time and energy are needed to switch between active and inactive modes, and this is different depending on how many components are switched off in the radio chip, thus motivating the use of intermediate modes, with some components still active, when a transmission is likely to happen again soon. This also means that the choice is not independent of the transmission decisions (the past ones, and some indication of the possible future ones), so that in our control application it must be related to the event-based control law.

Hereafter, we will denote by N_1 the number of transmission levels and by N_2 the number of non-transmitting modes, so that the total number of radio-modes is given by $N = N_1 + N_2$. Moreover, the *feedback law* will refer to the feedback control law that is used to compute the control input $\hat{u}_k$ possibly sent to the receiver node and applied to the system, while the *switching policy* will refer to the node's decision to switch to a given radio-mode. The switching decision is denoted $v_k \in \{1, 2, \ldots, N\}$, where $v_k = i$ means that the radio-mode is switched to mode i at time k.

3.2.1 System Model

The system to be controlled is a linear discrete-time system with an additive zero-mean white Gaussian noise, described by Eq. (3.1):

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{3.1}$$

where $x_k \in \mathbb{R}^{n_x}$ is the system state, $u_k \in \mathbb{R}^{n_u}$ is the control input, and $w_k \sim \mathcal{N}(\mathbf{0}, W)$ is the system noise. For simplicity, we assume that the state is fully observed, namely the sensor measures x_k (more precisely, the measurement should be x_k plus the measurement noise, but we can assume that the term w_k already models both the system and measurement noises). A and B have appropriate dimensions and the system is controllable, and may be unstable.

Our goal is to stabilize this system around the origin, i.e., $x_k = \mathbf{0}$, while saving energy.

3.2.2 Radio Chip Model

The state of the radio chip is the mode at time k, m_k:

$$m_k \in \mathbb{M} \triangleq \mathbb{M}_1 \cup \mathbb{M}_2$$

with

$\mathbb{M}_1 \triangleq \{1, 2, \ldots, N_1\}$ being the set of transmitting modes and

$\mathbb{M}_2 \triangleq \{N_1 + 1, N_1 + 2, \ldots, N\}$ the set of non-transmitting modes.

The radio-mode is updated according to the switching decision: $m_{k+1} = v_k$.

The consumption of the radio chip during any sampling interval, called the transition cost and denoted θ_{m_k, v_k}, depends on the radio-mode m_k and on the switching decision v_k. The transition costs are introduced in details in Sect. 3.2.5.

The amount of energy E consumed since the commissioning can be computed as follows:

$$\begin{aligned} E_0 &= 0, \\ E_{k+1} &= E_k + \theta_{m_k, v_k} = E_k + \theta_{m_k, m_{k+1}}. \end{aligned}$$

3.2.3 Channel Model

As it is done in [19], the channel is modeled as a simplified memoryless erasure channel where the message $\hat{u}_k$ is dropped with probability $\varepsilon(m_k)$ for $m_k \in \mathbb{M}_1$, and

otherwise is correctly received. We consider a model where the dropout concerns the real-valued message $\hat{u}_k$, not single bits or packets.

The dropout probabilities depend on the transmission power used by the radio chip, i.e. , the transmitting mode. Higher transmission power implies higher success probability, i.e. , $\varepsilon(1) < \varepsilon(2) < \cdots < \varepsilon(N_1)$. The dropouts are modeled by Bernoulli random variables β_k, where 1 denotes success and 0 dropout, and where

$$\mathbb{P}\{\beta_k = 0 | m_k = m\} = \varepsilon(m),$$
$$\mathbb{P}\{\beta_k = 1 | m_k = m\} = 1 - \varepsilon(m).$$

Given the mode m_k, β_k is conditionally independent of the past $\{\beta_h\}_{h<k}$, $\{w_k\}_{h<k}$; the mapping $\varepsilon(m)$, $m \in \mathbb{M}_1$, is known for design purposes.[1]

Acknowledgments (ACKs) are sent by the receiver node to confirm to the smart sensor that a new control update has been applied to the system. These ACKs are assumed to be reliable, i.e. , always correctly received by the smart node. This is a reasonable assumption since we are not considering energy restrictions at the actuator node, so that the ACKs can be transmitted with enough power. Moreover, an ACK is a short message (possibly reduced to a single bit), and hence it is possible to encode it with a very large redundancy for error correction.

If ACKs were not reliable, then the smart node would not know if the control update has been applied when no ACK is received. This would result in having the control memory at the smart node possibly different from the actual control input applied to the system, and the associated decision not optimal. However, the control memory at the smart node side and the actual value applied to the system will be synchronized again as soon as a control update transmission is successfully sent and acknowledged.[2]

3.2.4 Switching Policy and Feedback Law

The sensor node embeds a switching policy η (whose joint-design with the feedback law μ will be described hereafter) to assign the radio-mode. The decision to switch between modes is based on the current system output x_k, the last control input u_{k-1} applied to the system, and the current radio-mode, denoted m_k. Introducing $\tilde{u}_k = u_{k-1}$, the memory of the last control input, the switching decision is given by $v_k = \eta(x_k, \tilde{u}_k, m_k)$. The ACK sent by the actuator node when a transmission is successful

[1] It is reasonable to assume that the values of $\varepsilon(m)$ are known as motivated in [28].

[2] When considering unreliable ACKs, several behaviors can be considered for the smart sensor in the case where no ACK is received. They depend mostly on the system and thus will not be discussed in details here. In a few words, depending on the criticality of the system, the difference between the sensor-node memory and the actual control input can be ignored, or the system can enforce a transmission until an ACK is successfully received. Again, a trade-off appears between energy consumption and closed-loop performance.

lets the smart sensor have perfect knowledge of the last control input applied to the system.

The control input applied to the system, denoted u_k, depends on the arrival of the update $\hat{u}_k$, which depends on the transmission success and on the decision to send an update, as described by Eq. (3.2). If an update is received, then the control law is the optimal law computed by the smart sensor, denoted $\hat{u}_k = \mu(x_k, \tilde{u}_k, m_k)$ and derived in the next sections. Otherwise, the control input is held to its previous value as long as no update is received from the smart node:

$$u_k = \begin{cases} \beta_k \hat{u}_k + (1 - \beta_k) u_{k-1}, & \text{in case of transmission,} \\ u_{k-1}, & \text{otherwise.} \end{cases} \tag{3.2}$$

Thus, we have a Jump Linear System (JLS), where we decide upon the transition probabilities. This stands in contrast to many Markov JLS approaches, where transition probabilities are assumed given; see, e.g., [14, 30].

3.2.5 Transition Costs Model

A radio chip can be modeled by an automaton where:

- The states describe the radio modes (e.g., `Transmitting`, `Idle`, `Sleep`),
- And the transitions model events, which in our case are the radio-mode switching requests.

The number of states of the automaton is the number of radio-modes N. The modes that allow a transmission are numbered 1 to N_1 and the non-transmitting ones are numbered $N_1 + 1$ to $N = N_1 + N_2$.

Concerning the transitions, we consider that the state of the automaton only changes at the sampling instant. In other words, the radio chip stays in a given state for the whole sampling interval and may only switch periodically at the sampling instants. The costs associated to the transitions have to take into account the current and time constraints imposed by the radio chip, as they are provided in the radio chip datasheet.[3] The transition costs are denoted $\theta_{i,j}$, and describe the energy needed to switch from mode i to mode j and then stay in mode j until the next sampling instant. Note that $\theta_{i,i}$ gives the cost to stay in mode i.

The costs $\theta_{i,j}$ are derived in order to respect the properties of the radio chip. For instance, there is no extra cost to change the transmission level, which implies that the cost to switch from a transmitting mode Tx_1 to another transmitting mode Tx_2, denoted θ_{Tx_1,Tx_2}, is the same as the cost to stay in the transmitting mode Tx_2, denoted θ_{Tx_2,Tx_2}. Also, it is often impossible to switch between modes that are not consecutive at the radio chip level, for instance, in order to switch from the

[3] See [27] for the case of the CC1100 Low-Power Sub-1 GHz RF Transreceiver from Texas Instruments.

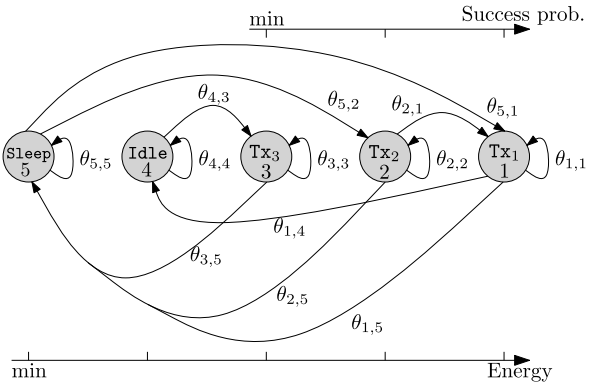


Fig. 3.3 Illustration of the transition costs with $N_1 = 3$ and $N_2 = 2$. `Idle` is an intermediate mode between the transmitting modes and the `Sleep` mode. Modes are ordered according to their energy consumption, e.g., $\theta_{5,5} < \theta_{4,4}$. Transmitting modes with higher energy consumption have a higher probability to transmit successfully. The *arrows* represent the transition costs. A more detailed figure would make every transition $\theta_{i,j}$ from any mode i to any mode j appear

`Sleep` mode to any transmitting mode, the radio needs to switch first to the `Idle` mode. However, the costs $\theta_{i,j}$ describe the consumption of the radio over a sampling interval, which may include several state switchings, as long as the sampling interval is large enough to satisfy the time constraints of the radio chip.

This model assumes that the mode transition time-constraints are smaller than the sampling interval. This implies that the smart node has enough time to switch to the desired mode (and possibly transmit the control input) before the next sampling time. An impossible transition can be modeled by an infinite cost.

An illustration of the transition costs with 3 transmitting modes and 2 non-transmitting modes is given in Fig. 3.3.

3.2.5.1 Scenario in which to Apply Our Technique

Our approach considers several non-transmitting modes in order to save further energy than a simple `On`/`Off` pattern. However, because of transition costs and time-constraints, switching to intermediate non-transmitting mode may result in more energy waste than holding a transmitting mode. Although the goal of this contribution is to derive a switching policy that offers actual energy savings, our approach reduces the expended energy only under some assumptions on the transition costs $\theta_{i,j}$.

Assumption 3.1 $\theta_{i,i} > \theta_{j,j} \geq 0$ for all $i \in \mathbb{M}_1$ and $j \in \mathbb{M}_2$. This means that the transmitting modes consume more than the non-transmitting modes and that the cost to stay in a given mode for a sampling period is always positive for the transmitting modes, and can be null for the non-transmitting modes. Moreover, it holds that $\theta_{i,j} > 0$ for all $i, j \in \mathbb{M}$ such that $i \neq j$. This means that any mode transition has a positive cost.

We introduce the convention to number the modes according to their energy costs, namely $\theta_{1,1} \geq \theta_{2,2} \geq \cdots \geq \theta_{N,N}$. The amount of energy that can be saved when using the non-transmitting modes is directly related to the difference between $\theta_{1,1}$ and $\theta_{N,N}$.

Assumption 3.2 For any $j_1 \in \mathbb{M}_2$ and $j_2 \in \mathbb{M}_2$, if $j_1 < j_2$ then $0 < \theta_{i,j_1} \leq \theta_{i,j_2}$ for all $i \in \mathbb{M}_1$. This means that the transition cost from a transmitting mode to a non-transmitting mode is larger for deeper non-transmitting mode. The symmetric condition (from a non-transmitting mode to a transmitting one) is also assumed to hold, i.e. , for any $j_1 \in \mathbb{M}_2$ and $j_2 \in \mathbb{M}_2$, if $j_1 < j_2$ then $0 < \theta_{j_1,i} \leq \theta_{j_2,i}$ for all $i \in \mathbb{M}_1$.

The way radio chips are designed enforces that the transition to (or from) a non-transmitting mode has a larger cost when the non-transmitting mode saves more energy (i.e. , there are more radio-chip components to be turned off or on during the transition). However, this does not imply that Assumption 3.2 holds because the cost represented by $\theta_{i,j}$ accounts for the energy consumed over an entire sampling interval, including both the cost of the transition and the cost to remain in the reached mode after having finished the transition.

3.2.6 Switched Model

The evolution of the system under the different choices of radio-mode is now formulated as a switched linear system, with as many systems as the number of modes N. From a control point of view, the different modes are actually reduced to two cases: when a successful transmission occurs (i.e. , the control loop is closed) and when the system runs open loop. The different modes affect the energy consumption and also the success probability (in case of transmission).

Choosing the switching policy at time k is equivalent to choosing the radio-mode. The evolution of the switched system depends on x_k, the state of the system, on $\tilde{u}_k = u_{k-1}$, the memory keeping track of the last applied control input, and on m_k the mode of the radio chip. We define z_k as the system state augmented with the control memory:

$$z_k = \begin{bmatrix} x_k \\ \tilde{u}_k \end{bmatrix} \in \mathbb{R}^{n_x+n_u}, \quad \text{and also} \quad \omega_k = \begin{bmatrix} w_k \\ \mathbf{0} \end{bmatrix}.$$

Then, the state of the switched system is the following:

$$(z_k, m_k) \in \mathbb{X} \triangleq \mathbb{R}^{n_x+n_u} \times \mathbb{M}.$$

The evolution of the system given in Eq. (3.1) with the feedback law μ described in Eq. (3.2), together with the radio-mode switching policy η, give rise to the following

switched system:

$$\begin{cases} z_{k+1} = f_{v_k}(z_k, \hat{u}_k, \beta_k, \omega_k), \\ m_{k+1} = v_k = \eta(z_k, m_k), \\ \hat{u}_k = \mu(z_k, m_k), \end{cases} \tag{3.3}$$

where the function f_{v_k} is defined as

$$f_{v_k}(z_k, \hat{u}_k, \beta_k, \omega_k) = \Phi_{v_k}(\beta_k) z_k + \Gamma_{v_k}(\beta_k)\hat{u}_k + \omega_k,$$

and the matrices $\Phi_{v_k}(\beta_k)$, $\Gamma_{v_k}(\beta_k)$, for $v_k \in \mathbb{M}$, are as follows:

1. If $v_k \in \mathbb{M}_1$, i.e. , if there is a transmission, then

$$\Phi_{v_k}(\beta_k) = \begin{bmatrix} A & (1-\beta_k)B \\ \mathbf{0} & (1-\beta_k)\mathbf{I} \end{bmatrix} = \begin{cases} \Phi_{CL} = \begin{bmatrix} A & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \text{if } \beta_k = 1, \\ \Phi_{OL} = \begin{bmatrix} A & B \\ \mathbf{0} & \mathbf{I} \end{bmatrix} & \text{if } \beta_k = 0, \end{cases}$$

$$\Gamma_{v_k}(\beta_k) = \beta_k \begin{bmatrix} B \\ \mathbf{I} \end{bmatrix} = \begin{cases} \Gamma_{CL} = \begin{bmatrix} B \\ \mathbf{I} \end{bmatrix} & \text{if } \beta_k = 1, \\ \Gamma_{OL} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} & \text{if } \beta_k = 0. \end{cases}$$

2. If $v_k \in \mathbb{M}_2$, i.e. , if there is no transmission, then

$$\Phi_{v_k}(\beta_k) = \Phi_{OL} \quad \forall \beta_k,$$
$$\Gamma_{v_k}(\beta_k) = \Gamma_{OL} \quad \forall \beta_k.$$

3.3 Derivation of the Optimal Joint Control Law and Switching Policy

The goal of the jointly designed switching policy and feedback law is to obtain a trade-off between the energy consumption and the feedback performance. The framework of Optimal Control is chosen to derive the joint policy since it solves an optimal problem based on a cost function. The cost function that we consider explicitly accounts for the two criteria under focus, namely the feedback performance and the energy consumption. First, we define the cost-to-go, also called step cost, denoted ℓ, as the cost that is payed by the closed-loop system over one sampling interval. This cost-to-go depends on the state of the system x_k, on the state of the radio chip m_k, on the control input u_k, and on the switching decision v_k:

$$\ell_{v_k}(x_k, m_k, u_k) = \underbrace{x_k^\top \bar{Q} x_k}_{\text{performance}} + \underbrace{u_k^\top \bar{R} u_k}_{\text{control energy}} + \underbrace{\theta_{m_k, v_k}}_{\text{transmission energy}}. \tag{3.4}$$

$\bar{Q}$ and $\bar{R}$ are symmetric positive definite matrices which can be tuned to give different trade-offs between the feedback performance and energy consumption.

As detailed in Eq. (3.2), the control input depends on the switching decision v_k and on the success of the transmission as described by $\beta_k \in \{0, 1\}$. Thus the cost-to-go can be written as follows:

$$\ell_{v_k}(x_k, m_k, u_k, \beta_k) = \begin{cases} x_k^\top \bar{Q} x_k + \beta_k \hat{u}_k^\top \bar{R} \hat{u}_k + (1-\beta_k)\tilde{u}_k^\top \bar{R} \tilde{u}_k + \theta_{m_k, v_k} & \text{if } v_k \in \mathbb{M}_1, \\ x_k^\top \bar{Q} x_k + \tilde{u}_k^\top \bar{R} \tilde{u}_k + \theta_{m_k, v_k}, & \text{otherwise.} \end{cases}$$

Finally, using the notation introduced in Sect. 3.2.6:

$$\ell_{v_k}(z_k, m_k, \hat{u}_k, \beta_k) = z_k^\top Q_{v_k}(\beta_k) z_k + \hat{u}_k^\top R_{v_k}(\beta_k) \hat{u}_k + \theta_{m_k, v_k}, \tag{3.5}$$

where the matrices $Q_{v_k}(\beta_k)$ and $R_{v_k}(\beta_k)$, for $v_k \in \mathbb{M}$, are defined as follows:

1. If $v_k \in \mathbb{M}_1$, i.e. , if $u_k = \beta_k \hat{u}_k + (1-\beta_k)\tilde{u}_k$, then

$$Q_{v_k}(\beta_k) = \begin{bmatrix} \bar{Q} & \mathbf{0} \\ \mathbf{0} & (1-\beta_k)\bar{R} \end{bmatrix} = \begin{cases} Q_{CL} = \begin{bmatrix} \bar{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \text{if } \beta_k = 1, \\ Q_{OL} = \begin{bmatrix} \bar{Q} & \mathbf{0} \\ \mathbf{0} & \bar{R} \end{bmatrix} & \text{if } \beta_k = 0, \end{cases}$$

$$R_{v_k}(\beta_k) = \beta_k \bar{R} = \begin{cases} R_{CL} = \bar{R} & \text{if } \beta_k = 1, \\ R_{OL} = \mathbf{0} & \text{if } \beta_k = 0. \end{cases}$$

2. If $v_k \in \mathbb{M}_2$, i.e. , if $u_k = \tilde{u}_k$, then

$$Q_{v_k}(\beta_k) = Q_{OL} \quad \forall \beta_k,$$
$$R_{v_k}(\beta_k) = R_{OL} \quad \forall \beta_k.$$

The cost function $J_{\mu,\eta}$ accounts for the expected cost the system has to pay when controlling the system (3.1) with the policy (μ, η) over an infinite time horizon with the initial condition z_0, m_0:

$$J_{\mu,\eta}(z_0, m_0) = \lim_{H \to \infty} \mathsf{E}_{\{\beta_k, \omega_k\}_{k=0}^{H-1}} \left[\sum_{k=0}^{H-1} \lambda^k \ell_{v_k}(z_k, m_k, \hat{u}_k, \beta_k) \right], \tag{3.6}$$

where $\hat{u}_k = \mu(z_k, m_k)$, $v_k = \eta(z_k, m_k)$, $z_{k+1} = f_{v_k}(z_k, \hat{u}_k, \beta_k, \omega_k)$, and $\lambda \in [0, 1)$ is a discount factor discussed hereafter. $J_{\mu,\eta}$ is a function of the initial conditions only, defined as an expectation with respect to the random sequence of noises and channel erasures encountered along the evolution of the dynamical system, where the dynamics are those obtained by applying the policies μ and η for the control feedback law and for the radio-mode switching decision, respectively. The time horizon is considered infinite by taking the limit when the horizon length H goes to infinity.

The optimization problem we consider consists in finding a feedback law μ^* and a switching policy η^* such that

$$J^*(z_0, m_0) \triangleq J_{\mu^*,\eta^*}(z_0, m_0) = \min_{\mu,\eta} J_{\mu,\eta}(z_0, m_0). \tag{3.7}$$

Note that the minimum is indeed attained, as proved in Sect. 3.3.1.1.

Remark 3.1 We are only searching for a stationary policy (μ^*, η^*) because a time-dependent policy (μ_k, η_k) on an infinite time-horizon would not be implementable. Fortunately, as it is explained in Sect. 3.3.1, a stationary policy exists, which is optimal among all policies.

Remark 3.2 According to Assumptions 3.1 and 3.2, the optimization problem is considered under the non-triviality assumption that transmissions have a non-zero cost whatever the previous mode, i.e., $\theta_{i,j} > 0 \ \forall i \in \mathbb{M}, \ \forall j \in \mathbb{M}_1$. Indeed, in the case where a mode allows transmission free of cost, the optimal radio-management consists in always transmitting.

It is only relevant to consider a solution of the optimal problem (3.7) that leads to a finite cost. Indeed, if the cost function $J_{\mu,\eta}(z, m)$ is always infinite for any policy (μ, η) and any state (z, m), then any policy is considered as optimal in the sense that any policy leads to $J^* = \min_{\mu,\eta} J_{\mu,\eta}(z, m) = +\infty$.

The purpose of the discount factor is precisely to force the cost function to be finite for some policies. It weights the importance of immediate actions versus long-term decisions. It is generally not used (i.e., taken equal to 1) when the cost function is naturally finite for some policies.

Unfortunately, given the structure of the cost-to-go (3.5), one cannot ensure that such cost function is finite if $\lambda = 1$. Indeed, when $\lambda = 1$, one can distinguish two cases from the basic formulation of the cost-to-go (3.4):

- The first case considers that an infinite number of transmissions is scheduled over the infinite horizon. The term related to the energy cost of the radio chip, θ_{m_k,v_k}, is a non-zero additive term in infinitely many time steps. Indeed, $\theta_{i,j} > 0 \ \forall i \in \mathbb{M}, \ j \in \mathbb{M}_1$ thanks to Assumption 3.1. This makes the sum infinite.
- In the second case, the number of transmissions is finite, which implies that there exists a time index k_0 such that, for all $k > k_0$, the system runs open loop since no more transmissions update the control input. From that observation, the conditions for the cost function to converge to a finite value are that the system is open-loop stable despite the noise and that one of the radio-modes has a null cost that is scheduled an infinite number of times.

In order to be as general as possible, and particularly to consider open-loop unstable systems, the discount factor is taken such that $\lambda < 1$ in order to admit policies that make the cost function finite.

However, it has to be noticed that introducing the discount factor $\lambda < 1$ prevents from proving stability with the standard argument in Linear Quadratic (LQ) optimal

control, using the cost-to-go as a Lyapunov function. To the authors' knowledge, a stability proof in discounted problems solved with Dynamic Programming is an open issue. The authors of [5, 16] discuss the stability of systems controlled with feedback laws derived using Dynamic Programming approaches.

Finally, we notice that the control space is constrained by the switching decision. Indeed, if no transmission is scheduled, then the control input is forced to the memory value $\tilde{u}$. We thus define the control space $\mathbb{U}(\tilde{u})$ as follows:

Definition 3.1

$$\mathbb{U}(\tilde{u}) \triangleq \left\{(u, v) : u \in \mathbb{R}^{n_u},\ v \in \mathbb{M}_1\right\} \cup \left\{(\tilde{u}, v) : v \in \mathbb{M}_2\right\}.$$

For ease of notation, we will use $\mathbb{U}(z)$ instead of $\mathbb{U}(\tilde{u})$. The joint policy (μ, η) must take its values in $\mathbb{U}(z)$.

3.3.1 Optimal Solution with the Value Iteration Method and Convergence Proof

The framework of Dynamic Programming provides methods to solve optimization problems which can be decomposed into nested sub-problems. We consider the Value Iteration method to solve the optimization problem described in Eq. (3.6) and (3.7).This method is based on Bellman's Principle of Optimality [2], stated as follows:

> An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

In order to introduce the Value Iteration method, we consider that the horizon length is finite, and called H, i.e., that no limit is taken in Eq. (3.6). Note also that in the finite horizon case, one often considers a final cost, denoted $\ell_F(z, m)$, that accounts for the state of the system at the end of the horizon. This final cost is simply added to the right-hand side of Eq. (3.6).

The Value Iteration method exploits the fact that if we know all the optimal paths from time $k+1$ to time H, then every optimal path from time k must use one of the optimal paths from time $k+1$. We denote $J^*_{H-i}(z, m)$ the optimal cost function for the same problem starting at time $k = H - i$ over an horizon i, $0 < i \leq H$. The Bellman's Principle of Optimality leads to the following relation:

$$J^*_{H-i}(z, m) = \min_{(\hat{u}, v) \in \mathbb{U}(z)} \left\{\mathsf{E}_{\beta,\omega}\left[\lambda^{H-i} \ell_v(z, m, \hat{u}, \beta) + J^*_{H-i+1}\left(f_v(z, \hat{u}, \omega, \beta), v\right)\right]\right\}. \tag{3.8}$$

The Value Iteration method consists in computing backward in time (starting from the final cost at time H, $J^*_H(z_H, m_H) = \ell_F(z_H, m_H)$) the optimal cost

$J^* = J_0^*$. The Value Iteration method is a recursion that computes the so-called Value Function, denoted $V_i(z, m)$, at each time step i from H to 0. The Value Function is related to the cost function as follows:

$$V_i(z, m) = \frac{J_{H-i}^*(z, m)}{\lambda^{H-i}}.$$

Based on the recursion on the cost function given by Eq. (3.8), the recursion on the Value Function, up to iteration H, $\forall (z, m) \in \mathbb{X}$, is given by (see [4, Proposition 1.3.1]):

$$\begin{aligned} V_0(z, m) &\triangleq g_F(z, m), \\ V_{i+1}(z, m) &= \min_{(\hat{u}, v) \in \mathbb{U}(z)} \left\{ \mathsf{E}_{\beta,\omega}\left[\ell_v(z, m, \hat{u}, \beta) + \lambda V_i\left(f_v(z, \hat{u}, \omega, \beta), v \right) \right] \right\}, \\ V_H(z, m) &= J_0^*(z, m) = J^*(z, m). \end{aligned} \tag{3.9}$$

The re-scaled final cost $g_F(z, m)$ is defined as a function of the final cost $\ell_F(z, m)$:

$$g_F(z, m) = \frac{\ell_F(z, m)}{\lambda^H}.$$

Note that the notion of final cost is only used for the initialization of the recursion, it will be dropped shortly in the infinite horizon formulation. Note also that the iteration index of the Value Function i goes backward in time.

This scheme provides the optimal cost function J_0^* (also denoted J_{H-H}^* when using the notation previously introduced) over a finite horizon H. Then, under proper assumption on the initialization stage, the computation of the optimal cost function over an horizon $H + 1$ is directly obtained from Eq. (3.9), when the optimal cost function over an horizon H is already computed.

This provides the intuition that taking the limit of the recursion as H goes to infinity converges to the optimal cost function J^* over an infinite horizon. The actual Value Iteration method consists indeed in iterating the Value Function V_i according to Eq. (3.9), but initializing V_0 to the null function. This algorithm provides an optimal joint policy (μ^*, η^*) and the optimal cost J^*, as stated in Theorem 3.1. The proofs are given in the next subsection. Note that the optimal joint policy derived with the Value Iteration method is not necessarily unique; if several joint policies lead to the optimal cost, this method allows deriving one of them.

The Value Iteration method consists in the following iterative computation of the functions $V_i(z, m)$, $(z, m) \in \mathbb{X}$:

$$\begin{aligned} V_0(z, m) &= 0, \\ V_{i+1}(z, m) &= \min_{(\hat{u}, v) \in \mathbb{U}(z)} \left\{ \mathsf{E}_{\beta,\omega}\left[\ell_v(z, m, \hat{u}, \beta) + \lambda V_i\left(f_v(z, \hat{u}, \beta, \omega), v \right) \right] \right\}, \end{aligned} \tag{3.10}$$

$$\begin{aligned} &\left(\mu_{i+1}(z, m), \eta_{i+1}(z, m) \right) \\ &\quad \triangleq \arg\min_{(\hat{u}, v) \in \mathbb{U}(z)} \left\{ \mathsf{E}_{\beta,\omega}\left[\ell_v(z, m, \hat{u}, \beta) + \lambda V_i\left(f_v(z, \hat{u}, \beta, \omega), v \right) \right] \right\}. \end{aligned} \tag{3.11}$$

Theorem 3.1 *Given the switched system* (3.3) *and the cost function* $J_{\mu,\eta}(z,m)$ (*see Eq.* (3.6)), *then the Value Iteration* (3.10) *converges to the optimal cost* $J^*(z,m) \triangleq \min_{\mu,\eta} J_{\mu,\eta}(z,m)$, *i.e.*,

$$J^*(z,m) = \lim_{i\to\infty} V_i(z,m),$$

and, moreover, any stationary policy $(\bar{\mu},\bar{\eta})$ *obtained as a limit point of the policies* $\{\mu_i,\eta_i\}$ *computed with the Value Iteration method* (3.11) *is optimal, i.e.*,

$$J^*(z,m) \triangleq J_{\bar{\mu},\bar{\eta}}(z,m).$$

Comprehensive details about the Value Iteration method both in the finite and infinite cases can be found in [4, 6, 21].

3.3.1.1 Proof of the Convergence of the Value Iteration Method

This subsection proves Theorem 3.1. The proof is based on [6, Chap. 3]. However, the problem that is considered here is slightly different. Indeed, in our case, the state vector is composed of a real vector z_k and a discrete variable m_k taking values in a finite set, whereas the state space in [6, Chap. 3] is only $\mathbb{R}^n$. Moreover, in addition to the system noise ω, we consider another random variable β (describing the message dropout) and we deal with two control variables, i.e., the discrete switching decision v and the continuous feedback control input u.

For these reasons, we need to prove that the Value Iteration method given in [6, Chap. 3] and reformulated for our setting in Eqs. (3.10) and (3.11) actually provides the optimal cost and policy. The reasoning that we use to prove our main result is similar to the one in [6, Chap. 3], and in this section we refer to the propositions and corollaries from this book. In spite of the differences between our problem and the one considered in this book, some propositions from the book have proofs that can be extended to our case without any difficulty and hence will be stated here without proof. On the contrary, we will give a detailed proof of the results which require a careful new arrangement of ideas and techniques still inspired from [6], in particular due to the necessity to combine continuity arguments for the vector-valued part of the state and arguments involving the finiteness of the number of modes.

Another comment has to be made about the nature of the noise ω_k in Eq. (3.3). We have defined it as a Gaussian noise, whose realizations are vectors of reals, whereas noise takes values in a countable space in the book [6], on which the proof is based. For simplicity, we present here the proof for the case of noise taking values in a countable space, following [6], but our results can be extended to a more general setting using the results from [7, Chap. 9].

After proving that the cost function is well defined, we show that the optimal solution is the smallest fixed point of the Bellman's equation, and that the Value Iteration method converges to the smallest fixed point the Bellman's equation; therefore, the Value Iteration method converges to the optimal value.

As a shorthand notation, we introduce the operator T, defined as follows. Given a function $V(z,m)$, $TV(z,m)$ is given by:

$$TV(z,m) = \min_{(\hat{u},v)\in\mathbb{U}(z)} \mathsf{E}_{\beta,\omega}\big[\ell_v(z,m,\hat{u},\beta) + \lambda V\big(f_v(z,\hat{u},\beta,\omega),v\big)\big],$$

$$T_{\mu,\eta}V(z,m) = \mathsf{E}_{\beta,\omega}\big[\ell_{\eta(z,m)}\big(z,m,\mu(z,m),\beta\big) + \lambda V\big(f_{\eta(z,m)}\big(z,\mu(z,m),\beta,\omega\big),\eta(z,m)\big)\big].$$

To ease the notation, we can write $T_{u,v}V(z,m)$ where u and v are real values rather than functions.

For the operators T and $T_{\mu,\eta}$, we will denote by the "power" notation T^i and $T^i_{\mu,\eta}$ the composition, e.g., $T^2V(z,m) = (T(TV))(z,m)$ and $T^0V(z,m) = V(z,m)$.

We recall that a fixed point for the operator T is a function $F(z,m)$ such that $F = TF$.

Our problem satisfies the Positivity Assumption, as defined in [6, Chap. 3.1] by:

Assumption 3.3 (Positivity) The cost per stage ℓ_v satisfies

$$0 \le \ell_v(z,m,\hat{u},\beta), \quad \text{for all } (z,m)\in\mathbb{X},\ \hat{u}\in\mathbb{R}^{n_u},\ v\in\mathbb{M} \text{ and } \beta\in\{0,1\}.$$

Remark 3.3 The Positivity Assumption implies the monotonicity of T and $T_{\mu,\eta}$:

$$\begin{aligned} &V(z,m) \le V'(z,m) \quad \Rightarrow \\ &TV(z,m) \le TV'(z,m) \quad \text{and} \quad T_{\mu,\eta}V(z,m) \le T_{\mu,\eta}V'(z,m). \end{aligned} \tag{3.12}$$

We define V_0 as the zero function on $\mathbb{X}$,

$$V_0(z,m) = 0, \quad \forall (z,m)\in\mathbb{X}.$$

We notice that positivity implies

$$V_0(z,m) \le TV_0(z,m) \le \cdots \le T^iV_0(z,m), \quad \forall i > 1. \tag{3.13}$$

Indeed, $TV_0(z,m) \ge 0 = V_0(z,m)$, and then the other inequalities follow by recursively applying the monotonicity property (3.12).

We define the function V_∞ as the following limit:

$$V_\infty(z,m) = \lim_{i\to\infty} T^iV_0(z,m).$$

The limit exists thanks to monotonicity, see Eq. (3.13). However, notice that V_∞ might take the value ∞.

The proof of the convergence of the Value Iteration method follows five steps:

1. J^* is a fixed point of T (Proposition 3.1);
2. J^* is the smallest fixed point of T (Proposition 3.2);
3. V_∞ is a fixed point of T (Proposition 3.3);

4. $V_\infty = J^*$ (Corollary 3.2);
5. The Value Iteration method also provides an optimal stationary policy (Proposition 3.5).

As a first step, we recall Bellman's equation, which states that J^* is a fixed point of T:

Proposition 3.1 (Proposition 3.1.1 from [6]) *The optimal cost function J^* is a fixed point for T, i.e., J^* satisfies*

$$J^*(z,m) = \min_{(\hat{u},v)\in\mathbb{U}(z)} \mathsf{E}_{\beta,\omega}\big[\ell_v(z,m,\hat{u},\beta) + \lambda J^*\big(f_v(z,\hat{u},\beta,\omega),v\big)\big], \quad \forall (z,m)\in\mathbb{X}$$

or, equivalently,

$$J^* = TJ^*.$$

Proof See proof of Proposition 3.1.1 in [6]. □

In the second step, we establish that J^* is the smallest fixed point of T.

Proposition 3.2 (Proposition 3.1.2 from [6]) *If $\tilde{J}:\mathbb{X}\to(-\infty,\infty]$ satisfies $\tilde{J}\geq T\tilde{J}$ and $\tilde{J}\geq 0$, then $\tilde{J}\geq J^*$.*

Proof See proof of Proposition 3.1.2 in [6]. □

This proposition leads to the following corollary:

Corollary 3.1 (Corollary 3.1.2.1 from [6]) *If $\tilde{J}:\mathbb{X}\to(-\infty,\infty]$ satisfies $\tilde{J}\geq T_{\mu,\eta}\tilde{J}$ and $\tilde{J}\geq 0$, then $\tilde{J}\geq J_{\mu,\eta}$.*

Proof The proof is obtained with a simple trick, introduced in [6] in the proof of Corollary 1.1.2.1 and used to prove Corollary 3.1.2.1. The idea is that the above proposition can be applied to a modified optimization problem, where the minimization in (3.7) is performed over a restricted set of policies, containing only one policy μ,η. □

In the third step, we show that V_∞ is a fixed point of T, i.e., that V_∞ satisfies $V_\infty = TV_\infty$. This is the part of the proof where some care is needed in adapting the proofs from [6] to our setting where part of the state is a vector of reals and part of the state is finite. Before stating and proving the main proposition, we make a remark about continuity, which will be used in the proof.

Remark 3.4 If $V(z,m)$ a continuous function of z, for any fixed $v\in\mathbb{M}_1$, then $\mathsf{E}_{\beta,\omega}[\ell_v(z,m,\hat{u},\beta)+\lambda V(f_v(z,\hat{u},\beta,\omega),v)]$ is a continuous function of $\hat{u}$, which tends to $+\infty$ for $\|\hat{u}\|\to\infty$.

Moreover, given $V_0(z,m)=0\ \forall(z,m)$, it holds that, for all $k\geq 0$, $T^kV_0(z,m)$ is a continuous function of z.

Proposition 3.3 *V_∞ is a fixed point of T, i.e.,*

$$V_\infty = T V_\infty.$$

Moreover, letting $(\bar{\mu}, \bar{\eta})$ be the stationary policy corresponding to any limit point of the optimal policies along the Value Iterations,

$$V_\infty = T_{\bar{\mu},\bar{\eta}} V_\infty.$$

Proof This proof is inspired by the proofs of [6, Proposition 3.1.6] and [6, Proposition 3.1.7], adapted to our setting.

First, we prove that $V_\infty \leq T V_\infty$. By (3.13), we have

$$V_0 \leq T V_0 \leq T^2 V_0 \leq \cdots \leq T^i V_0 \leq \cdots \leq V_\infty.$$

In particular, for all $i \geq 0$, $T^i V_0 \leq V_\infty$, from which, by applying T to both sides and recalling monotonicity (see Remark 3.3), we get

$$T^{i+1} V_0 \leq T V_\infty.$$

Taking the limit when i goes to infinity, we obtain

$$V_\infty \leq T V_\infty.$$

Then, in the second and main part of the proof, we will show that any stationary policy $(\bar{\mu}, \bar{\eta})$ obtained as a limit point of the Value Iteration method is such that $V_\infty \geq T_{\bar{\mu},\bar{\eta}} V_\infty$.

Consider any $(z, m) \in \mathbb{X}$. If $V_\infty(z, m) = \infty$, then trivially $V_\infty(z, m) \geq T_{\bar{\mu},\bar{\eta}} V_\infty(z, m)$ for any policy $(\bar{\mu}, \bar{\eta})$. If $V_\infty(z, m) < \infty$, then consider, for all $i \geq 0$,

$$(\hat{u}_i, v_i) = \underset{(\hat{u},v)\in\mathbb{U}(\bar{z})}{\arg\min} \; \mathsf{E}_{\beta,\omega}\big[\ell_v(\bar{z}, \bar{m}, \hat{u}, \beta) + \lambda T^i V_0\big(f_v(\bar{z}, \hat{u}, \beta, \omega), v\big)\big].$$

Notice that this is one of the possibly multiple minimizers and that its existence is ensured by Remark 3.4.

Now we want to prove that:

- There exists a subsequence of $\{(\hat{u}_i, v_i)\}$ which is convergent,
- And for all limit points $(\hat{\bar{u}}, \bar{v})$ of $\{(\hat{u}_i, v_i)\}$, $T_{\hat{\bar{u}},\bar{v}} V_\infty(z, m) \leq V_\infty(z, m)$.

Notice that v_i is eventually constant on any convergent subsequence, and there exist subsequences with constant v_i.

For all $j \geq 0$ such that $j \leq i$, from Eq. (3.13), it holds that

$$T^j V_0(z, m) \leq T^i V_0(z, m).$$

The monotonicity of $T_{\hat{u}_i,v_i}$ yields

$$\big(T_{\hat{u}_i,v_i}\big(T^j V_0\big)\big)(z, m) \leq \big(T_{\hat{u}_i,v_i}\big(T^i V_0\big)\big)(z, m).$$

We notice that

$$\left(T_{\hat{u}_i,v_i}\left(T^i V_0\right)\right)(z,m) = T^{i+1} V_0(z,m),$$

and that

$$T^{i+1} V_0(z,m) \leq V_\infty(z,m),$$

according to Eq. (3.13). This implies the following:

$$\left(T_{\hat{u}_i,v_i}\left(T^j V_0\right)\right)(z,m) \leq V_\infty(z,m).$$

Look at any subsequence with $v_i = \bar{v}$, take $j = 0$, and notice that

$$T_{\hat{u}_i,\bar{v}} V_0(z,m) \leq V_\infty(z,m)$$

implies the existence of a subsequence with $i \in \mathbb{K}$, where $\mathbb{K}$ is an infinite subset of integers, such that

$$\lim_{i\to\infty, i\in\mathbb{K}} \hat{u}_i = \bar{\hat{u}}.$$

Indeed, the set $\{u : T_{u,\bar{v}} V_0(z,m) \leq V_\infty(z,m)\}$ is compact, and thus such a $\bar{\hat{u}}$ exists.

Take any limit point $(\bar{\hat{u}}, v)$, take any subsequence $\{(\hat{u}_i, \bar{v})\}_{i\in\mathbb{K}}$ converging to $(\bar{\hat{u}}, v)$, for all j, for all $i \in \mathbb{K}$, $i \geq j$, then

$$T_{\hat{u}_i,\bar{v}} T^j V_0(z,m) \leq V_\infty(z,m). \tag{3.14}$$

Taking the limit of Eq. (3.14) as i goes to infinity, $i \in \mathbb{K}$, $i \geq j$,

$$\forall j: \quad T_{\bar{\hat{u}},\bar{v}} T^j V_0(z,m) \leq V_\infty(z,m).$$

Now taking the limit as j goes to infinity, it holds that

$$T_{\bar{\hat{u}},\bar{v}} V_\infty(z,m) \leq V_\infty(z,m).$$

If we define a stationary policy (μ, η) as follows:

$$\left(\bar{\mu}(z,m), \bar{\eta}(z,m)\right) = \begin{cases} (\bar{\hat{u}}, \bar{v}) \text{ as defined above if } V_\infty(z,m) < \infty, \\ \text{an arbitrary } (\bar{\hat{u}}, \bar{v}) \text{ if } V_\infty(z,m) = \infty, \end{cases}$$

we have proved that $T_{\bar{\mu},\bar{\eta}} V_\infty \leq V_\infty$.

For the final part of the proof, we will show that $V_\infty = T_{\bar{\mu},\bar{\eta}} V_\infty = T V_\infty$.

By noticing that the optimal policy leads to the smallest Value Function, i.e., that

$$T V_\infty \leq T_{\bar{\mu},\bar{\eta}} V_\infty,$$

and by using the first and second part of the proof, we have the following chain of inequalities

$$V_\infty \leq T V_\infty \leq T_{\bar{\mu},\bar{\eta}} V_\infty \leq V_\infty,$$

which implies that all the above inequalities are indeed equalities. This ends the proof of Proposition 3.3. □

The fourth step consists in proving that V_∞ is indeed the optimal cost function. We know that V_∞ is a fixed point of T, but we need to prove that it is the smallest one, to conclude that $V_\infty = J^*$. To this end, we can use the following result.

Proposition 3.4 (Proposition 3.1.5 from [6]) *If $V_0 \leq V \leq J^*$ and*

$$V_\infty(z, m) = T V_\infty(z, m), \quad \forall (z, m) \in \mathbb{X},$$

then

$$\lim_{i \to \infty} T^i V(z, m) = J^*(z, m), \quad \forall (z, m) \in \mathbb{X}.$$

Proof See the proof of [6, Proposition 3.1.5]; $V_\infty = J^*$ follows by taking $V(z, m) = V_0(z, m) = 0$ for all $(z, m) \in \mathbb{X}$. □

Choosing $V(z, m) = V_0(z, m) = 0$, Proposition 3.3 ensures that the assumptions of Proposition 3.4 are satisfied, which immediately proves the following result.

Corollary 3.2 *The value iteration converges to the optimal cost, i.e.,*

$$V_\infty = J^*.$$

In the last step, we prove that the stationary policy derived in the proof of Proposition 3.3 is an optimal policy.

Proposition 3.5 *The policy $(\bar{\mu}, \bar{\eta})$ derived in Proposition 3.3 is optimal, i.e.,*

$$J_{\bar{\mu},\bar{\eta}}(z, m) = J^*(z, m), \quad \forall (z, m) \in \mathbb{X}.$$

Proof From Proposition 3.3, $V_\infty = T_{\bar{\mu},\bar{\eta}}$ and, by Corollary 3.2 ,$V_\infty = J^*$, so that

$$J^* = T_{\bar{\mu},\bar{\eta}} J^*.$$

This result allows applying Corollary 3.1 with $\tilde{J} = J^*$, which leads to

$$J^* \geq J_{\bar{\mu},\bar{\eta}}.$$

On the other hand, the optimal cost function is the smallest possible cost, i.e., $J^* \leq J_{\bar{\mu},\bar{\eta}}$. This implies that $J_{\bar{\mu},\bar{\eta}} = J^*$. □

3.3.2 Numerical Computation of the Optimal Solution

While the previous section provides an iterative algorithm to derive an optimal joint control law and switching policy, this section discusses the implementation issues of such a method. First, we highlight the practical procedure to obtain an optimal solution from the Value Iteration method. Then we study the particular case of deterministic formulation which introduces an analytical expression of the Value Function to derive the optimal solution.

3.3.2.1 Implementation of the Optimal Solution

The method we have presented to solve the optimization problem is composed of two parts.

The first part is run offline, and it provides the optimal joint policy to switch between radio-modes and to control the feedback loop. It consists in iterating the Value Function $V_i(z, m)$ to converge to the optimal cost function, which provides the solution of our optimization problem. A caveat is that, at each iteration, we need to compute a function of (z, m), where (z, m) takes values in an uncountable space $\mathbb{X}$. A common way to implement such iterations in practice is to partition (a portion of) $\mathbb{X}$ in a grid, then compute the Value Function $V_{i+1}(z, m)$ at the grid points only, by using interpolation to find $V_i(f_v(z))$ when $f_v(z)$ is not on the grid. The subset of $\mathbb{X}$ is chosen to fit the domain of interest the system is supposed to lie in. The number of grid points is chosen as a trade-off between the precision needed for the final application and the computation burden it implies. This approach has been taken, e.g., in [31], and provides a look-up table for $\eta_i(z, m)$ at all grid points.

Computationally, it is very heavy, although this is not a major issue since the long computations are done offline. Some minor drawbacks of this approach are that the solution is limited to a finite domain, and that the numerical approximations and interpolation might introduce some errors giving a sub-optimal solution.

The second part, run online on the smart sensor, consists in computing the optimal switching decision, v_k^*, only for the current switched system state (z_k, m_k) at time k, according to the switching policy derived offline. If the switching decision schedules a transmission, then the update of the control law is computed for the current state, also from the feedback law derived offline.

Depending on the way the feedback law and the switching policy are derived in the offline computation, some computations may be needed online to *extract* the values for a given state (z_k, m_k). This is the case in the deterministic case presented in Sect. 3.3.2.2. To limit the computation burden online in the smart sensor, it is possible to compute offline the optimal switching decisions and feedback control inputs for a finite number of states, on a given grid on a portion of the state space. This ends up with a look-up table, easily implementable on nodes with limited computational capacity.

Practical Convergence of the Value Iteration Method The Value Function iteration method consists in running a recursion an infinite number of times in order to obtain the optimal policy. Obviously, the recursion is not run an infinite number of times in practice, and this raises the question of defining a stopping criterion. Let us assume that the iteration is run using a discretization scheme on a grid on a portion of $\mathbb{X}$, denoted $\bar{\mathbb{X}}$. We propose a stopping criterion by introducing a scalar ξ whose choice depends on the precision needed:

If $\|V_{i+1}(z,m) - V_i(z,m)\| < \xi \quad \forall (z,m) \in \bar{\mathbb{X}}$,

then we consider that the Value Function is close enough to the optimal value.

3.3.2.2 The Deterministic Case

As we explained in Sect. 3.3.2.1, it is difficult to find an analytical expression of the Value Function which can be used to compute the iterations. However, in the case where the system is deterministic and no channel dropouts are considered, a quasi-analytical expression can be derived. In this case, we can exploit a structure that the Value Function preserves along the iterations. This is inspired both by the results in classic LQ optimal control and by the work in [21]. As a reminder, in LQ control, for any i the Value Function is a quadratic function of z, i.e., $V_i(z) = z^\top \Pi_i z$.

In the rest of this subsection, we consider the deterministic system, i.e., $w_k = 0$ and $\beta_k = 1$ for all $k \geq 0$, namely that the system is not affected by noise and that the communication channel never loses messages. This implies that the control input $\hat{u}$ computed at the sensor side is always the one applied to the system at the actuator side u. Hereafter we drop the notation $\hat{u}$ to use only u. Also, to ease the notation, we use $f_v(z,u)$ and $\ell_v(z,m,u)$ rather than $f_v(z,u,1,0)$ and $\ell_v(z,m,u,1)$, respectively. Finally, note that in the case where there are no message dropouts, there is no need for several transmission power levels. Then, the number of transmitting modes is limited to 1 in the rest of this subsection.

For our problem, the structure is more involved than in the LQ case: as we show hereafter, $V_i(z,m)$ is the minimum on some finite set of quadratic functions of the form $z^\top \Pi z + \pi_m$. Indeed, if $V_0(z,m) \equiv 0 \ \forall (z,m)$, then the iterations (3.10) give Value Functions $V_i(z,m)$ such that

$$V_i(z,m) = \min_{(\Pi,\boldsymbol{\pi}) \in \mathcal{P}_i} \left\{ z^\top \Pi z + \pi_m \right\},$$

where the set $\mathcal{P}_i$ is composed of elements $(\Pi, \boldsymbol{\pi})$, where Π is a symmetric matrix and $\boldsymbol{\pi} = \left[\pi_1, \pi_2, \ldots, \pi_N\right] \in \mathbb{R}^N$ is a vector of non-negative scalars, and π_m represents the mth component of $\boldsymbol{\pi}$.

The correctness of the above expressions for $\mathcal{P}_i$ can be proved using mathematical induction, which also gives an explicit recursive construction of the set $\mathcal{P}_i$:

(a) Initial Step By definition, $V_0(z,m)=0$ for all $(z,m)\in\mathbb{X}$. $V_1(z,m)$ is computed using Eq. (3.10):

$$\begin{aligned}V_1(z,m) &= \min_{(u,v)\in\mathbb{U}(z)}\left\{\ell_v(z,m,u)+\lambda V_0\big(f_v(z,u),v\big)\right\}\\ &= \min_{(u,v)\in\mathbb{U}(z)}\left\{z^\top Q_v z+u^\top R_v u+\theta_{m,v}\right\}\\ &= \min\Big\{\min_{u\in\mathbb{R}^{n_u}}\left\{z^\top Q_{CL}z+u^\top R_{CL}u+\theta_{m,1}\right\};\ \min_{v\in\mathbb{M}_2}\left\{z^\top Q_{OL}z+\theta_{m,v}\right\}\Big\}.\end{aligned} \tag{3.15}$$

One can explicitly find the value u^* that minimizes the first expression in Eq. (3.15), which is clearly $u^*=0$. This yields

$$V_1(z,m)=\min\Big\{z^\top Q_{CL}z+\theta_{m,1};\, z^\top Q_{OL}z+\min_{v\in\mathbb{M}_2}\{\theta_{m,v}\}\Big\},$$

$$V_1(z,m)=\min_{(\Pi,\boldsymbol{\pi})\in\mathcal{P}_1}\left\{z^\top\Pi z+\pi_m\right\},$$

where

$$\mathcal{P}_1=\left\{\begin{array}{l}\big(Q_{CL},\big[\theta_{1,1}\quad\theta_{2,1}\quad\cdots\quad\theta_{N,1}\big]\big),\\ \big(Q_{OL},\big[\min_{v\in\mathbb{M}_2}\{\theta_{1,v}\}\quad\min_{v\in\mathbb{M}_2}\{\theta_{2,v}\}\quad\cdots\quad\min_{v\in\mathbb{M}_2}\{\theta_{N,v}\}\big]\big)\end{array}\right\}.$$

(b) Inductive Step We assume that the Value Function at the ith iteration is given by

$$V_i(z,m)=\min_{(\Pi,\boldsymbol{\pi})\in\mathcal{P}_i}\left\{z^\top\Pi z+\pi_m\right\}.$$

The computation of the Value Function at the next iteration, using Eq. (3.10), is very similar to the base case:

$$\begin{aligned}&V_{i+1}(z,m)\\ &\quad=\min_{(u,v)\in\mathbb{U}(z)}\left\{\ell_v(z,m,u)+\lambda V_i\big(f_v(z,u),v\big)\right\}\\ &\quad=\min_{(u,v)\in\mathbb{U}(z)}\Big\{z^\top Q_v z+u^\top R_v u+\theta_{m,v}\\ &\qquad\qquad+\lambda\min_{(\Pi,\boldsymbol{\pi})\in\mathcal{P}_i}\left\{z^\top\Phi_v^\top\Pi\Phi_v z+u^\top\Gamma_v^\top\Pi\Gamma_v u+2u^\top\Gamma_v^\top\Pi\Phi_v z+\pi_v\right\}\Big\}\\ &\quad=\min\Big\{\min_{u\in\mathbb{R}^{n_u},(\Pi,\boldsymbol{\pi})\in\mathcal{P}_i}\left\{z^\top\big(Q_{CL}+\lambda\Phi_{CL}^\top\Pi\Phi_{CL}\big)z+u^\top\big(R_{CL}+\lambda\Gamma_{CL}^\top\Pi\Gamma_{CL}\big)u\right.\\ &\qquad\left.+\theta_{m,1}+2\lambda u^\top\Gamma_{CL}^\top\Pi\Phi_{CL}z+\lambda\pi_1\right\};\\ &\qquad\min_{v\in\mathbb{M}_2,(\Pi,\boldsymbol{\pi})\in\mathcal{P}_i}\left\{z^\top\big(Q_{OL}+\lambda\Phi_{OL}^\top\Pi\Phi_{OL}\big)z+\theta_{m,v}+\lambda\pi_v\right\}\Big\}.\end{aligned} \tag{3.16}$$

Note that $z\Phi_v^\top \Pi \Gamma_v u = u^\top \Gamma_v^\top \Pi \Phi_v z$ since the Π matrices in $\mathcal{P}_j$ are symmetric for any admissible j, see Lemma 3.1.

We denote $\Psi(u) = z^\top (Q_{CL} + \lambda \Phi_{CL}^\top \Pi \Phi_{CL}) z + u^\top (R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL}) u + \theta_{m,1} + 2\lambda u^\top \Gamma_{CL}^\top \Pi \Phi_{CL} z$. One can compute the value of u that minimizes the previous equation, denoted u^*, as follows:

$$\frac{\partial \Psi(u)}{\partial u} = \big(2\big(R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL}\big)u + 2\lambda \Gamma_{CL}^\top \Pi \Phi_{CL} z\big),$$

$$\left.\frac{\partial \Psi(u)}{\partial u}\right|_{u^*} = 0 \quad \Rightarrow \quad u^* = -\big(R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL}\big)^{-1} \lambda \Gamma_{CL}^\top \Pi \Phi_{CL} z \triangleq -\kappa_\Pi z.$$

We can check that u^* actually exists and is a minimum. First of all, one should notice that $(R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL})$ is positive definite, as the sum of the positive definite matrices $\Gamma_{CL}^\top \Pi \Gamma_{CL}$ and R_{CL}, then the inverse exists, so as κ_Π. The same argument can be used to prove that the extremum is actually a minimum since one can check that the Hessian matrix of $\Psi(u^*)$ is positive definite. Note also that the minimum in Eq. (3.15) is indeed attained for u^*.

We can also check that the minimization on u provides a state feedback depending only on z. Assuming that Π can be written $\left[\begin{smallmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12} & \Pi_{22} \end{smallmatrix}\right]$, and denoting $\hat{R}_{CL} \triangleq \bar{R} + \lambda(B^\top \Pi_{11} B + B^\top \Pi_{12} + \Pi_{12} B + \Pi_{22})$, we have

$$\begin{aligned} \kappa_\Pi &= \big(R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL}\big)^{-1} \lambda \Gamma_{CL}^\top \Pi \Phi_{CL} \\ &= (\hat{R}_{CL})^{-1} \lambda \big[(B^\top \Pi_{11} + \Pi_{12}) A \quad 0\big]. \end{aligned}$$

From the form of κ_Π, one sees that only the component on x from z is used to compute u^*.

We notice that $\Psi(u^*)$ can be written as follows:

$$\begin{aligned} \Psi\big(\hat{u}^*\big) &= z^\top \big(Q_{CL} + \lambda \Phi_{CL}^\top \Pi \Phi_{CL}\big) z + u^\top \overbrace{\big[\big(R_{CL} + \lambda \Gamma_{CL}^\top \Pi \Gamma_{CL}\big)u + \lambda \Gamma_{CL}^\top \Pi \Phi_{CL} z\big]}^{=0} \\ &\quad - \lambda u^\top \Gamma_{CL}^\top \Pi \Phi_{CL} z + \theta_{m,1}, \\ \Psi\big(\hat{u}^*\big) &= z^\top \big(Q_{CL} + \lambda \Phi_{CL}^\top \Pi \Phi_{CL}\big) z - \lambda z^\top \kappa_\Pi^\top \Gamma_{CL}^\top \Pi \Phi_{CL} z + \theta_{m,1}. \end{aligned} \tag{3.17}$$

Equations (3.16) and (3.17) yield

$$\begin{aligned} V_{i+1}(z,m) = \min\Big\{ & z^\top \big(Q_{CL} + \lambda \Phi_{CL}^\top \Pi \Phi_{CL} - \lambda \kappa_\Pi^\top \Gamma_{CL}^\top \Pi \Phi_{CL}\big) z + \theta_{m,1} + \lambda \pi_1; \\ & z^\top \big(Q_{OL} + \lambda \Phi_{OL}^\top \Pi \Phi_{OL}\big) z + \min_{v \in \mathbb{M}_2} \{\theta_{m,v} + \lambda \pi_v\} \Big\}, \end{aligned}$$

$$V_{i+1}(z,m) = \min_{(\Pi, \boldsymbol{\pi}) \in \mathcal{P}_{i+1}} \big\{ z^\top \Pi z + \pi_m \big\},$$

where

$$\mathcal{P}_{i+1} = \mathcal{P}_{i+1}^{(1)} \cup \mathcal{P}_{i+1}^{(2)},$$

$$\mathcal{P}_{i+1}^{(1)} \triangleq \Big\{\big(Q_{CL} + \lambda \Phi_{CL}^{\top} \Pi \Phi_{CL} - \lambda \kappa_{\Pi}^{\top} \Gamma_{CL}^{\top} \Pi \Phi_{CL},$$
$$\begin{bmatrix}(\theta_{1,1} + \lambda\pi_1) & (\theta_{2,1} + \lambda\pi_1) & \cdots & (\theta_{N,1} + \lambda\pi_1)\end{bmatrix}\big)$$
$$\text{such that } (\Pi, \boldsymbol{\pi}) \in \mathcal{P}_i \text{ and } \kappa_{\Pi} = \big(R_{CL} + \lambda \Gamma_{CL}^{\top} \Pi \Gamma_{CL}\big)^{-1} \lambda \Gamma_{CL}^{\top} \Pi \Phi_{CL}\Big\},$$

$$\mathcal{P}_{i+1}^{(2)} \triangleq \left\{ \left(Q_{OL} + \lambda \Phi_{OL}^{\top} \Pi \Phi_{OL}, \begin{bmatrix} \min_{v \in \mathbb{M}_2}\{\theta_{1,v} + \lambda\pi_v\} \\ \min_{v \in \mathbb{M}_2}\{\theta_{2,v} + \lambda\pi_v\} \\ \vdots \\ \min_{v \in \mathbb{M}_2}\{\theta_{N,v} + \lambda\pi_v\} \end{bmatrix}^{\top} \right) \text{such that } (\Pi, \boldsymbol{\pi}) \in \mathcal{P}_i \right\}. \tag{3.18}$$

Lemma 3.1 *The Π matrices in $\mathcal{P}_i$ are symmetric positive definite for any $i > 0$.*

Proof The recursion used to compute the Π matrices is given in Eq. (3.18). Each Π matrix in $\mathcal{P}_i$ generates one matrix in $\mathcal{P}_{i+1}^{(1)}$ and another in $\mathcal{P}_{i+1}^{(2)}$, denoted hereafter $\Pi_{+}^{(1)}$ and $\Pi_{+}^{(2)}$, respectively. Equation (3.18) yields

$$\begin{aligned}
\Pi_{+}^{(1)} &= Q_{CL} + \lambda \Phi_{CL}^{\top} \Pi \Phi_{CL} - \lambda \kappa_{\Pi}^{\top} \Gamma_{CL}^{\top} \Pi \Phi_{CL} \\
&= Q_{CL} + \lambda \Phi_{CL}^{\top} \Pi \Phi_{CL} + \kappa_{\Pi}^{\top} R_{CL} \kappa_{\Pi} + \lambda \kappa_{\Pi}^{\top} \Gamma_{CL}^{\top} \Pi \Gamma_{CL} \kappa_{\Pi} - 2\lambda \kappa_{\Pi}^{\top} \Gamma_{CL}^{\top} \Pi \Phi_{CL} \\
&\quad \text{(thanks to Eq. (3.17))} \\
&= Q_{CL} + \lambda \Phi'^{\top} \Pi \Phi' + \kappa_{\Pi}^{\top} R_{CL} \kappa_{\Pi} \\
&\quad \text{where } \Phi' = \Phi_{CL} - \kappa_{\Pi} \Gamma_{CL}
\end{aligned}$$

and

$$\Pi_{+}^{(2)} = Q_{OL} + \lambda \Phi_{OL}^{\top} \Pi \Phi_{OL}.$$

We recall that Q_{CL}, Q_{OL}, and R_{CL} are symmetric positive definite. Also, for any matrices Λ and Ξ of appropriate dimensions, if Λ is symmetric positive definite, then $\Xi^{\top} \Lambda \Xi$ is also symmetric positive definite.

Then, if the Π matrices in $\mathcal{P}_i$ are symmetric positive definite, we have proven that the Π matrices in $\mathcal{P}_{i+1}$ are also symmetric positive definite. Since the Π matrices in $\mathcal{P}_1$ are Q_{CL} and Q_{OL} (which are symmetric positive definite) then this proves that the Π matrices in $\mathcal{P}_i$ are symmetric positive definite for any $i \geq 0$. □

Implementation of the Optimal Solution in the Deterministic Case As explained in Sect. 3.3.2.2, the iterative algorithm in the deterministic case can use the analytical formulation of the Value Function to compute an optimal solution. In this case, the implementation still comprises a two-phases computation process, which is given as follows.

Offline Computation The first stage, run offline, consists in computing the Value Function $V_i(z, m)$ given by

$$V_i(z, m) = \min_{(\Pi, \boldsymbol{\pi}) \in \mathcal{P}_i} \left\{ z^\top \Pi z + \pi_m \right\},$$

where the computation of $\mathcal{P}_i$ is done recursively. After initializing the Value Function to the zero function, the recursion is given by Eq. (3.18). The Value Function converges to the optimal cost function as i goes to infinity. In the case of practical computations, the recursion is stopped after a sufficient number of iterations that can be denoted by I, large enough to assume convergence of the iterative scheme, and the Value Function at iteration I is considered as optimal. In the following subsection, we call $\mathcal{P}_\infty$ the set $\mathcal{P}_I$.

Online Computation The second stage is run online. It consists in computing the optimal switching decision at time k, v_k^*, as a function of the current state of the system (z_k, m_k), as given by the optimal switching policy: $v_k^* = \eta^*(z_k, m_k)$. In order to compute v_k^*, one first needs to compute the value of the Value Function for the current system state, or more precisely to determine the couple (Π_k, π_k) that results from the minimization of the Value Function:

$$(\Pi_k, \boldsymbol{\pi}_k) \triangleq \operatorname*{arg\,min}_{(\Pi, \boldsymbol{\pi}) \in \mathcal{P}_\infty} \left\{ z_k^\top \Pi z_k + \pi_{m_k} \right\}.$$

Then the optimal switching decision is computed as follows:

$$v_k^* = \eta^*(z_k, m_k) \triangleq \begin{cases} 1 & \text{if } (\Pi_k, \boldsymbol{\pi}_k) \in \mathcal{P}_\infty^{(1)}, \\ \arg\min_{v \in \mathbb{M}_2} \{\lambda \boldsymbol{\pi}_k|_v + \theta_{m_k, v}\} & \text{if } (\Pi_k, \boldsymbol{\pi}_k) \in \mathcal{P}_\infty^{(2)}. \end{cases}$$

The notation $\boldsymbol{\pi}_k|_v$ refers to the vth element of the vector $\boldsymbol{\pi}_k$. Notice also that, by construction, $\mathcal{P}_\infty$ is the union of two subsets, $\mathcal{P}_\infty^{(1)}$ and $\mathcal{P}_\infty^{(2)}$.

If a transmission is scheduled at time k, i.e., $v_k^* = 1$, then the optimal feedback control is given by

$$u_k^* = \mu^*(z_k, m_k) \triangleq -\kappa_{\Pi_k} z_k = -\left(R_{CL} + \lambda \Gamma_{CL}^T \Pi_k \Gamma_{CL}\right)^{-1} \lambda \Gamma_{CL}^T \Pi_k \Phi_{CL} z_k.$$

The definitions of η^* and μ^* are obtained naturally from the recursion (3.18).

Stopping Criterion in the Deterministic Case In the deterministic case, where we can implement the recursion on the sets $\mathcal{P}_i$ instead of the grid approach, having the set $\mathcal{P}_i$ converging to a fixed set would provide a stopping criterion. However, the number of elements in this set increases exponentially. Despite $\mathcal{P}_i$ not being convergent to any set $\mathcal{P}_\infty$, clearly the Value Function $V_i(z, m)$ is still ensured to converge to the optimal cost $J^*(z, m)$ by Theorem 3.1. Hence we will simply implement the same stopping criterion as in the general case.

3.4 Simulations

To illustrate the proposed method, we present here a simple example of a first-order unstable system, with the following parameters:

$$x_{k+1} = 1.074x_k - 1.4808u_k + w_k, \qquad T_s = 0.05 \text{ s},$$

where the variance of the noise is $W = 0.02$.

In order to simplify the simulation, and to focus on the radio-mode switching policy, we decide to use a static state feedback. This implies that the control law is fixed and given by

$$\hat{u}_k = \mu(x_k, \tilde{u}_k, m_k) = -Kx_k, \quad \forall k \geq 0,$$

with $K = -0.23$.

The parameters of the cost function are as follows:

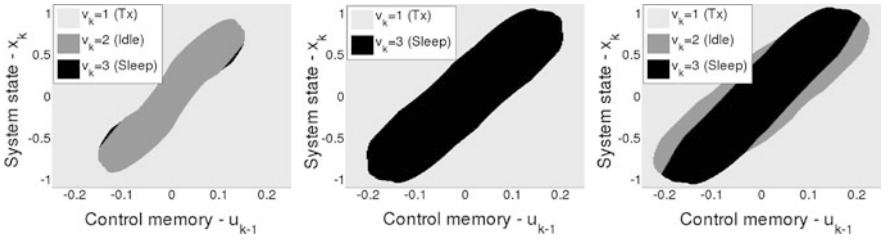
$$\bar{Q} = 0.01, \qquad \bar{R} = 0.1, \qquad \lambda = 0.8.$$

We consider the radio chip Texas Instruments CC1100. We consider one transmitting mode, $N_1 = 1$, namely, the `Tx` mode (mode number 1), and two non-transmitting modes, $N_2 = 2$, the `Idle` and `Sleep` modes, numbered mode 2 and 3, respectively. The values of the transition costs are given in [mJ] and are computed from the datasheet of the radio chip:

$$\begin{bmatrix} \theta_{11} = 2.85 & \theta_{12} = 1.8 & \theta_{13} = 1.9 \\ \theta_{21} = 3.2 & \theta_{22} = 1.4 & \theta_{23} = 6 \times 10^{-5} \\ \theta_{31} = 3.5 & \theta_{32} = 3.7 \times 10^{-3} & \theta_{33} = 6 \times 10^{-5} \end{bmatrix}.$$

Finally, we consider that 30 % of the messages are dropped by the erasure channel, i.e., $\varepsilon = 0.3$.

Our choice of this simple example is motivated by the fact that a first-order system allows a clear pictorial representation of the offline optimization because in this case the mode-switching law can be plotted in a simple figure, as it will be shown in Sect. 3.4.1. Despite the simplifying assumption that the control law is a pre-defined static feedback law $\hat{u}_k = -Kx_k$, this example still illustrates the idea that the mode-switching decision is designed taking into account the control problem. Indeed, the switching policy is derived by solving the optimization problem in Eqs. (3.6) and (3.7) (where now optimization is w.r.t. η only, with fixed control law μ), which takes into account the control problem. Roughly speaking, the intuitive idea is that a transmission is triggered only if this is necessary (no transmission when the state is already near to zero, nor when the memory contains a value similar to the computed control input), and that, moreover, when not transmitting the choice between `Idle` and `Sleep` depends on how likely a new transmission will happen soon. This is in contrast with pre-defined schedules where the choice between modes is done irrespective of the current measured state.



(a) Current mode is `Tx` ($m_k = 1$) (b) Current mode is `Idle` ($m_k = 2$) (c) Current mode is `Sleep` ($m_k = 3$)

Fig. 3.4 Optimal switching policy derived from the offline computation, *light grey* $\Leftrightarrow$ switch to `Tx` ($v_k = 1$), *dark grey* $\Leftrightarrow$ switch to `Idle` ($v_k = 2$), and *black* $\Leftrightarrow$ switch to `Sleep` ($v_k = 3$), the control memory $\tilde{u}_k = u_{k-1}$ is on the x-axis and the system output x_k on the y-axis

3.4.1 Offline Results

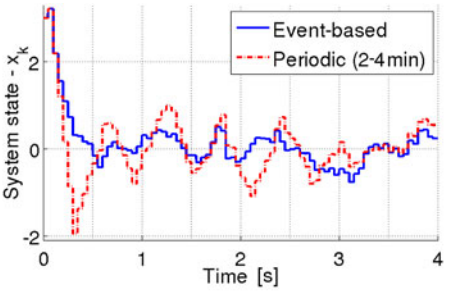
The offline computation provides the switching policy $v = \eta^*(x, \tilde{u}, m)$ which is the optimal for the problem described in Eqs. (3.6) and (3.7), where in our simplified example the optimization is w.r.t. the mode-switching η only, while the control law is fixed to state-feedback $\hat{u}_k = -Kx_k$. The solution is obtained with the Value Iteration Method, i.e., by iterating Eq. (3.11). The policy η^* is stationary, i.e., at any time k, it gives the optimal switching decision $v^* \in \mathbb{M}$ as a function of the current measurement x_k, of the last control input applied to the system $u_{k-1} = \tilde{u}_k$ and of the current radio-mode m_k. This function is depicted in Fig. 3.4, with a sub-figure per each current mode m_k.

We observe that the regions where the radio is switched to non-transmitting modes (colors dark grey and black in Fig. 3.4) are finite sets around the origin $(x_k, \tilde{u}_k) = (0, 0)$, and follow the direction $\tilde{u}_k = -Kx_k$ ($K < 0$ in our example). Outside of these regions, a transmission is forced. This means that when the last control input applied to the system is close to what the state-feedback law would have decided, the switching policy does not send an update.

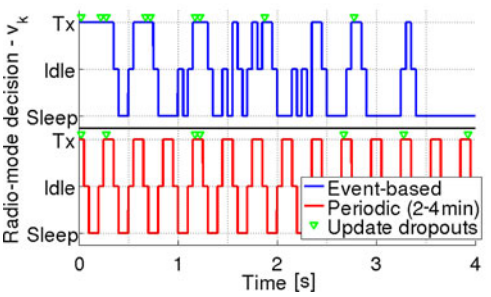
Note that we obtain an event-based radio-mode switching policy. Indeed, a switching occurs only when the state of the system crosses one of the regions in Fig. 3.4.

3.4.2 Online Results

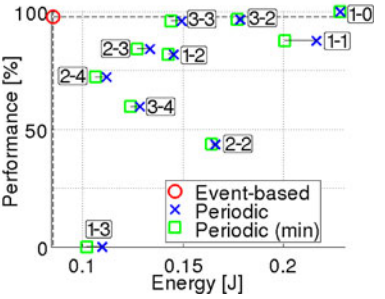
After deriving the switching policy, we run online temporal simulations to observe the behavior of the system. In Fig. 3.5, we compare our event-based switching policy with some periodical ones using the same state feedback law $u = -Kx$, where the radio is alternatively switched to `Tx` and low-consuming modes. We consider various periodic patterns for the radio mode: we will denote by *periodic* i–j a sequence with period $i + j$ where the mode is `Tx` for i consecutive sampling intervals and



(a) State of the system x_k.



(b) Switching decision v_k.



(c) Performance vs. energy consumption E_{Tmax} for several periodic schemes.

Fig. 3.5 Online simulations comparing our event-based switching policy with periodic ones. Additive zero-mean white Gaussian noise and channel dropouts are considered. The *green triangles* in (**b**) indicate time instants where a transmission was intended but a dropout occurred

then is `Sleep` for j intervals. We will then denote by *periodic i–j min* a sequence with the same period and the same `Tx` intervals (so that the control performance is unchanged), but where the mode for the non-transmitting intervals is chosen in $\mathbb{M}_2$ so as to minimize the energy consumption. The online simulations include channel dropouts and additive output noise on the system.

Figure 3.5(a) shows that the system is stabilized in a set around the origin in both cases. In Fig. 3.5(b), one can see the switching decisions for both cases. The green triangles indicate time instants where a transmission was intended but a dropout occurred. When an update is dropped, the Event-Based Control (EBC) scheme holds the `Tx` mode to actually transmit a new control update, while the periodic scheme is not taking dropouts into account. Moreover, the EBC scheme may hold the `Sleep` mode for a long time interval when transmissions cost more than the deviation observed on the state.

Finally, Fig. 3.5(c) compares the EBC scheme to several periodic patterns. In this figure, the performance of the closed loop system is computed from the cost function $(\sum_{i=0}^{\text{Tmax}}(x_i^\top \bar{Q} x_i + u_i^\top \bar{R} u_i)$ where Tmax $= 4$ s) and normalized such that 100 % and 0 % are the best and the worst performances, respectively. The EBC

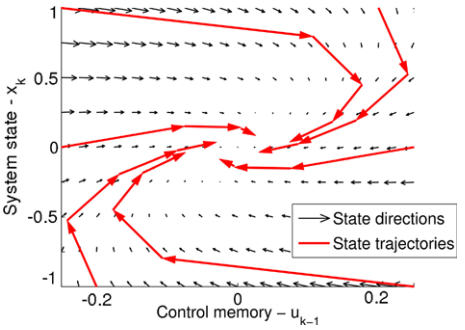


Fig. 3.6 Stability check—Average trajectories with zero-noise for some different initial states are plotted in a vector field giving the initial mean directions of the augmented state z, omitting the radio mode

scheme is very close to the best, although not exactly the best, but it offers the least energy consumption, and especially the best trade-off.

3.4.3 A Posteriori Stability

It is worth mentioning that the commonly used definition for stability is to asymptotically drive the state of the system to the origin as time goes to infinity. This kind of stability cannot be achieved in our setup when the plant is open-loop unstable. Indeed, our scheme bases its energy saving on the action to turn off the radio chip when the transmission cost dominates the cost associated with the control performance. This makes the system run open loop and deviate from the origin. The policy schedules a transmission when the cost associated with the control performance dominates the transmission cost. In our simulations, the system naturally oscillates around the origin without leaving a ball around the origin. Considering this definition, one could expect to check a posteriori that the obtained switching policy is actually stable, as done in [16].

In this section, we check that, on a given domain around the origin, the direction given by our policy leads toward the equilibrium. However, this is not a formal proof but only an illustration. In all our figures, we consider the same system and policy as described and illustrated online and offline in the beginning of this section, and we plot trajectories with zero noise w and averaged w.r.t. to the packet loss β.

Figure 3.6 shows a vector field of the directions of the augmented state $z = [x\ \tilde{u}]^\top$ over a sampling interval. The vector at a point z is given by

$$\frac{1}{\alpha}\big((1-\varepsilon) f_{\eta^*(z,m)}(z, -Kz, 1, \omega) + \varepsilon f_{\eta^*(z,m)}(z, -Kz, 0, \omega) - z\big),$$

where $\omega = 0$ and $m = 2$ (a very similar figure is obtained if $m = 1$ or $m = 3$ are considered instead). The scaling factor α is introduced to avoid intricate figure. We recall that the switched system described by f_v is given in Eq. (3.3) and the switching law η^* is the one minimizing the cost described in Eq. (3.6).

This figure also shows some (average) state trajectories, i.e., for some initial conditions, the bold red arrows connect points which are obtained from an

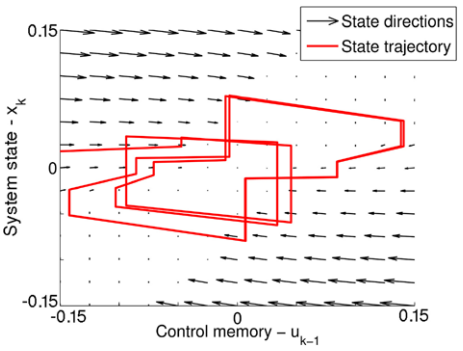


Fig. 3.7 Stability check—Same plot as in Fig. 3.6 zoomed near the origin for only one trajectory over an important number of sampling periods. This figure illustrates that the system keeps oscillating around the origin in a constrained region

initial condition (z_0, m_0) by applying over some consecutive time steps the law $f_{\eta^*(z_i,m_i)}(z_i, -Kz_i, \beta, 0)$. Note that the mode also varies along the trajectory, but is not depicted here. It can be seen that all these trajectories are driven toward the origin.

However, this does not ensure that the system is kept around the origin for the next sampling periods. Figure 3.7 depicts what actually happens around the origin for one trajectory, on a zoomed area. One can observe that the system keeps oscillating around the origin because of the switchings, it deviates when switched to the unstable open-loop behavior, and it converges when transmissions occur.

3.5 Conclusion

In this chapter, we have studied the optimal management of the radio-chip modes of a wireless smart sensor in a networked control problem. The novelty of this work is to introduce the use of more than two radio-modes in a control problem, whereas previous related control-theoretic literature was focused on the choice between two options (`Tx`/`Sleep`). We have considered a networked control problem where the system to be stabilized is linear and can be unstable. The closed-loop includes a single wireless smart sensor whose transmissions to the actuator are performed with an optimal choice of the radio-mode.

For this problem, we have defined a suitable cost function, which describes a trade-off between the control performance and the energy consumption, and whose minimum can be computed with an iterative Dynamic Programming algorithm (the Value Iteration method). We obtained an event-based policy to switch between radio-modes. Limiting the amount of communication to save energy naturally decreases the closed-loop performance, but we show on an example that an event-based approach permits keeping the performance good and saving a larger amount of energy than with a periodic approach.

This work is a first step in the direction of understanding the advantages of radio-mode management in more general networked control problems. Some further work in a similar direction has been addressed in [9], where we have considered a finite-horizon approach to solve the same problem as the one considered in this chapter. In

that work, the optimization is performed over a finite receding horizon, similarly to the framework of Model Predictive Control, an approach which has the advantage of providing tools to prove stability in the practical-Input-to-State stability sense. However, the problem of radio-mode management in the context of closed-loop systems is still open, in particular in the case of multi-sensor and multi-actuator systems.

Acknowledgement This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Al-Karaki, J.N., Kamal, A.E.: Routing techniques in wireless sensor networks: a survey. IEEE Wirel. Commun. **11**(6), 6–28 (2004)
2. Bellman, R.E.: Dynamic Programming. Princeton University Press, Princeton (1957)
3. Bernardini, D., Bemporad, A.: Energy-aware robust Model Predictive Control based on wireless sensor feedback. In: 47th IEEE Conference on Decision and Control, pp. 3342–3347 (2008)
4. Bertsekas, D.P.: In: Dynamic Programming and Optimal Control, vol. 1 (2005)
5. Bertsekas, D.P.: Dynamic programming and suboptimal control: a survey from ADP to MPC. In: European Journal of Control, Fundamental Issues in Control, vol. 11 (2005)
6. Bertsekas, D.P.: In: Dynamic Programming and Optimal Control, vol. 2 (2007)
7. Bertsekas, D.P., Shreve, S.E.: Stochastic Optimal Control: The Discrete Time Case. Athena Scientific, Nashua (1978)
8. Brownfield, M.I., Nelson, T., Midkiff, S., Davis, N.J.: Wireless sensor network radio power management and simulation models. Open. Electr. Electron. Eng. J. **4**, 21–31 (2010)
9. Cardoso de Castro, N.: Energy-aware control and communication co-design in wireless networked control systems. Ph.D. thesis, Grenoble INP–GIPSA-lab–Inria (2012)
10. Cardoso de Castro, N., Canudas de Wit, C., Johansson, K.H.: On energy-aware communication and control co-design in wireless networked control systems. In: 2nd IFAC Workshop on Estimation and Control of Networked Systems, NecSys'10 (2010)
11. Chen, Y., Liestman, A., Liu, J.: A hierarchical energy-efficient framework for data aggregation in wireless sensor networks. IEEE Trans. Veh. Technol. **55**(3), 789–796 (2006)
12. Colandairaj, J., Irwin, G., Scanlon, W.: Wireless networked control systems with QoS-based sampling. IET Control Theory Appl. **1**(1), 430–438 (2007)
13. Correia, L.H., Macedo, D.F., dos Santos, A.L., Loureiro, A.A., Nogueira, J.M.S.: Transmission power control techniques for wireless sensor networks. Comput. Netw. **51**(17), 4765–4779 (2007)
14. Do Valle Costa, O., Fragoso, M., Marques, R.: Discrete-Time Markov Jump Linear Systems. Probability and Its Applications. Springer, Berlin (2005)
15. Fischione, C., Johansson, K.H., Graziosi, F., Santucci, F.: Distributed cooperative processing and control over wireless sensor networks. In: Proceedings of the 2006 International Conference on Wireless Communications and Mobile Computing, pp. 1311–1316 (2006)
16. Görges, D., Izak, M., Liu, S.: Optimal control and scheduling of switched systems. IEEE Trans. Autom. Control **56**(1), 135–140 (2011)
17. Heemels, W., Donkers, M., Teel, A.: Periodic event-triggered control based on state feedback. In: 50th IEEE Conference on Decision and Control and European Control Conference (2011)
18. Hespanha, J., Naghshtabrizi, P., Xu, Y.: A survey of recent results in networked control systems. Proc. IEEE **95**(1), 138–162 (2007)
19. Imer, O.C., Yüksel, S., Başar, T.: Optimal control of LTI systems over unreliable communication links. Automatica **42**, 1429–1439 (2006)

20. Lin, C., He, Y.X., Xiong, N.: An energy-efficient dynamic power management in wireless sensor networks. In: International Symposium on Parallel and Distributed Computing, pp. 148–154 (2006)
21. Lincoln, B.: Dynamic programming and time-varying delay systems. Ph.D. thesis, Department of Automatic Control, Lund Institute of Technology, Sweden (2003)
22. Liu, X., Goldsmith, A.: Wireless network design for distributed control. In: 43rd IEEE Conference on Decision and Control, vol. 3, pp. 2823–2829 (2004)
23. Marco, P.D., Park, P., Fischione, C., Johansson, K.H.: Trend: a timely, reliable, energy-efficient dynamic WSN protocol for control application. In: IEEE International Conference on Communications (2010)
24. Nair, G., Fagnani, F., Zampieri, S., Evans, R.: Feedback control under data rate constraints: an overview. Proc. IEEE **95**(1), 108–137 (2007)
25. Park, P., Fischione, C., Bonivento, A., Johansson, K., Sangiovanni-Vincentelli, A.: Breath: a self-adapting protocol for wireless sensor networks in control and automation. In: 5th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks, pp. 323–331 (2008)
26. Pottie, G.J., Kaiser, W.J.: Wireless integrated network sensors. Commun. ACM **43**(5), 51–58 (2000)
27. Products, C.: CC1100 Low-Power Sub-1 GHz RF Transceiver Texas Instruments
28. Quevedo, D., Ahlén, A., Ostergaard, J.: Energy efficient state estimation with wireless sensors through the use of predictive power control and coding. IEEE Trans. Signal Process. **58**(9), 4811–4823 (2010)
29. Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M., Sastry, S.: Kalman filtering with intermittent observations. IEEE Trans. Autom. Control **49**(9), 1453–1464 (2004)
30. Smith, S., Seiler, P.: Estimation with lossy measurements: jump estimators for jump systems. IEEE Trans. Autom. Control **48**(12), 2163–2171 (2003)
31. Sundström, O., Ambühl, D., Guzzella, L.: On implementation of dynamic programming for optimal control problems with final state constraints. Oil Gas Sci. Technol., Rev. IFP **65**(1), 91–102 (2010)
32. Tang, Z., Glover, I., Evans, A., He, J.: An energy-efficient adaptive DSC scheme for wireless sensor networks. IEEE Trans. Signal Process. **87**(12), 2896–2910 (2007). Special Section: Information Processing and Data Management in Wireless Sensor Networks

Part II
Stochastic Networked Control and Estimation

Decentralized stochastic control arises in multi-stage decision-making with multiple decision-makers having different information and a common objective. This second part of the book collects three contributions investigating fundamental relations between control and information-theoretic quantities.

In Chap. 4, Nayyar et al. present the common-information approach to decentralized stochastic control. The key idea behind this approach is to formulate an equivalent centralized stochastic control problem from the point of view of a fictitious coordinator that observes only the information that is commonly available to all decision-makers. The optimal control problem for the fictitious coordinator is shown to be a partially observable Markov decision process which can be solved using techniques from Markov decision theory. The authors describe this approach for a general model and illustrate it by examples from real-time communication, networked control systems, paging and registration in cellular systems, and multi-access broadcast systems.

Chapter 5 by Asnani et al. reviews some of the recent literature on relations between information- and estimation-theoretic quantities. The chapter begins by exploring the connections between mutual information and causal/non-causal, matched/mismatched estimation for the setting of a continuous-time source corrupted by white Gaussian noise. Relations involving causal estimation, in both matched and mismatched cases, and mutual information persist in the presence of feedback. The authors present a new unified framework, based on Girsanov theory and Itô's Calculus, to derive these relations, and conclude by deriving some new results using this framework.

In Chap. 6, Yüksel discusses the properties of information/measurement channels for stabilization and optimization problems in networked control. First, the chapter considers a finite horizon optimal control problem, and investigates structural and topological properties of such a problem over the space of information channels. The existence of optimal channels is studied, and the structure and existence of optimal quantization policies are investigated, first for static settings and then for dynamic settings. Then, the stabilization problem of open-loop unstable linear systems controlled over communication channels is discussed and tight necessary and sufficient conditions for stochastic stabilizability of such systems driven by Gaussian noise over channels are presented.

Chapter 4
The Common-Information Approach to Decentralized Stochastic Control

Ashutosh Nayyar, Aditya Mahajan, and Demosthenis Teneketzis

4.1 Introduction

Many modern technological systems, such as cyber-physical systems, communication, transportation and social networks, smart grids, sensing and surveillance systems are informationally decentralized. A key feature of informationally decentralized systems is that decisions are made by multiple decision makers that have access to different information. This feature violates the fundamental assumption upon which centralized stochastic control theory is based, namely, that all decisions are made by a centralized decision maker who has access to all the information and perfectly recalls all past observations and decisions/actions. Consequently, techniques from centralized stochastic control cannot be directly applied to decentralized stochastic control problem primarily for the following reason. In centralized stochastic control, the controller's belief on the current state of the system is a sufficient statistic for decision making. A similar sufficient statistic does not work for decentralized stochastic control because controllers have different information and hence their beliefs on the state of the system are not consistent.

Nevertheless, two general approaches that use ideas from centralized stochastic control theory have been used for the solution of decentralized control problems:

A. Nayyar (✉)
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA
e-mail: anayyar@berkeley.edu

A. Mahajan
Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada
e-mail: aditya.mahajan@mcgill.ca

D. Teneketzis
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, USA
e-mail: teneket@umich.edu

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_4,

(i) *the person-by-person approach*; and (ii) *the designer's approach*. A detailed discussion of the features and merits of these approaches, as well as their application to various classes of problems appears in [20]. Here we briefly present the key characteristics of each approach.

The person-by-person approach investigates the decentralized control problem from the viewpoint of one decision-maker, say the ith decision-maker and proceeds as follows: (i) arbitrarily fixes the strategy of all other decision-makers; and (ii) uses centralized stochastic control to derive structural properties for the optimal best-response strategy of the ith decision-maker. The person-by-person approach can, in several problem instances [9, 13, 16–18, 26–30, 33–37, 41, 42], identify qualitative properties of globally optimal control strategies; furthermore, it provides an iterative method to obtain person-by-person optimal strategies [7] which, in general, are not globally optimal.

The designer's approach looks at the decentralized control problem from the point of view of a system designer who knows the system model and the statistics of the primitive random variables, and chooses control/decision strategies for all decision makers. This approach leads to a centralized planning problem whose solution results in globally optimal control strategies. Such strategies are determined by a dynamic program where each step is a functional optimization problem (in contrast to the usual centralized dynamic program where each step is a parameter optimization problem). Thus, the determination of globally optimal strategies via the designer's approach is a computationally formidable problem [40].

In several instances of decentralized control problems [13, 36, 37], the person-by-person approach is used first to identify qualitative properties of globally optimal strategies; then, the designer's approach is employed to determine globally optimal strategies with the identified qualitative property.

In addition to the above mentioned approaches, other methods that exploit system's information structure have been developed for the solution of decentralized control problems. Specifically, solution approaches for systems with partially nested information structure have appeared in [4, 8, 10, 11, 23]; a generalization of partial nestedness called stochastic nestedness was defined and studied in [44]. In [8], it was shown that for linear quadratic Gaussian (LQG) control problems with partially nested information structure, there is an affine control strategy that is globally optimal. In general, the problem of determining optimal control strategies within the class of affine control policies may not be a convex optimization problem; conditions under which it is convex were identified in [2, 24].

Decentralized stochastic control problems with specific models of information sharing among controllers, such as delayed information sharing [1, 19, 31, 43], periodic information sharing [22], broadcast information structure [42], control sharing [3, 12], and systems with common and private observations [14] have also been investigated in the literature.

In [20], a new general model of decentralized stochastic control, called *partial history sharing* information structure, was presented. In this model, it is assumed that: (i) controllers sequentially share part of their past data (observations and control actions) with one another by means of a shared memory; and (ii) all controllers

have perfect recall of the commonly available data (also called the common information). This model subsumes a large class of decentralized control models where information is shared among the controllers. A solution methodology for this model was presented in [20]. This solution methodology is based on the *common information approach* developed in [15] which is applicable to all sequential decision making problems. The common information approach provides a unified framework for several decentralized control problems that had previously been addressed using problem specific solution techniques. The key idea behind this approach is the reformulation of the original decentralized control problem into an equivalent centralized problem from the perspective of a coordinator. The coordinator knows the common information and selects prescriptions that map each controller's local information to its control actions. The optimal control problem at the coordinator is a partially observable Markov decision process (POMDP) that can solved using techniques from Markov decision theory. This approach provides: (i) structural results (qualitative properties) for optimal strategies; and (ii) a dynamic program for obtaining *globally optimal* strategies for all controllers in the original decentralized problem. Notably, the structural results of optimal control strategies obtained by the common information approach cannot be obtained by the person-by-person approach (see [20, Sect. III-A]; and the dynamic program obtained by the common information approach is simpler than that obtained by the designer's approach (see [20, Sect. III-A].

In this chapter, we present the common information approach to decentralized stochastic control. Our objective is to demonstrate that this approach is conceptually powerful as it overcomes some of the fundamental difficulties in decentralized decision making, it has broad applicability, it can resolve a long-standing open problem in decentralized stochastic control, and it can simplify the search for globally optimal strategies.

This chapter is organized as follows. In Sect. 4.2, we first present two examples, one for a one stage decentralized control problem (static team) and the other for a two stage decentralized control problem (dynamic team), that illustrate how the common information approach simplifies the search of globally optimal strategies; then we describe the key steps of the approach. In Sect. 4.3, we present a brief recap of partially observed Markov decision processes (POMDPs) which play a key role in the common information approach. In Sect. 4.4, we illustrate how the common information approach can be used to solve problems that arise in control, communication, and queueing systems. In Sect. 4.5, we demonstrate how our approach can resolve a long-standing open problem [39] in decentralized stochastic control. We conclude in Sect. 4.6 by discussing how the common information circumvents the conceptual difficulties associated with decentralized stochastic control.

4.1.1 Terminology

Decentralized stochastic control problems are also referred to as *team problems* and further classified as *static* and *dynamic* teams. In dynamic teams, the information

observed by a decision maker depends on the control actions of other decision makers, while in static teams it does not; see [7, 8] for details. Decentralized stochastic control problems are typically dynamic team problems.

4.1.2 Notation

Random variables are denoted by upper case letters; their realization by the corresponding lower case letter. For integers $a \leq b$ and $c \leq d$, $X_{a:b}$ is a short hand for the vector $(X_a, X_{a+1}, \ldots, X_b)$. When $a > b$, $X_{a:b}$ equals the empty set. In general, subscripts are used as time index while superscripts are used to index controllers. $\mathbb{P}(\cdot)$ is the probability of an event, $\mathbb{E}[\cdot]$ is the expectation of a random variable. For a collection of functions $\boldsymbol{g}$, we use $\mathbb{P}^{\boldsymbol{g}}(\cdot)$ and $\mathbb{E}^{\boldsymbol{g}}[\cdot]$ to denote that the probability measure/expectation depends on the choice of functions in $\boldsymbol{g}$.

4.2 The Common Information Approach to Decentralized Stochastic Control

The main idea of the common information approach to decentralized stochastic control is to formulate and analyze an alternative but equivalent centralized stochastic control problem. To illustrate this idea, we start with two of the simplest examples of decentralized stochastic control: (i) a two controller static team problem; and (ii) a two controller two-stage dynamic team problem. For both these examples, we show how the common information approach works and simplifies the search of globally optimal strategies. After presenting these examples, we present a high-level description of the main steps of the common information approach.

4.2.1 Illustrative Example 1: A Two Controller Static Team

The following example, which is adapted from [20], illustrates how the common information approach decomposes a static team problem into several smaller subproblems that are easier to solve.

Consider a two controller static team. Nature selects a random variable W. Controller i, $i = 1, 2$, observes a common observation C and a local observation M^i. The observations (C, M^1, M^2) are a function of W.

The controllers select their control actions U^1 and U^2 using control laws g^1 and g^2 of the form

$$U^1 = g^1\big(C, M^1\big), \qquad U^2 = g^2\big(C, M^2\big).$$

The system incurs a loss $\ell(W, U^1, U^2)$.

Suppose all system variables are finite valued and W, C, M^i, U^i take values in finite sets $\mathcal{W}$, $\mathcal{C}$, $\mathcal{M}^i$, and $\mathcal{U}^i$, $i = 1, 2$, respectively. The objective is to choose control laws

$$g^1 : \mathcal{C} \times \mathcal{M}^1 \mapsto \mathcal{U}^1, \qquad g^2 : \mathcal{C} \times \mathcal{M}^2 \mapsto \mathcal{U}^2$$

to minimize

$$J\big(g^1, g^2\big) = \mathbb{E}^{(g^1, g^2)}\big[\ell\big(W, U^1, U^2\big)\big].$$

Since all system variables are finite valued, one solution approach is to find the globally optimal control laws (g^1, g^2) by a brute force search over all possible $\prod_{i=1}^{2} |\mathcal{U}^i|^{|\mathcal{C}||\mathcal{M}^i|}$ control laws. For example, if all system variables are binary valued, we need to search over $2^4 \times 2^4 = 256$ possibilities.

The common information approach reduces the number of possibilities that need to be searched. The main idea of this approach is that instead of specifying the control laws (g^1, g^2) directly, we specify them indirectly as follows. Consider an alternative *coordinated system* in which a *coordinator* observes the common information C and chooses prescriptions (Γ^1, Γ^2), where Γ^i is a mapping from *local information* M^i to control action U^i, according to a *coordination law* ψ that is of the form

$$\big(\Gamma^1, \Gamma^2\big) = \psi(C).$$

The coordinator communicates these prescriptions (Γ^1, Γ^2) to the controllers who use them to generate control actions as follows:

$$U^1 = \Gamma^1\big(M^1\big), \qquad U^2 = \Gamma^2\big(M^2\big).$$

The objective of the coordinated system is to find a coordination law ψ to minimize

$$\tilde{J}(\psi) = \mathbb{E}^{\psi}\big[\ell\big(W, U^1, U^2\big)\big].$$

It is easy to verify that there is an one-to-one correspondence between the control laws (g^1, g^2) of the original system and the coordination law ψ of the coordinated system. The optimization problem at the coordinator is a centralized stochastic optimization problem in which the coordinator is the only decision-maker. To solve this centralized stochastic optimization problem, consider any coordination law ψ and for any $c \in \mathcal{C}$, let $(\gamma_c^1, \gamma_c^2) = \psi(c)$. Write the expected loss $\tilde{J}(\psi)$ as

$$\sum_{c \in \mathcal{C}} \mathbb{P}(C = c) \mathbb{E}\big[\ell\big(W, \gamma_c^1\big(M^1\big), \gamma_c^2\big(M^2\big)\big) \,\big|\, C = c\big].$$

Minimizing $\tilde{J}(\psi)$ is equivalent to separately minimizing, for each value of $c \in \mathcal{C}$, the expected conditional loss $\mathbb{E}[\ell(W, \gamma_c^1(M^1), \gamma_c^2(M^2)) \mid C = c]$ over the choice of (γ_c^1, γ_c^2). One solution approach to solve each of these latter minimizations is by a brute force search over all possible $\prod_{i=1}^{2} |\mathcal{U}^i|^{|\mathcal{M}^i|}$ possibilities. Thus, this approach requires searching over $|\mathcal{C}| \prod_{i=1}^{2} |\mathcal{U}^i|^{|\mathcal{M}^i|}$ possibilities. For example, if all system

variables are binary valued, we need to search over $2 \times 2^2 \times 2^2 = 32$ possibilities. Contrast this by the 256 possibilities that need to be evaluated for a brute force search in the original setup. In general, for this example, the common information approach provides an exponential simplification by reducing the search complexity from

$$\left(\prod_{i=1}^{2} |\mathcal{U}^i|^{|\mathcal{M}^i|}\right)^{|\mathcal{C}|} \quad \text{to } |\mathcal{C}| \prod_{i=1}^{2} |\mathcal{U}^i|^{|\mathcal{M}^i|}.$$

4.2.2 *Illustrative Example 2: A Two-Stage Two-Controller Dynamic Team*

The following example illustrates how the common information approach provides a dynamic programming decomposition in a multi-stage dynamic team problem. Consider a two-stage two-controller dynamic team that evolves as follows.

- At $t = 1$, nature selects a random variable W_1. Controller i, $i = 1, 2$, observes a common observation C_1 and a local observation M_1^i. The observations (C_1, M_1^1, M_1^2) are a function of W_1.

  The controllers select their control actions U_1^1 and U_1^2 using control laws g_1^1 and g_1^2 of the form

$$U_1^1 = g_1^1\big(C_1, M_1^1\big), \qquad U_1^2 = g_1^2\big(C_1, M_1^2\big).$$

- At $t = 2$, nature selects a random variable W_2 that may be correlated with W_1. As in stage 1, controller i, $i = 1, 2$, observes a common observation C_2 and a local observation M_2^i. The difference from stage 1 is that the observations (C_2, M_2^1, M_2^2) are a function of (W_2, U_1^1, U_1^2).

  The controllers select their control actions U_2^1 and U_2^2 using control laws g_2^1 and g_2^2 of the form

$$U_2^1 = g_2^1\big(C_1, C_2, M_1^1, M_2^1\big), \qquad U_2^2 = g_2^2\big(C_1, C_2, M_1^2, M_2^2\big).$$

- At the end of the two stages, the system incurs a loss $\ell(W_1, W_2, U_1^1, U_1^2, U_2^1, U_2^2)$.

Suppose all system variables are finite valued and W_t, C_t, M_t^i, U_t^i take values in finite sets $\mathcal{W}_t$, $\mathcal{C}_t$, $\mathcal{M}_t^i$, and $\mathcal{U}_t^i$, $i = 1, 2$, $t = 1, 2$. The objective is to choose control laws

$$g_1^i : \mathcal{C}_1 \times \mathcal{M}_1^i \mapsto \mathcal{U}_1^i, \qquad g_2^i : \mathcal{C}_1 \times \mathcal{C}_2 \times \mathcal{M}_1^i \times \mathcal{M}_2^i \mapsto \mathcal{U}_1^i, \quad i = 1, 2$$

to minimize

$$J\big(g_1^1, g_2^1, g_1^2, g_2^2\big) = \mathbb{E}^{(g_1^1, g_2^1, g_1^2, g_2^2)}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big)\big].$$

Since all system variables are finite valued, one solution approach is to find globally optimal control strategies $(g_1^1, g_2^1, g_1^2, g_2^2)$ by a brute force search over all possible

$$\prod_{i=1}^{2} |\mathcal{U}_1^i|^{|\mathcal{C}_1||\mathcal{M}_1^i|} |\mathcal{U}_2^i|^{|\mathcal{C}_1||\mathcal{C}_2||\mathcal{M}_1^i||\mathcal{M}_2^i|}$$

control strategies. For example, if all system variables are binary valued, we need to search over $(2^4 \times 2^{16})^2 = 2^{40}$ possibilities.

The common information approach enables us to decompose the above multi-stage optimization problem using a dynamic program. As in the static case, the main idea of the common information approach is that instead of specifying the control strategies $(g_1^1, g_2^1, g_1^2, g_2^2)$ directly, we specify them indirectly as follows. Consider an alternative two-stage *coordinated system* in which a coordinator *with perfect recall* observes the common information C_t at time t and chooses prescriptions (Γ_t^1, Γ_t^2) where Γ_1^i is a mapping from *local information* M_1^i to control action U_1^i while Γ_2^i is a mapping from *local information* (M_1^i, M_2^i) to control action U_2^i. These prescriptions are chosen according to a *coordination strategy* (ψ_1, ψ_2) that is of the form

$$\big(\Gamma_1^1, \Gamma_1^2\big) = \psi_1(C_1), \qquad \big(\Gamma_2^1, \Gamma_2^2\big) = \psi_2(C_1, C_2).$$

At time t, the coordinator communicates prescriptions (Γ_t^1, Γ_t^2) to the controllers who use them to generate control actions as follows:

$$U_1^i = \Gamma_1^i\big(M_1^i\big), \qquad U_2^i = \Gamma_2^i\big(M_1^i, M_2^i\big), \quad i = 1, 2.$$

The objective of the coordinated system is to find coordination strategy (ψ_1, ψ_2) to minimize

$$\tilde{J}(\psi_1, \psi_2) = \mathbb{E}^{(\psi_1, \psi_2)}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big)\big].$$

It is easy to verify that there is a one-to-one correspondence between the control strategies $(g_1^1, g_2^1, g_1^2, g_2^2)$ of the original system and the coordination strategy (ψ_1, ψ_2) of the coordinated system. The multi-stage optimization problem at the coordinator is a centralized stochastic control problem in which the coordinator is the only decision maker and has perfect recall. To solve this centralized stochastic control problem, proceed as follows. Consider any coordination strategy (ψ_1, ψ_2) and any realization $(c_1, c_2) \in \mathcal{C}_1 \times \mathcal{C}_2$ of the common information. Suppose the prescriptions $(\gamma_1^1, \gamma_1^2) = \psi_1(c_1)$ are fixed. Given this information, what is the best choice of the prescriptions $(\gamma_2^1, \gamma_2^2) = \psi_2(c_1, c_2)$ at time $t = 2$? For any choice $(\tilde{\gamma}_2^1, \tilde{\gamma}_2^2)$ of the prescriptions at time $t = 2$, the expected conditional loss is given by

$$\mathbb{E}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big) \mid c_1, c_2, \gamma_1^1, \gamma_1^2, \tilde{\gamma}_2^1, \tilde{\gamma}_2^2\big].$$

Since all the prescriptions are specified, the control actions $(U_1^1, U_1^2, U_1^2, U_2^2)$ are well-defined random variables, and the above conditional expectation is well-defined. To obtain the best choice of the prescriptions at time $t = 2$, minimize the

above conditional expectation over all possible choices of $(\tilde{\gamma}_2^1, \tilde{\gamma}_2^2)$ and define the minimum value as

$$V(c_1, c_2, \gamma_1^1, \gamma_1^2) = \min_{\tilde{\gamma}_2^1, \tilde{\gamma}_2^2} \mathbb{E}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big) \,\big|\, c_1, c_2, \gamma_1^1, \gamma_1^2, \tilde{\gamma}_2^1, \tilde{\gamma}_2^2\big]. \tag{4.1}$$

One solution approach to solve the above minimization is by a brute force search over all possible $\prod_{i=1}^2 |\mathcal{U}_2^i|^{|\mathcal{M}_1^i||\mathcal{M}_2^i|}$ prescription pairs. To find the optimal coordination law ψ_2, we need to solve the above minimization problem *for all possible realizations of the common information* (c_1, c_2) *and choices of past prescription* (γ_1^1, γ_1^2). Thus, we need to solve $|\mathcal{C}_1||\mathcal{C}_2| \prod_{i=1}^2 |\mathcal{U}_1^i|^{|\mathcal{M}_1^i|}$ minimization problems, each requiring the evaluation of $\prod_{i=1}^2 |\mathcal{U}_2^i|^{|\mathcal{M}_1^i||\mathcal{M}_2^i|}$ conditional expectations.

Now that we know how the coordinator selects optimal prescriptions at time $t = 2$, what is the best choice of prescriptions (γ_1^1, γ_1^2) at time $t = 1$? For any realization $c_1 \in \mathcal{C}_1$ and any choice of coordination law $\tilde{\psi}_2$, the expected conditional loss at the coordinator when the prescriptions at time $t = 1$ are $(\tilde{\gamma}_1^1, \tilde{\gamma}_1^2)$ is given as

$$\begin{aligned}
&\mathbb{E}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big) \,\big|\, c_1, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2\big] \\
&\quad = \mathbb{E}\big[\mathbb{E}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big) \,\big|\, c_1, C_2, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2, \tilde{\psi}_2\big] \,\big|\, c_1, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2\big].
\end{aligned} \tag{4.2}$$

Use the optimal prescription at time $t = 2$, which is given by (4.1), to lower bound the conditional expected cost in (4.2) as follows:

$$\begin{aligned}
&\mathbb{E}\big[\mathbb{E}\big[\ell\big(W_1, W_2, U_1^1, U_1^2, U_1^2, U_2^2\big) \,\big|\, c_1, C_2, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2, \tilde{\psi}_2\big] \,\big|\, c_1, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2\big] \\
&\quad \geq \mathbb{E}\big[V\big(c_1, C_2, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2\big) \,\big|\, c_1, \tilde{\gamma}_1^1, \tilde{\gamma}_1^2\big]
\end{aligned} \tag{4.3}$$

with equality if the coordinator uses the optimal prescriptions at time $t = 2$, which are given by (4.1).

One solution approach to select the best prescriptions at time $t = 1$ is to evaluate the conditional expectation in (4.3) for all $\prod_{i=1}^2 |\mathcal{U}_1^i|^{|\mathcal{M}_1^i|}$ choices of $(\tilde{\gamma}_1^1, \tilde{\gamma}_1^2)$. To find the optimal coordination law ψ_1, we need to solve the above minimization problem *for all possible realizations of* c_1. Thus, we need to solve $|\mathcal{C}_1|$ minimization problems, each requiring the evaluation of $\prod_{i=1}^2 |\mathcal{U}_1^i|^{|\mathcal{M}_1^i|}$ conditional expectations.

The above *dynamic program* based on the common information approach requires

$$|\mathcal{C}_1| \prod_{i=1}^2 |\mathcal{U}_1^i|^{|\mathcal{M}_1^i|} + |\mathcal{C}_1||\mathcal{C}_2| \prod_{i=1}^2 |\mathcal{U}_2^i|^{|\mathcal{M}_1^i||\mathcal{M}_2^i|}$$

evaluations.[1] For example, when all variables are binary valued, we need to evaluate 2^{14} conditional expectations. Contrast this with 2^{40} possibilities that need to be evaluated for a brute force search in the original setup. In general, for this example, the common information approach provides an exponential simplification by reducing the search complexity from

$$\left(\prod_{i=1}^{2}|\mathcal{U}_1^i|^{|\mathcal{M}_1^i|}\left(|\mathcal{U}_2^i|^{|\mathcal{M}_1^i||\mathcal{M}_2^i|}\right)^{|\mathcal{C}_2|}\right)^{|\mathcal{C}_1|}$$

to

$$|\mathcal{C}_1|\left(\prod_{i=1}^{2}|\mathcal{U}_1^i|^{|\mathcal{M}_1^i|}+|\mathcal{C}_2|\prod_{i=1}^{2}|\mathcal{U}_2^i|^{|\mathcal{M}_1^i||\mathcal{M}_2^i|}\right).$$

In general, it is possible to improve the computational advantage of the common information approach by:

1. *Identifying irrelevant information at the controllers:* One way of reducing the complexity of coordinator's problem is to show that part of local information is irrelevant for controllers. If this can be established (often by using the person-by-person approach described in Sect. 4.1), then the coordinator's prescription are mappings from the reduced local information to control actions. This reduces the number of possible prescription choices to be considered by the coordinator.
2. *Identifying an information state for the coordinator:* An information state serves as a sufficient statistic for the data available to the coordinator. Instead of finding best prescriptions for all possible realizations of coordinator's data, we only need to find best prescriptions for each realization of coordinator's information state. If the coordinator's decision problem can be shown to be equivalent to some known models of centralized stochastic control (such as Markov decision problems or partially observed Markov decision problem), then we can use stochastic control techniques to find an information state for the coordinator.

4.2.3 The Common Information Approach

The previous two examples illustrate how the common information approach works for simple static and dynamic teams. We generalize this approach to a broad class of decentralized stochastic control systems by proceeding as follows:

[1] We assume that evaluating expected loss or expected conditional loss requires the same computational effort irrespective of the cost function and the probability measure. This analysis is meant to provide a general idea of reduction in complexity, and is not a strict evaluation of the computational benefits of the common information approach.

1. *Construct a coordinated system*: The first step of the approach is to identify the common information at the controllers. The common information at time t must be known to all controllers at t. Define the local information at a controller to be the information left after subtracting the common information from all the data available at that controller. If the common information is non-empty, construct a coordinated system in which at each time a coordinator has access to the common information at that time and selects a set of prescriptions that map each controllers' local information to its control action. The loss function of the coordinated system is the same as the loss function of the original system. The objective of the coordinator is to choose a coordination strategy (i.e., a sequence of coordination laws) to minimize the expected total loss.
2. *Formulate the coordinated system as a POMDP*: If the system model is such that the data available at the coordinator—the common information—is increasing with time, then the decision problem at the coordinator is centralized stochastic control problem. The second step of the approach is to formulate this centralized stochastic control problem as a partially observable Markov decision process (POMDP). To do so, we need to identify the (unobserved) state for input–output mapping for the coordinated system. In general, the vector consisting of the state of the original system and the local information of all controllers (or an appropriate subset of this vector) is a state for input–output mapping for the coordinated system.
3. *Solve the resultant POMDP*: The third step of the approach is to use Markov decision theory to identify the structure of optimal coordination strategies in the coordinated system and to identify a dynamic program to obtain an optimal coordination strategy with such structure.
4. *Show equivalence between the original system and the coordinated system*: The fourth step of the approach is to show that the two models are equivalent. In particular, for any coordination strategy in the coordinated system, there exists a control strategy in the original system that yields the same expected loss, and vice-versa.
5. *Translate the solution of the coordinated system to the original system*: The fifth step of the approach is to use the equivalence of the fourth step to translate the structural results and the dynamic program obtained in the third step for the coordinated system to structural results and dynamic program for the original system.

In Sects. 4.4 and 4.5, we illustrate how the above methodology applies to problems is communication, control, and queueing systems. Before we present these applications, we briefly review the POMDP model and results.

4.3 A Brief Recap of Partially Observable Markov Decision Processes (POMDPs)

A partially observable Markov decision process (POMDP) is a model of centralized (single decision-maker) stochastic control. It consists of a state process $\{S_t\}_{t=1}^{T}$,

an observation process $\{O_t\}_{t=1}^T$, and an action process $\{A_t\}_{t=1}^T$. For simplicity, assume that all system variables are finite valued and S_t, O_t, A_t takes value in time-homogeneous finite sets $\mathcal{S}$, $\mathcal{O}$, and $\mathcal{A}$. A POMDP has the following features:

1. The decision maker perfectly recalls its past observations and actions and chooses the action as a function of its observation and action history, that is,

$$A_t = d_t(O_{1:t}, A_{1:t-1}),$$

   where d_t is the decision rule at time t.
2. The state, observation, and action processes satisfy the following controlled Markov property

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, O_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t).$$

3. At each time, the system incurs an instantaneous cost $\ell(S_t, A_t)$.
4. The objective of the decision-maker is to choose a decision strategy $\mathbf{d} := (d_1, \dots, d_T)$ to minimize a total cost which is given by

$$J(\mathbf{d}) = \mathbb{E}\left[\sum_{t=1}^{T} \ell(S_t, A_t)\right].$$

The following standard result from Markov decision theory identifies the structure of globally optimal decision strategies and a dynamic program to find optimal strategies with that structure; see [38] for details.

Theorem 4.1 (POMDP Result) *Let Θ_t be the conditional probability distribution of the state S_t at time t given the observations $O_{1:t}$ and actions $A_{1:t-1}$,*

$$\Theta_t(s) := \mathbb{P}(S_t = s \mid O_{1:t}, A_{1:t-1}), \quad s \in \mathcal{S}.$$

Then,

(a) *$\Theta_{t+1} = \eta_t(\Theta_t, A_t, O_{t+1})$, where η_t is the standard nonlinear filter described as follows: If θ_t, a_t, o_{t+1} are the realizations of Θ_t, A_t and O_{t+1}, then the realization of s^{th} element of the vector Θ_{t+1} is*

$$\begin{aligned}\theta_{t+1}(s) &= \frac{\sum_{s'} \theta_t(s') \mathbb{P}(S_{t+1} = s, O_{t+1} = o_{t+1} \mid S_t = s', A_t = a_t)}{\sum_{s'', \tilde{s}} \theta_t(s'') \mathbb{P}(S_{t+1} = \tilde{s}, O_{t+1} = o_{t+1} | S_t = s'', A_t = a_t)} \\ &=: \eta_t^s(\theta_t, a_t, o_{t+1}).\end{aligned}$$

   The function $\eta_t(\theta_t, a_t, o_{t+1})$ is the vector of functions $(\eta_t^s(\theta_t, a_t, o_{t+1}))_{s \in \mathcal{S}}$.

(b) *There exists an optimal decision strategy of the form*

$$A_t = \hat{d}_t(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy: Define

$$V_T(\theta) := \min_a \mathbb{E}\big[\ell(S_T, a) \mid \Theta_T = \theta\big],$$

and for $t = T-1, T-2, \ldots, 1$, recursively define

$$V_t(\theta) := \min_a \mathbb{E}\big[\ell(S_t, a) + V_{t+1}\big(\eta_t(\theta, a, O_{t+1})\big) \mid \Theta_t = \theta, A_t = a\big].$$

Then, for each time t and each realization of θ of Θ_t, the optimal action $\hat{d}_t(\theta)$ is the minimizer in the definition of $V_t(\theta)$.

4.4 Applications of the Common Information Approach to Communication, Networked Control, and Queueing Systems

In this section, we illustrate how the common information approach provides a unified framework for solving problems that arise in various disciplines such as communication, networked control, and queueing systems. These problems have been previous investigated using problem specific solution techniques.

4.4.1 Point-to-Point Real-Time Communication with Feedback

Communication problems can be thought of as team problems with the encoders and the decoders as the decision-makers in the team. Point-to-point feedback communication, in particular, is a dynamic team problem because: (a) the encoder (and in some cases the decoder as well) has to make decisions over time based on information that is changing with time, and (b) the decoder's information is directly affected by the decisions (i.e., the transmitted symbols) selected at the encoder. We will consider the point-to-point feedback communication with the real-time constraint, that is, we will require the decoder to produce estimates of the current state of the source in real-time. We describe the model and the common information approach below.

4.4.1.1 Problem Description

Consider the model of real-time communication with noiseless feedback, shown in Fig. 4.1, that was investigated in [37]. The source $X_t \in \mathcal{X}$, $t = 1, 2, \ldots, T$ is a discrete-time, finite state Markov chain with a fixed transition probability matrix, $P^S(\cdot|\cdot)$, and a fixed distribution on the initial state. At each time instant, the encoder can send a symbol $Z_t \in \mathcal{Z}$ to the decoder over a memoryless noisy channel that is characterized by the transition probability matrix $P^C(\cdot|\cdot)$. The received symbol

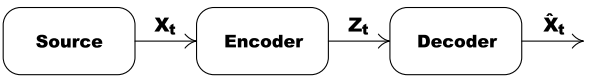


Fig. 4.1 A real-time communication system with noiseless feedback

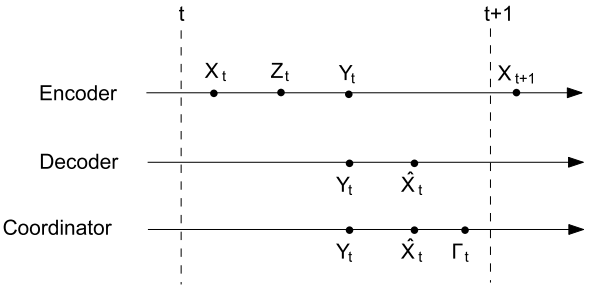


Fig. 4.2 Timing diagram for the real-time communication system

$Y_t \in \mathcal{Y}$ at the decoder is fed back noiselessly to the encoder. At the end of each time instant t, the decoder produces an estimate $\hat{X}_t \in \mathcal{X}$ of the current state of the Markov source. A distortion metric $\rho(X_t, \hat{X}_t)$ measures the accuracy of the decoder's estimate. The order of events at time t is the following (see Fig. 4.2): (i) the state X_t of the Markov source is generated, (ii) the encoder transmits Z_t over the channel, (iii) the channel outputs Y_t to the receiver, (iv) Y_t is fed back to the encoder, and (v) the decoder produces the estimate $\hat{X}_t$.

The encoder and the decoder are the two decision makers in this system. The encoder selects the symbol Z_t to be transmitted according to

$$Z_t = f_t(X_{1:t}, Y_{1:t-1}, Z_{1:t-1}),$$

where f_t is the encoder's decision rule at time t and $\mathbf{f} := (f_1, f_2, \ldots, f_T)$ is the encoder's strategy. The decoder selects its estimate according to

$$\hat{X}_t = g_t(Y_{1:t}),$$

where g_t is the decoder's decision rule at time t and $\mathbf{g} := (g_1, g_2, \ldots, g_T)$ is the decoder's strategy. The objective is to select $\mathbf{f}, \mathbf{g}$ so as to minimize

$$J(\mathbf{f}, \mathbf{g}) := \mathbb{E}\left[\sum_{t=1}^{T} \rho(X_t, \hat{X}_t)\right]. \tag{4.4}$$

4.4.1.2 Preliminary Result: Ignoring Irrelevant Information

As explained in Sect. 4.2.2, one way to extend the scope of the common information approach is to combine it with the person-by-person approach so as to identify and ignore irrelevant information at the decision makers. For the above example, a person-by-person approach was used in [37] to show that irrespective of the decoder's strategy, there is no loss of performance in restricting attention to encoding

strategies of the form

$$Z_t = f_t(X_t, Y_{1:t-1}). \tag{4.5}$$

This result is a consequence of the Markovian nature of the source and the real-time nature of the distortion function. After restricting attention to encoding strategies of the form in (4.5), we proceed with the common information approach.

4.4.1.3 Applying the Common Information Approach

We follow the five-step outline in Sect. 4.2.3.

1. *Construct a coordinated system*: After the time of reception of Y_t at the decoder, the information at the decoder is $I_t^d := \{Y_{1:t}\}$. Just before the time of transmission of Z_{t+1} by the encoder, the information at the encoder is $I_t^e := \{X_{t+1}, Y_{1:t}\}$. Between the time of reception of Y_t and the time of transmission of Z_{t+1}, the common information is then defined as

$$C_t = I_t^d \cap I_t^e = \{Y_{1:t}\}.$$

   The local information at the encoder is $I_t^e \setminus C_t = X_{t+1}$ and the local information at the decoder is $I_t^d \setminus C_t = \emptyset$.

   The first step of the approach is to construct a coordinated system in which a coordinator observes the common information and selects the prescriptions for the encoder and decoder that map their respective local information to their decisions. Since the decoder has no local information, the coordinator's prescription is simply a prescribed decision $\hat{X}_t$ for the decoder. The prescription for encoder, Γ_t, is a mapping from $\mathcal{X}$ to $\mathcal{Z}$. For each possible value of encoder's local information x_{t+1}, the prescription Γ_t prescribes a decision $z_{t+1} = \Gamma_t(x_{t+1})$. The coordinator selects its prescriptions according to a coordination strategy $(\psi_1^e, \psi_1^d), \ldots, (\psi_T^e, \psi_T^d)$ so that

$$\Gamma_t = \psi_t^e(Y_{1:t}), \qquad \hat{X}_t = \psi_t^d(Y_{1:t}). \tag{4.6}$$

   For this coordinated system, the source dynamics, the distortion metric, and the problem objective are the same as in the original system.
2. *Formulate the coordinated system as a POMDP*: The second step of the approach is to formulate the decision problem for the coordinator as a POMDP. In order to do so, we need to identify a state for input–output mapping for the coordinated system. As suggested in Step 2 of the common information approach, the state for input–output mapping is a subset of the state of the original dynamic system (in this case, the source) and the local information at each decision maker. In this example, the state of the source X_t is sufficient for input–output mapping. In particular, define the state, action, and observation processes for the coordinator as:

$$S_t := X_t, \qquad A_t = (\Gamma_t, \hat{X}_t), \qquad O_t := Y_t.$$

It is easy to verify that

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, O_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t). \tag{4.7}$$

Furthermore, for specific realization of the random variables involved, the right hand side of (4.7) can be written as

$$\mathbb{P}(x_{t+1}, y_{t+1} \mid x_t, \gamma_t, \hat{x}_t) = P^C\big(y_{t+1} \mid \gamma_t(x_{t+1})\big) P^S(x_{t+1}|x_t)$$

and the distortion cost can be written as

$$\rho(X_t, \hat{X}_t) = \tilde{\rho}(S_t, A_t),$$

with a suitably defined $\tilde{\rho}$. Thus, the coordinator's decision problem can be viewed as an instance of the POMDP model of Sect. 4.3.

3. *Solve the resultant POMDP*: The third step of the common information approach is to solve the resultant POMDP at the coordinated system. Using Theorem 4.1 for the coordinated system, we get the following structural result and dynamic programming decomposition.

Theorem 4.2 *Let Θ_t be the conditional probability distribution of the state X_t at time t given the coordinator's observations $Y_{1:t}$ and actions $\Gamma_{1:t-1}, \hat{X}_{1:t-1}$, i.e.,*

$$\Theta_t(x) = \mathbb{P}(X_t = x | Y_{1:t}, \Gamma_{1:t-1}, \hat{X}_{1:t-1}), \quad x \in \mathcal{X}.$$

Then,

(a) *If $\theta_t, \gamma_t, \hat{x}_t, y_{t+1}$ are the realizations of $\Theta_t, \Gamma_t, \hat{X}_t$ and Y_{t+1}, the realization of x^{th} element of the vector Θ_{t+1} is*

$$\begin{aligned}
\theta_{t+1}(x) &= \frac{\sum_{x'} \theta_t(x') \mathbb{P}(X_{t+1} = x, Y_{t+1} = y_{t+1} | X_t = x', \Gamma_t = \gamma_t, \hat{X}_t = \hat{x}_t)}{\sum_{x'', \tilde{x}} \theta_t(x'') \mathbb{P}(X_{t+1} = \tilde{x}, Y_{t+1} = y_{t+1} | X_t = x'', \Gamma_t = \gamma_t, \hat{X}_t = \hat{x}_t)} \\
&= \frac{\sum_{x'} \theta_t(x') P^C(Y_{t+1} = y_{t+1} | Z_{t+1} = \gamma_t(x)) P^S(X_{t+1} = x | X_t = x')}{\sum_{x'', \tilde{x}} \theta_t(x'') P^C(Y_{t+1} = y_{t+1} | Z_{t+1} = \gamma_t(\tilde{x})) P^S(X_{t+1} = \tilde{x} | X_t = x'')} \\
&=: \eta_t^x(\theta_t, \gamma_t, y_{t+1}).
\end{aligned} \tag{4.8}$$

Therefore, we have that $\theta_{t+1} = \eta_t(\theta_t, \gamma_t, y_{t+1})$ where $\eta_t(\theta_t, \gamma_t, y_{t+1})$ is the vector of functions $(\eta_t^x(\theta_t, \gamma_t, y_{t+1}))_{x \in \mathcal{X}}$.

(b) *There exists an optimal coordinator strategy of the form*

$$\Gamma_t = \psi_t^e(\Theta_t), \qquad \hat{X}_t = \psi_t^d(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy. Define:

$$V_T(\theta) := \min_{\hat{x}} \mathbb{E}\big[\rho(X_T, \hat{x}) \,\big|\, \Theta_T = \theta\big],$$

and for $t = T-1, T-2, \ldots, 1$, *recursively define*

$$V_t(\theta) := \min_{\hat{x},\gamma} \mathbb{E}\big[\rho(X_t, \hat{x}) + V_{t+1}\big(\eta_t(\theta, \gamma, Y_{t+1})\big) \,\big|\, \Theta_t = \theta, \Gamma_t = \gamma\big].$$

Then, for each time t *and each realization of* θ *of* Θ_t, *the optimal prescriptions* $\psi_t^e(\theta)$, $\psi_t^d(\theta)$ *are the minimizers in the definition of* $V_t(\theta)$.

4. *Show equivalence between the original system and the coordinated system*: The fourth step of the common information approach is to show the equivalence between the original system and the coordinated system. To show this equivalence, we show that any strategy for the coordinator can be implemented in the original system and vice versa.

   Let ψ_t^e, ψ_t^d, $t = 1, 2, \ldots, T$, be the coordinator's strategy of the form (4.6). Define the strategies for the encoder and decoder in the original system as follows:

$$f_{t+1}(\cdot, Y_{1:t}) := \psi_t^e(Y_{1:t}), \qquad g_t(Y_{1:t}) := \psi_t^d(Y_{1:t}). \tag{4.9}$$

   For each realization of the common information $y_{1:t}$ and each realization of the source state x_{t+1}, the encoder and decoder strategies as defined by (4.9) result in the same symbol z_{t+1} being transmitted and same estimate $\hat{x}_t$ being produced as in the coordinated system. Thus, the strategies for the encoder and the decoder defined by (4.9) will achieve the same expected cost as the coordinator's strategies ψ_t^e, ψ_t^d, $t = 1, 2, \ldots, T$.

   Conversely, given any strategies $\mathbf{f} = (f_1, \ldots, f_T)$, $\mathbf{g} = (g_1, \ldots, g_T)$, for the encoder and the decoder in the original system, we can construct strategies for the coordinator that achieve the same expected cost. Simply reverse (4.9) and define the coordinator's strategy as:

$$\psi_t^e(Y_{1:t}) := f_{t+1}(\cdot, Y_{1:t}), \qquad \psi_t^d(Y_{1:t}) := g_t(Y_{1:t}). \tag{4.10}$$

   Then, for each realization of the common information $y_{1:t}$ and each realization of the source state x_{t+1}, the coordinator strategies as defined by (4.10) will result in the same symbol z_{t+1} being transmitted and same estimate $\hat{x}_t$ being produced as in the original system. Thus, the coordinator's strategies defined by (4.10) will achieve the same expected cost as the strategies $(\mathbf{f}, \mathbf{g})$.

   Consequently, the original system is equivalent to the coordinated system. The equivalence between the two systems implies that translating a globally optimal strategy for the coordinator to the original system (using (4.9)) will give globally optimal strategies for the original system.

5. *Translate the solution of the coordinated system to the original system*: The last step of the approach is to translate the result of Theorem 4.2 to the original system, which gives the following:

Theorem 4.3 *For the real-time communication problem formulated above, there exist globally optimal encoding and decoding strategies of the form*

$$Z_{t+1} = f_{t+1}^*(X_{t+1}, \Theta_t), \qquad \hat{X}_t = g_t^*(\Theta_t),$$

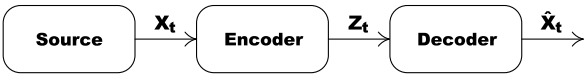


Fig. 4.3 A networked control system with communication over a rate-limited channel

where $\Theta_t = \mathbb{P}^{f_{1:t}}(X_t|Y_{1:t})$ *and* Θ_t *evolves according to* (4.8). *Furthermore, if* $(\psi_t^{e*}, \psi_t^{d*})$ *is an optimal coordination strategy* (*i.e., the solution of dynamic program of Theorem* 4.2), *then the optimal encoding and decoding strategies are given by*

$$f_{t+1}^*(\cdot, \Theta_t) = \psi_t^{e*}(\Theta_t), \qquad g_t^*(\Theta_t) = \psi_t^{d*}(\Theta_t).$$

The result of Theorem 4.3 is equivalent to the result of [37, Theorem 2 and (4.4)].

4.4.2 Networked Control Systems

In networked control systems, the controller relies on a communication network to gather information from the sensors at the plant and/or to send control actions to actuators at the plant. Communication related imperfections such as rate limited channels, delays and noise can affect the performance of the control system. A key question in such systems is whether the communication system and the control system can be jointly designed for improved performance. We consider a basic model of such a system where a sensor needs to communicate with the controller over a rate limited channel.

4.4.2.1 Problem Description

The structure of the problem above bears considerable similarity to the real-time communication problem formulated in Sect. 4.4.1.

Consider the model of networked control system with communication over a rate-limited channel shown in Fig. 4.3. A related problem was first considered in [36]. The state of the plant $X_t \in \mathcal{X}$, $t = 1, 2, \ldots, T$ is a discrete-time, finite state controlled Markov chain that evolves according to the equation

$$X_{t+1} = h_t(X_t, U_t, W_t),$$

where U_t is the control action applied by the controller and W_t is the random noise. The sensor observes the state of the plant and sends a symbol $Z_t \in \mathcal{Z}$ to the controller. We assume that $\mathcal{Z}$ is finite; thus, the communication link between the sensor and the controller is a rate limited communication link. At the end of each time instant t, the controller selects a control action U_t that is applied to the system. The order of events at time instant t is the following: (i) the state X_t is generated, (ii) the sensor transmits Z_t over the channel, (iii) the controller generates U_t.

The sensor selects the symbol to be transmitted Z_t according to

$$Z_t = f_t(X_{1:t}, Z_{1:t-1}),$$

the controller selects its action according to

$$U_t = g_t(Z_{1:t}).$$

At each time an instantaneous cost $\ell(X_t, U_t)$ is incurred. The objective is to select $\mathbf{f} = (f_1, \ldots, f_T), \mathbf{g} = (g_1, \ldots, g_T)$ so as to minimize

$$J(\mathbf{f}, \mathbf{g}) := \mathbb{E}\left[\sum_{t=1}^{T} \ell(X_t, U_t)\right]. \tag{4.11}$$

4.4.2.2 Preliminary Result: Ignoring Irrelevant Information

The structure of the problem above bears considerable similarity to the real-time communication problem formulated in Sect. 4.4.1. As in that example, using a person-by-person approach, we can show that irrespective of the controller's strategy, there is no loss of performance in restricting attention to sensor strategies of the form

$$Z_t = f_t(X_t, Z_{1:t-1}).$$

This result is analogous to structural result of encoder's strategies in (4.5) and is derived using similar arguments; see [36] for a proof of a similar result for a slightly different channel model.

4.4.2.3 Applying the Common Information Approach

We follow the five-step outline in Sect. 4.2.3.

1. *Construct a coordinated system*: Between the time of reception of Z_t at the controller and the time of transmission of Z_{t+1} by the sensor, the information at the controller is $I_t^c := \{Z_{1:t}\}$, and the information at the sensor is $I_t^s := \{X_{t+1}, Z_{1:t}\}$. The common information is then defined as

$$C_t = I_t^c \cap I_t^s = \{Z_{1:t}\}.$$

The local information at the sensor is $I_t^s \setminus C_t = X_{t+1}$ and the local information at the controller is $I_t^c \setminus C_t = \emptyset$.

The first step of the approach is to construct a coordinated system in which a coordinator observes the common information and selects the prescriptions for the sensor and the controller that map their respective local information to their

decisions. Since the controller has no local information, the coordinator's prescription is simply a prescribed action U_t for the controller. The prescription for the sensor, Γ_t, is a mapping from $\mathcal{X}$ to $\mathcal{Z}$. For each possible value of sensor's local information x_{t+1}, the prescription Γ_t prescribes a decision $z_{t+1} = \Gamma_t(x_{t+1})$. The coordinator selects its prescriptions according to a coordination strategy $(\psi_1^s, \psi_1^c), \ldots, (\psi_T^s, \psi_T^c)$ so that

$$\Gamma_t = \psi_t^s(Z_{1:t}), \qquad U_t = \psi_t^c(Z_{1:t}). \tag{4.12}$$

For this coordinated system, the plant dynamics, the loss function and the problem objective are the same as in the original system.

2. *Formulate the coordinated system as a POMDP*: The second step of the approach is to formulate the decision problem for the coordinator as a POMDP. As in Sect. 4.4.1, define the state, action, and observation processes for the POMDP as

$$S_t := X_t, \qquad A_t := (\Gamma_t, U_t), \qquad O_t := Z_t.$$

It is easy to verify that

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, O_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t)$$

and the instantaneous cost can be written as

$$\ell(X_t, U_t) = \tilde{\ell}(S_t, A_t),$$

with a suitably defined $\tilde{\ell}$. Thus, the coordinator's decision problem can be viewed as an instance of the POMDP model of Sect. 4.3.

3. *Solve the resultant POMDP*: The third step of the common information approach is to solve the resultant POMDP at the coordinated system. Using Theorem 4.1 for the coordinated system, we get the following structural result and dynamic programming decomposition.

Theorem 4.4 *Let Θ_t be the conditional probability distribution of the state X_t at time t given the coordinator's observations $Z_{1:t}$ and actions $\Gamma_{1:t-1}, \hat{X}_{1:t-1}$, i.e.,*

$$\Theta_t(x) = \mathbb{P}(X_t = x \mid Z_{1:t}, \Gamma_{1:t-1}, U_{1:t-1}), \quad x \in \mathcal{X}.$$

Then,

(a) *The realization θ_t of Θ_t updates according to a nonlinear filtering equation,*

$$\theta_{t+1} = \eta_t(\theta_t, \gamma_t, u_t, z_{t+1}).$$

(b) *There exists an optimal decision strategy of the form*

$$\Gamma_t = \psi_t^s(\Theta_t), \qquad \hat{X}_t = \psi_t^c(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy: *Define*

$$V_T(\theta) := \min_u \mathbb{E}\big[\ell(X_T, u) \mid \Theta_T = \theta\big],$$

and for $t = T-1, T-2, \ldots, 1$, *recursively define*

$$V_t(\theta) := \min_{u,\gamma} \mathbb{E}\big[\ell(X_t, u) + V_{t+1}\big(\eta_t(\theta, \gamma, Z_{t+1})\big) \mid \Theta_t = \theta, \Gamma_t = \gamma\big].$$

Then, for each time t *and each realization of* θ *of* Θ_t, *the optimal prescriptions* $(\psi_t^s(\theta), \psi_t^c(\theta))$ *are the minimizers in the definition of* $V_t(\theta)$.

4. *Show equivalence between the original system and the coordinated system*: The fourth step of the common information approach is to show the equivalence between the original system and the coordinated system. This equivalence follows from the same argument used in Sect. 4.4.1. In particular, the optimal strategy for the coordinator can be translated to optimal strategies for the sensor and the controller in the original system.
5. *Translate the solution of the coordinated system to the original system*: The last step of the approach is to translate the result of Step 3 to the original system, which gives the following result.

Theorem 4.5 *For the networked control problem formulated above, there exist globally optimal strategies for the sensor and the controller of the form*

$$Z_{t+1} = f_{t+1}^*(X_{t+1}, \Theta_t), \qquad U_t = g_t^*(\Theta_t)$$

where $\Theta_t = \mathbb{P}^{f_{1:t}}(X_t | Z_{1:t})$. *Furthermore, if* ψ_t^{s*}, ψ_t^{c*} *is an optimal coordination strategy* (*i.e., a solution of the dynamic program of Theorem* 4.4), *then the optimal sensor and controller strategies are given by*

$$f_{t+1}^*(\cdot, \Theta_t) = \psi_t^{s*}(\Theta_t), \qquad g_t^*(\Theta_t) = \psi_t^{c*}(\Theta_t).$$

The result of Theorem 4.5 is equivalent to the result of [36, Theorem 3.2] when specialized to the above model.

4.4.3 Paging and Registration in Cellular Networks

In cellular networks, the network needs to keep track of the location of a mobile station. This tracking may be done in two ways: the network may either page the mobile station, or the mobile station may register its location with the network. Both operations have an associated cost. The problem of finding optimal paging and registration strategies can be viewed as a team problem with the mobile station and the network operator as the decision-makers.

4.4.3.1 Problem Description

Consider a cellular network consisting of one mobile station (MS) and one network operator (N). The mobile station's motion is described by a discrete-time, finite state Markov chain $X_t \in \mathcal{X}$, $t = 1, 2, \ldots$, with known transition probability matrix. Each state represent a cell of the cellular network. At each time instant t, the MS may or may not register with the network. The cost of registration is r. If the MS registers with the network at time t, the network learns its location X_t. At each time t, the network may receive an exogenous paging request to seek MS's location. The exogenous paging request is an i.i.d. binary process which is independent of the motion of MS. The probability of a paging request at any time t is p. If a paging request arrives, the network operator must decide an order in which the cells are to be searched in order to locate the MS. We assume that if the MS is present in the cell being searched, the network successfully finds it. Further, we assume that the time it takes to search one cell is negligible compared to the time step of MS's motion, so that the paging request is completed within one time step. The cost of paging depends on the number of cells that are searched before MS is located. This model was investigated in [5].

The order of events at time instant t is the following: (i) The MS moves to location X_t according to a probability distribution that depends on its previous location; (ii) A paging request arrives with probability p; (iii) If a paging request arrives, the network operator must decide an order in which the cells are to be searched; (iv) If no paging request is made, the MS decides whether or not to register its location with the network.

Define a random variable Y_t as

$$Y_t = \begin{cases} X_{t-1} & \text{if the network learns MS location either by a paging} \\ & \text{request or by MS registration at time } t-1, \\ \varepsilon & \text{otherwise.} \end{cases}$$

Let $\sigma(\mathcal{X})$ denote the set of all permutations of the locations in $\mathcal{X}$. At the beginning of time t, if the network received a paging request, it selects $U_t^{\mathrm{N}} \in \sigma(\mathcal{X})$ according to

$$U_t^{\mathrm{N}} = g_t(Y_{1:t}).$$

If a paging request does not arrive, the MS makes a decision $U_t^{\mathrm{MS}} \in \{0, 1\}$ according to

$$U_t^{\mathrm{MS}} = f_t(X_{1:t}, Y_{1:t}),$$

where $U_t^{\mathrm{MS}} = 1$ represents a decision to register and $U_t^{\mathrm{MS}} = 0$ represents a decision to not register with the network. The collection of functions $\mathbf{f} := (f_1, f_2, \ldots, f_T)$ and $\mathbf{g} := (g_1, g_2, \ldots, g_T)$ are the strategies of the MS and the network, respectively.

The objective is to select $\mathbf{f}, \mathbf{g}$ so as to minimize

$$J(\mathbf{f}, \mathbf{g}) := \mathbb{E}\left[\sum_{t=1}^{T}(1-p)rU_t^{\mathrm{MS}} + pk\tau\left(X_t, U_t^{\mathrm{N}}\right)\right], \tag{4.13}$$

where p is the probability of paging request arrival, r is the cost of registration by MS, k is the cost of searching one cell, $\tau(x, u^{\mathrm{N}})$ is the position of x in the permutation specified by u^{N} and, therefore, $\tau(X_t, U_t^{\mathrm{N}})$ is the number of cells searched by the network before MS is located at time t.

4.4.3.2 Preliminary Result: Ignoring Irrelevant Information

For the above example, we may use an argument similar to the argument based on the person-by-person approach used in Sect. 4.4.1 to show that irrespective of the strategy of the network, there is no loss of performance in restricting attention to the strategies of the MS of the form

$$U_t^{\mathrm{MS}} = f_t(X_t, Y_{1:t}). \tag{4.14}$$

This result is a consequence of the Markovian nature of the MS motion and the fact that a paging request is completed within one time step. After restricting attention to MS strategies of the form in (4.14), we proceed with the common information approach.

4.4.3.3 Applying the Common Information Approach

We follow the five-step outline in Sect. 4.2.3.

1. *Construct a coordinated system*: At the beginning of time t, the information at the network is $I_t^{\mathrm{N}} := \{Y_{1:t}\}$, and the information at the MS is $I_t^{\mathrm{MS}} := \{X_t, Y_{1:t}\}$. The common information at time t is

$$C_t = I_t^{\mathrm{N}} \cap I_t^{\mathrm{MS}} = \{Y_{1:t}\}.$$

The local information at the network is $I_t^{\mathrm{N}} \setminus C_t = \emptyset$ and the local information at the MS is $I_t^{\mathrm{MS}} \setminus C_t = X_t$.

The first step of the approach is to construct a coordinated system in which a coordinator observes the common information and selects the prescriptions for the network and the MS. Since the network has no local information, the coordinator's prescription is simply a prescribed order U_t^{N} in which to search the cells if a paging request arrives. The prescription Γ_t for the MS is a mapping from $\mathcal{X}$ to $\{0, 1\}$. If a paging request does not arrive, the prescription Γ_t prescribes a registration decision $u_t^{\mathrm{MS}} = \Gamma_t(x_t)$ for each possible value of MS

location x_t. The coordinator selects its prescriptions according to a coordination strategy $(\psi_1^{\mathrm{MS}}, \psi_1^{\mathrm{N}}), \ldots, (\psi_T^{\mathrm{MS}}, \psi_T^{\mathrm{N}})$ so that

$$\Gamma_t = \psi_t^{\mathrm{MS}}(Y_{1:t}), \qquad U_t^{\mathrm{N}} = \psi_t^{\mathrm{N}}(Y_{1:t}).$$

For this coordinated system, the MS motion dynamics, the cost function and the problem objective are the same as in the original system.

2. *Formulate the coordinated system as a POMDP*: The second step of the approach is to formulate the decision problem for the coordinator as a POMDP. In order to do so, we define the state, action and observation processes of the POMDP as:

$$S_t := X_t, \qquad A_t = \big(\Gamma_t, U_t^{\mathrm{N}}\big), \qquad O_t := Y_t.$$

It is easy to verify that

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, O_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t)$$

and that the instantaneous cost $(1-p)rU_t^{\mathrm{MS}} + pk\tau(X_t, U_t^{\mathrm{N}})$ can be written as a function of S_t and A_t. Thus, the coordinator's decision problem can be viewed as an instance of the POMDP model of Sect. 4.3.

3. *Solve the resultant POMDP*: The third step of the common information approach is to solve the resultant POMDP at the coordinated system. Using Theorem 4.1 for the coordinated system, we get the following structural result and dynamic programming decomposition.

Theorem 4.6 *Let Θ_t be the conditional probability distribution of the state X_t at time t given the coordinator's observations $Y_{1:t}$ and actions $\Gamma_{1:t-1}$, i.e.,*

$$\Theta_t(x) := \mathbb{P}(X_t = x \mid Y_{1:t}, \Gamma_{1:t-1}), \quad x \in \mathcal{X}.$$

Then,

(a) *The realization θ_t of Θ_t updates according to a nonlinear filtering equation,*

$$\theta_{t+1} = \eta_t(\theta_t, \gamma_t, y_{t+1}).$$

(b) *There exists an optimal decision strategy of the form*

$$\Gamma_t = \psi_t^{\mathrm{MS}}(\Theta_t), \qquad U_t^{\mathrm{N}} = \psi_t^{\mathrm{N}}(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy: Define

$$V_T(\theta) := \min_{\gamma, u^{\mathrm{N}}} \mathbb{E}\big[(1-p)r\Gamma_T(X_T) + pk\tau\big(X_T, U_T^{\mathrm{N}}\big) \,\big|\, \Theta_T = \theta,$$

$$\Gamma_T = \gamma, U_T^n = u^{\mathrm{N}}\big],$$

and for $t = T-1, T-2, \ldots, 1$, *recursively define*

$$V_t(\theta) := \min_{\gamma, u^{\mathrm{N}}} \mathbb{E}\big[(1-p) r \Gamma_t(X_t) + pk\tau\big(X_t, U_t^{\mathrm{N}}\big) + V_{t+1}\big(\eta_t(\theta, \gamma, Y_{t+1})\big) \,\big|\, \Theta_t = \theta, \Gamma_t = \gamma, U_t^{\mathrm{N}} = u^{\mathrm{N}}\big].$$

Then, for each time t *and each realization of* θ *of* Θ_t, *the optimal prescriptions* $(\psi_t^{\mathrm{MS}}(\theta), \psi_t^{\mathrm{N}}(\theta))$ *are the minimizers in the definition of* $V_t(\theta)$.

4. *Show equivalence between the original system and the coordinated system*: The fourth step of the common information approach is to show the equivalence between the original system and the coordinated system. This equivalence follows from the same argument used in Sect. 4.4.1. In particular, the optimal strategy for the coordinator can be translated to optimal strategies for the MS and the network in the original system.
5. *Translate the solution of the coordinated system to the original system*: The last step of the approach is to translate the result of Step 3 to the original system, which gives the following result.

Theorem 4.7 *For the paging and registration problem formulated above, there exist globally optimal strategies of the form*

$$U_t^{\mathrm{N}} = g_t^*(\Theta_t), \qquad U_t^{\mathrm{MS}} = f_t^*(X_t, \Theta_t),$$

where $\Theta_t = \mathbb{P}^{f_{1:t}}(X_t | Y_{1:t})$. *Furthermore, if* $(\psi_t^{\mathrm{MS}*}, \psi_t^{\mathrm{N}*})$ *is an optimal coordination strategy* (*i.e., the solution of dynamic program of Theorem* 4.6), *then the optimal paging and registration strategies* $(\mathbf{f}^*, \mathbf{g}^*)$ *are given by*

$$f_t^*(\cdot, \Theta_t) = \psi_t^{\mathrm{MS}*}(\Theta_t), \qquad g_t^*(\Theta_t) = \psi_t^{\mathrm{N}*}(\Theta_t).$$

The result of Theorem 4.7 is equivalent to the result of [5, Sect. III-C]. The dynamic program was using in [5] to identify further structural properties of the optimal paging and registration strategies when the motion of the MS follows a symmetric random walk.

4.4.4 Multiaccess Broadcast Systems

In a multiaccess broadcast system, multiple users communicate to a common receiver over a broadcast medium. If more than one user transmits at a time, the transmissions "collide" and the receiver cannot decode the packets due to interference. Such systems can be viewed as team problems in which all users must cooperate to maximize system throughput. In this section, we consider a specific variation of a two-user multiaccess broadcast system.

4.4.4.1 Problem Description

Consider a two-user multiaccess broadcast system. At time t, $W_t^i \in \{0, 1\}$ packets arrive at each user according to independent Bernoulli processes with $\mathbb{P}(W_t^i = 1) = p^i$, $i = 1, 2$. Each user may store only $X_t^i \in \{0, 1\}$ packets in a buffer. If a packet arrives when the user-buffer is full, the packet is dropped.

Both users may transmit $U_t^i \in \{0, 1\}$ packets over a shared broadcast medium. If only one user transmits at a time, the transmission is successful and the transmitted packet is removed from the queue. If both users transmit simultaneously, packets "collide" and remain in the queue. Thus, the state update for user 1 is given by

$$X_{t+1}^1 = \max\big(X_t^1 - U_t^1 \cdot \big(1 - U_t^2\big) + W_t^1, 1\big).$$

The state update rule for user 2 is symmetric dual of the above.

Due to the broadcast nature of the communication medium, each user knows the control action of the other user after one-step delay. Thus, each user chooses a transmission decision as

$$U_t = g_t^i\big(X_{1:t}^i, \mathbf{U}_{1:t-1}\big)$$

where $\mathbf{U}_t = (U_t^1, U_t^2)$. A user can transmit only if it has a packet, thus only actions $U_t^i \le X_t^i$ are feasible.

Instead of costs, it is more natural to work with rewards in this example. The objective is to maximize throughput, or the number of successful packet transmissions. Thus, the per unit reward is $r(\mathbf{X}, \mathbf{U}) = U^1 \oplus U^2$, where $\oplus$ means binary XOR. The objective is to maximize

$$J(\mathbf{g}) = \mathbb{E}\left[\sum_{t=1}^{T} U_t^1 \oplus U_t^2\right]$$

which corresponds to the total throughput.

When the arrival rates at both users are the same ($p^1 = p^2$), the above model corresponds to the two-user multiaccess broadcast system considered in [6, 12, 14, 21]. Slight variations of the above model were considered in [25, 32].

4.4.4.2 Preliminary Result: Ignoring Irrelevant Information

As suggested in Sect. 4.2.2, we may use the person-by-person approach to identify and ignore irrelevant information at the decision makers before applying the common information approach. For the above example, a person-by-person approach was used in [12] to show that there is no loss of performance in restricting attention to the transmission strategies of the form

$$U_t^i = g_t^i\big(X_t^i, \mathbf{U}_{1:t-1}\big). \tag{4.15}$$

This result is a consequence of the fact that irrespective of the transmission strategies, the processes $\{X_t^1\}$ and $\{X_t^2\}$ are conditionally independent given $\mathbf{U}_{1:t-1}$. After restricting attention to transmission strategies of the form (4.15), we proceed with the common information approach.

4.4.4.3 Applying the Common Information Approach

We follow the five-step outline in Sect. 4.2.3.

1. *Construct a coordinated system*: At the beginning of time t, the information at user i is $I_t^i = \{X_t^i, \mathbf{U}_{1:t-1}\}$. Thus, the common information at time t is

$$C_t = I_t^1 \cap I_t^2 = \{\mathbf{U}_{1:t-1}\}.$$

   The local information at user i is $I_t^i \setminus C_t = X_t^i$.

   The first step of the common information approach is to construct a coordinated system in which a coordinator observes the common information and selects the prescriptions that map each user's local information to its actions. The prescription Γ_t^i for user i is a mapping from $\mathcal{X}^i$ to $\mathcal{U}^i$. For each realization x_t^i of the local information, the prescription Γ_t^i prescribes a decision $u_t^i = \Gamma_t^i(x_t^i)$. Since, $\Gamma_t^i(0) = 0$, the prescription Γ_t^i is completely specified by $\Gamma_t^i(1)$, which we denote by $Y_t \in \{0, 1\}$. Then, the control action is $U_t^i = X_t^i Y_t^i$. The coordinator selects its prescriptions according to a coordination strategy $(\psi_1, \dots \psi_T)$, so that

$$\left(Y_t^1, Y_t^2\right) = \psi_t(\mathbf{U}_{1:t-1}).$$

   For this coordinated system, the queue dynamics, the reward function, and the problem objective are the same as the original system.
2. *Formulate the coordinated system as a POMDP*: The second step of the approach is to formulate the decision problem for the coordinator as a POMDP. Define the state, action, and observation processes for the POMDP as

$$S_t := \left(X_t^1, X_t^2\right), \qquad A_t := \left(Y_t^1, Y_t^2\right), \qquad O_t := \left(U_{t-1}^1, U_{t-1}^2\right).$$

   It is easy to verify that

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, O_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t)$$

   and the instantaneous reward function can be written as

$$r(X_t, U_t) = U_t^1 \oplus U_t^2 = X^1 Y^1 \oplus X^2 Y^2 =: \tilde{r}(S_t, A_t).$$

   Thus, the coordinator's decision problem can be viewed as an instance of the POMDP model of Sect. 4.3.
3. *Solve the resultant POMDP*: The third step of the common information approach is to solve the resultant POMDP at the coordinated system. Using Theorem 4.1 for the coordinated system, we get the following structural result and dynamic programming decomposition.

Theorem 4.8 *Let Θ_t be the conditional probability distribution of the state $\mathbf{X}_t = (X_t^1, X_t^2)$ given the coordinator's observations $\mathbf{U}_{1:t-1}$ and actions $\mathbf{Y}_{1:t-1}$, i.e.,*

$$\Theta_t(\mathbf{x}) = \mathbb{P}(\mathbf{X}_t = \mathbf{x} \mid \mathbf{U}_{1:t-1}, \mathbf{Y}_{1:t-1}).$$

Then,

(a) *The realization θ_t of Θ_t updates according to a nonlinear filtering equation,*

$$\theta_{t+1} = \eta_t\big(\theta_t, y_t^1, y_t^2, u_t^1, u_t^2\big).$$

(b) *There exists an optimal decision strategy of the form*

$$\big(Y_t^1, Y_t^2\big) = \psi_t(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy: Define

$$V_T(\theta) := \min_{(y^1, y^2)} \mathbb{E}\big[X_T^1 y^1 \oplus X_T^2 y^2 \;\big|\; \Theta_T = \theta\big],$$

and for $t = T-1, T-2, \ldots, 1$, recursively define

$$V_t(\theta) := \min_{(y^1, y^2)} \mathbb{E}\big[X_t^1 y^1 \oplus X_t^2 y^2 + V_{t+1}(\psi_t\big(\theta, y^1, y^2, X_t^1 y^1, X_t^2 y^2\big) \;\big|\; \Theta_t = \theta\big]$$

Then, for each time t and each realization θ of Θ_t, the optimal prescription $\psi_t(\theta)$ is the minimizer in the definition of $V_t(\theta)$.

4. *Show equivalence between the original system and the coordinated system*: The fourth step of the common information approach is to show the equivalence between the original system and the coordinated system. This equivalence follows from the same argument used in Sect. 4.4.1. In particular, the optimal strategy for the coordinator can be translated to optimal transmission strategies in the original system.
5. *Translate the solution of the coordinated system to the original system*: The last step of the approach is to translate the result of Step 3 to the original system, which gives the following result.

Theorem 4.9 *For the two-user multiaccess broadcast system formulated above, there exists optimal transmission strategies of the form*

$$U_t^i = X_t^i \cdot \psi_t^i(\Theta_t)$$

where $\Theta_t = \mathbb{P}^{\psi_{1:t}}(\mathbf{X}_t | \mathbf{U}_{1:t-1})$. Furthermore, an optimal $\psi_t^* = (\psi^{*,1}, \psi^{*,2})$ is given by the solution of the dynamic program in Theorem 4.8.

The result of Theorem 4.9 is equivalent to the result of [12, Proposition 14]. The dynamic program (extended to infinite horizon average reward setup) was used in [12] to explicitly characterize the optimal transmission strategies when $p_1 = p_2$.

4.5 Application to Delayed Sharing Information Structures

In this section, we present the result of [19], in which we use the common information approach to solve a long standing open problem associated with delayed sharing information structures.

In a decentralized control system with delayed sharing information structure, the controllers sharing their observations and control actions with each other after a fixed delay. The delayed sharing information structure is a link between classical information structure, which may be viewed as a degenerate decentralized control system in which controllers instantaneous sharing their observations and control actions, and a completely decentralized information structure, where there is no "lateral" sharing of information.

This information structure was proposed by Witsenhausen in a seminal paper [39] where he conjectured the structure of the globally optimal control strategies. Later Varaiya and Walrand [32] showed that Witsenhausen's assertion is true when the delay in the sharing of information is one (called one-step delayed sharing), but false for larger sharing delay; see [19] for a more detailed history of the problem.

4.5.1 Problem Description

The delayed-sharing information structure consists of n controllers. Let X_t denote the state of the system, Y_t^i denote the observations of controller i, and U_t^i denote the control action of controller i. The system evolves according to

$$X_{t+1} = f_t^i\big(X_t, \mathbf{U}_t, W_t^0\big)$$

where $\mathbf{U}_t = (U_t^1, \ldots, U_t^n)$ and $\{W_t^0\}_{t=1}^T$ is an i.i.d. noise process that is independent of the initial state X_1. The observations of the controllers are given by

$$Y_t^i = h_t^i\big(X_t, W_t^i\big), \quad i = 1, \ldots, n$$

where $\{W_t^i\}_{t=1}^T$, $i = 1, \ldots, n$, are i.i.d. noise process that are independent of each other and also independent of $\{W_t^0\}_{t=1}^T$ and X_1.

The controllers share their observations and control actions with each other after a k-step delay. Thus, the control actions are selected as follows:

$$U_t^i = g_t^i\big(\mathbf{Y}_{1:t-k}, \mathbf{U}_{1:t-k}, Y_{t-k+1:t}^i, U_{t-k+1:t-1}^i\big)$$

where $\mathbf{Y}_t = (Y_t^1, \ldots, Y_t^n)$.

The instantaneous loss function is given by $\ell(X_t, \mathbf{U}_t)$.

For simplicity, assume that all system variables are finite valued and X_t, Y_t^i, U_t^i, W_t^i take values in time-homogeneous finite sets $\mathcal{X}$, $\mathcal{Y}^i$, $\mathcal{U}^i$, and $\mathcal{W}^i$, respectively.

The objective is to choose control strategies $\mathbf{g}^{1:n}$ where $\mathbf{g}^i = (g_1^i, \dots, g_T^i)$, to minimize the expected total loss

$$J\left(\mathbf{g}^{1:n}\right) = \mathbb{E}^{(\mathbf{g}^{1:n})}\left[\sum_{t=1}^{T} \ell(X_t, \mathbf{U}_t)\right].$$

4.5.2 Applying the Common Information Approach

1. *Construct a coordinated system*: The first step of the approach is to construct the coordinated system. At the beginning of time t, the information at controller i is

$$I_t^i = \left(\mathbf{Y}_{1:t-k}, \mathbf{U}_{1:t-k}, Y_{t-k+1:t}^i, U_{t-k+1:t-1}^i\right).$$

   Thus, the common information at all controllers is

$$C_t = \bigcap_{i=1}^{n} I_t^i = (\mathbf{Y}_{1:t-k}, \mathbf{U}_{1:t-k})$$

   and the local information at controller i is $L_t^i = (Y_{t-k+1:t}^i, U_{t-k+1:t-1}^i)$.

   Consider a coordinated system where the coordinator observes the common information and selects prescriptions $(\Gamma_t^1, \dots, \Gamma_t^n)$ for the controllers where Γ_t^i maps the local information L_t^i to control action U_t^i, i.e., for each possible value l_t^i of the local information L_t^i, the prescription Γ_t^i prescribes a control action $u_t^i = \Gamma_t^i(l_t^i)$. For convenience, define $Z_t = (\mathbf{Y}_{t-k}, \mathbf{U}_{t-k})$ so that $C_t = Z_{1:t}$. The coordinator selects its prescriptions according to a coordination law ψ_t so that

$$\left(\Gamma_t^1, \dots, \Gamma_t^n\right) = \psi_t(C_t) = \psi_t(Z_{1:t}).$$

   For this coordinated system, the source dynamics, the loss function, and the problem objective are the same as the original problem.
2. *Formulate the coordinated system as a POMDP*: The second step of the approach is to formulate the coordinated system as a POMDP. In order to do so, define the state, observation, and action processes of the POMDP as

$$S_t = (X_t, \mathbf{L}_t), \qquad O_t = Z_t, \qquad A_t = \left(\Gamma_t^1, \dots, \Gamma_t^n\right).$$

   It is easy to verify that

$$\mathbb{P}(S_{t+1}, O_{t+1} \mid S_{1:t}, A_{1:t}) = \mathbb{P}(S_{t+1}, O_{t+1} \mid S_t, A_t)$$

   and that the instantaneous loss $\ell(X_t, \mathbf{U}_t) = \tilde{\ell}(S_t, A_t)$ for an appropriately defined $\tilde{\ell}$. Hence, the decision problem at the coordinator is a POMDP.

3. *Solve the resultant POMDP*: The third step of the common information approach is to solve the resultant POMDP at the coordinated system. Using Theorem 4.1 for coordinated system defined above, we get the following structural result and dynamic programming decomposition.

Theorem 4.10 *Let Θ_t be the conditional probability distribution of the state S_t given the coordinator's history of observations C_t and actions $(\Gamma^1_{1:t-1}, \dots, \Gamma^n_{1:t-1})$, i.e., for any realization s of S_t*

$$\Theta_t(s) = \mathbb{P}\big(S_t = s \;\big|\; C_t, \Gamma^1_{1:t-1}, \dots, \Gamma^n_{1:t-1}\big).$$

Then,

(a) *The realization θ_t of Θ_t updates according to a nonlinear filtering equation,*

$$\theta_{t+1} = \eta_t\big(\theta_t, z_{t+1}, \gamma_t^1, \dots, \gamma_t^n\big)$$

where $z_{t+1} = (\mathbf{y}_{t-k+1}, \mathbf{u}_{t-k+1})$.

(b) *There exists an optimal coordination strategy of the form*

$$\big(\Gamma_t^1, \dots, \Gamma_t^n\big) = \psi_t(\Theta_t).$$

Furthermore, the following dynamic program determines such an optimal strategy (recall that $U_t^i = \Gamma_t^i(L_t^i)$): Define

$$V_T(\theta) = \min_{(\gamma_T^1, \dots, \gamma_T^n)} \mathbb{E}\big[\ell(X_T, \mathbf{U}_T) \;\big|\; \Theta_T = \theta, \Gamma_T^1 = \gamma_T^1, \dots, \Gamma_T^n = \gamma_T^n\big]$$

and for $t = T-1, T-2, \dots, 1$, recursively define

$$V_t(\theta) = \min_{(\gamma_t^1, \dots, \gamma_t^n)} \mathbb{E}\big[\ell(X_t, \mathbf{U}_t) + V_{t+1}\big(\eta_t\big(\theta, Z_{t+1}, \gamma_t^1, \dots, \gamma_t^n\big)\big) \;\big|\; \Theta_t = \theta,$$
$$\Gamma_t^1 = \gamma_t^1, \dots, \Gamma_t^n = \gamma_t^n\big]. \tag{4.16}$$

Then, for each time t and each realization θ to Θ_t, the optimal prescription $(\gamma_t^1, \dots, \gamma_t^n)$ is the minimizer in the definition of $V_t(\theta)$.

4. *Show equivalence between the original system and the coordinated system*: The fourth step of the common information approach is to show the equivalence between the original system and the coordinated system. This equivalence follows from the same argument used in Sect. 4.4. As a consequence, we can translate an optimal coordination strategy for the coordinated system to an optimal control strategy for the original system.
5. *Translate the solution of the coordinated system to the original system*: The last step of the approach is to translate the results of Step 3 to the original system, which gives the following result.

Theorem 4.11 *For the delayed sharing information structure, there exists optimal control strategies of the form*

$$U_t^i = g_t^i\big(L_t^i, \Theta_t\big)$$

where $\Theta_t = \mathbb{P}(X_t, \mathbf{L}_t \mid C_t)$. *Furthermore, if* ψ^* *is the optimal coordination strategy (i.e., the solution to the dynamic program of Theorem 4.10), and* $\psi^{*,i}$ *denote its ith component, then the optimal control strategy* $\mathbf{g}^*_{1:T}$ *is given by*

$$g_t^{*,i}(\cdot, \theta) = \psi_t^{*,i}(\theta).$$

4.6 Conclusion

In centralized stochastic control, the controller's belief on the current state of the system plays a fundamental role for predicting future costs. If the control strategy for the future is fixed as a function of future beliefs, then the current belief is a sufficient statistic for future costs under any choice of current action. Hence, the optimal action at any time t is only a function of the controller's belief on the system state at time t. In decentralized problems, where there are many controllers with different information interacting with each other, the controllers' belief on the system state and their predictions of future costs are not expected to be consistent. Furthermore, since the costs depend both on system state as well as other controllers' actions, any controller's prediction of future costs must involve a belief on system state along with a prediction of other controllers' actions. The above discussion describes the difficulties that arise if one attempts to use a controller's belief on the system state for decision-making in decentralized systems.

The common information approach attempts to address the above difficulties based on two key observations: (i) Beliefs based on common information are consistent among all controllers and can serve as a consistent sufficient statistic. (ii) Even though controllers cannot accurately predict each other's control actions, for any realization of common information they can know the exact mapping used by each controller to map its local information to its control actions. These observations motivate the creation of a coordinated system with a fictitious coordinator which observes only the common information, forms its beliefs based on the common information, selects prescriptions (described in Sects. 4.4 and 4.5) and has the same objective as the original decentralized stochastic control problem. If the system model is such that the data available at the coordinator—the common information—is increasing with time, then the decision problem at the coordinator is a centralized stochastic control problem. This centralized problem is equivalent to the original decentralized stochastic control problem. This equivalence allows the use of results obtained from centralized stochastic control theory to obtain: (i) qualitative properties of optimal strategies for the controllers in the original decentralized stochastic control problem, and (ii) a dynamic program for determining optimal strategies for all controllers. The fictitious coordinator is invented purely for conceptual clarity.

It is important to realize that the coordinator's problem can be solved by each controller in the original system. Thus, the presence of the coordinator is not necessary. Nevertheless, its presence allows one to look at the original optimization problem from the view point of a "higher level authority" and simultaneously determine how each controller maps its local information to its action for the given realization of common information.

A key assumption in the common information approach is that common information is increasing with time. This assumption ensures that the coordinator has perfect recall and connects the coordinator's problem with centralized stochastic control and POMDPs. The connection between the coordinator's problem and POMDPs can be used for computational purposes as well. The dynamic program obtained for the coordinator is essentially similar to that for POMDPs. In particular, just as in POMDPs, the value-functions can be shown to be piecewise linear and concave function of the coordinator's belief. This characterization of value functions is utilized to find computationally efficient algorithms for POMDPs. Such algorithmic solutions to general POMDPs are well-studied and can be employed here. We refer the reader to [45] and references therein for a review of algorithms to solve POMDPs.

This chapter illustrates how common information approach can be used to solve decentralized stochastic control/decision-making problems that arise in control, communication and queueing systems and to resolve a long-standing theoretical problem on the structure of optimal control strategies in delayed sharing information structures.

As is the case for the designer's approach discussed in Sect. 4.1, the common information approach may be combined with the person-by-person approach as follows. First, use the person-by-person approach to identify qualitative properties of globally optimal strategies (e.g., identifying irrelevant information at controllers). Then, use the common information approach to further refine the qualitative properties and determine globally optimal strategies with those properties. In fact, all the examples of Sect. 4.4 used such a combined approach.

In this chapter, and in [20], it is assumed that the system has a partial history sharing information structure in which: (i) part of the past data (observations and control actions) of each controller is commonly available to all controllers; and (ii) all controllers have perfect recall of this commonly available data. Although this particular information structure makes it easier to describe the common information approach, it is not necessary for the approach to work. In particular, the common information approach applies to all sequential decision making problems (see [15] for a complete exposition).

Acknowledgements A. Mahajan's work was partially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN 402753-11 and D. Teneketzis's work was partially supported by National Science Foundation (NSF) Grant CCF-1111061 and by NASA through the grant NNX12A054G.

This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Aicardi, M., Davoli, F., Minciardi, R.: Decentralized optimal control of Markov chains with a common past information set. IEEE Trans. Autom. Control **32**(11), 1028–1031 (1987)
2. Bamieh, B., Voulgaris, P.: A convex characterization of distributed control problems in spatially invariant systems with communication constraints. Syst. Control Lett. **54**(6), 575–583 (2005)
3. 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)
4. Gattami, A.: Control and estimation problems under partially nested information pattern. In: Proceedings of 48th IEEE Conference on Decision and Control, pp. 5415–5419 (2009)
5. Hajek, B., Mitzel, K., Yang, S.: Paging and registration in cellular networks: jointly optimal policies and an iterative algorithm. IEEE Trans. Inf. Theory **64**, 608–622 (2008)
6. Hluchyj, M.G., Gallager, R.G.: Multiaccess of a slotted channel by finitely many users. In: Proceedings of National Telecommunication Conference, pp. D.4.2.1–D.4.2.7 (1981)
7. Ho, Y.C.: Team decision theory and information structures. Proc. IEEE **68**(6), 644–654 (1980)
8. Ho, Y.C., Chu, K.C.: Team decision theory and information structures in optimal control problems—Part I. IEEE Trans. Autom. Control **17**(1), 15–22 (1972)
9. Kaspi, Y., Merhav, N.: Structure theorem for real-time variable-rate lossy source encoders and memory-limited decoders with side information. In: Proceedings of the IEEE Symposium on Information Theory, Austin, TX (2010)
10. Kim, J., Lall, S.: A unifying condition for separable two player optimal control problems. In: Proceedings of 50th IEEE Conference on Decision and Control (2011)
11. Lessard, L., Lall, S.: A state-space solution to the two-player decentralized optimal control problem. In: Proceedings of 49th Annual Allerton Conference on Communication, Control and Computing (2011)
12. Mahajan, A.: Optimal decentralized control of coupled subsystems with control sharing. In: Proc. 50th IEEE Conf. Decision and Control and European Control Conf. (CDC-ECC), Orlando, FL, pp. 5726–5731 (2011)
13. Mahajan, A., Teneketzis, D.: On the design of globally optimal communication strategies for real-time noisy communication systems with noisy feedback. IEEE J. Sel. Areas Commun. **26**(4), 580–595 (2008)
14. Mahajan, A., Nayyar, A., Teneketzis, D.: Identifying tractable decentralized control problems on the basis of information structure. In: Proc. 46th Annual Allerton Conf. Communication, Control, and Computing, Monticello, IL, pp. 1440–1449 (2008)
15. Nayyar, A.: Sequential decision making in decentralized systems. Ph.D. thesis, University of Michigan, Ann Arbor, MI (2011)
16. Nayyar, A., Teneketzis, D.: On jointly optimal real-time encoding and decoding strategies in multiterminal communication systems. In: Proceedings of 47th IEEE Conference of Decision and Control (2008)
17. Nayyar, A., Teneketzis, D.: Decentralized detection with signaling. In: Proceeding of the Workshop on the Mathematical Theory of Networks and Systems (MTNS) (2010)
18. Nayyar, A., Teneketzis, D.: Sequential problems in decentralized detection with communication. IEEE Trans. Inf. Theory **57**(8), 5410–5435 (2011)
19. Nayyar, A., Mahajan, A., Teneketzis, D.: Optimal control strategies in delayed sharing information structures. IEEE Trans. Autom. Control **56**(7), 1606–1620 (2011)
20. Nayyar, A., Mahajan, A., Teneketzis, D.: Decentralized stochastic control with partial history sharing information structures: a common information approach. IEEE Trans. Autom. Control **58**(7), 1644–1658 (2013)
21. Ooi, J.M., Wornell, G.W.: Decentralized control of a multiple access broadcast channel: performance bounds. In: Proceedings of the 35th IEEE Conference on Decision and Control, Kobe, Japan, pp. 293–298 (1996)

22. Ooi, J.M., Verbout, S.M., Ludwig, J.T., Wornell, G.W.: A separation theorem for periodic sharing information patterns in decentralized control. IEEE Trans. Autom. Control **42**(11), 1546–1550 (1997)
23. Rantzer, A.: Linear quadratic team theory revisited. In: Proceedings of the American Control Conference, pp. 1637–1641 (2006)
24. Rotkowitz, M., Lall, S.: A characterization of convex problems in decentralized control. IEEE Trans. Autom. Control **51**(2), 274–286 (2006)
25. Schoute, F.C.: Decentralized control in packet switched satellite communication. IEEE Trans. Autom. Control **AC-23**(2), 362–371 (1976)
26. Teneketzis, D.: On the structure of optimal real-time encoders and decoders in noisy communication. IEEE Trans. Inf. Theory 4017–4035 (2006)
27. Teneketzis, D., Ho, Y.C.: The Decentralized Wald problem. Inf. Comput. **73**, 23–44 (1987)
28. Teneketzis, D., Varaiya, P.: The decentralized quickest detection problem. IEEE Trans. Autom. Control **AC-29**(7), 641–644 (1984)
29. Tenney, R.R., Sandell, N.R., Jr.: Detection with distributed sensors. IEEE Trans. Aerosp. Electron. Syst. **AES-17**(4), 501–510 (1981)
30. Tsitsiklis, J.N.: Decentralized detection. In: Advances in Statistical Signal Processing, pp. 297–344. JAI Press, London (1993)
31. Varaiya, P., Walrand, J.: On delayed sharing patterns. IEEE Trans. Autom. Control **23**(3), 443–445 (1978)
32. Varaiya, P., Walrand, J.: Decentralized control in packet switched satellite communication. IEEE Trans. Autom. Control **AC-24**(5), 794–796 (1979)
33. Veeravalli, V.V.: Decentralized quickest change detection. IEEE Trans. Inf. Theory **47**(4), 1657–1665 (2001)
34. Veeravalli, V.V., Basar, T., Poor, H.: Decentralized sequential detection with a fusion center performing the sequential test. IEEE Trans. Inf. Theory **39**, 433–442 (1993)
35. Veeravalli, V.V., Basar, T., Poor, H.: Decentralized sequential detection with sensors performing sequential tests. Math. Control Signals Syst. **7**(4), 292–305 (1994)
36. Walrand, J.C., Varaiya, P.: Causal coding and control of Markov chains. Syst. Control Lett. **3**, 189–192 (1983)
37. Walrand, J.C., Varaiya, P.: Optimal causal coding-decoding problems. IEEE Trans. Inf. Theory **29**(6), 814–820 (1983)
38. Whittle, P.: Optimization over Time. Wiley Series in Probability and Mathematical Statistics, vol. 2. Wiley, New York (1983)
39. Witsenhausen, H.S.: Separation of estimation and control for discrete time systems. Proc. IEEE **59**(11), 1557–1566 (1971)
40. Witsenhausen, H.S.: A standard form for sequential stochastic control. Math. Syst. Theory **7**(1), 5–11 (1973)
41. Witsenhausen, H.S.: On the structure of real-time source coders. Bell Syst. Tech. J. **58**(6), 1437–1451 (1979)
42. Wu, J., Lall, S.: A dynamic programming algorithm for decentralized Markov decision processes with a broadcast structure. In: Proceedings of the 49th IEEE Conference on Decision and Control, pp. 6143–6148 (2010)
43. Yoshikawa, T.: Dynamic programming approach to decentralized stochastic control problems. IEEE Trans. Autom. Control **20**(6), 796–797 (1975)
44. Yüksel, S.: Stochastic nestedness and the belief sharing information pattern. IEEE Trans. Autom. Control 2773–2786 (2009)
45. Zhang, H.: Partially observable Markov decision processes: a geometric technique and analysis. Oper. Res. **58**(1), 214–228 (2010)

Chapter 5
Relations Between Information and Estimation in the Presence of Feedback

Himanshu Asnani, Kartik Venkat, and Tsachy Weissman

5.1 Introduction

In this chapter, we present and discuss relations between Information and Estimation for signals corrupted by Gaussian noise. While providing an exposition of previously established results in this context, we will present a new unified framework for understanding the existing relations and deriving new ones. As we illustrate in the exposition, this framework allows us to understand whether and how these results carry over to accommodate the presence of feedback, a natural element in control and communication. Interpretations of information-theoretic quantities in terms of minimum mean loss also hold for other channels, such as the Poisson channel [1] (under an appropriate loss function), and many of the results in this discussion have parallels therein. For concreteness and to keep the exposition to a reasonable scope, in this presentation, we focus exclusively on the Gaussian channel under mean squared loss.

We begin by introducing the scalar Gaussian channel. This problem is characterized by an underlying clean signal X (which follows a law P_X) and its AWGN corrupted version Y_γ measured at a given 'signal-to-noise ratio' γ, which is to say

$$Y_\gamma | X \sim \mathcal{N}\left(\sqrt{\gamma} X, 1\right), \tag{5.1}$$

where $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian distribution with mean μ and variance σ^2.

In a communication setting, we are interested in the mutual information between the input X and the output Y_γ, denoted by $I(X; Y_\gamma)$. It quantifies the exponential

H. Asnani (✉) · K. Venkat · T. Weissman
Electrical Engineering Department, Stanford University, Stanford, CA, USA
e-mail: asnani@stanford.edu

K. Venkat
e-mail: kvenkat@stanford.edu

T. Weissman
e-mail: tsachy@stanford.edu

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_5,

growth rate of the maximum number of distinguishable messages that can be transmitted across multiple uses of the channel. In an estimation setting, one would be interested in using the observed output to estimate the underlying input signal optimally with respect to a given loss function. Define $\text{mmse}(\gamma)$ to be the minimum mean square error at 'signal-to-noise ratio' γ

$$\text{mmse}(\gamma) \triangleq \mathbb{E}\big[\big(X - \mathbb{E}[X|Y_\gamma]\big)^2\big]. \tag{5.2}$$

Intriguing ties have been discovered between the input–output mutual information and the mean squared estimation loss for the Gaussian channel. Before delving into the main results, we begin with some classical definitions (cf., e.g., [4] for more on these and a broader context).

Definition 5.1 (Entropy) For a discrete random variable X with probability mass function p, we define the entropy as

$$H(X) = -\sum_x p(x) \log p(x). \tag{5.3}$$

When X is a continuous random variable with support S, and admits a density f—the corresponding quantity known as differential entropy is defined as

$$h(X) = -\int_S f(x) \log f(x)\, dx. \tag{5.4}$$

Definition 5.2 (Relative entropy) For a given measurable space $(\Omega, \mathcal{F})$ and on it defined two probability measures P and Q, where P is absolutely continuous with respect to Q. The relative entropy between P and Q is defined as

$$D(P \parallel Q) = \int \log \frac{dP}{dQ}\, dP. \tag{5.5}$$

Definition 5.3 (Mutual information) For a given probability space $(\Omega, \mathcal{F}, P)$ with jointly distributed random variables (X, Y), the mutual information between X and Y is defined as

$$I(X;Y) = D(P_{X,Y} \parallel P_X \otimes P_Y), \tag{5.6}$$

where $P_X \otimes P_Y$ denotes the product measure of the marginals.

In [7], Guo et al. discovered the I-MMSE relationship, namely, for the additive Gaussian channel, an elegant functional relationship holds between the minimum mean square error and the mutual information between input X and output Y_{snr} (the subscript making the 'signal-to-noise ratio' explicit):

Theorem 5.1 (I-MMSE) *Let X be a random variable with finite variance, and Y_{snr} related to X as in (5.1). Then,*

$$\frac{d}{d\text{snr}} I(X; Y_{\text{snr}}) = \frac{1}{2} \text{mmse}(\text{snr}), \tag{5.7}$$

or writing (5.7) in its integral form,

$$I(X; Y_{\text{snr}}) = \frac{1}{2} \int_0^{\text{snr}} \text{mmse}(\gamma)\, d\gamma. \tag{5.8}$$

The aforementioned relationship holds for all input distributions of the random variable X. It is quite intriguing that a simple interpretation of the mutual information, in terms of the mean squared error, holds universally. One of the proofs of (5.7) relies on establishing its equivalence with de Bruijn's identity which was presented in Stam's paper [16]. This identity presents the derivative of the differential entropy of a Gaussian noise corrupted observation as a Fisher information. An integral version of de Bruijn's identity, while imposing only a finite second moment constraint on X was presented by Barron in [2]. An interesting and alternative proof route for Theorem 5.1 presented in [7], which is referred to as the 'incremental channel technique', proves (5.7) by analyzing the decrease in mutual information in the presence of an infinitesimal amount of additional Gaussian noise. In this exposition, we will not focus on the intriguing regularity properties associated with the MMSE and mutual information functionals, a detailed treatment of which can be obtained in [22].

Beyond elegance, the statement in Theorem 5.1 has several implications and applications in problems of estimation, and information theory alike. For one example, note that from (5.8), one can immediately read off the following relationship in the infinite snr limit, when X is discrete:

$$H(X) = \frac{1}{2} \int_0^{\infty} \text{mmse}(\gamma)\, d\gamma, \tag{5.9}$$

which is quite striking since, while the left hand side is obviously invariant to one-to-one transformations of X, such an invariance would be hard to deduce from the expression in the right side. Another immediate implication of Theorem 5.1 is the concavity of the mutual information as a function of the signal to noise ratio. Noting that Gaussian signals are the hardest to estimate [8, Proposition 15], (5.8) directly shows that Gaussian inputs maximize the mutual information and are thus capacity achieving under AWGN. The I-MMSE relationship and its extensions have similar implications in multi-terminal problems in information theory, and applications in communication problems [8]. There have also been extensions for vector channels and general additive channels which are non-Gaussian, such as in [13].

An interesting generalization of the I-MMSE relationship (5.8) to the scenario of mismatched estimation was revealed by Verdú in [19]. For the same observation model, the underlying clean signal X is distributed according to P, while the decoder or estimator is optimized assuming the law is Q. Verdú in [19] presents the following relationship between the relative entropy of the true and mismatched

output laws, and the difference between the mismatched and matched estimation losses:

Theorem 5.2 (Mismatched estimation and relative entropy) *For a random variable X with finite variance under laws P and Q,*

$$D\big(P * \mathcal{N}(0, 1/\text{snr}) \,\big\|\, Q * \mathcal{N}(0, 1/\text{snr})\big) = \frac{1}{2}\int_0^{\text{snr}} \text{mse}_{P,Q}(\gamma) - \text{mse}_{P,P}(\gamma)\, d\gamma, \tag{5.10}$$

where $*$ denotes the convolution operation, and $\text{mse}_{P,Q}(\gamma)$ is defined as

$$\text{mse}_{P,Q}(\gamma) = \mathbb{E}_P\big[\big(X - \mathbb{E}_Q[X|Y_\gamma]\big)^2\big]. \tag{5.11}$$

The above relations between fundamental quantities in information and estimation, which hold regardless of the specific underlying distributions, give general insights into both information measures and estimation loss. Having introduced the scalar channel and the interesting links between relative entropy and the cost of mismatch, we now proceed to the continuous-time Gaussian channel.

Let $X_0^T = \{X_t, 0 \le t \le T\}$ be the underlying stochastic process to be estimated.

Remark 5.1 (Stipulations) Throughout our treatment of continuous-time signals in this work, we impose the condition of square integrability on $X_{(\cdot)}$, which assumes $\mathbb{E}[\int_0^T X_s^2\, ds] < \infty$. For a given probability space $(\Omega, \mathcal{F}, P)$, we also assume that the mapping $X(t, \omega)$ on the set $[0, T] \times \Omega$ is (jointly) measurable with respect to the product σ-algebra $\mathcal{B}[0, T] \times \mathcal{F}$.

The continuous-time channel is characterized by the following relationship between the input and output processes at signal-to-noise ratio snr,

$$dY_t = \sqrt{\text{snr}}\, X_t\, dt + dW_t, \tag{5.12}$$

where $\{W_t\}_{t\ge 0}$ is a standard Brownian motion, independent of X_0^T (for more discussion on the properties of the Brownian motion, the reader is referred to [10]). Let $I(X_0^T; Y_0^T)$ denote the mutual information between the input and the output process, observed for a time duration $[0, T]$. We are interested in studying the relationship between filtering and smoothing errors, and the mutual information. To do this, we first denote the instantaneous filtering and smoothing errors at time $t \in [0, T]$ as follows:

$$\text{cmmse}(t, \text{snr}) \triangleq \mathbb{E}\big[\big(X_t - \mathbb{E}\big[X_t \big| Y_0^t\big]\big)^2\big] \tag{5.13}$$

$$\text{mmse}(t, \text{snr}) \triangleq \mathbb{E}\big[\big(X_t - \mathbb{E}\big[X_t \big| Y_0^T\big]\big)^2\big]. \tag{5.14}$$

Denote the time averaged filtering and smoothing squared errors by

$$\text{cmmse}_{\text{snr}}(T) \triangleq \int_0^T \text{cmmse}(t, \text{snr})\, dt, \tag{5.15}$$

$$\mathrm{mmse}_{\mathrm{snr}}(T) \triangleq \int_0^T \mathrm{mmse}(t, \mathrm{snr})\, dt. \tag{5.16}$$

In [5], Duncan proved the equivalence of input–output mutual information to the filtering squared error, of a square integrable continuous-time signal X_t, corrupted according to (5.12) to yield the process Y_t.

Theorem 5.3 (Mutual information and causal squared error) *Let X_0^T be a continuous-time stochastic process observed through the Gaussian channel at signal to noise ratio* snr, *as in* (5.12). *Then, we have*

$$I\left(X_0^T; Y_0^T\right) = \frac{\mathrm{snr}}{2} \mathrm{cmmse}_{\mathrm{snr}}(T). \tag{5.17}$$

Thus, the causal filtering error bears a direct equivalence with the input–output mutual information regardless of the input distribution. Further, the invariance of the mutual information to time-reversal, indicates the remarkable equality of the time-averaged causal and the average anti-causal error, i.e., when we consider the causal estimation of the time reversed signal. Note that the optimal anti-causal filter will have a different structure than the optimal causal filter, in general. In the continuous-time setting, the authors of [7] also established the I-MMSE relationship for processes, where we have a representation of the mutual information in terms of the non-causal estimation error. In [23], Zakai derived this relationship among other relationships in continuous-time, using Malliavin calculus. This result, which can be viewed as an extension of Theorem 5.1 for continuous-time processes, is stated below.

Theorem 5.4 (I-MMSE for processes) *For a continuous-time process X_0^T, corrupted by the Gaussian channel in* (5.12), *the input–output mutual information satisfies the following relationship*:

$$I\left(X_0^T; Y_0^T\right) = \frac{1}{2} \int_0^{\mathrm{snr}} \mathrm{mmse}_{\gamma}(T)\, d\gamma. \tag{5.18}$$

From both the above results, one gets the remarkable relationship between the filtering and smoothing errors (cf. [7]) in a crisp distribution independent form, which can be stated as:

$$\mathrm{cmmse}_{\mathrm{snr}}(T) = \frac{1}{\mathrm{snr}} \int_0^{\mathrm{snr}} \mathrm{mmse}_{\gamma}(T)\, d\gamma. \tag{5.19}$$

That two such fundamental quantities are related in such a simple manner, and bridged further by the mutual information, is indeed striking, particularly since the optimal filtering and smoothing filters may in general be highly nonlinear. Thus, the mutual information emerges as a bridge to help understand two purely estimation theoretic quantities.

The pioneering result by Duncan in 1970 was followed shortly by its extension to incorporate the presence of feedback by Kadota et al. in [9]. The channel input ϕ_t is a function of the underlying process X_t as well as the past outputs of the channel Y_0^t, in an additive Gaussian noise setting. The observation window is $t \in [0, T]$. The channel can be represented as

$$Y_t = \int_0^t \phi_s\big(Y_0^s, X_s\big)\, ds + W_t, \tag{5.20}$$

where, as usual, the standard Brownian motion $W_{(\cdot)}$ is independent of the underlying process $X_{(\cdot)}$. In differential form, (and using shorthand to represent $\phi_t(Y_0^t, X_t)$) we can rewrite (5.20) as

$$dY_t = \phi_t\, dt + dW_t. \tag{5.21}$$

We denote the causal estimate of ϕ_t based on observations up until t by

$$\hat{\phi}_t = \mathbb{E}\big[\phi_t \big| Y_0^t\big]. \tag{5.22}$$

Under mild regularity conditions on ϕ_t (cf. [9]), the mutual information between the input and output is equal to half the causal mean squared error. With our notation, the main result of [9] is expressed as follows:

Theorem 5.5 (Channels with feedback) *For the Gaussian channel with feedback, as in* (5.20)*, the input–output mutual information satisfies*

$$I\big(X_0^T; Y_0^T\big) = \frac{1}{2}\int_0^T \mathbb{E}\big[(\phi_t - \hat{\phi}_t)^2\big]\, dt. \tag{5.23}$$

That Duncan's result (Theorem 5.3) extends so naturally to incorporate feedback in the continuous-time channel, is quite satisfying. Further, in the recent [21], Theorem 5.3 is extended to more general scenarios involving the presence of feedback, and it is shown that (5.17) remains true in these more general cases upon replacing the mutual information on the left-hand side with directed information. Note that one can simply ignore the feedback by setting $\phi_t = X_t$, in which case Theorem 5.5 recovers the result stated in Theorem 5.3, when $\mathrm{snr} = 1$.

One may be tempted to conjecture a similar extension to the presence of feedback for the relation between mutual information and the smoothing error in Theorem 5.4. However, such an extension does not hold. This is primarily because such an extension to feedback depends on the adaptedness of the decoder's estimate to the filtration induced by the observations.[1] This feature is maintained in the presence of feedback in causal but not in non-causal estimation.

[1] Define the filtration $\mathcal{F}_t^Y = \sigma\{Y(B) : B \subseteq \{s : s < t\}\}$. Note that in the setting of Theorem 5.5, the encoder ϕ_t is measurable w.r.t. the σ-algebra $\mathcal{F}_T^X \vee \mathcal{F}_t^Y$, and the estimate $\hat{\phi}_t$ is measurable w.r.t. (or *adapted to the filtration*) $\mathcal{F}_t^Y$.

Inspired by the relationship between mismatched estimation and relative entropy in the scalar Gaussian channel in Theorem 5.2, Weissman [20] proved a generalization of Duncan's relationship for the mismatched scenario in continuous-time. The result establishes that the relative entropy between the output laws can be expressed as the cost of mismatch in estimation, much like in the scalar setting. In particular, this extension, not only recovers the relationship of Duncan, but also holds in the presence of feedback, as in the general observation model in (5.20).

Let us now consider the setting in [20] in detail, where a continuous-time signal X_t, distributed according to a law P, is observed through additive Gaussian noise, and is estimated by an estimator that would have been optimal if the signal were governed by the law Q. In this general setting, the main result in [20] shows that the relative entropy between the laws of the output for the two different underlying distributions (P and Q) is exactly half the difference between the mismatched and matched filtering errors. Let Y_t be the continuous-time AWGN corrupted version of X_t as given by (5.12). Let $P_{Y_0^T}$ and $Q_{Y_0^T}$ be the output distributions when the underlying signal X_0^T has law P and Q, respectively. As before, T denotes the time duration for which the process is observed. We denote the mismatched causal mean squared error

$$\mathrm{cmse}_{P,Q}(T) = \int_0^T \mathbb{E}_P\big[\big(X_t - \mathbb{E}_Q\big[X_t \big| Y_0^T\big]\big)^2\big]\, dt. \tag{5.24}$$

In this setting, then next theorem [20] informs us that the relative entropy between the output distributions is half the difference between the mismatched and matched filtering errors.

Theorem 5.6 (Mismatched filtering and relative entropy) *Let X_0^T be a continuous-time process in $[0, T]$, under laws P and Q. Then, for the Gaussian channel in (5.12), we have*

$$D(P_{Y_0^T} \parallel Q_{Y_0^T}) = \frac{1}{2}\big[\mathrm{cmse}_{P,Q}(T) - \mathrm{cmse}_{P,P}(T)\big]. \tag{5.25}$$

The above results for mismatched estimation, also give insights into the decision-theoretic formulation of minimax statistical estimation. This is in contrast to the Bayesian framework where the underlying signal has a known distribution. In minimax filtering, the source is known to belong to a class of possible sources, and the goal is to find the best filter that would minimize the worst case difference between its MSE and the MMSE of the active source. For details on the application of Theorem 5.6 to this setting, the reader is referred to the recent [12], where the authors show that the optimal minimax filter is, in fact, a Bayesian filter, under a particular "least favorable" prior, which turns out to coincide with the capacity achieving distribution for a certain channel.

5.2 Pointwise Extensions and Identities

Thus far, we have explored important and fundamental links between measures of information and estimation loss, for the Gaussian channel in scalar and continuous-time. Recently, the authors in [18], presented a pointwise analysis of the information-estimation identities discussed in Sect. 5.1. By characterizing the aforementioned identities as identities of expectations over random quantities, such an analysis not only generalizes and gives previously established results as corollaries, but also presents new and intriguing relations between central quantities in information and estimation. The Girsanov theorem and Itô calculus emerge as tools to understand the pointwise behavior of these random quantities, and to explore their properties. Here, we provide a brief illustration of the pointwise approach, by applying the analysis to Duncan's theorem for the continuous-time Gaussian channel.

Recall the continuous-time Gaussian channel in (5.12), where for simplicity we set $\mathrm{snr} = 1$,

$$dY_t = X_t\,dt + dW_t. \tag{5.26}$$

Denoting the time averaged filtering squared error

$$\mathrm{cmmse}(T) = \int_0^T \mathbb{E}\big[\big(X_t - \mathbb{E}\big[X_t \big| Y_0^T\big]\big)^2\big]\,dt \tag{5.27}$$

and letting $I(X_0^T; Y_0^T)$ denote the input–output mutual information for the channel (5.26), Duncan's theorem (Theorem 5.3), as we have already seen in the previous section, tells us that

$$I\big(X_0^T; Y_0^T\big) = \frac{1}{2}\,\mathrm{cmmse}(T). \tag{5.28}$$

Consider the random variable $\mathcal{E}$ defined as

$$\mathcal{E} = \log \frac{dP_{Y_0^T|X_0^T}}{dP_{Y_0^T}} - \frac{1}{2}\int_0^T \big(X_t - \mathbb{E}\big[X_t \big| Y_0^T\big]\big)^2\,dt. \tag{5.29}$$

Thus $\mathcal{E}$ is the difference of two quantities. The information density, defined as the log Radon–Nikodym derivative of the conditional distribution of the output given the input, with respect to the law of the output, is a quantity of fundamental interest in itself, for its mean is the mutual information between the input and the output processes. Further, its variance under the capacity achieving input prior is referred to in the communication literature as the channel dispersion, and emerges in the characterization of fundamental limits in the finite block-length regime, cf. [15]. The second quantity is half the cumulative squared error in estimating the input process in a causal manner. Duncan's theorem tells us that these two quantities are equal in expectation, and is therefore equivalently expressed as

$$E[\mathcal{E}] = 0. \tag{5.30}$$

The quantity $\mathcal{E}$ referred to as the 'tracking error' is the difference between the underlying random quantities which appear in Duncan's theorem. In the following proposition from [18], this quantity is explicitly characterized:

Proposition 5.1 *Let $\mathcal{E}$ be as defined in* (5.29). *Then*,

$$\mathcal{E} = \int_0^T \left(X_t - \mathbb{E}\left[X_t \middle| Y_0^T\right]\right) \cdot dW_t \quad a.s. \tag{5.31}$$

Note that on the right side of (5.31) is a stochastic integral with respect to the Brownian motion $W.$, driving the noise in the channel. With this representation, Duncan's theorem follows from the mere fact that this stochastic integral is a martingale and, in particular, has zero expectation.

On applying another basic property of the stochastic integral one gets the following interesting result for the variance of $\mathcal{E}$.

Theorem 5.7 (Variance of tracking error) *For a continuous-time signal X_0^T, $\mathcal{E}$ as defined in* (5.29) *satisfies*

$$\mathrm{Var}(\mathcal{E}) = \mathrm{cmmse}(T). \tag{5.32}$$

In conjunction with Duncan's theorem (5.17), we get the following relationship:

$$\mathrm{Var}(\mathcal{E}) = \mathrm{cmmse}(T) = 2\, I\left(X_0^T; Y_0^T\right). \tag{5.33}$$

Thus, the difference of the information density and half the cumulative squared error, a random variable which we know to have zero mean via Duncan's pioneering result in 1970, turns out to be expressible as a martingale. Further, what is striking is that the variance of this random variable is equal to the filtering error, which in turn is given by twice the mutual information by invoking Duncan's theorem. That this statement is true regardless of the distribution of the input gives further insight into the general structure of the underlying links between information and estimation. In [18], the authors present such a pointwise extension of other fundamental identities in both scalar and continuous-time. Of particular interest in the context of feedback is the pointwise extension of the relationship given by Kadota et al. [9] for the white Gaussian noise channel with feedback. With our notation, the main result of [9] is expressed in Theorem 5.5 as

$$I\left(X_0^T; Y_0^T\right) = \frac{1}{2}\int_0^T \mathbb{E}\left[(\phi_t - \hat{\phi}_t)^2\right] dt. \tag{5.34}$$

Define

$$\mathcal{E}_\phi \triangleq \log \frac{dP_{Y_0^T|X_0^T}}{dP_{Y^T}} - \frac{1}{2}\int_0^T (\phi_t - \hat{\phi}_t)^2\, dt. \tag{5.35}$$

In [18], the authors present the following pointwise characterization of Theorem 5.5:

Theorem 5.8 (A pointwise information-estimation result)

$$\mathcal{E}_\phi = \int_0^T (\phi_t - \hat{\phi}_t) \cdot dW_t \quad a.s., \tag{5.36}$$

where $\mathcal{E}_\phi$ *is as defined in* (5.35).

Parallel to the discovery in the pointwise treatment of Duncan's theorem, one can use Theorem 5.8 to deduce various results. Note from (5.36) that $\mathcal{E}_\phi$ is a martingale. Therefore,

$$\mathbb{E}[\mathcal{E}_\phi] = 0, \tag{5.37}$$

recovering the main result of [9], namely Theorem 5.5. Using Itô's Isometry, one also immediately obtains

$$\mathrm{Var}(\mathcal{E}_\phi) = \int_0^T \mathbb{E}\big[(\phi_t - \hat{\phi}_t)^2\big]\, dt. \tag{5.38}$$

Thus, even for the generalized setting of communication over channels with feedback, one can characterize how closely the information density and squared filtering error track each other. In particular, these results may have applications in approximating the mutual information via estimation theoretic quantities, for channels with feedback. In the special case when $\phi_t = X_t$, we recover the results obtained in the pointwise treatment of Duncan's theorem (5.31). Venkat and Weissman [18] also extend this analysis to incorporate mismatch.

The key ingredients that have emerged in the pointwise analysis of the identities discussed in Sect. 5.1 include the use of Girsanov's theorem to characterize the information density via a change of measure argument, in conjunction with Itô's formula for the time evolution of a martingale. Further, the specific structure of the Gaussian noise allows one to express the log-Radon–Nikodym derivative as a time integral of the squared estimation loss, plus a zero mean martingale. Equipped with these tools, we now proceed to illustrate their utility in providing a unified understanding of the above results, as well as obtaining new ones.

5.3 Applications

In this section, we illustrate how Girsanov theory and Itô's rule are at the heart of the basic relations between information and estimation in Gaussian noise. In doing so, we employ these tools to recover the I-MMSE and the mismatched estimation relationship. Along with emphasizing the simplicity of this proof route, we show that these tools also yield new and hitherto unknown results in relations between information and estimation.

We organize this section into three parts. The first part introduces a generalized divergence between two probability measures, a special case of which is the relative entropy. We observe that for the full generality of the Gaussian channel with feedback, our tools reveal an elegant relationship between the divergence and an associated filtering error of the channel input. In the second part, we show how the same approach gives rise to simple proofs of previously known results. For brevity, we only present the results for the scalar Gaussian channel, though we stress that essentially the same technique carries through in continuous-time as well. Finally, we discuss and present an informational quantity referred to as Lautum Information, introduced by Palomar and Verdu in [14], and present a new conservation principle exhibited by this quantity and the mutual information.

5.3.1 *f*-Information via Itô

We begin by considering a generalized divergence between two probability measures defined on an appropriate measurable space. Given a convex function f satisfying $f(1) = 0$, the f-divergence between measures P and Q is defined as

$$D_f(P\|Q) = \int f\left(\frac{dQ}{dP}\right) dP. \tag{5.39}$$

Note this is a generalization of relative entropy, as the latter is recovered by taking $f = -\log$. This divergence retains key properties of relative entropy such as non-negativity, monotonicity, convexity. For more discussion on this topic, the reader is referred to [11].

There is also a corresponding generalized notion of mutual information,

$$I_f(X;Y) \triangleq D_f(P_{X,Y} \parallel P_X \times P_Y) = \int D_f(P_{Y|X=x} \parallel P_Y)\, dP_X(x). \tag{5.40}$$

In this section, we shall look at the generalized mutual information and how we can use Itô calculus to represent it in terms of estimation theoretic quantities of interest.

The channel input is a function of an underlying process X_t as well as the past outputs of the channel Y_t, in an additive white Gaussian noise setting:

$$Y_t = \int_0^t \phi_s\left(Y_0^s, X_s\right) ds + W_t. \tag{5.41}$$

Note that we can use the Girsanov Theorem [6] to write the Radon–Nikodym derivative of the conditional law $P_{Y_0^t|X_0^t}$ with respect to the marginal law $P_{Y_0^t}$ as follows;

$$\Lambda_t = \frac{dP_{Y_0^t|X_0^t}}{dP_{Y_0^t}} = \exp\left\{\frac{1}{2}\int_0^t (\phi_s - \hat{\phi}_s)^2\, ds + \int_0^t (\phi_s - \hat{\phi}_s)\cdot dW_s\right\}, \tag{5.42}$$

where $\hat{\phi}_t = E[\phi_t | Y_0^t]$.

Let

$$Z_t \triangleq \frac{1}{2}\int_0^t (\phi_s - \hat{\phi}_s)^2\, ds + \int_0^t (\phi_s - \hat{\phi}_s)\cdot dW_s. \tag{5.43}$$

Let g be a concave function corresponding to any general convex function f, related as $g(x) = f(\frac{1}{x})$. We can represent the generalized f-information

$$\begin{aligned} I_f\big(X_0^t, Y_0^t\big) &= \mathbb{E}_P\left[g\left(\exp\left\{\frac{1}{2}\int_0^t (\phi_s - \hat{\phi}_s)^2\, ds + \int_0^t (\phi_s - \hat{\phi}_s)\cdot dW_s\right\}\right)\right] \\ &= \mathbb{E}_P\big[g\big(e^{Z_t}\big)\big]. \end{aligned} \tag{5.44}$$

We now proceed to write the Itô evolution equation (cf. [17]) for $g(e^{Z_t})$

$$\begin{aligned} g\big(e^{Z_t}\big) = g(1) &+ \int_0^t g'\big(e^{Z_s}\big)e^{Z_s}\left[\frac{1}{2}(\phi_s - \hat{\phi}_s)^2\, ds + (\phi_s - \hat{\phi}_s)\cdot dW_s\right] \\ &+ \frac{1}{2}\int_0^t \big[g''\big(e^{Z_s}\big)e^{2Z_s} + g'\big(e^{Z_s}\big)e^{Z_s}\big](\phi_s - \hat{\phi}_s)^2\, ds. \end{aligned} \tag{5.45}$$

On taking expectation with respect to P, and simplifying we get the following expression (for arbitrary f) for the generalized f-information

$$I_f = \mathbb{E}\big[g\big(e^{Z_t}\big)\big] = g(1) + \int_0^t \mathbb{E}\left[\left(g'\big(e^{Z_s}\big)e^{Z_s} + \frac{1}{2}g''\big(e^{Z_s}\big)e^{2Z_s}\right)(\phi_s - \hat{\phi}_s)^2\right] ds. \tag{5.46}$$

Note that for the special choice of $g(x) = \log(x)$ the above expression simplifies to give the known Kadota–Ziv–Zakai relationship (Theorem 5). Note in this case that the term in the large parentheses in the r.h.s. simplifies to $\frac{1}{2}$. Another sanity check is choosing $g(x) = 1/x$ in which case the integrand in the r.h.s. becomes 0. In this case, we recover the identity $\mathbb{E}[dQ/dP] = 1$.

5.3.2 Information and Estimation via Itô

In this subsection, we show how a simple application of Itô's rule yields a relation from which we derive, as corollaries, relations such as the heat equation [3], the de Bruijn identity [16], I-MMSE [7], and D-MMSE [20]. Consider a scalar Itô's drift–diffusion process, with zero drift and diffusion coefficient unity,

$$dY_t = dB_t, \qquad Y_0 = x, \tag{5.47}$$

where B_t is the standard Brownian motion. Let $f(t, y, x)$ be sufficiently smooth and denote $f_t = \frac{\partial}{\partial t} f$, $f_y = \frac{\partial}{\partial y} f$, $f_{yy} = \frac{\partial^2}{\partial y^2} f$, etc. Then Itô's rule reads

$$f(t, Y_t, x) = f(0, Y_0, x) + \int_0^t f_y(s, Y_s, x)\, dB_s + \int_0^t f_t(s, Y_s, x)\, ds + \frac{1}{2}\int_0^t f_{yy}(s, Y_s, x)\, ds \tag{5.48}$$

almost surely. Taking expectations (and noting that the stochastic integral above has zero mean), and then differentiating with respect to t gives

$$\frac{\partial}{\partial t} E\big[f(t, Y_t, x)\big] = E\big[f_t(t, Y_t, x)\big] + \frac{1}{2} E\big[f_{yy}(t, Y_t, x)\big]. \tag{5.49}$$

Now if x is a realization of the random variable X, then taking expectation with respect to X in Eq. (5.49) (assuming the expectations remain well-defined), we obtain

$$\frac{\partial}{\partial t} E\big[f(t, Y_t, X)\big] = E\big[f_t(t, Y_t, X)\big] + \frac{1}{2} E\big[f_{yy}(t, Y_t, X)\big], \tag{5.50}$$

which is the heat equation, as derived in Theorem 1 of [3]. Using the above Eq. (5.50), we have the following immediate results:

- *I-MMSE*, i.e., the I-MMSE relation in [7],

$$\frac{\partial}{\partial \gamma} I(X; \sqrt{\gamma} X + W) = \frac{1}{2} \operatorname{mmse}(X; \sqrt{\gamma} X + W), \tag{5.51}$$

  where $W \sim \mathcal{N}(0, 1)$ independent of X. To derive this, let f_Y denote the density of the output process Y. Take $f(t, y, x) = -\log f_Y(y)$. With this substitution, we obtain de Brujn's Identity [16],

$$\frac{\partial}{\partial t} h(Y_t) = \frac{1}{2} J(Y_t), \tag{5.52}$$

  where $h(\cdot)$ stands for the differential entropy and $J(\cdot)$ for Fisher information. With slight abuse of notation we denote $f_Y(Y_t)$ by f_{Y_t}. Let $Z_\gamma = \sqrt{\gamma} X + W$. First note that a simple manipulation yields

$$\frac{\frac{\partial}{\partial y} f_{Y_t}}{f_{Y_t}} = \frac{\partial}{\partial y} \log f_{Y_t} = \frac{1}{t}\big(\mathbb{E}[X|Y_t] - Y_t\big), \tag{5.53}$$

  which implies

$$\begin{aligned} J(Y_t) &= \mathbb{E}\left[\left(\frac{\partial}{\partial y} \log f_{Y_t}\right)^2\right] \\ &= \frac{1}{t^2}\mathbb{E}\big[\big(\mathbb{E}[X|Y_t] - Y_t\big)^2\big] \end{aligned}$$

$$= \frac{1}{t^2}\mathbb{E}\big[\big(Y_t - X + X - E[X|Y_t]\big)^2\big]$$

$$= \frac{1}{t^2}\big[t + 2\mathbb{E}\big[(Y_t - X)\big(X - \mathbb{E}[X|Y_t]\big)\big] + \mathbb{E}\big[\big(X - \mathbb{E}[X|Y_t]\big)^2\big]\big]. \quad (5.54)$$

As any function of Y_t is orthogonal to the estimation error $X - \mathbb{E}[X|Y_t]$, we have $\mathbb{E}[(Y_t - \mathbb{E}[X|Y_t])(X - \mathbb{E}[X|Y_t])] = 0$ and equation (5.54) becomes

$$J(Y_t) = \frac{1}{t^2}\big[t - \mathbb{E}\big[\big(X - \mathbb{E}[X|Y_t]\big)^2\big]\big]$$

$$= \frac{1}{t^2}[t - \text{MMSE}]. \quad (5.55)$$

Now, from the definition of mutual information, we can write

$$\frac{\partial}{\partial\gamma} I(X; Z_\gamma) = \frac{\partial}{\partial\gamma}\big[h(Z_\gamma) - h(Z_\gamma|X)\big]$$

$$= \frac{\partial}{\partial\gamma}\big[h(Z_\gamma) - h(Z_\gamma - \sqrt{\gamma}X|X)\big].$$

Substituting $Z_\gamma - \sqrt{\gamma}X$ by W, we obtain

$$\frac{\partial}{\partial\gamma} I(X; Z_\gamma) = \frac{\partial}{\partial\gamma}\big[h(Z_\gamma) - h(W)\big].$$

Since $h(W)$ does not depend on γ, the partial derivative of $h(W)$ with respect to γ is zero. Therefore, we have

$$\frac{\partial}{\partial\gamma} I(X; Z_\gamma) = \frac{\partial}{\partial\gamma} h(Z_\gamma) = \frac{\partial}{\partial\gamma} h(\sqrt{\gamma}Y_{\frac{1}{\gamma}}) \overset{(a)}{=} \frac{\partial}{\partial\gamma}\left[h(Y_{\frac{1}{\gamma}}) + \frac{1}{2}\log\gamma\right]. \quad (5.56)$$

The equality (a) follows from the scaling property of differential entropy. Let

$$p(t) = h(Y_t),$$

$$q(t) = J(Y_t).$$

The equation (5.56) can be rewritten as

$$\frac{\partial}{\partial\gamma} I(X; Z_\gamma) = -\frac{1}{\gamma^2} p'\left(\frac{1}{\gamma}\right) + \frac{1}{2\gamma}.$$

By applying de Bruijn's Identity, we obtain

$$\frac{\partial}{\partial\gamma} I(X; Z_\gamma) = -\frac{1}{2\gamma}\cdot\frac{1}{\gamma} q\left(\frac{1}{\gamma}\right) + \frac{1}{2\gamma} = \frac{1}{2}\left[\frac{1}{\gamma} - \frac{1}{\gamma^2} J(Y_{\frac{1}{\gamma}})\right]. \quad (5.57)$$

Comparing Eqs. (5.55) and (5.57), we obtain the I-MMSE relation.

- *D-MMSE*, i.e., the relation,

$$\frac{\partial}{\partial t} D(P_{Y_t} \parallel Q_{Y_t}) = \frac{1}{2t^2}\big(\mathrm{mse}_{P,P}(t) - \mathrm{mse}_{P,Q}(t)\big), \tag{5.58}$$

where $D(\cdot \parallel \cdot)$ is the Kullback–Leiber divergence and

$$\mathrm{mse}_{P,Q} = \mathbb{E}_P\big[\big(X - \mathbb{E}_Q[X|Y_t]\big)^2\big]. \tag{5.59}$$

To obtain this, let $f_{Y_t}^P$ and $f_{Y_t}^Q$ denote the density of Y_t when X is distributed as P and Q, respectively. Also, let $\hat{X}_P$ and $\hat{X}_Q$ stand for $\mathbb{E}_P[X|Y_t]$ and $\mathbb{E}_Q[X|Y_t]$, respectively. We substitute

$$\begin{aligned} f(t,y,x) = \log \frac{f_Y^P(y)}{f_Y^Q(y)} - \frac{2}{t^2}\int_0^y \int_0^v \big(\mathbb{E}_Q[X|Y=u]\big)^2 - \mathbb{E}_Q\big[X^2\big|Y=u\big]\,du\,dv \\ - \frac{2}{t^2}\int_0^y \int_0^v (u-x)\mathbb{E}_Q[X|Y=u]\,du\,dv, \end{aligned} \tag{5.60}$$

in Eq. (5.50) to obtain (note again with slight abuse of notation, $f_Y^Q(Y_t)$ and $f_Y^P(Y_t)$ are denoted by $f_{Y_t}^Q$ and $f_{Y_t}^P$, respectively)

$$\begin{aligned} \frac{\partial}{\partial t} D(P_{Y_t} \parallel Q_{Y_t}) &= \mathbb{E}_P\left[\frac{\frac{\partial}{\partial t} f_{Y_t}^P}{f_{Y_t}^P}\right] - \mathbb{E}_P\left[\frac{\frac{\partial}{\partial t} f_{Y_t}^Q}{f_{Y_t}^Q}\right] \\ &\quad - \frac{1}{2}\mathbb{E}_P\left[\left(\frac{\frac{\partial}{\partial y} f_{Y_t}^P}{f_{Y_t}^P}\right)^2\right] + \frac{1}{2}E_P\left[\left(\frac{\frac{\partial}{\partial y} f_{Y_t}^Q}{f_{Y_t}^Q}\right)^2\right] \\ &\quad + \frac{1}{2}\mathbb{E}_P\left[\frac{\frac{\partial^2}{\partial y^2} f_{Y_t}^P}{f_{Y_t}^P}\right] - \frac{1}{2}\mathbb{E}_P\left[\frac{\frac{\partial^2}{\partial y^2} f_{Y_t}^Q}{f_{Y_t}^Q}\right] \\ &\quad - \frac{1}{t^2}\mathbb{E}_P\big[\big(\mathbb{E}_Q[X|Y_t]\big)^2 - \mathbb{E}_Q\big[X^2\big|Y_t\big] + (Y_t - X)\mathbb{E}_Q[X|Y_t]\big]. \end{aligned} \tag{5.61}$$

A simple differentiation yields the heat equation for the density, $f_{Y_t}^Q$, i.e.,

$$\mathbb{E}_P\left[\frac{\frac{\partial}{\partial t} f_{Y_t}^Q}{f_{Y_t}^Q}\right] = \frac{1}{2}\mathbb{E}_P\left[\frac{\frac{\partial^2}{\partial y^2} f_{Y_t}^Q}{f_{Y_t}^Q}\right].$$

Using this, simplifies Eq. (5.61) to

$$\frac{\partial}{\partial t} D(P_{Y_t} \parallel Q_{Y_t}) = 2\mathbb{E}_P\left[\frac{\frac{\partial}{\partial t} f_{Y_t}^P}{f_{Y_t}^P}\right] - 2\mathbb{E}_P\left[\frac{\frac{\partial}{\partial t} f_{Y_t}^Q}{f_{Y_t}^Q}\right]$$

$$-\frac{1}{2}\mathbb{E}_P\left[\left(\frac{\frac{\partial}{\partial y}f^P_{Y_t}}{f^P_{Y_t}}\right)^2\right]+\frac{1}{2}E_P\left[\left(\frac{\frac{\partial}{\partial y}f^Q_{Y_t}}{f^Q_{Y_t}}\right)^2\right]$$
$$-\frac{1}{t^2}\mathbb{E}_P\big[\big(\mathbb{E}_Q[X|Y_t]\big)^2-\mathbb{E}_Q\big[X^2|Y_t\big]+(Y_t-X)\mathbb{E}_Q[X|Y_t]\big]. \tag{5.62}$$

What is left is to show that the expression on the right-hand side of the above equation equals $\frac{1}{2t^2}(\mathrm{mse}_{P,P}(t)-\mathrm{mse}_{P,Q}(t))$. Note from Eq. (5.53), we have

$$\frac{\frac{\partial}{\partial y}f^Q_{Y_t}}{f^Q_{Y_t}}=\frac{1}{t}\big(\mathbb{E}_Q[X|Y_t]-Y_t\big). \tag{5.63}$$

A similar algebra yields

$$\frac{\frac{\partial}{\partial t}f^Q_{Y_t}}{f^Q_{Y_t}}=\frac{1}{2t^2}\mathbb{E}_Q\big[(X-Y_t)^2\big|Y_t\big]. \tag{5.64}$$

Substituting the above partial derivatives in Eq. (5.62), we obtain

$$\begin{aligned}\frac{\partial}{\partial t}D(P_{Y_t}\parallel Q_{Y_t})&=-\frac{1}{2t^2}\big(\mathbb{E}_Q[X|Y_t]-\mathbb{E}_P[X|Y_t]\big)^2\\&=\frac{1}{2t^2}\big(\mathrm{mse}_{P,P}(t)-\mathrm{mse}_{P,Q}(t)\big).\end{aligned} \tag{5.65}$$

5.3.3 *Lautum Information and Estimation*

The Lautum information between X and Y is defined as

$$L(X;Y)=D(P_X\times P_Y\parallel P_{X,Y}), \tag{5.66}$$

i.e., we switch the order of the joint and the product distributions in the relative entropy from the definition of mutual information. We refer to recent work by Palomar and Verdu [14] for more about this measure of dependence and its significance.

Consider now our usual setting where $dY_t=X_t\,dt+dW_t$ (for simplicity, though what follows carries over to the more general setting of Sect. 5.3.1). Assuming X_0^T is a zero mean process,

$$\begin{aligned}L\big(X_0^t;Y_0^t\big)&\overset{(a)}{=}\mathbb{E}_{P_{X_0^t}\times P_{Y_0^t}}\left[-\frac{1}{2}\int_0^t\big(X_s-\hat{X}_s\big(Y^s\big)\big)^2\,ds-\int_0^t\big(X_s-\hat{X}_s\big(Y^s\big)\big)\,dW_s\right]\\&=\mathbb{E}_{P_{X_0^t}\times P_{Y_0^t}}\left[-\frac{1}{2}\int_0^t\big(X_s-\hat{X}_s\big(Y^s\big)\big)^2\,ds\right.\end{aligned}$$

$$
\begin{aligned}
&\quad - \int_0^t \big(X_s - \hat{X}_s\big(Y^s\big)\big)(dY_s - X_s\,ds)\Bigg] \\
&\overset{\text{(b)}}{=} \mathbb{E}_{P_{X_0^t}\times P_{Y_0^t}}\Bigg[-\frac{1}{2}\int_0^t X_s^2 + \hat{X}_s\big(Y^s\big)^2\,ds + \int_0^t X_s^2\,ds + \int_0^t \hat{X}_s\big(Y^s\big)\,dY_s\Bigg] \\
&\overset{\text{(c)}}{=} \mathbb{E}_{P_{X_0^t,Y_0^t}}\Bigg[-\frac{1}{2}\int_0^t X_s^2 + \hat{X}_s\big(Y^s\big)^2\,ds + \int_0^t X_s^2\,ds + \int_0^t \hat{X}_s\big(Y^s\big)\,dY_s\Bigg] \\
&\overset{\text{(d)}}{=} \mathbb{E}_{P_{X_0^t,Y_0^t}}\Bigg[-\frac{1}{2}\int_0^t X_s^2 + \hat{X}_s\big(Y^s\big)^2\,ds + \int_0^t X_s^2\,ds + \int_0^t \hat{X}_s\big(Y^s\big)X_s\,ds\Bigg] \\
&= \mathbb{E}_{P_{X_0^t,Y_0^t}}\Bigg[-\frac{1}{2}\int_0^t \big(X_s - \hat{X}_s\big(Y^s\big)\big)^2\,ds + \int_0^t X_s^2\,ds\Bigg], \qquad (5.67)
\end{aligned}
$$

where

- (a) follows using the following equation:

$$\Lambda_t = \frac{dP_{Y_0^t|X_0^t}}{dP_{Y_0^t}} = \exp\left\{\frac{1}{2}\int_0^t (\phi_s - \hat{\phi}_s)^2\,ds + \int_0^t (\phi_s - \hat{\phi}_s)\,dW_s\right\}, \qquad (5.68)$$

- (b) follows since under $P_{X_0^T}\times P_{Y_0^T}$ we have $\mathbb{E}[X_s\,\hat{X}_s] = \mathbb{E}[X_s]\,E[\hat{X}_s] = 0$, $\mathbb{E}[\hat{X}_s] = 0$ and similarly $\mathbb{E}[X_s\,dY_s] = 0$ and $\mathbb{E}[\hat{X}_s\,X_s] = 0$,
- (c) follows because the expectation of the expression in the square brackets depends on the joint distribution of (X_0^t, Y_0^t) only through the marginals of X_0^t and Y_0^t,
- (d) follows since under $P_{X_0^t,Y_0^t}$ we have

$$\mathbb{E}\Bigg[\int_0^t \hat{X}_s\big(Y^s\big)\,dY_s\Bigg] = \mathbb{E}\Bigg[\int_0^t \hat{X}_s\big(Y^s\big)X_s\,ds\Bigg].$$

Thus, for a general, not necessarily zero mean process we get

$$L\big(X_0^t; Y_0^t\big) = \mathbb{E}\Bigg[-\frac{1}{2}\int_0^t \big(X_s - \hat{X}_s\big(Y^s\big)\big)^2\,ds + \int_0^t \mathrm{Var}(X_s)\,ds\Bigg]. \qquad (5.69)$$

Or, more succinctly, and recalling also Duncan's relation, we get

$$L\big(X_0^t; Y_0^t\big) = \int_0^t \mathrm{Var}(X_s)\,ds - \frac{1}{2}\,\mathrm{cmmse}_t = \int_0^t \mathrm{Var}(X_s)\,ds - I\big(X_0^t; Y_0^t\big). \qquad (5.70)$$

In other words, the mutual and Lautum information satisfy the "conservation law"

$$I\big(X_0^t; Y_0^t\big) + L\big(X_0^t; Y_0^t\big) = \int_0^t \mathrm{Var}(X_s)\,ds. \qquad (5.71)$$

As usual, of course, all this specializes to the scalar setting of X observed through and additive Gaussian at SNR level snr, for which we get

$$I(\text{snr}) + L(\text{snr}) = \text{snr} \cdot \text{Var}(X), \tag{5.72}$$

and by differentiating,

$$L'(\text{snr}) = \text{Var}(X) - I'(\text{snr}) = \text{Var}(X) - \frac{1}{2}\,\text{mmse}(\text{snr}). \tag{5.73}$$

Similar to the Eq. (5.72), for the case of a vector source corrupted with Gaussian noise, a result appears in [13] (cf. Eq. (114)) relating Lautum and mutual information between input and output. In (5.71), we present a new analogous result in continuous-time.

5.4 Conclusion

We revisited the problem of mean-squared estimation of a source corrupted by white Gaussian noise, and explored relations between causal and non-causal estimation errors and the mutual information between input and output process. In both matched and mismatched settings, the relations involving causal estimation are observed to persist even in the presence of feedback. Girsanov Theory and Itô's calculus give the required tools to explore the pointwise nature of identities between information and estimation, and forms the basis of the approach presented in this chapter, that allows to derive the existing results as corollaries, as well as obtain new ones.

Acknowledgement This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Atar, R., Weissman, T.: Mutual information, relative entropy, and estimation in the Poisson channel. IEEE Trans. Inf. Theory **58**(3), 1302–1318 (2012)
2. Barron, A.R.: Entropy and the central limit theorem. Ann. Probab. **14**(1), 336–342 (1986)
3. Brown, L., Dasgupta, A., Haff, L.R., Strawderman, W.E.: The heat equation and Stein's identity: connections, applications. J. Stat. Plan. Inference **136**(7), 2254–2278 (2006)
4. Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd edn. Wiley, New York (2006)
5. Duncan, T.E.: On the calculation of mutual information. SIAM J. Appl. Math. **19**, 215–220 (1970)
6. Girsanov, I.V.: On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. Theory Probab. Appl. **5**, 285–301 (1960)
7. Guo, D., Shamai, S., Verdú, S.: Mutual information and minimum mean-square error in Gaussian channels. IEEE Trans. Inf. Theory **IT-51**(4), 1261–1283 (2005)
8. Guo, D., Wu, Y., Shamai (Shitz), S., Verdú, S.: Estimation in Gaussian noise: properties of the minimum mean-square error. IEEE Trans. Inf. Theory **57**(4), 2371–2385 (2011)

9. Kadota, T.T., Zakai, M., Ziv, J.: Mutual information of the white. Gaussian channel with and without feedback. IEEE Trans. Inf. Theory **IT-17**(4), 368–371 (1971)
10. Karatzas, I., Shreve, A.E.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York (1988)
11. Merhav, N.: Data processing theorems and the second law of thermodynamics. IEEE Trans. Inf. Theory **57**(8), 4926–4939 (2011)
12. No, A., Weissman, T.: Minimax filtering regret via relations between information and estimation. In: 2013 IEEE International Symposium on Information Theory Proceedings (ISIT), 7–12 July 2013, pp. 444–448 (2013)
13. Palomar, D., Verdú, S.: Representation of mutual information via input estimates. IEEE Trans. Inf. Theory **53**(2), 453–470 (2007)
14. Palomar, D.P., Verdu, S.: Lautum information. IEEE Trans. Inf. Theory **54**(3), 964–975 (2008)
15. Polyanskiy, Y., Poor, H.V., Verdú, S.: New channel coding achievability bounds. In: IEEE Int. Symposium on Information Theory 2008, Toronto, Ontario, Canada, 6–11 July 2008
16. Stam, A.J.: Some inequalities satisfied by the quantities of information of Fisher and Shannon. Inf. Control **2**(2), 101–112 (1959)
17. Steele, J.M.: Stochastic Calculus and Financial Applications. Springer, Berlin (2010)
18. Venkat, K., Weissman, T.: Pointwise relations between information and estimation in Gaussian noise. IEEE Trans. Inf. Theory **58**(10), 6264–6281 (2012)
19. Verdú, S.: Mismatched estimation and relative entropy. IEEE Trans. Inf. Theory **56**(8), 3712–3720 (2010)
20. Weissman, T.: The relationship between causal and noncausal mismatched estimation in continuous-time AWGN channels. IEEE Trans. Inf. Theory **56**(9), 4256–4273 (2010)
21. Weissman, T., Kim, Y.-H., Permuter, H.H.: Directed information, causal estimation, and communication in continuous time. IEEE Trans. Inf. Theory **59**(3), 1271–1287 (2012)
22. Wu, Y., Verdu, S.: Functional properties of MMSE and mutual information. IEEE Trans. Inf. Theory **58**(3), 1289–1301 (2012)
23. Zakai, M.: On mutual information, likelihood ratios, and estimation error for the additive Gaussian channel. IEEE Trans. Inf. Theory **51**(9), 3017–3024 (2005)

Chapter 6
Design of Information Channels for Optimization and Stabilization in Networked Control

Serdar Yüksel

6.1 Introduction and the Information Structure Design Problem

In stochastic control, typically a partial observation model/channel is given and one looks for a control policy for optimization or stabilization. Consider a single-agent dynamical system described by the discrete-time equations

$$x_{t+1} = f(x_t, u_t, w_t), \tag{6.1}$$

$$y_t = g(x_t, v_t), \quad t \geq 0, \tag{6.2}$$

for (Borel 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 (6.2), 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 $A \in \mathcal{B}(\mathbb{Y})$.

In networked control systems, the observation channel described above itself is also subject to design. In a more general setting, we can shape the channel input by coding and decoding. This chapter is concerned with design and optimization of such channels.

We will consider a controlled Markov model given by (6.1). The observation channel model is described as follows: This system is connected over a noisy channel with a finite capacity to a controller, as shown in Fig. 6.1. The controller has

S. Yüksel (✉)
Department of Mathematics and Statistics, Queen's University, Kingston, Ontario, Canada, K7L 3N6
e-mail: yuksel@mast.queensu.ca

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_6,

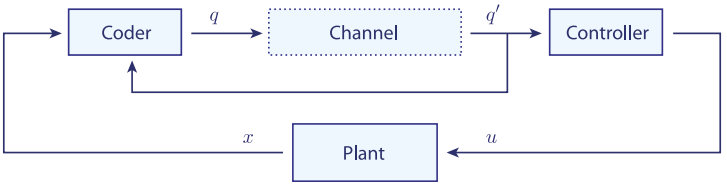


Fig. 6.1 Control over a noisy channel with feedback. The quantizer and the channel encoder form the coder in the figure

access to the information it has received through the channel. A quantizer 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 assume that the channel is a discrete channel with input alphabet $\mathcal{M}$ and output alphabet $\mathcal{M}'$. Hence, the channel maps $q \in \mathcal{M}$ to channel outputs $q' \in \mathcal{M}'$ probabilistically so that $P(q'|q)$ is a stochastic kernel. Further probabilistic properties can be imposed on the channels depending on the particular application.

We refer by a *Composite Coding Policy* Π^{comp}, a sequence of functions $\{Q_t^{\text{comp}}, t \geq 0\}$ which are causal such that the quantization output (channel input) at time t, $q_t \in \mathcal{M}$, under Π^{comp} is generated by a function of its local information, that is, a mapping measurable on the sigma-algebra generated by

$$\mathcal{I}_t^e = \left\{x_{[0,t]}, q'_{[0,t-1]}\right\}$$

to a finite set $\mathcal{M}$, the quantization output alphabet given by

$$\mathcal{M} := \{1, 2, \ldots, M\},$$

for $0 \leq t \leq T-1$ and $i = 1, 2$. Here, we have the notation for $t \geq 1$:

$$x_{[0,t-1]} = \{x_s, 0 \leq s \leq t-1\}.$$

The receiver/controller, upon receiving the information from the encoders, generates its decision at time t, also causally: An admissible causal controller policy is a sequence of functions $\gamma = \{\gamma_t\}$ such that

$$\gamma_t : \mathcal{M}'^{t+1} \to \mathbb{R}^m, \quad t \geq 0,$$

so that $u_t = \gamma_t(q'_{[0,t]})$.

We call such encoding and control policies, *causal* or *admissible*.

Two problems will be considered.

Problem P1: Value and Design of Information Channels for Optimization Given a controlled dynamical system (6.1), find solutions to minimization problem

$$\inf_{\Pi^{\text{comp}},\gamma} E_P^{\Pi^{\text{comp}},\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{6.3}$$

over the set of all admissible coding and control policies, given $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}_+$, a cost function.

Problem P2: Value of Information Channels for Stabilization The second problem concerns stabilization. In this setting, we replace (6.1) with an n-dimensional linear system of the form

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad t \geq 0 \tag{6.4}$$

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, at least one eigenvalue has magnitude greater than 1.

The stabilization problem is as follows: Given a system of the form (6.4) controlled over a channel, find the set of channels $\mathcal{Q}$ for which there exists a policy (both control and encoding) such that $\{x_t\}$ is stable. Stochastic stability notions will be ergodicity and existence of finite moments, to be specified later.

The literature on such problems is rather long and references will be cited as they are particularly relevant. We refer the reader to [56, 63], and [58] for a detailed literature review.

6.2 Problem P1: Channel Design for Optimization

In this section, we consider the optimization problem. We will first consider a single state problem and investigate topological properties of measurement channels.

6.2.1 Measurement Channels as Information Structures

6.2.1.1 Topological Characterization of Measurement Channels

Let, as in (6.2), g induce a stochastic kernel Q, P be the probability measure on the initial state, and 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 via

$$PQ(A) = \int_A Q(dy|x)P(dx), \quad A \in \mathcal{B}(\mathbb{X} \times \mathbb{Y}).$$

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.

Consider the following cost function:

$$J(P, Q, \gamma) = E_P^{Q,\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right], \tag{6.5}$$

over the set of all admissible policies γ, where $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is a Borel measurable stagewise cost (loss) function and $E_P^{Q,\gamma}$ denotes the expectation with initial state probability measure given by P, under policy γ and given channel Q.

Here, 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 $\mathbb{Y}$.

Let $\{\mu_n, n \in \mathbb{N}\}$ be a sequence in $\mathcal{P}(\mathbb{R}^n)$, where $\mathcal{P}(\mathbb{R}^n)$ is the set of probability measures on $\mathbb{R}^n$. Recall that $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^n)$ *weakly* [5] if

$$\int_{\mathbb{R}^n} c(x)\mu_n(dx) \to \int_{\mathbb{R}^n} c(x)\mu(dx)$$

for every continuous and bounded $c : \mathbb{R}^n \to \mathbb{R}$. The sequence $\{\mu_n\}$ is said to converge to $\mu \in \mathcal{P}(\mathbb{R}^n)$ *setwise* if

$$\mu_n(A) \to \mu(A), \quad \text{for all } A \in \mathcal{B}\big(\mathbb{R}^n\big)$$

For two probability measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$, the *total variation* metric is given by

$$\begin{aligned} \|\mu - \nu\|_{TV} &:= 2 \sup_{B \in \mathcal{B}(\mathbb{R}^n)} \big|\mu(B) - \nu(B)\big| \\ &= \sup_{f : \|f\|_\infty \le 1} \left| \int f(x)\mu(dx) - \int f(x)\nu(dx) \right|, \end{aligned}$$

where the infimum 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 $\|\mu_n - \mu\|_{TV} \to 0$.

These three convergence notions are in increasing order of strength: convergence in total variation implies setwise convergence, which in turn implies weak convergence.

Given these definitions, we have the following.

Definition 6.1 (Convergence of Channels [63])

(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 6.1 only induce the corresponding topologies on the resulting equivalence classes in $\mathcal{Q}$, instead of $\mathcal{Q}$. Let

$$J(P, Q) := \inf_\gamma E_P^{Q,\gamma}\left[\sum_{t=0}^{T-1} c\big(x_t, \gamma_t(y_{[0,t]})\big)\right].$$

In the following, we will discuss the following problems.

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

$$\inf_{\gamma \in \Gamma} J(P, Q_n, \gamma) \to \inf_{\gamma \in \Gamma} J(P, Q, \gamma)?$$

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

$$\inf_{Q \in \mathcal{Q}} \inf_{\gamma} E_P^{Q,\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right] \quad \text{and} \quad \sup_{Q \in \mathcal{Q}} \inf_{\gamma} E_P^{Q,\gamma}\left[\sum_{t=0}^{T-1} c(x_t, u_t)\right]. \tag{6.6}$$

If solutions to these problems exist, are they unique?

Before proceeding further, however, we will obtain in the next section a structural result on such optimization problems.

6.2.1.2 Concavity of the Measurement Channel Design Problem and Blackwell's Comparison of Information Structures

We first present the following concavity results.

Theorem 6.1 [61] *Let $T = 1$ and let the integral $\int c(x, \gamma(y)) P Q(dx, dy)$ exist for all $\gamma \in \boldsymbol{\Gamma}$ and $Q \in \mathcal{Q}$. Then, the function*

$$J(P, Q) = \inf_{\gamma \in \boldsymbol{\Gamma}} E_P^{Q,\gamma}\big[c(x, u)\big]$$

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

$$\begin{aligned} J(P, Q) &= J\big(P, \alpha Q' + (1 - \alpha) Q''\big) = \inf_{\gamma \in \boldsymbol{\Gamma}} E_P^{Q,\gamma}\big[c(x, u)\big] \\ &= \inf_{\gamma \in \boldsymbol{\Gamma}} \int c\big(x, \gamma(y)\big) P Q(dx, dy) \\ &= \inf_{\gamma \in \boldsymbol{\Gamma}} \Big(\alpha \int c\big(x, \gamma(y)\big) P Q'(dx, dy) \end{aligned}$$

$$
\begin{aligned}
&\quad + (1-\alpha)\int c\big(x,\gamma(y)\big)P\,Q''(dx,dy)\Big)\\
&\geq \inf_{\gamma\in\boldsymbol{\Gamma}}\Big(\alpha\int c\big(x,\gamma(y)\big)P\,Q'(dx,dy)\Big)\\
&\quad + \inf_{\gamma\in\boldsymbol{\Gamma}}\Big((1-\alpha)\int c\big(x,\gamma(y)\big)P\,Q''(dx,dy)\Big)\\
&= \alpha J\big(P,Q'\big)+(1-\alpha)J\big(P,Q''\big), \qquad (6.7)
\end{aligned}
$$

proving that $J(P,Q)$ is concave in Q. □

Proposition 6.1 [61] *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, due to David Blackwell [4]. In a single decision maker setup, we refer to the probability space induced on $\mathbb{X}\times\mathbb{Y}$ as an information structure.

Definition 6.2 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

$$Q_2(B|x)=\int_{\mathbb{Y}} Q'(B|y)Q_1(dy|x), \quad B\in\mathcal{B}(\mathbb{Y}),\ \ x\in\mathbb{X}.$$

We have the following.

Theorem 6.2 (Blackwell [4]) *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*

$$\inf_{\gamma} E_P^{Q_2,\gamma}\big[c(x,u)\big]\geq \inf_{\gamma} E_P^{Q_1,\gamma}\big[c(x,u)\big],$$

for all measurable and bounded cost functions c.

Proof The proof follows from [61]. 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 $P(x|dy^2,y^1)=P(x|y^1)$. Thus we have

$$J(P,Q_2)=\int V\big(P\big(\cdot|y^2\big)\big)P\big(dy^2\big)$$

$$
\begin{aligned}
&= \int V\left(\int P(\cdot|y^1)P(dy^1|y^2)\right)P(dy^2)\\
&\geq \int\left(\int P(dy^1|y^2)V\left(P(\cdot|y^1)\right)\right)P(dy^2)\\
&= \int V\left(P(\cdot|y^1)\right)\left(\int P(dy^1|y^2)P(dy^2)\right)\\
&= \int V\left(P(\cdot|y^1)\right)P(dy^1) = J(P, Q_1),
\end{aligned}
$$

where in arriving at the inequality, we used Proposition 6.1 and Jensen's inequality. □

Remark 6.1 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 [6] and [10].

The comparison argument applies also for the case $T > 1$.

Theorem 6.3 [60] *For the multi-stage problem* (6.5)*, 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* (6.5)

$$J(P, Q_1) \leq J(P, Q_2).$$

Remark 6.2 Blackwell's informativeness provides a partial order in the space of measurement channels; that is, not every pair of 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 in an appropriate sense, then one can obtain a total order on the space of channels.

6.2.1.3 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

$$
\begin{aligned}
J(P, Q) &= \inf_{\gamma} E_P^{Q,\gamma}\left[c(x_0, u_0)\right]\\
&= \inf_{\gamma\in\boldsymbol{\Gamma}} \int_{\mathbb{X}\times\mathbb{Y}} c\big(x, \gamma(y)\big)Q(dy|x)P(dx) \qquad (6.8)
\end{aligned}
$$

in the channel $Q \in \mathcal{Q}$, where $\boldsymbol{\Gamma}$ is the collection of all Borel measurable functions mapping $\mathbb{Y}$ into $\mathbb{U}$. Note that γ is an admissible one-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 6.1

A1. The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is non-negative, bounded, and continuous on $\mathbb{X} \times \mathbb{U}$.
A2. The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is non-negative, measurable, and bounded.
A3. The function $c : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ is non-negative, measurable, bounded, and continuous on $\mathbb{U}$ for every $x \in \mathbb{X}$.
A4. $\mathbb{U}$ is a compact 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.

Theorem 6.4 [63] *Suppose assumptions* A3 *and* A4 *hold. Then, there exists an optimal control policy for any channel* Q.

Theorem 6.5 [63]

(a) *J defined in (6.8) is not continuous under setwise or weak convergence even for continuous and bounded cost functions c.*
(b) *Suppose that c is continuous and bounded on $\mathbb{X} \times \mathbb{U}$, $\mathbb{U}$ is compact, and $\mathbb{U}$ is convex. If $\{Q_n\}$ is a sequence of channels converging weakly at input P to a channel Q, then J satisfies $\limsup_{n\to\infty} J(P, Q_n) \leq J(P, Q)$, that is, $J(P, Q)$ is upper semi-continuous under weak convergence.*
(c) *If c is bounded, measurable, then J is sequentially upper semi-continuous on $\mathcal{Q}$ under setwise convergence.*

We have continuity under the stronger notion of total variation.

Theorem 6.6 [63] *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.*

Thus, total variation, although a strong metric, is useful in establishing continuity. This will be useful in our analysis to follow for the existence of optimal quantization/coding policies.

In [63], (sequential) compactness conditions for a set of communication channels have been established. Given the continuity conditions, these may be used to identify conditions for the existence of best and worst channels for (6.6) when $T = 1$.

6.2.2 Quantizers as a Class of Channels

In this section, we consider the problem of convergence and optimization of quantizers.

We start with the definition of a quantizer.

Definition 6.3 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_i's are called the cells (or bins) of Q.

We allow for the possibility that some of the cells of the quantizer are empty. 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 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 $\mathcal{Q}_D(M) \subset \mathcal{Q}(M)$. We note also that elements of $\mathcal{Q}(M)$ are sometimes referred to as random quantizers.

Consider the set of probability measures

$$\Theta := \left\{\zeta \in P\left(\mathbb{R}^n \times \mathcal{M}\right) : \zeta = PQ, Q \in \mathcal{Q}\right\},$$

on $\mathbb{R}^n \times \mathcal{M}$ 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 $\mathbb{R}^n \times \mathcal{M}$ with a fixed input marginal P [9]. Borel measurability of Θ follows from [40] since set of probability measures on $\mathbb{R}^n \times \mathcal{M}$ 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.

Lemma 6.1 [63] *The set of quantizers $\mathcal{Q}_D(M)$ is setwise sequentially precompact at any input P.*

Proof The proof follows from 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 $\nu(\mathbb{R}^n \times$

$\{1,\dots,M\}) = M$). Then for any measurable $B \subset \mathbb{R}^n$ and $i = 1,\dots,M$, we have $\nu(B \times \{i\}) = P(B)\lambda(\{i\}) = P(B)$ and thus

$$PQ\big(B \times \{i\}\big) = P(B \cap B_i) \leq P(B) = \nu\big(B \times \{i\}\big).$$

Since any measurable $D \subset \mathbb{X} \times \{1,\dots,M\}$ can be written as the disjoint union of the sets $D_i \times \{i\}$, $i = 1,\dots,M$, with $D_i = \{x \in \mathcal{X} : (x,i) \in D\}$, the above implies $PQ(D) \leq \nu(D)$ and this domination leads to precompactness under setwise convergence (see [5, Theorem 4.7.25]). □

The following lemma provides a useful result.

Lemma 6.2 [63] *A sequence $\{Q_n\}$ in $\mathcal{Q}(M)$ converges to a Q in $\mathcal{Q}(M)$ setwise at input P if and only if*

$$\int_A Q_n(i|x)P(dx) \to \int_A Q(i|x)P(dx) \quad \textit{for all } A \in \mathcal{B}(\mathbb{X}) \textit{ and } i = 1,\dots,M.$$

Proof The lemma follows by noticing that for any $Q \in \mathcal{Q}(M)$ and measurable $D \subset \mathbb{X} \times \{1,\dots,M\}$,

$$PQ(D) = \int_D Q(dy|x)P(dx) = \sum_{i=1}^{M} \int_{D_i} Q(i|x)P(dx)$$

where $D_i = \{x \in \mathcal{X} : (x,i) \in D\}$. □

However, unfortunately, the space of quantizers $\mathcal{Q}_D(M)$ is not closed under setwise (and hence, weak) convergence, see [63] for an example. This will lead us to consider further restrictions in the class of quantizers considered below.

In the following, we show that an optimal channel can be replaced with an optimal quantizer without any loss in performance.

Proposition 6.2 [63] *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 For a policy $\gamma : \{1,\dots,M\} \to \mathbb{U} = \mathbb{X}$ (with finite cost) define for all i,

$$\bar{B}_i = \big\{x : c\big(x,\gamma(i)\big) \leq c\big(x,\gamma(j)\big),\ j = 1,\dots,M\big\}.$$

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 $Q' \in \mathcal{Q}_D(M)$. Then $E_P^{Q',\gamma}[c(x,u)] \leq E_P^{Q,\gamma}[c(x,u)]$ for any $Q \in \mathcal{Q}(M)$. □

The following shows that setwise convergence of quantizers implies convergence under total variation.

Theorem 6.7 [63] *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.*

Combined with Lemma 6.2, this theorem will be used to establish existence of optimal quantizers.

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 [24], there exist pairs of complementary closed half spaces $\{(H_{i,j}, H_{j,i}) : 1 \le i, j \le M, i \ne j\}$ 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 and consider a sequence $\{Q_n\}$ in $\mathcal{Q}_C(M)$. It can be shown (see the proof of Theorem 1 in [20]) that using an appropriate parametrization of the separating hyperplanes, a subsequence Q_{n_k} can be found 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 the following, we consider quantizers with convex codecells and an input distribution that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$ [20]. We note that such quantizers are commonly used in practice; 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; see [20] for further results on convexity of bins for entropy constrained quantization problems.

Theorem 6.8 [63] *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 6.9 [63] *Let P admit a density function 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.*

Remark 6.3 Regarding existence results, there have been few studies in the literature in addition to [63]. The authors of [1] and [41] have considered nearest neighbor encoding/decoding rules for norm based distortion measures. The L_2-norm leads to convex codecells for optimal design. We also note that the convexity assumption as well as the atomlessness property of the input measure can be relaxed in a class of settings, see [1] and Remark 4.9 in [60].

6.2.3 The Multi-stage Case

6.2.3.1 Static Channel/Coding

We now consider the general stochastic control problem in (6.5) with T stages. 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 the evolution of the controller's uncertainty on the state) at future time-stages. This is known as the dual effect of control [3]. The next theorem shows the continuity of the optimal cost in the measurement channel under some regularity conditions.

Definition 6.4 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., [15], then the resulting Q_n will converge (with probability one) uniformly in total variation [63].

Theorem 6.10 [63] *Consider the cost function* (6.5) *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*

$$\lim_{n\to\infty} J(P, Q_n) = J(P, Q).$$

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 multi-stage problems. Likewise, upper semi-continuity under weak convergence and setwise convergence require more stringent uniformity assumptions. On the other hand, the concavity property applies directly to the multi-stage case. That is, $J(P, Q)$ is concave in the space of channels; the proof of this result follows that of Theorem 6.1.

One further interesting problem regarding the multi-stage case is to consider adaptive observation channels. For example, one may aim to design optimal adaptive quantizers for a control problem. We consider this next.

6.2.3.2 Dynamic Channel and Optimal Vector Quantization

We consider a causal encoding problem where a sensor encodes an observed source to a receiver with zero-delay. Consider the source in (6.1). The source $\{x_t\}$ to be encoded is an $\mathbb{R}^n$-valued Markov process. The encoder encodes (quantizes) its information $\{x_t\}$ and transmits it to a receiver over a discrete noiseless channel with common input and output alphabet $\mathcal{M} := \{1, 2, \ldots, M\}$, where M is a positive integer, i.e., the encoder quantizes its information.

As in (6.5), for a finite horizon setting the goal is to minimize the cost

$$J_{\pi_0}\big(\Pi^{\text{comp}}, \gamma, T\big) := E_{\pi_0}^{\Pi^{\text{comp}},\gamma}\left[\frac{1}{T}\sum_{t=0}^{T-1} c_0(x_t, u_t)\right], \tag{6.9}$$

for some $T \geq 1$, where $c_0 : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_+$ is a (measurable) cost function and $E_{\pi_0}^{\Pi}[\cdot]$ denotes the expectation with initial state distribution π_0 and under the composite quantization policy Π^{comp} and receiver policy γ.

There are various structural results for such problems, primarily for control-free sources; see [25, 27, 49, 52, 53, 58] among others. In the following, we consider the case with control, which have been considered for finite-alphabet source and control action spaces in [51] and [27]. The result essentially follows from Witsenhausen [53].

Theorem 6.11 [57] *For the finite horizon problem, any* causal composite quantization policy *can be replaced without any loss in performance by one which, at time $t = 1, \ldots, T-1$, only uses, x_t and $q_{[0,t-1]}$, with the original control policy unaltered.*

Hereafter, let $\mathcal{P}(\mathbb{X})$ denote the space of probability measures on $\mathbb{X}$ endowed with weak convergence. Given a composite quantization policy Π^{comp}, let $\pi_t \in \mathcal{P}(\mathbb{R}^n)$ be the conditional probability measure defined by

$$\pi_t(A) := P(x_t \in A | q_{[0,t-1]})$$

for any Borel set A. Walrand and Varaiya [52] considered sources taking values in a finite set, and obtained the essence of the following result. For control-free sources, the result appears in [58] for $\mathbb{R}^n$-valued sources.

Theorem 6.12 [57] *For a finite horizon problem, any* causal composite quantization policy *can be replaced, without any loss in performance, by one which at any time $t = 1, \ldots, T-1$ only uses the conditional probability $P(dx_{t-1}|q_{[0,t-1]})$ and the state x_t. This can be expressed as a quantization policy which only uses (π_t, t) to generate a quantizer $Q_t : \mathbb{R}^n \to \mathcal{M}$, where the quantizer Q_t uses x_t to generate the quantization output as $q_t = Q_t(x_t)$ at time t.*

For any quantization policy in Π_W and any $T \geq 1$ we have

$$\inf_{\gamma} J_{\pi_0}\big(\Pi^{\text{comp}}, \gamma, T\big) = E^{\Pi^{\text{comp}}}_{\pi_0}\left[\frac{1}{T}\sum_{t=0}^{T-1} c(\pi_t, Q_t)\right],$$

where

$$c(\pi_t, Q_t) = \sum_{i=1}^{M} \inf_{u\in\mathbb{U}} \int_{Q_t^{-1}(i)} \pi_t(dx) c_0(x,u).$$

In the following, we consider the existence problem. However, to facilitate the analysis we will take the source to be control-free, and assume further structure on the source process. We have the following assumptions in the source $\{x_t\}$ and the cost function.

Assumption 6.2

(i) The evolution of the Markov source $\{x_t\}$ is given by

$$x_{t+1} = f(x_t) + w_t, \quad t \geq 0 \tag{6.10}$$

where $\{w_t\}$ is an independent and identically distributed zero-mean Gaussian vector noise sequence and $f : \mathbb{R}^n \to \mathbb{R}^n$ is measurable.

(ii) $\mathbb{U}$ is compact and $c_0 : \mathbb{R}^n \times \mathbb{U} \to \mathbb{R}_+$ is bounded and continuous.

(iii) The initial condition x_0 is zero-mean Gaussian.

We note that the class of quantization policies which admit the structure suggested in Theorem 6.12 is an important one. We henceforth define:

$$\begin{aligned}\Pi_W := \big\{\Pi^{\text{comp}} = \big\{Q_t^{\text{comp}}, t \geq 0\big\} : \exists \Upsilon_t : \mathcal{P}(\mathbb{X}) \to \mathcal{Q} \\ Q_t^{\text{comp}}\big(\mathcal{I}_t^e\big) = \big(\Upsilon_t(\pi_t)\big)(x_t), \forall \mathcal{I}_t^e\big\},\end{aligned} \tag{6.11}$$

to represent this class of policies. For a policy in this class, properties of conditional probability lead to the following expression for $\pi_t(dx)$:

$$\frac{\int \pi_{t-1}(dx_{t-1}) P(q_{t-1}|\pi_{t-1}, x_{t-1}) P(x_t \in dx | x_{t-1})}{\int\int \pi_{t-1}(dx_{t-1}) P(q_{t-1}|\pi_{t-1}, x_{t-1}) P(x_t \in dx | x_{t-1})}.$$

Here, $P(q_{t-1}|\pi_{t-1}, x_{t-1})$ is determined by the quantizer policy. The following follows from the proof of Theorem 2.5 of [58].

Theorem 6.13 *The sequence of conditional measures and quantizers* $\{(\pi_t, Q_t)\}$ *form a controlled Markov process in* $\mathcal{P}(\mathbb{R}^n) \times \mathcal{Q}$.

Theorem 6.14 *Under Assumption 6.2, an optimal receiver policy exists.*

Proof At any given time an optimal receiver will minimize $\int P(dx_t|q_{[0,t]})c(x_t, u_t)$. The existence of a minimizer then follows from Theorem 3.1 in [63]. □

Let Π_W^C be the set of coding policies in Π_W with quantizers having convex codecells (that is, $Q_t \in \mathcal{Q}_C(M)$). We have the following result on the existence of optimal quantization policies.

Theorem 6.15 [62] *For any $T \geq 1$, under Assumption 6.2, there exists a policy in Π_W^C such that*

$$\inf_{\Pi^{\text{comp}} \in \Pi_W^C} \inf_{\gamma} J_{\pi_0}\left(\Pi^{\text{comp}}, \gamma, T\right) \tag{6.12}$$

is achieved. Letting $J_T^T(\cdot) = 0$ and

$$J_0^T(\pi_0) := \min_{\Pi^{\text{comp}} \in \Pi_W^C, \gamma} J_{\pi_0}\left(\Pi^{\text{comp}}, \gamma, T\right),$$

the dynamic programming recursion

$$TJ_t^T(\pi_t) = \min_{Q \in \mathcal{Q}_C(M)} \left(c(\pi_t, Q_t) + TE\left[J_{t+1}^T(\pi_{t+1})|\pi_t, Q_t\right]\right)$$

holds for all $t = 0, 1, \ldots, T-1$.

We note that also for optimal multi-stage vector quantization, [8] has obtained existence results for an infinite horizon setup with discounted costs under a uniform boundedness assumption on the reconstruction levels.

6.2.3.3 The Linear Quadratic Gaussian (LQG) Case

There is a large literature on jointly optimal quantization for the LQG problem dating back to early 1960s (see, for example, [23] and [13]). References [2, 7, 17, 18, 31, 38, 48], and [57] considered the optimal LQG quantization and control, with various results on the optimality or the lack of optimality of the separation principle.

For controlled Markov sources, in the context of Linear Quadratic Gaussian (LQG) systems, existence of optimal policies has been established in [57] and [60], where it has been shown that without any loss, the control actions can be decoupled from the performance of the quantization policies, and a result similar to Theorem 6.15 for linear systems driven by Gaussian noise leads to the existence of an optimal quantization policy.

Structural results with control have also been studied by Walrand and Varaiya [51] in the context of finite control and action spaces and by Mahajan and Teneketzis [27] for control over noisy channels, also for finite state-action space settings.

6.2.3.4 Case with Noisy Channels with Noiseless Feedback

The results presented in this section apply also to coding over discrete memoryless (noisy) channels (DMCs) with feedback. The equivalent results of Theorems 6.11 and 6.12 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 $\pi_t \in \mathcal{P}(\mathbb{X})$ 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, $\pi_t(A) = P(x_t \in A | q'_{[0,t-1]})$, $A \in \mathcal{B}(\mathbb{X})$.

Theorem 6.16 [60] *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 6.17 [60] *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.*

Remark 6.4 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 under certain conditions on the memory update rules at the decoder. We refer the reader to [26, 49], and [25] for a class of such settings.

6.3 Problem P2: Characterization of Information Channels for Stabilization

In this section, we consider the stabilization problem over communication channels. The goal will be to identify conditions so that the controlled state is stochastically stable in the sense that

- $\{x_t\}$ is asymptotically mean stationary (AMS) and satisfies the requirements of Birkhoff's sample path ergodic theorem. This may also include the condition that the controlled (and possibly sampled) state and encoder parameters have a unique invariant probability measure.
- $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} |x_t|^2$ exists and is finite almost surely (this will be referred to as quadratic stability).

There is a very large literature on this problem. Particularly related references include [11, 28–30, 32, 37, 42, 43, 45, 46, 54, 55, 59]. In the context of discrete channels, many of these papers considered a bounded noise assumption, except notably [30, 36, 37, 55, 64], and [56]. We refer the reader to [32] and [60] for a detailed literature review.

In this section, we will present a proof program developed in [55] and [64] for stochastic stabilization of Markov chains with event-driven samplings applied to

networked control. Toward this end, we will first review few results from the theory of Markov chains. However, we first will establish fundamental bounds on information requirements for stabilization.

6.3.1 Fundamental Lower Bounds for Stabilization

We consider a scalar LTI discrete-time system and then later in Sect. 6.3.6 the multi-dimensional case. Here, the scalar system is described by

$$x_{t+1} = ax_t + bu_t + w_t, \quad t \geq 0 \tag{6.13}$$

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 with a finite capacity to a controller, as shown in Fig. 6.1, with the information structures described in Sect. 6.1.

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 6.5 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 $\mathbb{R}$ 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 $P^{n+1}_{\text{DMC}}(q'_{[0,n]}|q_{[0,n]}) = \prod_{k=0}^{n} P_{\text{DMC}}(q'_k|q_k)$, $\forall q_{[0,n]} \in \mathcal{M}^{n+1}$, $q'_{[0,n]} \in \mathcal{M}'^{n+1}$, where q_k, q'_k denote the kth component of the vectors $q_{[0,n]}, q'_{[0,n]}$, respectively.

Channels can also have memory. We state the following for both discrete and continuous-alphabet channels.

Definition 6.6 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 $\mathcal{M}'^{n+1}$.

In this chapter, while considering discrete channels, we will assume channels with finite alphabets.

Remark 6.5 Another setting involves continuous-alphabet channels. Such channels can be regarded as limits of discrete-channels: Note that the mutual information for

real valued random variables x, y is defined as

$$I(x;y) := \sup_{Q_1,Q_2} I\big(Q_1(x);Q_2(y)\big),$$

where Q_1 and Q_2 are quantizers with finitely many bins (see Chap. 5 in [19]). As a consequence, the discussion for discrete channels applies for continuous alphabet channels. On the other hand, the Gaussian channel is a very special channel which needs to be considered in its own right, especially in the context of linear quadratic Gaussian (LQG) systems and problems. A companion chapter deals with such channels, see [65], as well as [60].

Theorem 6.18 [56] *Suppose that a linear plant given as in* (6.13) *controlled over a DMC, under some admissible coding and controller policies, satisfies the condition*

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \leq 0, \tag{6.14}$$

where h denotes the entropy function. Then, the channel capacity C must satisfy

$$C \geq \log_2\big(|a|\big).$$

Remark 6.6 Condition (6.14) is a weak one. For example, a stochastic process whose second moment grows subexponentially in time, namely,

$$\liminf_{T\to\infty} \frac{\log(E[x_T^2])}{T} \leq 0,$$

satisfies this condition.

We now present a supporting result due to Matveev.

Proposition 6.3 [30] *Suppose that a linear plant given as in* (6.13) *is controlled over a DMC. If*

$$C < \log_2\big(|a|\big),$$

then

$$\limsup_{T\to\infty} P\big(|x_T| \leq b(T)\big) \leq \frac{C}{\log_2(|a|)},$$

for all $b(T) > 0$ such that $\lim_{T\to\infty} \frac{1}{T}\log_2(b(T)) = 0$

We note that similar characterizations have also been considered in [29, 43], and [32], for systems driven by bounded noise.

Theorem 6.19 [60] *Suppose that a linear plant given as in* (6.13) *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*

$$C \geq \log_2\bigl(|a|\bigr).$$

In the following, we will observe that the condition $C \geq \log_2(|a|)$ in Theorems 6.18 and 6.19 is almost sufficient as well for stability in the AMS sense. Furthermore, the result applies to multi-dimensional systems. Toward this goal, we first discuss the erasure channel with feedback (which includes the noiseless channel as a special case), and then consider more general DMCs, followed by a class of channels with memory. We will also investigate quadratic stability. We discuss an essential ingredient in the proof program next.

6.3.2 Stochastic Stability and Random-Time State-Dependent Drift Approach

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 $f : \mathbb{X} \to \mathbb{R}$ and measures μ on $\mathcal{B}(\mathbb{X})$ via

$$Pf(x) := \int_{\mathbb{X}} P(x, dy) f(y), \quad x \in \mathbb{X},$$

$$\mu P(A) := \int_{\mathbb{X}} \mu(dx) P(x, A), \quad A \in \mathcal{B}(\mathbb{X}).$$

A probability measure π on $\mathcal{B}(\mathbb{X})$ is called invariant if $\pi P = \pi$. That is,

$$\int \pi(dx) P(x, A) = \pi(A), \quad A \in \mathcal{B}(\mathbb{X}).$$

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. A comprehensive treatment of Markov chains can be found in [35].

Throughout this subsection, the sequence of stopping times $\{\mathcal{T}_i : i \in \mathbb{N}_+\}$ is assumed to be non-decreasing, with $\mathcal{T}_0 = 0$, measurable on the filtration generated by the state process. Additional assumptions are made in the results that follow.

Before proceeding further, we note that 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, $P^n(x, B) \geq \mu(B), \forall x \in A$, and $B \in \mathcal{B}(\mathbb{X})$. A small set leads to the construction of an accessible atom and to an invariant probability measure [35]. In many practical settings, compact sets are small; sufficient conditions on when a compact set is small has been presented in [60] and [35].

Theorem 6.20 [64] *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*:

$$\begin{aligned} &E\left[V(\phi_{\mathcal{T}_{z+1}}) \,\middle|\, \mathcal{F}_{\mathcal{T}_z}\right] \leq V(\phi_{\mathcal{T}_z}) - \delta(\phi_{\mathcal{T}_z}) + b1_{\{\phi_{\mathcal{T}_z} \in C\}}, \\ &E\left[\sum_{k=\mathcal{T}_z}^{\mathcal{T}_{z+1}-1} f(\phi_k) \mid \mathcal{F}_{\mathcal{T}_z}\right] \leq \delta(\phi_{\mathcal{T}_z}), \quad z \geq 0. \end{aligned} \tag{6.15}$$

Then the following hold:

(i) ϕ *is positive Harris recurrent, with unique invariant distribution* π

(ii) $\pi(f) := \int f(\phi)\pi(d\phi) < \infty$

(iii) *For any function* g *that is bounded by* f, *in the sense that* $\sup_\phi |g(\phi)|/f(\phi) < \infty$, *we have convergence of moments in the mean, and the Law of Large Numbers holds*:

$$\lim_{t\to\infty} E_\phi\left[g(\phi_t)\right] = \pi(g),$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=0}^{N-1} g(\phi_t) = \pi(g) \quad a.s., \ \phi \in \mathbb{X}.$$

This theorem will be important for the stability analysis to follow.

6.3.3 Noiseless and Erasure Channels

We begin with erasure channels (which contain discrete noiseless channels as a special case), before discussing more general noisy channels. Before discussing the multi-dimensional case in Sect. 6.3.3.2, we first discuss the scalar version described by (6.13).

The details of the erasure channel are specified as follows: The channel source consists of state values from $\mathbb{R}$. The source output is, as before, quantized. We consider the following uniform quantizer class. A *modified uniform quantizer* Q_K^Δ : $\mathbb{R} \to \mathbb{R}$ with step size Δ and $K + 1$ (with K even) number of bins satisfies the

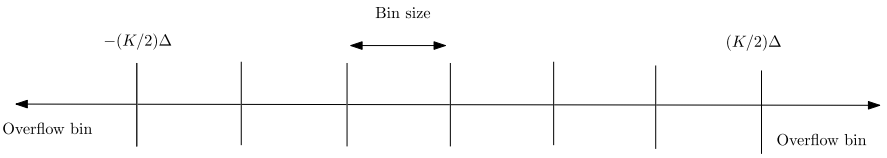


Fig. 6.2 A modified uniform quantizer. There is a single overflow bin

following for $k = 1, 2 \ldots, K$ (see Fig. 6.2):

$$Q_K^{\Delta}(x) = \begin{cases} (k - \frac{1}{2}(K+1))\Delta & \text{if } x \in [(k-1-\frac{1}{2}K)\Delta, (k-\frac{1}{2}K)\Delta), \\ (\frac{1}{2}(K-1))\Delta & \text{if } x = \frac{1}{2}K\Delta, \\ 0 & \text{if } x \notin [-\frac{1}{2}K\Delta, \frac{1}{2}K\Delta]. \end{cases} \tag{6.16}$$

where we have $\mathcal{M} = \{1, 2, \ldots, K+1\}$. 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. 6.2). We will refer to this quantizer as a *modified uniform quantizer*, since the overflow region is assigned a single bin.

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

$$P\big(q' = i | q = i\big) = p, \qquad P\big(q' = e | q = i\big) = 1 - p, \quad i \in \mathcal{M}.$$

At each time $t \geq 0$, the controller/decoder applies a mapping $\mathcal{D}_t : \mathcal{M} \cup \{e\} \to \mathbb{R}$, given by

$$\mathcal{D}_t\big(q_t'\big) = \mathcal{E}_t^{-1}\big(q_t'\big) \times 1_{\{q_t' \neq e\}} + 0 \times 1_{\{q_t' = e\}}.$$

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

$$R := \log_2(K+1), \qquad R' = \log_2(K). \tag{6.17}$$

We fix positive scalars δ, α satisfying

$$|a|2^{-R'} < \alpha < 1 \tag{6.18}$$

and

$$\alpha\big(|a| + \delta\big)^{p^{-1}-1} < 1. \tag{6.19}$$

We consider the following update rules. For $t \in \mathbb{Z}_+$ and with $\Delta_0 \in \mathbb{R}$ selected arbitrarily, consider

$$\begin{aligned} u_t &= -\frac{a}{b}\hat{x}_t, \\ \hat{x}_t &= \mathcal{D}_t\big(q_t'\big) = \Upsilon_t Q_K^{\Delta_t}(x_t), \\ \Delta_{t+1} &= \Delta_t \bar{Q}\left(\Delta_t, \left|\frac{x_t}{\Delta_t 2^{R'-1}}\right|, \Upsilon_t\right). \end{aligned} \tag{6.20}$$

Here, $\bar{Q} : \mathbb{R} \times \mathbb{R} \times \{0, 1\} \to \mathbb{R}$ is defined below, where $L > 0$ is a constant;

$$\begin{aligned} &\bar{Q}(\Delta, h, p) = |a| + \delta \quad \text{if } |h| > 1, \text{ or } p = 0, \\ &\bar{Q}(\Delta, h, p) = \alpha \quad \text{if } 0 \le |h| \le 1, \ p = 1, \ \Delta > L, \\ &\bar{Q}(\Delta, h, p) = 1 \quad \text{if } 0 \le |h| \le 1, \ p = 1, \ \Delta \le L. \end{aligned}$$

The update equations above imply that

$$\Delta_t \ge L\alpha =: L'. \tag{6.21}$$

Without any loss of generality, we assume that $L' \ge 1$.

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. As depicted in Fig. 6.1, the encoder has access to channel outputs, that is, there is noiseless feedback.

Lemma 6.3 *Under* (6.20)*, the process* (x_t, Δ_t) *is a Markov chain.*

Proof The system's state evolution can be expressed

$$x_{t+1} = ax_t - a\hat{x}_t + w_t,$$

where $\hat{x}_t = \Upsilon_t Q_K^{\Delta_t}(x_t)$. It follows that the process (x_t, Δ_t) evolves as a nonlinear state space model:

$$\begin{aligned} x_{t+1} &= a\big(x_t - \Upsilon_t Q_K^{\Delta_t}(x_t)\big) + w_t, \\ \Delta_{t+1} &= \Delta_t \bar{Q}\left(\Delta_t, \left|\frac{x_t}{2^{R'-1}\Delta_t}\right|, \Upsilon_t\right). \end{aligned} \tag{6.22}$$

in which (w_t, Υ_t) is i.i.d. Thus, the pair (x_t, Δ_t) forms a Markov chain. □

Let for a Borel set S, $\tau_S = \inf(k > 0 : (x_k, \Delta_k) \in S)$ and $E_{x,\Delta}$, $P_{(x,\Delta)}$ denote the expectation and probabilities conditioned on $(x_0, \Delta_0) = (x, \Delta)$.

Proposition 6.4 [64] *If* (6.18)–(6.19) *hold, then there exists a compact set* $A \times B \subset \mathbb{R}^2$ *satisfying the recurrence condition*

$$\sup_{(x,\Delta)\in A\times B} E_{x,\Delta}[\tau_{A\times B}] < \infty$$

and the recurrence condition $P_{(x,\Delta)}(\tau_{A\times B} < \infty) = 1$ *for any admissible* (x, Δ).

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 6.4.

Theorem 6.21 [64] *For an adaptive quantizer satisfying* (6.18)–(6.19), *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, Δ_t) *forms a positive Harris recurrent Markov chain. If the integers taken are relatively prime* (*that is they share no common divisors except for* 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 $\log_2(K+1)p$ [12]. From (6.18)–(6.19), the following is obtained.

Theorem 6.22 *If* $\log_2(K)p > \log_2(|a|)$, *then* α, δ *exist such that Theorem* 6.21 *is satisfied.*

Remark 6.7 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 a more generalized interpretation of stationarity, this result applies to a large class of memoryless channels and a class of channels with memory as to be seen later in this chapter (see Theorem 6.27): There is a direct relationship between the existence of a stationary measure and the Shannon capacity of the channel used in the system.

Under slightly stronger conditions we obtain a finite second moment:

Theorem 6.23 [64] *Suppose that the assumptions of Theorem 6.21 hold, and in addition the following bound holds*:

$$a^2\left(1-p+\frac{p}{(2^R-1)^2}\right)<1. \tag{6.23}$$

Then, for each initial condition, $\lim_{t\to\infty} E[x_t^2] = E_\pi[x_0^2] < \infty$.

Remark 6.8 We note from Minero et al. [36] that a necessary condition for mean square stability is $a^2(1-p+\frac{p}{(2^R)^2}) < 1$. Thus, the sufficiency condition in Theorem 6.23 almost meets this bound except for the additional symbol sent for the under-zoom events. We note that the average rates can be made arbitrarily close to zero 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, yet n-ergodic, quadratically stable, and asymptotic mean stationary (AMS).

6.3.3.1 Connections with Random-Time Drift Criteria

We point out the connection of the results above with random-time drift criteria in Theorem 6.20.

By Lemma 6.3, the process (x_t, Δ_t) forms 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

$$\mathcal{T}_0 = 0, \qquad \mathcal{T}_{z+1} = \inf\{k > \mathcal{T}_z : |h_k| \le 1, \Upsilon_k = 1\}, \quad z \in \mathbb{N}.$$

We can apply Theorem 6.20 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. The following lemma is key:

Lemma 6.4 *The discrete probability measure* $P(\mathcal{T}_{z+1} - \mathcal{T}_z = k \mid x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z})$ *has the upper bound*

$$P(\mathcal{T}_{z+1} - \mathcal{T}_z \ge k \mid x_{\mathcal{T}_z}, \Delta_{\mathcal{T}_z}) \le (1-p)^{k-1} + G_k(\Delta_{\mathcal{T}_z}),$$

where $G_k(\Delta_{\mathcal{T}_z}) \to 0$ *as* $\Delta_{\mathcal{T}_z} \to \infty$ *uniformly in* $x_{\mathcal{T}_z}$.

In view of Lemma 6.20, first without an irreducibility assumption, we can 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 $B_0 > 0$. One can establish the irreducibility

of the Markov chain by imposing a countability condition on the set of admissible bin sizes. A similar discussion, with a quadratic Lyapunov function, applies for finite moment analysis.

6.3.3.2 The Multi-dimensional Case

The result for the scalar problem has a natural counterpart in the multi-dimensional setting. Consider the linear system described by

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{6.24}$$

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. We also assume at this point that the eigenvalues are real, even though the extension to the complex case is primarily technical. 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 also assume that B is invertible (if B is not invertible, a sampled system can be made to have an invertible control matrix, with a periodic scheme with period at most n).

We restrict the analysis to noiseless channel in this section. The scheme proposed in the previous section is also applicable to the multi-dimensional setup. 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 each of 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 modes, 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, as we discuss now.

Consider the following system:

$$\begin{bmatrix} x_{t+1}^1 \\ x_{t+1}^2 \end{bmatrix} = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} x_t^1 \\ x_t^2 \end{bmatrix} + B \begin{bmatrix} u_t^1 \\ u_t^2 \end{bmatrix} + \begin{bmatrix} w_t^1 \\ w_t^2 \end{bmatrix}. \tag{6.25}$$

The approach entails quantizing the components in the system according to the adaptive quantization rule provided earlier for scalar systems: For $i = 1, 2$, let $R' = R_i' = \log_2(2^{R_i} - 1) = \log_2(K_i)$ (that is, the same rate is used for quantizing the components with the same eigenvalue). For $t \geq 0$ and with $\Delta_0^1, \Delta_0^2 \in \mathbb{R}$, con-

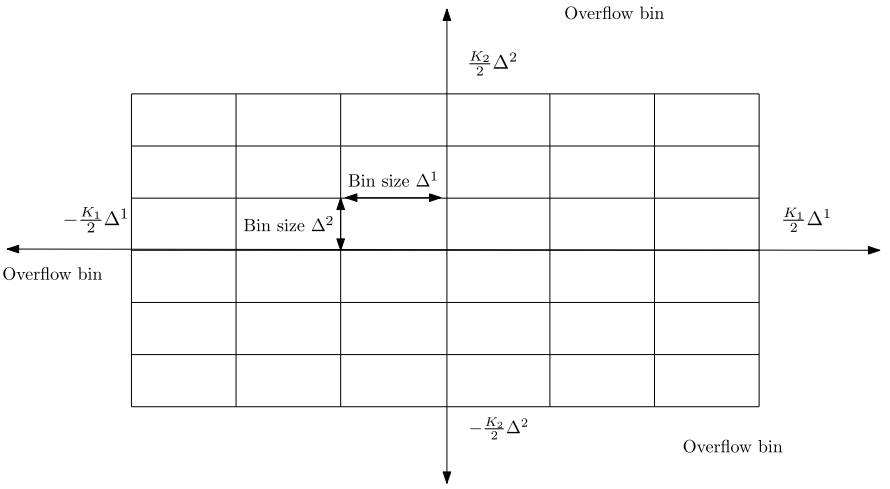


Fig. 6.3 A uniform vector quantizer. There is a single overflow bin

sider:

$$u_t = -B^{-1}A\hat{x}_t,$$

$$\begin{bmatrix} \hat{x}_t^1 \\ \hat{x}_t^2 \end{bmatrix} = \begin{bmatrix} Q_{K_1}^{\Delta_t^1}(x_t^1) \\ Q_{K_2}^{\Delta_t^2}(x_t^2) \end{bmatrix}, \tag{6.26}$$

$$\Delta_{t+1}^1 = \Delta_t^1 \bar{Q}\big(|h_t^1|, |h_t^2|, \Delta_t^1\big), \qquad \Delta_{t+1}^2 = \Delta_t^2 \bar{Q}\big(|h_t^1|, |h_t^2|, \Delta_t^2\big), \tag{6.27}$$

with, for $i = 1, 2$, $\delta^i, \varepsilon^i, \eta^i > 0$, $\eta^i < \varepsilon^i$ and $L^i > 0$ such that

$$\bar{Q}(x, y, \Delta) = |\lambda| + \delta^i \quad \text{if } |x| > 1, \text{ or } |y| > 1,$$

$$\bar{Q}(x, y, \Delta) = \frac{|\lambda|}{2^{R_i'} - \eta^i} \quad \text{if } 0 \le |x| \le 1,\ |y| \le 1,\ \Delta^i > L^i,$$

$$\bar{Q}(x, y, \Delta) = 1 \quad \text{if } 0 \le |x| \le 1,\ |y| \le 1,\ \Delta^i \le L^i.$$

Note that the above imply that $\Delta_t^i \ge L^i \frac{|\lambda|}{2^{R_i'} - \eta^i} =: L'^i$. 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 $t \ge 0$. See Fig. 6.3 for a depiction of the quantizer used at a particular time. The sequence of stopping times is now defined as follows:

$$\mathcal{T}_0 = 0, \qquad \mathcal{T}_{z+1} = \inf\big\{k > \mathcal{T}_z : |h_k^i| \le 1, i \in \{1, 2, \ldots, n\}\big\}, \quad z \in \mathbb{Z}_+,$$

where $h_k^i = \frac{x_t^i}{\Delta_t^i 2^{R_i'-1}}$. 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.

Theorem 6.24 [21, 60] *Consider the multi-dimensional system* (6.24). *If the system is controlled over a discrete-noiseless channel with capacity*

$$C > \sum_{|\lambda_i|>1} \log_2\big(|\lambda_i|\big),$$

there exists a stabilizing scheme leading to a Markov chain with a bounded second moment in the sense that $\limsup_{t\to\infty} E[|x_t|_2^2] < \infty$.

Extensions of such settings also apply to systems with decentralized multiple sensors. We refer the reader to [22] and [60].

6.3.4 Stochastic Stabilization over Noisy Channels with Noiseless Feedback

In this subsection, we consider discrete noisy channels with noiseless feedback. We first investigate Discrete Memoryless Channels (DMCs).

6.3.4.1 Asymptotic Mean Stationarity and *n*-Ergodicity

The condition $C \geq \log_2(|a|)$ in Theorem 6.18 is almost sufficient for establishing ergodicity and stability, as captured by the following discussion.

Consider the following update algorithm. 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. 6.1). Define the uniform quantizer as follows: For $k = 1, 2 \ldots, K(n)$,

$$Q_{K(n)}^{\Delta}(x) = \begin{cases} (k - \frac{1}{2}(K(n)+1))\Delta & \text{if } x \in [(k-1-\frac{1}{2}K(n))\Delta, (k-\frac{1}{2}K(n))\Delta), \\ (\frac{1}{2}(K(n)-1))\Delta & \text{if } x = \frac{1}{2}K(n)\Delta, \\ \mathcal{Z} & \text{if } x \notin [-\frac{1}{2}K(n)\Delta, \frac{1}{2}K(n)\Delta], \end{cases}$$

where $\mathcal{Z}$ is the overflow symbol in the quantizer. Let $\{x : Q_{K(n)}^{\Delta}(x) \neq \mathcal{Z}\}$ be the *granular region* of the quantizer.

At every sampling instant $t = kn, k = 0, 1, 2, \ldots$, the source coder $\mathcal{E}_t^s$ quantizes output symbols in $\mathbb{R} \cup \{\mathcal{Z}\}$ to a set $\mathcal{M}(n) = \{1, 2, \ldots, K(n)+1\}$. A channel encoder $\mathcal{E}_t^c$ maps the elements in $\mathcal{M}(n)$ to corresponding channel inputs $q_{[kn,(k+1)n-1]} \in$

$\mathcal{M}^n$. 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

$$c'_{(k+1)n-1} = \mathcal{D}_{kn}\big(q'_{[kn,(k+1)n-1]}\big).$$

Finally, the controller runs an estimator

$$\hat{x}_{kn} = \big(\mathcal{E}^s_{kn}\big)^{-1}\big(c'_{(k+1)n-1}\big) \times 1_{\{c'_{(k+1)n-1} \neq \mathcal{Z}\}} + 0 \times 1_{\{c'_{(k+1)n-1} = \mathcal{Z}\}}.$$

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

$$R'(n) > n \log_2\left(\frac{|a|}{\alpha}\right),$$

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 $t = kn$, $k \in \mathbb{N}$,

$$\begin{aligned} u_t &= -1_{\{t=(k+1)n-1\}} \frac{a^n}{b} \hat{x}_{kn}, \\ \Delta_{(k+1)n} &= \Delta_{kn} \bar{Q}\big(\Delta_{kn}, c'_{(k+1)n-1}\big), \end{aligned} \tag{6.28}$$

where c' denotes the decoder output variable. If we use $\delta > 0$ and $L > 0$ such that

$$\begin{aligned} &\bar{Q}\big(\Delta, c'\big) = (|a| + \delta)^n \quad \text{if } c' = \mathcal{Z}, \\ &\bar{Q}\big(\Delta, c'\big) = \alpha^n \quad \text{if } c' \neq \mathcal{Z},\ \Delta \geq L, \\ &\bar{Q}\big(\Delta, c'\big) = 1 \quad \text{if } c' \neq \mathcal{Z},\ \Delta < L, \end{aligned} \tag{6.29}$$

we can show 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 process space countable and as a result establish the irreducibility of the sampled process (x_{tn}, Δ_{tn}).

Theorem 6.25 [56] *For the existence of a compact coordinate recurrent set, the following is sufficient*: *The channel capacity C satisfies*: $C > \log_2(|a|)$.

Theorem 6.26 *For an adaptive quantizer satisfying the conditions of Theorem* 6.25, *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 space of admissible quantizer bins and state values*.

Theorem 6.27 [56] *Under the conditions of Theorems* 6.25 *and* 6.26, *the process* $\{x_t, \Delta_t\}$ *is* n*-stationary*, n*-ergodic*, *and hence asymptotically mean stationary* (*AMS*).

Proof Sketch The proof follows from the observation that a positive Harris recurrent Markov chain is recurrent and stationary and that if a sampled process is a positive Harris recurrent Markov chain, and if the intersampling time is fixed, with a time-homogeneous update in the inter-sampling times, then the process is mixing, n-ergodic and n-stationary. □

Remark 6.9 Converse Results for Quadratic Stability For quadratic stability, that is, the condition that $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} |x_t|^2$, exists and is finite almost surely; more restrictive conditions are needed and Shannon capacity is not sufficient (see [43] and [56]). We note that for erasure channels and noiseless channels, one can obtain tight converse theorems using Theorem 6.18 (see [36] and [64]). For general DMCs, however, a tight converse result on quadratic stabilizability is not yet available. One reason for this is that the error exponents of fixed length block codes with noiseless feedback for general DMCs are not currently known. It is worth noting that the error exponent of DMCs is typically improved with feedback, unlike the capacity of DMCs. Some partial results have been reported in [16] (e.g., the sphere packing upper bound is tight for a class of symmetric channels for rates above a critical rate even with feedback). Related references addressing partial results include [33] and [34] 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 [42] and [43]). Further discussion on this topic is available in [60] and [59].

6.3.5 Channels with Memory and Noiseless Feedback

Definition 6.7 Channels are said to be of *Class A* type, if

- 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

- Their capacity with feedback is given by

$$C = \lim_{T\to\infty} \max_{\{P(q_t|q_{[0,t-1]},q'_{[0,t-1]}),\quad 0\le t\le T-1\}} \frac{1}{T} I\left(q_{[0,T-1]} \to q'_{[0,T-1]}\right),$$

where the directed mutual information is defined by

$$I\left(q_{[0,T-1]} \to q'_{[0,T-1]}\right) = \sum_{t=1}^{T-1} I\left(q_{[0,t]}; q'_t | q'_{[0,t-1]}\right) + I\left(q_0; q'_0\right).$$

DMCs naturally belong to this class. For DMCs, feedback does not increase the capacity [12]. Such a class also includes finite state stationary Markov channels which are indecomposable [39], and non-Markov channels which satisfy certain symmetry properties [44]. Further examples can be found in [47] and in [14].

Theorem 6.28 [56] *Suppose that a linear plant given by* (6.13) *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*

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \le 0,$$

cannot be satisfied under any policy, and (*ii*) *the state process cannot be AMS under any policy.*

Remark 6.10 The result above is *negative*, but one can also obtain a positive result: 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 [39]), 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). We also note that if the channel is not information stable, then information spectrum methods lead to pessimistic realizations of capacity (known as the lim inf *in probability* of the normalized information density, see [47, 50]).

6.3.6 Higher-Order Plants

The result for the scalar problem has a natural counterpart in the multi-dimensional setting. Consider the linear system described by (6.24). In the following, we assume that all eigenvalues $\{\lambda_i, 1 \le i \le N\}$ 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.

Theorem 6.29 [60] *For such a system controlled over a Class A type noisy channel with feedback, if the channel capacity (with feedback) satisfies*

$$C < \sum_i \log_2\big(|\lambda_i|\big),$$

there does not exist a stabilizing coding and control scheme with the property

$$\liminf_{T\to\infty} \frac{1}{T} h(x_T) \le 0.$$

Proposition 6.5 [60] *For such a system controlled over a Class A type noisy channel with feedback, if*

$$C < \log_2\big(|A|\big),$$

then

$$\limsup_{T\to\infty} P\big(|x_T| \le b(T)\big) \le \frac{C}{\log_2(|A|)} > 0,$$

for all $b(T) > 0$ *such that* $\lim_{T\to\infty} \frac{1}{T} \log_2(b(T)) = 0$.

With this lemma, we state the following.

Theorem 6.30 [60] *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*

$$C \ge \log_2\big(|A|\big).$$

For sufficiency, we will assume that A is a diagonalizable matrix (a sufficient condition for which is that its eigenvalues are distinct real).

Theorem 6.31 [56] *Consider a multi-dimensional system with a diagonalizable matrix A. If the Shannon capacity of the DMC used in the controlled system satisfies*

$$C > \sum_{|\lambda_i|>1} \log_2\big(|\lambda_i|\big),$$

there exists a stabilizing scheme in the AMS sense.

On achievability of AMS stabilization over channels with memory, the discussions in Remark 6.10 also apply for this setting.

Remark 6.11 Theorem 6.31 can be extended to the case where the matrix A is not diagonalizable, in the same spirit as in Theorem 6.24, by constructing stopping times in view of the coupling between modes sharing a common eigenvalue [60].

6.4 Conclusion

In this chapter, we considered the optimization of information channels in networked control systems. We made the observation that quantizers can be viewed as a special class of channels and established existence results for optimal quantization and coding policies. Comparison of information channels for optimization has been presented. On stabilization, the relation between ergodicity and Shannon capacity has been discussed.

The value of information channels in optimization and control problems require further analysis. Particularly, further research from the information theory community for optimal non-asymptotic or finite delay coding will lead to useful applications in networked control. Error exponents with fixed block-length and feedback is currently an unresolved problem, which may lead to converse theorems for quadratic stabilization over noisy communication channels.

Acknowledgements This research has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC). Some of the figures in this chapter have appeared in [64] and [60].

The author is grateful to Giacomo Como, Bo Bernhardsson, and Anders Rantzer for hosting the workshop that took place in Lund University, which led to the publication of this book chapter. Some of the research reported in this chapter are results of the author's collaborations with Tamer Başar, Tamás Linder, Sean Meyn, and Andrew Johnston. Discussions with Giacomo Como, Aditya Mahajan, Nuno Martins, Maxim Raginsky, Anant Sahai, Naci Saldi, Sekhar Tatikonda, Demos Teneketzis, and Tsachy Weissman are gratefully acknowledged.

This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

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. 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)
3. Bar-Shalom, Y., Tse, E.: Dual effect, certainty equivalence and separation in stochastic control. IEEE Trans. Autom. Control **19**, 494–500 (1974)
4. Blackwell, D.: Equivalent comparison of experiments. Ann. Math. Stat. **24**, 265–272 (1953)
5. Bogachev, V.I.: Measure Theory. Springer, Berlin (2007)
6. Boll, C.: Comparison of experiments in the infinite case. PhD Dissertation, Stanford University (1955)
7. Borkar, V.S., Mitter, S.K.: LQG control with communication constraints. In: Kailath Festschrift. Kluwer Academic, Boston (1997)
8. Borkar, V.S., Mitter, S.K., Tatikonda, S.: Optimal sequential vector quantization of Markov sources. SIAM J. Control Optim. **40**, 135–148 (2001)
9. Borkar, V., Mitter, S., Sahai, A., Tatikonda, S.: Sequential source coding: an optimization viewpoint. In: Proceedings of the IEEE Conference on Decision and Control, pp. 1035–1042 (2005)
10. Cam, L.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 Ser. (1996)

11. 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)
12. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
13. Curry, R.E.: Estimation and Control with Quantized Measurements. MIT Press, Cambridge (1969)
14. Dabora, R., Goldsmith, A.: On the capacity of indecomposable finite-state channels with feedback. In: Proceedings of the Allerton Conf. Commun. Control Comput., pp. 1045–1052 (2008)
15. Devroye, L., Györfi, L.: Non-parametric Density Estimation: The L_1 View. Wiley, New York (1985)
16. Dobrushin, R.L.: An asymptotic bound for the probability error of information transmission through a channel without memory using the feedback. Probl. Kibern. **8**, 161–168 (1962)
17. Fischer, T.R.: Optimal quantized control. IEEE Trans. Autom. Control **27**, 996–998 (1982)
18. Fu, M.: Lack of separation principle for quantized linear quadratic Gaussian control. IEEE Trans. Autom. Control **57**, 2385–2390 (2012)
19. Gray, R.M.: Entropy and Information Theory. Springer, New York (1990)
20. György, A., Linder, T.: Codecell convexity in optimal entropy-constrained vector quantization. IEEE Trans. Inf. Theory **49**, 1821–1828 (2003)
21. Johnston, A., Yüksel, S.: Stochastic stabilization of partially observed and multi-sensor systems driven by Gaussian noise under fixed-rate information constraints. IEEE Trans. Autom. Control (2012, under review). arXiv:1209.4365
22. Johnston, A., Yüksel, S.: Stochastic stabilization of partially observed and multi-sensor systems driven by Gaussian noise under fixed-rate information constraints. In: Proceedings of the IEEE Conference on Decision and Control, Hawaii (2012)
23. Lewis, J.B., Tou, J.T.: Optimum sampled-data systems with quantized control signals. IEEE Trans. Ind. Appl. **82**, 229–233 (1965)
24. Luenberger, D.: Linear and Nonlinear Programming. Addison–Wesley, Reading (1984)
25. 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)
26. Mahajan, A., Teneketzis, D.: Optimal design of sequential real-time communication systems. IEEE Trans. Inf. Theory **55**, 5317–5338 (2009)
27. Mahajan, A., Teneketzis, D.: Optimal performance of networked control systems with non-classical information structures. SIAM J. Control Optim. **48**, 1377–1404 (2009)
28. 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)
29. 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)
30. Matveev, A.S.: State estimation via limited capacity noisy communication channels. Math. Control Signals Syst. **20**, 1–35 (2008)
31. 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)
32. Matveev, A.S., Savkin, A.V.: Estimation and Control over Communication Networks. Birkhäuser, Boston (2008)
33. Merhav, N.: On optimum parameter modulation-estimation from a large deviations perspective. IEEE Trans. Inf. Theory **58**, 172–182 (2012)
34. Merhav, N.: Exponential error bounds on parameter modulation-estimation for discrete memoryless channels. IEEE Trans. Inf. Theory (under review). arXiv:1212.4649
35. Meyn, S.P., Tweedie, R.: Markov Chains and Stochastic Stability. Springer, London (1993)
36. 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)
37. Nair, G.N., Evans, R.J.: Stabilizability of stochastic linear systems with finite feedback data rates. SIAM J. Control Optim. **43**, 413–436 (2004)

38. Nair, G.N., Fagnani, F., Zampieri, S., Evans, J.R.: Feedback control under data constraints: an overview. In: Proceedings of the IEEE, pp. 108–137 (2007)
39. Permuter, H.H., Weissman, T., Goldsmith, A.J.: Finite state channels with time-invariant deterministic feedback. IEEE Trans. Inf. Theory **55**(2), 644–662 (2009)
40. Phelps, R.: Lectures on Choquet's Theorem. Van Nostrand, New York (1966)
41. Pollard, D.: Quantization and the method of k-means. IEEE Trans. Inf. Theory **28**, 199–205 (1982)
42. Sahai, A.: Anytime Information Theory. Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA (2001)
43. 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. Theory **52**(8), 3369–3395 (2006)
44. Şen, N., Alajaji, F., Yüksel, S.: Feedback capacity of a class of symmetric finite-state Markov channels. IEEE Trans. Inf. Theory **56**, 4110–4122 (2011)
45. Silva, E.I., Derpich, M.S., Østergaard, J.: A framework for control system design subject to average data-rate constraints. IEEE Trans. Autom. Control **56**, 1886–1899 (2011)
46. Tatikonda, S., Mitter, S.: Control under communication constraints. IEEE Trans. Autom. Control **49**(7), 1056–1068 (2004)
47. Tatikonda, S., Mitter, S.: The capacity of channels with feedback. IEEE Trans. Inf. Theory **55**(1), 323–349 (2009)
48. Tatikonda, S., Sahai, A., Mitter, S.: Stochastic linear control over a communication channels. IEEE Trans. Autom. Control **49**, 1549–1561 (2004)
49. Teneketzis, D.: On the structure of optimal real-time encoders and decoders in noisy communication. IEEE Trans. Inf. Theory **52**, 4017–4035 (2006)
50. Verdú, S., Han, T.S.: A general formula for channel capacity. IEEE Trans. Inf. Theory **40**, 1147–1157 (1994)
51. Walrand, J.C., Varaiya, P.: Causal coding and control of Markov chains. Syst. Control Lett. **3**, 189–192 (1983)
52. Walrand, J.C., Varaiya, P.: Optimal causal coding-decoding problems. IEEE Trans. Inf. Theory **19**, 814–820 (1983)
53. Witsenhausen, H.S.: On the structure of real-time source coders. Bell Syst. Tech. J. **58**, 1437–1451 (1979)
54. 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)
55. Yüksel, S.: Stochastic stabilization of noisy linear systems with fixed-rate limited feedback. IEEE Trans. Autom. Control **55**, 2847–2853 (2010)
56. Yüksel, S.: Characterization of information channels for asymptotic mean stationarity and stochastic stability of non-stationary/unstable linear systems. IEEE Trans. Inf. Theory **58**, 6332–6354 (2012)
57. 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)
58. Yüksel, S.: On optimal causal coding of partially observed Markov sources in single and multi-terminal settings. IEEE Trans. Inf. Theory **59**, 424–437 (2013)
59. Yüksel, S., Başar, T.: Control over noisy forward and reverse channels. IEEE Trans. Autom. Control **56**, 1014–1029 (2011)
60. Yüksel, S., Başar, T.: Stochastic Networked Control Systems: Stabilization and Optimization Under Information Constraints. Birkhäuser, Boston (2013)
61. Yüksel, S., Linder, T.: Optimization and Convergence of Observation Channels in Stochastic Control, pp. 637–642. American Control Conference, San Francisco (2011)
62. Yüksel, S., Linder, T.: On optimal zero-delay quantization of vector Markov sources. In: Proceedings of the IEEE Conference on Decision and Control, Hawaii (2012)

63. Yüksel, S., Linder, T.: Optimization and convergence of observation channels in stochastic control. SIAM J. Control Optim. **50**, 864–887 (2012)
64. 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)
65. Zaidi, A.A., Oechtering, T.J., Yüksel, S., Skoglund, M.: Stabilization and control over Gaussian networks. In: Como, G., Bernhardsson, B., Rantzer, A. (eds.) Information and Control in Networks. Springer, Berlin (2013)

Part III
Information in Networks

Information in networks is the focus of very active current research. It is well known that information does not generally behave like a conservative fluid flow in communication networks with multiple sources and sinks. This phenomenon is even more pronounced in distributed computing systems whereby decentralized units perform computations and dynamically exchange information through a communication network. Another topic of great interest is Markov random fields, i.e., undirected graphical models describing spatial dependencies among countable sets of random variables. This third part of the book collects three contributions in these fields.

In Chap. 7, Nair addresses the question of whether there is a nontrivial class of network topologies for which associating separate data streams with each source–sink pair, with only routing and no coding performed at the network nodes, does not cause any loss of optimality. The chapter considers possibly cyclic, directed, errorless networks with n source–sink pairs and mutually independent source signals. The concept of *downward dominance* is introduced and it is shown that, if the network topology is downward dominated, then the achievability of a given combination of source signals and channel capacities implies the existence of a feasible multicommodity flow.

In Chap. 8, Elia et al. take the unifying view of systems interacting over communication networks as distributed computing systems and propose to study them as networked control systems. First, the chapter points out how a popular and well-behaved algorithm for distributed averaging can instead generate a collective global complex behavior when the inter-agent communication happens over unreliable links. Then, to mitigate the effects of the unreliable information exchange, the authors propose a new distributed averaging algorithm robust to noise and intermittent communication. The algorithm and the control perspective are the basis for the development of new distributed optimization systems that we can analyze and design as networked control systems.

Finally, Chap. 9 by Sedghi and Jonckheere discusses a problem of fast detection of faulty events and false data injection in the smart grid. The authors' approach is based on modeling the phasor measurement units as a Gaussian Markov random field, and performing a conditional covariance test.

Chapter 7
Structural Routability of n-Pairs Information Networks

Girish N. Nair

7.1 Introduction

In an *n-pairs* or *multiple unicast* communication network, n source signals must be conveyed to their corresponding sinks without exceeding any channel capacities. Until quite recently, the belief was that this was possible iff there existed a *routing* solution, i.e., if every symbol generated by a source could be carried without modification, over channels and through network nodes, until it reached the sink. At a macroscopic level, this is equivalent to presuming the existence of a feasible *multicommodity flow* [12].

However, in [2, 16], an example was constructed of a 2-pairs communication network that did not admit a routing solution, but became admissible if nodes could perform modulo-2 arithmetic on incoming bits. This counter-intuitive result started the field of *network coding*, in which nodes are permitted to not just route incoming symbols, but also to perform causal functions on them, so as to better exploit the network structure and the available channel capacities.

It is now known that the capacity regions for n-pairs networks are not generally given by feasible multicommodity flows. In [1], n-pairs networks were constructed with coding capacity much larger than the routing capacity. Other related work includes [9], in which a necessary and sufficient condition for broadcasting correlated sources over erroneous channels was found, and [13], in which linear network coding was shown to achieve capacity for a multicast network.

Notwithstanding the power of network codes, routing/multicommodity flow solutions are appealing in several respects. Most obviously they are simpler because network nodes are not required to perform extra mathematical operations on arriving bits. In addition, because different data streams are not 'hashed' together by means of some function, there is arguably less potential for cross-talk between different

A short, preliminary version without proofs was presented in the conference paper [14].

G.N. Nair (✉)
Dept. Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia
e-mail: gnair@unimelb.edu.au

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_7,

source–sink pairs, arising, for instance, from nonidealities during implementation in the physical layer. For similar reasons, routing may be preferred over network coding if security and privacy are important. Furthermore, being able to treat information as a conservative fluid flow could potentially provide a simple basis to analyze communication requirements in areas outside traditional multiterminal information theory, e.g., networked feedback control and multi-agent coordination/consensus problems—see, e.g., [3].

These considerations raise the natural questions of whether there is a general class of network topologies on which achievability is always equivalent to the existence of a feasible multicommodity flow. This chapter aims to answer this questions for possibly cyclic, directed, errorless networks with n source–sink pairs and mutually independent source signals, where the goal is to reconstruct source-signals perfectly at their respective sinks. The structural concept of *downward dominance* (Definition 7.8) is introduced, and the main result (Theorem 7.2) is that if the network topology is downward dominated then the existence of an achievable combination of source signals and channel capacities always implies the existence of a feasible multicommodity flow.

The proof relies on the iterative construction of an *entropically feasible* multicommodity flow (Definition 7.10). As downward dominance inheres solely in the topology of the network, this result suits situations where channels, switches, transceivers and interfaces are expensive to set up and difficult to move, or where channel capacities and source-signal statistics are unknown. On these structures, information can always be treated like a flow of conservative, immiscible fluids.

Downward dominance is a more general condition than the notion of 'triangularisability' that was introduced in the conference version [14] of this chapter. While it is not generally easy to verify in arbitrary n-pairs networks, Lemmas 7.2 and 7.3 give simpler, sufficient conditions for it to hold. Several examples are then provided in Sect. 7.6 to illustrate the applicability of Theorem 7.2 to various directed cyclic and acyclic examples, including cycles [10, 11], trees and trees of cycles, among others.

Although downward dominance is sufficient to guarantee that routing can always achieve the full coding capacity of a network, it is not necessary. For instance, bidirectional cycles can be shown to not be downward dominant, but routing achieves coding capacity on them [15]. The important question of finding a more general—or even tight—structural condition remains open. In the concluding section, potential directions for future work are outlined.

7.1.1 Notation and Basic Terminology

For convenience, the basic notation and terminology used in this chapter are described below.

- The set of nonnegative integers (i.e., whole numbers) is denoted by $\mathbb{W}$, the set of positive integers (i.e. natural numbers) by $\mathbb{N}$, and the set of positive reals by $\mathbb{R}_{>0}$.

- A contiguous set $\{i, i+1, \ldots, j\}$ of integers is denoted $[i : j]$.
- Other sets are usually written in boldface type.
- Random variables (rvs) are written in upper case and their realizations are indicated in corresponding lower case.
- The set operation $\mathbf{A} \setminus \mathbf{B}$ denotes $\mathbf{A} \cap \mathbf{B}^{\mathrm{c}}$.
- A discrete-time random signal or process $(F(k))_{k=0}^{\infty}$ is denoted F, and the finite sequence $(F(k))_{k=s}^{t}$ is denoted $F(s:t)$.
- Given a subscripted rv or signal F_j, with j belonging to a countable set $\mathbf{J}$, $F_{\mathbf{J}}$ denotes the tuple $(F_j)_{j\in\mathbf{J}}$, arranged according to the order on $\mathbf{J}$.
- The *entropy* of a discrete-valued rv E is denoted $\mathrm{H}[E] \geq 0$, and the conditional entropy of E given another rv F is $\mathrm{H}[E|F] := \mathrm{H}[E, F] - \mathrm{H}[F]$.
- The *mutual information* between rvs E and F is denoted $\mathrm{I}[E; F] := \mathrm{H}[E] - \mathrm{H}[E|F] \geq 0$, and the *conditional mutual information* between rvs E and F given G is denoted $\mathrm{I}[E; F|G] := \mathrm{H}[E|G] - \mathrm{H}[E|F, G]$.
- If E and F are random processes and E is discrete-valued, then the *entropy rates* of E, and the *conditional entropy rate* of E given (past and present) F are respectively defined as

$$\mathrm{H}_{\infty}[E] := \varliminf_{t\to\infty} \frac{\mathrm{H}[E(0:t)]}{t+1},$$

$$\mathrm{H}^{\infty}[E] := \varlimsup_{t\to\infty} \frac{\mathrm{H}[E(0:t)]}{t+1},$$

$$\mathrm{H}_{\infty}[E|F] := \varliminf_{t\to\infty} \frac{\mathrm{H}[E(0:t)|F(0:t)]}{t+1},$$

- If E, F and G are random processes, then the *mutual information rates* of E and F, and the *conditional mutual information rate* of E and F given (past and present) G are respectively defined as

$$\mathrm{I}^{\infty}[E; F] := \varlimsup_{t\to\infty} \frac{\mathrm{I}[E(0:t); F(0:t)]}{t+1},$$

$$\mathrm{I}_{\infty}[E; F] := \varliminf_{t\to\infty} \frac{\mathrm{I}[E(0:t); F(0:t)]}{t+1},$$

$$\mathrm{I}_{\infty}[E; F|G] := \varliminf_{t\to\infty} \frac{\mathrm{I}[E(0:t); F(0:t)|G(0:t)]}{t+1}.$$

- A *directed graph (digraph)* $(\mathbf{V}, \mathbf{A})$ consists of a set $\mathbf{V}$ of *vertices*, and a set $\mathbf{A}$ of *arcs* that each represent a directed link between a particular pair of vertices.
- The initial vertex of an arc is called its *tail* and the terminal vertex, its *head*.
- A *walk* in a digraph is an alternating sequence $\omega = (v_1, \alpha_1, v_2, \alpha_2, \ldots, \alpha_k, v_{k+1})$, $k \geq 0$, of vertices and arcs, beginning and ending in vertices s.t. each arc α_l connects the vertex v_l to v_{l+1}. Each vertex v_j and arc α_l in the sequence is said to *be in* the walk; with a minor abuse of notation, this is denoted $v_j \in \omega$.

- A *path* is a walk with no loops, i.e., it passes through no vertex more than once, including the initial one.
- An *undirected path* is an alternating sequence $\omega = (v_1, \alpha_1, v_2, \alpha_2, \ldots, \alpha_k, v_{k+1})$, $k \geq 0$, of vertices and arcs, beginning and ending in vertices s.t. no vertex is repeated and each arc α_l connects the vertex v_l to v_{l+1}, or v_{l+1} to v_l.
- A *cycle* is a walk in which the initial and final vertices are identical, but every other vertex occurs once.
- A *subpath* of a path $(v_1, \alpha_1, v_2, \alpha_2, \ldots, \alpha_k, v_{k+1})$ is a segment $(v_l, \alpha_l, v_{l+1}, \ldots, v_j)$ of it, where $1 \leq l \leq j \leq k+1$.
- A vertex v is said to be *reachable* from another vertex μ, denoted $\mu \rightsquigarrow v$, if $\exists$ a path leading from μ to v. Equivalently, it is said that μ *can reach* v. The same terminology and notation apply, with analogous meaning, for pairs of arcs as well as mixed pairs of arcs and vertices. For example, given an arc β, $\mu \rightsquigarrow \beta$ means that there is a path from the vertex μ to the tail of β.
- Similarly, a (vertex or arc) set $\mathbf{W}$ is said to be *reachable* from another set $\mathbf{U}$, denoted $\mathbf{U} \rightsquigarrow \mathbf{W}$, if there is an element of $\mathbf{W}$ that is reachable from an element of $\mathbf{U}$; equivalently, it is said that $\mathbf{U}$ *can reach* $\mathbf{W}$.
- For any vertex set $\mathbf{U} \subseteq \mathbf{V}$, $\mathrm{ARCS}(\mathbf{U}) \subseteq \mathbf{A}$ is the set of arcs with tails in $\mathbf{U}$.
- The notation $\mathrm{OUT}(\mathbf{U})$ ($\mathrm{IN}(\mathbf{U})$) represents the set of arcs in $\mathbf{A}$ that have tails (resp., heads) in a vertex set $\mathbf{U} \subseteq \mathbf{V}$ and heads (tails) $\in \mathbf{V} \setminus \mathbf{U}$. If $\mathrm{OUT}(\mathbf{U})$ ($\mathrm{IN}(\mathbf{U})$) consists of a single arc, this arc is denoted $\mathrm{out}(\mathbf{U})$ ($\mathrm{in}(\mathbf{U})$). When $\mathbf{U}$ is a singleton $\{\mu\}$, the braces are omitted.

7.2 Problem Formulation

A network of unidirectional, point-to-point channels may be modeled using a digraph $(\mathbf{V}, \mathbf{A})$, where the vertex set $\mathbf{V}$ represents information sources, sinks, repeaters, routers, etc., and the arc set $\mathbf{A}$ indicates the directions of any channels between nodes. As usual with digraphs, it is assumed that no arc leaves and enters the same vertex, and that at most one arc leads from the first to the second element of any given ordered pair of vertices. In other words, every arc in $\mathbf{A}$ may be uniquely identified with a tuple $(\mu, \nu) \in \mathbf{V}^2$, with $\mu \neq \nu$.[1] It is also assumed that the digraph is *connected*, i.e., there is an undirected path between any distinct pair of vertices.

In an n-pairs information network, the locations of sources and sinks are respectively represented by disjoint sets $\mathbf{S} = \{\sigma_1, \ldots, \sigma_n\}$ and $\mathbf{T} = \{\tau_1, \ldots, \tau_n\}$ of distinct vertices in $\mathbf{V}$, with each source σ_i aiming to communicate to exactly one sink τ_i. It is assumed that $\sigma_i \rightsquigarrow \tau_i$. Let $\mathbf{P}$ denote the sequence $((\sigma_i, \tau_i))_{i=1}^n$ of source–sink pairs, arranged in a specified order. Without loss of generality, it is assumed that every source (sink) has no in-coming (resp., out-going) arcs and exactly one out-going

[1] Such digraphs are sometimes called *simple*.

(in-coming) arc.[2] The *boundary* $\partial\mathbf{V}$ of the network is the set $\mathbf{S} \cup \mathbf{T}$ of source and sink vertices, and its *interior* is $\operatorname{int}\mathbf{V} := \mathbf{V} \setminus \partial\mathbf{V}$.

Each channel in the network can transfer bits errorlessly up to a maximum average rate, as specified by a positive *arc-capacity* $c_\alpha \in \mathbb{R}_{>0}$. In some situations, it may be natural to assign infinite capacity to certain arcs,[3] and the set of all such arcs is denoted $\mathbf{A}_\infty \subset \mathbf{A}$. In particular, the arcs leaving sources are by convention assigned infinite capacity. The set of finite-capacity arcs is written $\mathbf{A}_\mathrm{f} = \mathbf{A} \setminus \mathbf{A}_\infty$, with associated arc-capacity vector $c := (c_\alpha)_{\alpha\in\mathbf{A}_\mathrm{f}} \in \mathbb{R}_{>0}^{|\mathbf{A}_\mathrm{f}|}$. The *structure* of the n-pairs information network is defined as the tuple $\Sigma = (\mathbf{V}, \mathbf{A}_\mathrm{f}, \mathbf{A}_\infty, \mathbf{P})$.

The communication signals in the network are represented by a vector $S \equiv (S_\alpha)_{\alpha\in\mathbf{A}}$ of discrete-valued random processes called *arc signals*. In particular, the arc signals leaving sources and entering sinks respectively represent the exogenous inputs to and outputs from the network. For convenience, the input signal $S_{\mathrm{out}(\sigma_i)}$ generated by the ith source $\sigma_i \in \mathbf{S}$ is called X_i, and the output signal $S_{\mathrm{in}(\tau_i)}$ entering the ith sink $\tau_i \in \mathbf{T}$ is called Y_i. It is assumed throughout this chapter that the signals $X_1, \ldots, X_n$ are mutually independent processes with strictly positive entropy rates $\mathrm{H}_\infty[X_i] > 0$.

The arc-signal vector S is assumed to have the following property:

Definition 7.1 (Setwise Causality and Signal Graphs) An arc-signal vector S is called *setwise causal* on a structure $\Sigma = (\mathbf{V}, \mathbf{A}_\mathrm{f}, \mathbf{A}_\infty, \mathbf{P})$ if all arc signals leaving vertices in any internal vertex-set $\mathbf{U} \subseteq \operatorname{int}\mathbf{V}$ are causally determined by those entering $\mathbf{U}$ from outside it. That is, $\forall\mathbf{U} \subseteq \operatorname{int}\mathbf{V}$, $\exists$ an operator $g_\mathbf{U}$ s.t.

$$S_{\mathrm{ARCS}(\mathbf{U})}(t) = g_\mathbf{U}\big(t, S_{\mathrm{IN}(\mathbf{U})}(0:t)\big), \quad \forall t \in \mathbb{W}, \tag{7.1}$$

where $\mathrm{ARCS}(\mathbf{U}) \subseteq \mathbf{A}$ denotes the set of arcs leaving vertices of $\mathbf{U}$.

The tuple (Σ, S) is then called a *signal graph*.

Remark Setwise causality is a strengthened version of the basic concept of *well-posedness* [17] in feedback control theory. In a well-posed feedback system, the current values of all internal and output signals are uniquely determined by the past and present values of external inputs.[4] Setwise causality essentially imposes an analogous condition on any subcollection of nodes and associated signals, treated as a system. In acyclic digraphs (i.e., in which every walk is a path), it is equivalent to causality at every internal vertex. However, feedback signals may be present in cyclic digraphs, in which case vertex-wise causality cannot guarantee (7.1) without further assumptions, e.g., a positive time-delay at every vertex.

[2] If a source or sink were actually connected to multiple nodes in the network, it would be represented in the digraph by an auxiliary vertex connected by an arc (of infinite capacity) with a multiply-connected vertex.

[3] For instance, when a single network node is represented as two 'virtual' vertices connected by an arc of unbounded capacity.

[4] In the linear, time-invariant context of [17], this is equivalent to the corresponding transfer functions being well-defined and proper.

In the n-pairs network problem studied here, the objective is for each sink to perfectly reconstruct each source signal, block-by-block, using only causal operations and without exceeding any arc-capacities. This leads to the following definition:

Definition 7.2 (Achievability) Consider an n-pairs information network with structure Σ, source-signal vector X and arc-capacity vector $c \in \mathbb{R}_{>0}^{|\mathbf{A}_f|}$. The tuple (Σ, X, c) is called *achievable* if $\exists$ a setwise-causal arc-signal vector S (Definition 7.1) and a positive integer $m \in \mathbb{N}$ s.t.

$$S_{\text{out}(\sigma_i)} = X_i, \quad \forall i \in [1:n], \tag{7.2}$$

$$Y_i(km-1) = X_i\big((k-1)m : km-1\big), \quad \forall k \in \mathbb{N},\ i \in [1:n], \tag{7.3}$$

$$\mathrm{H}^{\infty}[S_\alpha] \le c_\alpha, \quad \forall \alpha \in \mathbf{A}_f. \tag{7.4}$$

Such an S is called a *solution* to the *n-pairs information network problem* (Σ, X, c). The arc-capacity vector c is called achievable on (Σ, X) and (X, c) is called achievable on Σ.

Remarks This differs from standard formulations of network coding in several respects. For instance, in [2, 5, 7, 11] and most of [10], the inequalities (7.4) are replaced by bounds either on the cardinalities of channel alphabets, or on block-coding rates over a period of time. In addition, in previous formulations, the sinks typically must reconstruct the source signals either perfectly and instantaneously [5, 7, 10], which corresponds to setting $m = 1$ in (7.3), or else with arbitrarily small probability of decoding error over blocks of sufficiently large length m [2, 11].

In this work, bounds are imposed directly on entropies, as in [10, Sect. VIII], in order to focus on the information-theoretic aspects of the problem. Errorless reconstruction is demanded so as to enable the graphical characterization of *informational dominance* from [10] to be used with very minor changes. However, perfect reconstruction is not required instantaneously in (7.3), but only in blocks of length m. This allows a solution S to be interpreted operationally in terms of variable bit-rate codes.[5] It is conjectured that the results in this chapter also apply if (7.3) is relaxed so that Y_i is causally determined by X_i, with $\mathrm{H}_\infty[Y_i] > 0$.

Finally, note that in articles on acyclic networks such as [5, 7], each source emits a vector of common dimension k and each channel carries a vector of identical dimension n, with all components taken from the same discrete alphabet. If $k = n$ a solution is said to exist and if $k = n = 1$, the solution is called *scalar*; otherwise,

[5] In other words, if S solves (Σ, X, c), then there exist variable bit-rate codes for each arc that yield errorless, block-by-block reconstruction of the source-signals at their sinks, with expected bit-rates at worst negligibly larger than arc-capacities. Conversely, if there exists a distributed entropy coding scheme that achieves perfect reconstruction of source-signals at their sinks in blocks of length m, and with expected bit-rates no larger than the arc-capacities, then this yields a solution S as defined above. However, these operational interpretations will not be used in this article.

if $k \neq n$ the solution is called *fractional*. The definition above is more general, and also applies to cyclic networks.

As mentioned in the introduction, it was once thought that a network was achievable[6] iff it admitted a routing solution. In the present context, this is equivalent to presuming the existence of an (X, c)-*feasible multicommodity flow*, i.e., of a non-negative tuple $f = (f_{\alpha,j})_{\alpha\in\mathbf{A},\, j\in[1:n]} \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|n}$, of bit-rates on each arc associated with every source–sink pair, s.t.

$$\sum_{j=1}^{n} f_{\alpha,j} \leq c_\alpha, \quad \forall \alpha \in \mathbf{A}_{\mathrm{f}} \qquad \text{(capacity bound)}, \tag{7.5}$$

$$f_{\mathrm{in}(\tau_j),j} = f_{\mathrm{out}(\sigma_j),j} = \mathrm{H}_\infty[X_j], \quad \forall j \in [1:n] \qquad \text{(supply equals demand)}, \tag{7.6}$$

$$\sum_{\alpha\in\mathrm{IN}(v)} f_{\alpha,j} = \sum_{\alpha\in\mathrm{OUT}(v)} f_{\alpha,j} \qquad \text{(conservation of flow)}, \tag{7.7}$$

for any $j \in [1:n]$ and $v \in \mathbf{V} \setminus (\{\sigma_j\} \cup \{\tau_j\})$. Via an explicit counterexample, the article [2] showed that this intuitive notion was incorrect, i.e., that although the existence of a feasible multicommodity flow is sufficient for achievability, it is not generally necessary. This laid the foundations for *(linear) network coding*, in which nodes are permitted to not just route incoming bits, but also to combine them using (linear) functions.

Nonetheless, routing/multicommodity-flow solutions have certain virtues, as discussed in Sect. 7.1. This chapter poses the question: Is there a general class of n-pairs information network structures $\mathcal{Σ}$ in which the achievability of (X, c) is equivalent to the existence of an (X, c)-feasible multicommodity flow f (7.5)–(7.7)?

Any n-pairs information network structure $\mathcal{Σ}$ can support (X, c)-feasible multicommodity flows if the arc-capacities are sufficiently larger than the source entropy rates, provided each sink is reachable from its source. However, there are examples of structures on which an (X, c)-feasible multicommodity flow does not exist if arc-capacities are reduced, even though (X, c) is still achievable (see Sect. 7.6).

The aim of this chapter is to isolate certain structural properties that ensure routability *over all achievable combinations of* (X, c). Such properties would inhere solely in $\mathcal{Σ}$, suiting situations in which channels, switches, transceivers and interfaces are expensive to set up and difficult to move, and/or where channel capacities and source-signal statistics are variable or unknown.

7.3 Preliminary Notions

Before proceeding, several existing graph-theoretic notions are needed. Throughout this section, $\mathcal{Σ} = (\mathbf{V}, \mathbf{A}, \mathbf{P}) \equiv (\mathbf{V}, \mathbf{A}_{\mathrm{f}}, \mathbf{A}_\infty, \mathbf{P})$ is the structure of an n-pairs informa-

[6] Ignoring differences in the definition of achievability.

tion network as described in Sect. 7.2, and $\Gamma = (\Sigma, S)$ is its setwise-causal signal graph (Definition 7.1), with source- and sink-signal vectors X and Y.

First, some largely familiar concepts are revisited. A (directed, acyclic) path in Σ that goes from a source σ_i to its sink τ_i is called here an *i-path*, and the set of all i-paths, an *i-bundle*. Given a set $\mathbf{J} \subseteq [1:n]$, the set of all i-paths with $i \in \mathbf{J}$ is called a **J**-*bundle* (not the same as the set of $\sigma_{\mathbf{J}} \rightsquigarrow \tau_{\mathbf{J}}$-paths, which contains it). Let $(\mathbf{V}^{\mathbf{J}}, \mathbf{A}^{\mathbf{J}})$ denote the subgraph formed by all the vertices and arcs in the **J**-bundle. In particular, $(\mathbf{V}^i, \mathbf{A}^i)$ is the subgraph formed by the i-bundle. A vertex set $\mathbf{U} \subset \mathbf{V}^i$ such that $\sigma_i \in \mathbf{U}$ and $\tau_i \notin \mathbf{U}$ is called an *i-cut*.

The following concepts are adapted from [10], with minor changes in terminology.

Definition 7.3 (Indirect i-Walks—Based on [10]) An *indirect i-walk (ii-walk)* ω is an alternating sequence $(\alpha_1, \beta_1, \ldots, \alpha_{j-1}, \beta_{j-1}, \alpha_j)$ of forward- and reverse-oriented paths in the n-pairs structure Σ such that

1. α_1 begins with the ith source vertex σ_i;
2. Both α_ℓ and β_ℓ end with the same vertex μ_ℓ, $\forall \ell \in [1 : j-1]$;
3. Both β_ℓ and $\alpha_{\ell+1}$ begin from the same source vertex, $\forall \ell \in [1 : j-1]$;
4. α_j ends with the sink vertex τ_i; and
5. Every arc and vertex in ω can reach τ_i.

An ii-walk ω is said to *bypass* an arc-set **C** if no arc in ω lies in **C**.

Remarks Note that the fifth condition above is equivalent to the requirement that each joint vertex μ_ℓ reaches τ_i, $\forall \ell \in [1 : j-1]$.

An ii-walk as defined above is, in the terminology of [10], an *indirect walk* from $\text{out}(\sigma_i)$ to $\text{in}(\tau_i)$ in a subgraph $G(\emptyset, i)$. Similarly, an ii-walk that bypasses **C** is an indirect walk from $\text{out}(\sigma_i)$ to $\text{in}(\tau_i)$ in a subgraph $G(\mathbf{C}, i)$; if such a bypass exists, then Y_i is not always fully determined by $S_{\mathbf{C}}$, even if all i-paths go through **C**. See Fig. 3 and [10, Definitions 10–11].

Indirect i-walks are related to the concept of *fd-separation* [11]. In particular, if $S_{\mathbf{C}}$ fd-separates X_i and Y_i for any setwise causal S (Definition 7.1), then all ii-walk's pass through **C**; that is, an ii-walk that bypasses **C** corresponds to an undirected path between X_i and Y_i in a *functional dependence* subgraph $\mathcal{G}_{X_i, S_{\mathbf{C}}, Y_i}$ constructed according to the procedure in [11].

However, the converse is not generally true, i.e., 'ii-separation' is a less stringent requirement. This is because paths connecting X_i and Y_i in $\mathcal{G}_{X_i, S_{\mathbf{C}}, Y_i}$ do not have to satisfy an analogue of the fifth condition, which arises from the requirement that each sink reproduce its source signal with perfect fidelity. For this to be possible, it turns out that each joint vertex μ_ℓ in an ii-walk must be able to reach τ_i.

Put another way, requiring $S_{\mathbf{C}}$ to fd-separate X_i and Y_i is equivalent to requiring that (a) **C** be an i-cut, and (b) for each $j \neq i$, either all $\sigma_j \rightsquigarrow \tau_i$-paths (if any) bypass **C**, or all pass through it. Under ii-separation, (a) must still hold, but (b) is relaxed: a source σ_j can have a path π to τ_i that bypasses **C** as well as another that passes through **C**, provided that π is not the last leg of an ii-walk that bypasses **C**.

Definition 7.4 (Structural Dominance—Based on [10]) For any arc-set $\mathbf{B} \subseteq \mathbf{A}$ in an n-pairs network, SDOM($\mathbf{B}$) is the smallest arc-set $\mathbf{C} \subseteq \mathbf{A}$ that satisfies the conditions below:

1. $\mathbf{C} \supseteq \mathbf{B}$
2. $\text{out}(\sigma_i) \in \mathbf{C}$ iff $\text{in}(\tau_i) \in \mathbf{C}$
3. If $\alpha \in \mathbf{A}$ is *downstream* from $\mathbf{C}$—i.e., all paths from sources to the tail of α pass through $\mathbf{C}$—then $\alpha \in \mathbf{C}$.
4. If all indirect i-walks (Definition 7.3) pass through $\mathbf{C}$ then $\text{out}(\sigma_i), \text{in}(\tau_i) \in \mathbf{C}$.

The arcs in SDOM($\mathbf{B}$) are said to be *structurally dominated by* $\mathbf{B}$.

Remarks Note that SDOM($\mathbf{B}$) is the smallest such arc-set in the sense of being contained by every $\mathbf{C} \subseteq \mathbf{A}$ that satisfies criteria 1–4.

As noted in [8, pp. 199–200], SDOM($\mathbf{B}$) can be constructed by setting $\mathbf{C} = \mathbf{B}$, letting $\mathbf{T} = \mathbf{A} \setminus \mathbf{B} \neq \emptyset$ be the set of arcs to be tested, and then following this greedy algorithm:

(i) Pick any arc $\alpha \in \mathbf{T}$.
(ii) If α satisfies any of the conditions 2–4 in Definition 7.4, update $\mathbf{C} \leftarrow \mathbf{C} \cup \{\alpha\}$ and then $\mathbf{T} \leftarrow \mathbf{A} \setminus \mathbf{C}$; else keep $\mathbf{C}$ the same and update $\mathbf{T} \leftarrow \mathbf{T} \setminus \{\alpha\}$.
(iii) If $\mathbf{T} = \emptyset$ then exit; else go to step (i).

The final set $\mathbf{C}$ is then SDOM($\mathbf{B}$). However, the following lemma gives two quicker conditions for guaranteeing that a specific arc lies in SDOM($\mathbf{B}$).

Lemma 7.1 (Based on [10])

1. *If an arc* $\alpha \in \mathbf{A}$ *is downstream from* $\mathbf{B}$, *then* $\alpha \in \text{SDOM}(\mathbf{B})$.
2. *If all indirect* i *walks* (*Definition* 7.3) *pass through* $\mathbf{B}$ *then* $\text{out}(\sigma_i), \text{in}(\tau_i) \in \text{SDOM}(\mathbf{B})$.

Proof If either of these criteria hold, then the relevant arcs—α, $\text{out}(\sigma_i)$, $\text{in}(\tau_i)$—must lie inside any arc-set $\mathbf{C}$ that satisfies the conditions 1–4 in Definition 7.4. As SDOM($\mathbf{B}$) is such a set, the lemma follows. □

The significance of structural dominance arises from the following result:

Theorem 7.1 (Informational Dominance—Based on [10]) *Consider any arc* $\alpha \in \mathbf{A}$ *and arc-set* $\mathbf{B} \subseteq \mathbf{A}$ *in an* n*-pairs network with structure* Σ. *If* $\alpha \in \text{SDOM}(\mathbf{B})$ (*Definition* 7.4), *then for any setwise causal arc-signal vector* S (*Definition* 7.1) *and positive integer* $m \in \mathbb{N}$ *that satisfy* (7.2) *and* (7.3), $\exists$ *a function* γ *such that*

$$S_\alpha(0:km) = \gamma\big(k, S_{\mathbf{B}}(0:km)\big), \quad \forall k \in \mathbb{N}. \tag{7.8}$$

Conversely, if for any setwise causal S *that meets* (7.2)–(7.3) *with block-length* m *there is a function* γ *ensuring that* (7.8) *holds, then* $\alpha \in \text{SDOM}(\mathbf{B})$.

Remarks The property specified in (7.8) is a version of the concept of *informational dominance* [10]; the important of this result lies in giving this functional concept a purely structural characterization. The proof follows similar lines as that of Theorem 10 in [10] and is omitted. Minor differences are that m is not constrained to be 1 here, and that cyclic networks are handled using the notion of setwise causality (Definition 7.1), rather than by introducing channel delays and 'unrolling' the network over time to yield an infinite directed acyclic graph.

7.4 Main Result

The main result of this chapter is presented in this section. In order to do so, several nonstandard graph-theoretic notions are needed. The reader is referred to Sect. 7.6 for examples that illustrate these notions. Recall again that $\Sigma = (\mathbf{V}, \mathbf{A}, \mathbf{P}) \equiv (\mathbf{V}, \mathbf{A}_{\mathrm{f}}, \mathbf{A}_{\infty}, \mathbf{P})$ is the structure of an n-pairs information network as described in Sect. 7.2, and $\Gamma = (\Sigma, S)$ is its setwise-causal signal graph (Definition 7.1), with source- and sink-signal vectors X and Y.

Definition 7.5 ($\mathbf{J}$-Disjointness) Given an index set $\mathbf{J} \subseteq [1:n]$, an arc set $\mathbf{B} \subseteq \mathbf{A}$ is $\mathbf{J}$-*disjoint* if each path in the $\mathbf{J}$-bundle passes through at most one arc in $\mathbf{B}$.

If $\mathbf{J} = \{i\}$ for some $i \in [1:n]$, then $\mathbf{B}$ is called i-*disjoint*.

Remarks It is easy to see that empty and singleton arc-sets are automatically $\mathbf{J}$-disjoint, that every $\mathbf{B} \subseteq \mathbf{A}$ is $\emptyset$-disjoint, and that every subset of a $\mathbf{J}$-disjoint set inherits its $\mathbf{J}$-disjointness. With a little effort, it can also be shown that $\mathbf{J}$-disjoint arc-sets satisfy the 'augmentation' property. Thus $\mathbf{J}$-disjoint sets form a *finite matroid* on $\mathbf{A}$.

Structural dominance (Definition 7.4) and $[1:i]$-disjointness are next used to define nested families of arc-sets with certain structural properties. These properties are needed later to inductively extract *entropically feasible* multicommodity flows (Definition 7.10). First, for any arc-set $\mathbf{E} \subseteq \mathbf{A}$ and $h \in [1:n]$ define the source-augmented set

$$\mathbf{E}_{*h} := \mathbf{E} \cup \mathrm{OUT}(\sigma_{\mathbf{J}^{h-1} \cup [h+1:n]}), \qquad \mathbf{J}^{h-1} \equiv \left\{ j \in [1:h-1] : \mathbf{E} \cap \mathbf{A}^j = \emptyset \right\}. \quad (7.9)$$

That is, $\mathbf{E}$ is augmented by those source-arcs that either have indices greater than h or that have indices less than h but no source-sink paths going through $\mathbf{E}$.

Definition 7.6 (i-Downward Dominated Sets) For each $i \in [1:n]$, the family $\mathcal{D}_i$ consists of all arc sets $\mathbf{E} \subseteq \mathbf{A}$ such that

1. $\mathbf{E}$ is $[1:i]$-disjoint (Definition 7.5), and
2. For each $h \in [1:i]$, either the h-bundle does not touch $\mathbf{E}$, i.e. $\mathbf{E} \cap \mathbf{A}^h = \emptyset$, or else the source-augmented arc-set $\mathbf{E}_{*h}$ (7.9) structurally dominates the source-arc $\mathrm{out}(\sigma_h)$ (Definition 7.4).

Every member-set of $\mathcal{D}_i$ is called *i-downward dominated*.

Remark Clearly, every $\mathcal{D}_i$-set is also in $\mathcal{D}_{i-1}$.

The next concept describes a class of i-cuts that have a special structure:

Definition 7.7 (Viable i-Cuts) Given an index $i \in [1:n]$, an i-cut $\mathbf{U} \subset \mathbf{V}^i$ is called *viable* under the following conditions:

1. Every arc leaving $\mathbf{U}$ in the i-bundle is finite-capacity, i.e., $\mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i \subseteq \mathbf{A}_{\mathrm{f}}$.
2. There is an i-path that leaves $\mathbf{U}$ without re-entering.
3. Each arc in $\mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i$ lies in an i-path that either exits $\mathbf{U}$ without re-entering or else lies in the $[1:i-1]$-bundle.
4. Every vertex $v \in \mathbf{U}$ lies on an undirected path π from σ_i to v such that
   (a) All vertices before v on π are in $\mathbf{U}$, and
   (b) Every reverse-oriented arc in π (i.e. pointing from v to σ_i) lies on an i-path that does not re-enter $\mathbf{U}$.

Remark Viable i-cuts correspond to possible *min-cuts* in a residual capacitated digraph that is used to prove the main result of this chapter (Theorem 7.2). Further investigation of these min-cuts may yield other structural properties to add to the list above; however, this is left for future work.

Definition 7.8 (Downward Dominance) A structure Σ is called *downward dominated* if for each $i \in [2:n]$ and viable i-cut $\mathbf{U}$ (Definition 7.7), the set $\mathbf{O}^i = \mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i$ of outgoing arcs in the i-bundle satisfies the following two conditions:

1. $\mathbf{O}^i \in \mathcal{D}_{i-1}$ (Definition 7.6), and
2. The source-augmented arc-set $\mathbf{O}^i_{*i}$ (7.9) structurally dominates the source-arc $\mathrm{out}(\sigma_i)$ (Definition 7.4).

Remark Note that 1-pair structures are automatically downward dominated, since the conditions above become empty.

A sequence of simpler and increasingly restrictive sufficient conditions for downward dominance can also be found by exploiting Lemma 7.1:

Lemma 7.2 (Simpler Condition 1) *Suppose that for each $i \in [2:n]$ and viable i-cut $\mathbf{U} \subset \mathbf{V}^i$ (Definition 7.7), the arc-set $\mathbf{O}^i = \mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i \subseteq \mathbf{A}_{\mathrm{f}}$ satisfies the following conditions*:

1. *$\mathbf{O}^i$ is $[1:i-1]$-disjoint (Definition 7.5), and*
2. *For each index $h \in [1:i]$ for which there is a h-path that passes through $\mathbf{O}^i$, i.e., $\mathbf{O}^i \cap \mathbf{A}^h \neq \emptyset$, all indirect h-walks (Definition 7.3) pass through the source-augmented arc-set $\mathbf{O}^i_{*h}$ (7.9).*

Then Σ is downward dominated (*Definition* 7.8).

Proof Follows immediately from applying Lemma 7.1 to Definitions 7.6 and 7.4. □

Lemma 7.3 (Simpler Condition 2) *Suppose that for each $i \in [2:n]$ and viable i-cut $\mathbf{U} \subset \mathbf{V}^i$ (Definition 7.7), the arc-set $\mathbf{O}^i = \mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i \subseteq \mathbf{A}_\mathrm{f}$ satisfies the following conditions*:

1. *$\mathbf{O}^i$ is $[1:i-1]$-disjoint (Definition 7.5).*
2. *For every $h \in [1:i]$ and $s \in [1:h]$ such that $\mathbf{O}^i \cap \mathbf{A}^h$ and $\mathbf{O}^i \cap \mathbf{A}^s \neq \emptyset$, all paths from σ_s to τ_h pass through $\mathbf{O}^i$.*

Then Σ is downward dominated (*Definition* 7.8).

Proof Let $\mathbf{O}^i \cap \mathbf{A}^h \neq \emptyset$ for some $h \in [1:i]$. It is asserted that all indirect h-walks (Definition 7.3) must pass through $\mathbf{O}^i_{*h}$.

To see this, suppose in contradiction that there is an indirect h-walk ω that does not pass through $\mathbf{O}^i_{*h} \equiv \mathbf{O}^i \cup \mathrm{OUT}(\sigma_{\mathbf{J}^{h-1} \cup [h+1:n]})$, where $\mathbf{J}^{h-1} \equiv \{j \in [1:h-1] : \mathbf{O}^i \cap \mathbf{A}^j = \emptyset\}$. Let σ_s be the last source vertex in ω, and let π be the subpath from σ_s to τ_h. Clearly, $s \notin \mathbf{J}^{h-1} \cup [h+1:n]$. In addition, $s \neq h$, since otherwise ω reduces to a path from σ_h to τ_h, which by the second condition above must pass through $\mathbf{O}^i \subseteq \mathbf{O}^i_{*h}$.

Thus $s \in [1:h-1] \setminus \mathbf{J}^{h-1}$, i.e. $\mathbf{O}^i \cap \mathbf{A}^s \neq \emptyset$. By the second condition above, all $\sigma_s \rightsquigarrow \tau_h$-paths must then pass through $\mathbf{O}^i$. As π is such a path, the indirect h-walk ω, of which it is a part, passes through $\mathbf{O}^i \subseteq \mathbf{O}^i_{*h}$, yielding a contradiction.

The result then follows from Lemma 7.2. □

Lemma 7.4 (Simpler Condition 3) *Suppose that for each $i \in [2:n]$ and every viable i-cut $\mathbf{U} \subset \mathbf{V}^i$ (Definition 7.7), there is exactly arc in $\mathbf{O}^i = \mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i \subseteq \mathbf{A}_\mathrm{f}$.*

Furthermore, suppose that for each $h \in [1:i]$ and $s \in [1:h]$ such that $\mathbf{O}^i \cap \mathbf{A}^h$ and $\mathbf{O}^i \cap \mathbf{A}^s \neq \emptyset$, all paths from σ_s to τ_h pass through $\mathbf{O}_i$, or none of them do.

Then Σ is downward dominated (*Definition* 7.8).

Proof Observe that $\mathbf{O}^i$ consists of a single arc α. Thus the first condition of Lemma 7.3 is trivially satisfied. To show that its second condition is also met, suppose that $\mathbf{O}^i \cap \mathbf{A}^h, \mathbf{O}^i \cap \mathbf{A}^s \neq \emptyset$ for some $h \in [1:i]$, $s \in [1:h]$. Thus α is on both an s-path and a h-path. Let π_1 be the subpath of the s-path from σ_s to the tail of α and π_2, the subpath of the h-path from the head of α to τ_h. Then the concatenation $\pi_1 \alpha \pi_2$ is a $\sigma_s \rightsquigarrow \tau_h$-path that passes through $\mathbf{O}^i$. By the all-or-nothing condition above, all $\sigma_s \rightsquigarrow \tau_h$-paths then pass through $\mathbf{O}^i$. The result then follows from Lemma 7.3. □

Remark It is often easier to check the conditions above on all i-cuts $\mathbf{U}$ having finite-capacity outgoing arcs in the i-bundle, rather than trying to identify the ones that are also viable.

The main result of this chapter can now be stated:

Theorem 7.2 (Downward Dominance $\Rightarrow$ Structural Routability) *If there is an ordering of the source–sink pairs in an n-pairs network so that the structure Σ is downward dominated* (*Definition* 7.8), *then the achievability of* (X, c) (*Definition* 7.2) *implies the existence of an* (X, c)*-feasible multicommodity flow* (7.5)–(7.7).

Conversely, if X is stationary and there exists an (X, c)-feasible multicommodity flow with (7.5) *holding in strict form, then* (Σ, X, c) *is achievable.* $\triangledown$

Remarks This result defines a non-trivial class of directed network structures for which achievability is essentially equivalent to the existence of a feasible multicommodity flow. On these structures, information can indeed be treated like an incompressible, immiscible fluid flow.

The proof of Theorem 7.2 is given in the next section. In Sect. 7.6, several network examples are discussed to illustrate the applicability of Theorem 7.2.

7.5 Proof of Theorem 7.2

In both the proofs of necessity and sufficiency, use will be made of the fact that $\forall i \in [1:n]$, any single-commodity flow q from σ_i to τ_i in the structure Σ can be decomposed into a superposition of *i-path flows* and *cycle flows* (see, e.g., [4, Theorem 3.3.1]). That is, if $\pi_{1,i}, \ldots, \pi_{p_i,i}$ are the distinct i-paths and $\gamma_1, \ldots, \gamma_g$, the distinct cycles, then $\exists$ numbers $u_{1,i}, \ldots, u_{p,i} \geq 0$ and $w_{1,i}, \ldots, w_{g,i} \geq 0$ s.t.

$$q_\alpha = \sum_{1 \leq k \leq p_i : \pi_{k,i} \ni \alpha} u_{k,i} + \sum_{1 \leq l \leq g : \gamma_l \ni \alpha} w_{l,i}. \tag{7.10}$$

If $w_{l,i} = 0$ for all $l \in [1:g]$, then the flow q is called *acyclic*.

The proof of sufficiency in Sect. 7.5.2 is relatively straightforward. Given an (X, c)-feasible multicommodity flow f (7.5)–(7.7) on Σ, the decomposition (7.10) is used directly to devise a routing solution S.

The proof of necessity in Sect. 7.5.1 is more difficult and involves induction, using the following building blocks.

Definition 7.9 (**J**-Flow) Given an index set $\mathbf{J} \subseteq [1:n]$, a nonnegative tuple $f = (f_{\alpha,j})_{\alpha \in \mathbf{A}, j \in \mathbf{J}} \in \mathbb{R}_{\geq 0}^{|\mathbf{A}||\mathbf{J}|}$ is called a **J**-*flow* on the structure Σ if $\forall j \in \mathbf{J}$ and $v \in \mathbf{V} \setminus \{\sigma_j\} \cup \{\tau_j\}$,

$$\sum_{\alpha \in \mathrm{IN}(v)} f_{\alpha,j} = \sum_{\alpha \in \mathrm{OUT}(v)} f_{\alpha,j} \quad (j\text{-flow conservation}), \tag{7.11}$$

As a convention, the $\emptyset$-flow is defined as the empty sequence ().

Remark A $\mathbf{J}$-flow is a (possibly infeasible) multicommodity flow with source-sink pairs (σ_j, τ_j), $j \in \mathbf{J}$. If each j-flow $f_{\mathbf{A},j}$ is acyclic, $\forall j \in \mathbf{J}$, then f is called an *acyclic* $\mathbf{J}$*-flow*.

The next concept is central to the proof of necessity. It defines a class of feasible $[1:i]$-flows that obey certain information-theoretic bounds when only the signals X_j, $j \in [1:i]$, need to be communicated.

Definition 7.10 (Entropic Feasibility) Given $i \in [1:n]$ and a solution S to (Σ, X, c) (Definition 7.2), a $[1:i]$-flow $f \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|i}$ (Definition 7.9) is called *entropically feasible* if it satisfies the following conditions:

(i) On every arc $\alpha \in \mathbf{A}_{\mathrm{f}}$,

$$\sum_{j=1}^{i} f_{j,\alpha} \leq c_\alpha. \tag{7.12}$$

(ii) On any i-downward dominated arc set $\mathbf{B}$ (Definition 7.6),

$$\sum_{\alpha \in \mathbf{B},\, j \in [1:i]} f_{\alpha,j} \leq \mathrm{H}^{\infty}[S_{\mathbf{B}} | X_{\mathbf{J}^i \cup [i+1:n]}], \tag{7.13}$$

where $\mathbf{J}^i = \{j \in [1:i] : \mathbf{B} \cap \mathbf{A}^j = \emptyset\}$.

(iii) On arcs entering sinks and leaving sources,

$$f_{\mathrm{in}(\tau_j),j} = f_{\mathrm{out}(\sigma_j),j} = \mathrm{H}_{\infty}[X_j], \quad \forall j \in [1:i]. \tag{7.14}$$

Remarks Note that the $\emptyset$-flow is entropically feasible, since the condition (7.14) disappears and (7.13) is trivially satisfied due to a zero left-hand side (LHS).

The proof of necessity in the next section proceeds by constructing an entropically feasible $[1:n]$-flow on (Σ, c, S), which automatically gives the desired (X, c)-feasible multicommodity flow (7.5)–(7.7).

7.5.1 Necessity Proof for Theorem 7.2

Let the arc-signal vector S be a solution (Definition 7.2) to the n-pairs information network problem (Σ, X, c). An entropically feasible $[1:n]$-flow (Definition 7.10) f^n will be constructed, using upward induction.

Let Σ be downward dominated (Definition 7.8) and suppose that $f^{i-1} = (f_{\alpha,j})_{\alpha \in \mathbf{A}, j \in [1:i-1]} \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|(i-1)}$ is an entropically feasible, acyclic $[1:i-1]$-flow for some $i \in [1:n]$, noting that the $\emptyset$-flow f^0 is entropically feasible. An i-flow $(f_{\alpha,i})_{\alpha \in \mathbf{A}} \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|}$ will be constructed in such a way that $f^i \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|i}$ will be an entropically feasible, acyclic $[1:i]$-flow.

On any arc $\alpha \in \mathbf{A}$, let

$$r_\alpha := \begin{cases} c_\alpha - \sum_{j=1}^{i-1} f_{\alpha,j} & \text{if } \alpha \in \mathbf{A}_{\mathrm{f}}, \\ \infty & \text{if } \alpha \in \mathbf{A}_\infty \equiv \mathbf{A} \setminus \mathbf{A}_{\mathrm{f}} \end{cases} \tag{7.15}$$

be the residual capacity after subtracting the relevant components of f^{i-1}. Note that $r_\alpha \overset{(7.12)}{\geq} 0$ since f^{i-1} is an entropically feasible $[1:i-1]$-flow. The next step is to find an acyclic i-flow (Definition 7.9) $q \in \mathbb{R}_{\geq 0}^{|\mathbf{A}|}$ from $\sigma_i \rightsquigarrow \tau_i$ that is *a)* $\leq$ the residual capacity on each arc, and *b)* $\geq \mathrm{H}_\infty[X_i]$ on the arc entering τ_i. There are two mutually exclusive cases to consider.

7.5.1.1 1st Case: ∃ an i-Path with No Finite-Capacity Arcs

Denote this i-path by π_e, noting that $r_\alpha = \infty, \forall \alpha \in \pi_e$ by the second line of (7.15). Set the i-path flows as

$$u_k = \begin{cases} \mathrm{H}_\infty[X_i] & \text{if } k = e, \\ 0 & \text{otherwise,} \end{cases} \qquad \forall k \in [1:p], \tag{7.16}$$

and the cycle flows equal to zero in the decomposition (7.10) (dropping the i-subscripts), so that

$$q_\alpha \overset{(7.10)}{=} \sum_{1 \leq k \leq p : \pi_k \ni \alpha} u_k, \quad \forall \alpha \in \mathbf{A}. \tag{7.17}$$

Evidently q is acyclic and meets the residual capacity constraint on all arcs in $\mathbf{A}$. Furthermore, since every i-path passes through the single arc entering τ_i,

$$q_{\mathrm{in}(\tau_i)} \overset{(7.17)}{=} \sum_{1 \leq k \leq p} u_k \overset{(7.16)}{=} u_e = \mathrm{H}_\infty[X_i], \tag{7.18}$$

satisfying the conditional information constraint.

7.5.1.2 2nd Case: Every i-Path Has One or More Finite-Capacity Arcs

Observe first that for any arc set $\mathbf{B} \subseteq \mathbf{A}$,

$$\begin{aligned} \sum_{\beta \in \mathbf{B}} c_\beta &\overset{(7.4)}{\geq} \sum_{\beta \in \mathbf{B}} \mathrm{H}^\infty[S_\beta] \equiv \sum_{\beta \in \mathbf{B}} \varlimsup_{t \to \infty} \frac{\mathrm{H}[S_\beta(0:t)]}{t+1} \\ &\geq \varlimsup_{t \to \infty} \frac{1}{t+1} \sum_{\beta \in \mathbf{B}} \mathrm{H}\big[S_\beta(0:t)\big] \geq \varlimsup_{t \to \infty} \frac{\mathrm{H}[S_{\mathbf{B}}(0:t)]}{t+1} \equiv \mathrm{H}^\infty[S_{\mathbf{B}}] \qquad (7.19) \\ &\geq \mathrm{H}^\infty[S_{\mathbf{B}} | X_{\mathbf{J}^{i-1} \cup [i+1:n]}] \qquad (7.20) \end{aligned}$$

$$
\begin{aligned}
&= \varlimsup_{t\to\infty}\left(\frac{\mathrm{H}[S_{\mathbf{B}}(0:t)|X_{\mathbf{J}^{i-1}\cup[i+1:n]}(0:t)] - \mathrm{H}[S_{\mathbf{B}}(0:t)|X_{\mathbf{J}^{i-1}\cup[i:n]}(0:t)]}{t+1}\right.\\
&\quad\left. + \frac{\mathrm{H}[S_{\mathbf{B}}(0:t)|X_{\mathbf{J}^{i-1}\cup[i:n]}(0:t)]}{t+1}\right)\\
&= \varlimsup_{t\to\infty}\left(\frac{\mathrm{I}[S_{\mathbf{B}}(0:t); X_i(0:t)|X_{\mathbf{J}^{i-1}\cup[i+1:n]}(0:t)]}{t+1}\right.\\
&\quad\left. + \frac{\mathrm{H}[S_{\mathbf{B}}(0:t)|X_{\mathbf{J}^{i-1}\cup[i:n]}(0:t)]}{t+1}\right)\\
&\geq \mathrm{I}_\infty[X_i; S_{\mathbf{B}}|X_{\mathbf{J}^{i-1}\cup[i+1:n]}] + \mathrm{H}^\infty[S_{\mathbf{B}}|X_{\mathbf{J}^{i-1}\cup[i:n]}]\\
&= \mathrm{I}_\infty[X_i; S_{\mathbf{B}}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}] + \mathrm{H}^\infty[S_{\mathbf{B}}|X_{\mathbf{J}^{i-1}\cup[i:n]}], \qquad (7.21)
\end{aligned}
$$

where (7.19) is due to the subadditivity of joint entropy, (7.20) holds because conditioning cannot increase entropy, and (7.21) arises from the mutual independence of $X_1, \ldots, X_n$.

Now, consider the residual capacitated digraph $(\mathbf{V}^i, \mathbf{A}^i, r_{\mathbf{A}^i})$ formed by the i-bundle.[7] Let q be an acyclic maximal flow on it under the constraints

$$0 \leq q_\alpha \leq r_\alpha, \quad \forall \alpha \in \mathbf{A}^i. \qquad (7.22)$$

By the *Max-Flow Min-Cut Theorem* (see, e.g., [4, Theorem 3.5.3]) $\exists$ an i-cut $\mathbf{U} \subset \mathbf{V}^i$, consisting of every vertex $v \in \mathbf{V}^i$ for which $\exists$ an undirected path π in $(\mathbf{V}^i, \mathbf{A}^i)$ from σ_i to v s.t.

- *(Forward Slack)* Every forward-oriented arc α in π (i.e., pointing from σ_i to v) has $q_\alpha < r_\alpha$, and
- *(Backward Flow)* Every backward-oriented arc α in π (pointing from v to σ_i) has $q_\alpha > 0$.

As a consequence of this,

$$q_\alpha = r_\alpha, \quad \forall \alpha \in \mathbf{O}^i := \mathrm{OUT}(\mathbf{U}) \cap \mathbf{A}^i, \qquad (7.23)$$

$$q_\alpha = 0, \quad \forall \alpha \in \mathbf{I}^i := \mathrm{IN}(\mathbf{U}) \cap \mathbf{A}^i. \qquad (7.24)$$

Note also that since the cyclic flow components $w_1, \ldots, w_j$ in (7.10) are zero,

$$q_\alpha = \sum_{1\leq k\leq p:\pi_k \ni \alpha} u_k, \quad \forall \alpha \in \mathbf{A}^i. \qquad (7.25)$$

The i-cut $\mathbf{U}$ evidently depends on the residual capacity vector r. However, the following *purely structural* statements may be made about it:

[7] Here, arcs are permitted to have $r_\alpha = 0$.

1. Every arc in $\mathbf{O}^i$ lies in $\mathbf{A}_\mathrm{f}$, i.e., is finite-capacity. Otherwise, $q_\alpha \overset{(7.23)}{=} r_\alpha \overset{(7.15)}{=} \infty$, implying by (7.25) that $u_k = \infty$ on some i-path π_k, which is impossible since every i-path in this case travels over at least one finite-capacity arc.
2. Every arc $\alpha \in \mathbf{O}^i$ is in an i-path that exits $\mathbf{U}$ without re-entering, or else α is in the $[1:i-1]$-bundle. To see this, suppose that every i-path π_k passing through α re-enters $\mathbf{U}$. Evidently, it must then pass through some arc $\beta \in \mathbf{I}^i$. By (7.24), $q_\beta = 0$, implying by virtue of (7.25) and nonnegativity that $u_k = 0$. From (7.23) and (7.25), this implies that $r_\alpha = 0$. As $c_\alpha > 0$, it must then hold that $f_{\alpha,j} > 0$ for some $j \in [1:i-1]$. As the j-flow $(f_{\alpha,j})_{\alpha\in\mathbf{A}}$ is acyclic by construction, α must then lie on a j-path, by (7.10).
3. There must be an i-path that leaves $\mathbf{U}$ without re-entering. To see this, suppose in contradiction that every i-path re-enters $\mathbf{U}$. By the preceding argument, all i-paths must then have associated acyclic flow components $u_k = 0$. Pick any i-path and let v be the last vertex in $\mathbf{U}$ that it traverses before leaving $\mathbf{U}$ without further re-entry. Let ω denote its subpath from $v \rightsquigarrow \tau_i$. By the definition of $\mathbf{U}$, there is an undirected path π from σ_i to v such that all forward-oriented arcs in it are slack and all backward-oriented arcs carry strictly positive q-flow. Note also that all vertices before v in π must also lie in $\mathbf{U}$, by construction. From (7.25), any backward arc in π would have to carry an i-path flow component $u_k > 0$, which would be a contradiction. Consequently, all the arcs in π must be forward-oriented, i.e., π is a directed path in $\mathbf{U}$ from $\sigma_i \rightsquigarrow v$. The concatenation of π with ω then yields an i-path that leaves $\mathbf{U}$ exactly once, a contradiction.
4. Finally, by construction of $\mathbf{U}$, every vertex v in it must lie on an undirected path π from σ_i to v such that
   (a) Every vertex before v in π is also in $\mathbf{U}$ (since the subpath from σ_i to v automatically satisfies the defining forward-slack and backward-flow properties), and
   (b) Every reverse-oriented arc in π lies on an i-path that does not re-enter $\mathbf{U}$ (since such arcs must by definition carry positive q-flow, and i-paths that re-enter $\mathbf{U}$ carry zero q-flow).

In other words, $\mathbf{U}$ is a *viable i-cut* (Definition 7.7). By downward dominance (Definition 7.8), $\mathbf{O}^i \cup \mathrm{OUT}(\sigma_{\mathbf{J}^{i-1}\cup[i+1:n]})$ structurally dominates $\mathrm{out}(\sigma_i)$ (Definition 7.4), and $\mathbf{O}^i$ is a $\mathcal{D}_{i-1}$-set (Definition 7.6). Using i-flow conservation,

$$
\begin{aligned}
q_{\mathrm{in}(\tau_i)} &= \sum_{\beta\in\mathbf{O}^i} q_\beta - \sum_{\alpha\in\mathbf{I}^i} q_\alpha \overset{(7.23),\,(7.24)}{=} \sum_{\beta\in\mathbf{O}^i} r_\beta \\
&\overset{(7.15)}{=} \sum_{\beta\in\mathbf{O}^i} c_\beta - \sum_{\beta\in\mathbf{O}^i,\, j\in[1:i-1]} f_{\beta,j} \\
&\overset{(7.21)}{\geq} \mathrm{I}_\infty[X_i; S_{\mathbf{O}^i}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}] + \mathrm{H}^\infty[S_{\mathbf{O}^i}|X_{\mathbf{J}^{i-1}\cup[i:n]}] - \sum_{\beta\in\mathbf{B},\, j\in[1:i-1]} f_{j,\beta} \\
&\overset{(7.13)}{\geq} \mathrm{I}_\infty[X_i; S_{\mathbf{O}^i}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}]. \qquad (7.26)
\end{aligned}
$$

As $\text{out}(\sigma_i) \in \text{SDOM}(\mathbf{O}^i \cup \text{OUT}(\sigma_{\mathbf{J}^{i-1}\cup[i+1:n]}))$, it follows that $X_i(0:km-1)$ is a function of $S_{\mathbf{O}^i}(0:km-1)$ and $X_{\mathbf{J}^{i-1}\cup[i+1:n]}(0:km-1)$. Consequently, $\forall k \in \mathbb{N}$,

$$\mathrm{I}\big[X_i(0:km-1); S_{\mathbf{O}^i}(0:km-1), X_{\mathbf{J}^{i-1}\cup[i+1:n]}(0:km-1)\big] = \mathrm{H}\big[X_i(0:km-1)\big].$$

As entropy and mutual information are monotonic, a sandwich argument with $k \to \infty$ then yields that the RHS of (7.26) is just $\mathrm{H}_\infty[X_i]$, so that

$$q_{\text{in}(\tau_i)} \geq \mathrm{I}_\infty[X_i; S_{\mathbf{O}^i}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}] = \mathrm{H}_\infty[X_i], \tag{7.27}$$

as desired.

7.5.1.3 Construction of f^i in Both Cases

For both cases above, let

$$f_{\alpha,i} := \underbrace{\frac{\mathrm{H}_\infty[X_i]}{q_{\text{in}(\tau_i)}}}_{=:v} q_\alpha \equiv v q_\alpha, \quad \forall \alpha \in \mathbf{A}, \tag{7.28}$$

where $v \in (0, 1]$ by (7.27). Clearly, $f_{\mathbf{A},i}$ is still an acyclic i-flow since it just a scaled version of q. Furthermore,

$$\sum_{j=1}^{i} f_{\alpha,j} \overset{(7.28)}{=} v q_\alpha + \sum_{j=1}^{i-1} f_{\alpha,j} \overset{(7.27)}{\leq} q_\alpha + \sum_{j=1}^{i-1} f_{\alpha,j} \overset{(7.15)}{\leq} c_\alpha.$$

The next step is to verify that $f^i = f_{\mathbf{A}\times[1:i]}$ satisfies the remaining conditions (7.13), (7.14) for an entropically feasible $[1:i]$-flow. First, (7.13) is checked. Let $\mathbf{E}$ be any arc-set in $\mathcal{D}_i$ (Definition 7.6). If $\mathbf{E} \cap \mathbf{A}^i = \emptyset$, then

$$\sum_{\eta\in\mathbf{E},\, j\in[1:i]} f_{\eta,j} = \sum_{\eta\in\mathbf{E},\, j\in[1:i-1]} f_{\eta,j} \overset{(7.13)}{\leq} \mathrm{H}^\infty[S_{\mathbf{E}}|X_{\mathbf{J}^{i-1}\cup[i:n]}] = \mathrm{H}^\infty[S_{\mathbf{E}}|X_{\mathbf{J}^{i}\cup[i+1:n]}]$$

since $\mathbf{E} \in \mathcal{D}_{i-1}$ automatically, and where the last equality follows because $\mathbf{J}^{i-1} \cup \{i\} = \mathbf{J}^i$. Else if $\mathbf{E} \cap \mathbf{A}^i \neq \emptyset$, write

$$\sum_{\eta\in\mathbf{E},\, j\in[1:i]} f_{\eta,j} = \sum_{\eta\in\mathbf{E}} f_{\eta,i} + \sum_{\eta\in\mathbf{E},\, j\in[1:i-1]} f_{\eta,j} \tag{7.29}$$

and bound each sum on the RHS as follows. First, note that since $\mathbf{E} \in \mathcal{D}_{i-1}$,

$$\sum_{\eta\in\mathbf{E},\, j\in[1:i-1]} f_{\eta,j} \overset{(7.13)}{\leq} \mathrm{H}^\infty[S_{\mathbf{E}}|X_{\mathbf{J}^{i-1}\cup[i:n]}]. \tag{7.30}$$

Then write

$$\begin{aligned}\sum_{\eta\in\mathbf{E}} f_{\eta,i} &\overset{(7.28)}{=} \sum_{\eta\in\mathbf{E}} vq_\eta \overset{(7.25),\,(7.17)}{=} v\sum_{\eta\in\mathbf{E}}\Bigg(\sum_{1\le k\le p:\pi_k\ni\eta} u_k\Bigg)\\ &= v\sum_{1\le k\le p} u_k\Bigg(\sum_{\eta\in\mathbf{E}:\eta\in\pi_k} 1\Bigg)\\ &\le v\sum_{1\le k\le p} u_k \equiv vq_{\mathrm{in}(\tau_i)} \overset{(7.28)}{=} \mathrm{H}_\infty[X_i], \end{aligned} \tag{7.31}$$

where the inequality arises because the i-path flows $u_1,\ldots,u_p\ge 0$ and each i-path π_k transits over at most one arc in $\mathbf{E}$. As $\mathrm{out}(\sigma_i)\in\mathrm{SDOM}(\mathbf{E}\cup\mathrm{OUT}(\sigma_{\mathbf{J}^{i-1}\cup[i+1:n]}))$, the same arguments that lead to the equality in (7.27) show that

$$\mathrm{H}_\infty[X_i] = \mathrm{I}_\infty[X_i; S_\mathbf{E}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}].$$

Substituting this into (7.31) and then combining with (7.29) and (7.30) yields

$$\begin{aligned}\sum_{\eta\in\mathbf{E},\,j\in[1:i]} f_{\eta,j} &\le \mathrm{I}_\infty[X_i; S_\mathbf{E}, X_{\mathbf{J}^{i-1}\cup[i+1:n]}] + \mathrm{H}^\infty[S_\mathbf{E}|X_{\mathbf{J}^{i-1}\cup[i:n]}]\\ &\overset{(7.20),\,(7.21)}{\le} \mathrm{H}^\infty[S_\mathbf{E}|X_{\mathbf{J}^{i-1}\cup[i+1:n]}] = \mathrm{H}^\infty[S_\mathbf{E}|X_{\mathbf{J}^{i}\cup[i+1:n]}], \end{aligned} \tag{7.32}$$

since $\mathbf{J}^i=\mathbf{J}^{i-1}$ in this case. This confirms that f^i satisfies (7.13). As f^{i-1} is an entropically feasible $[1:i-1]$-flow, (7.14) is satisfied $\forall j\in[1:i-1]$. Using flow conservation,

$$f_{\mathrm{out}(\sigma_i),i} = f_{\mathrm{in}(\tau_i),i} \overset{(7.28)}{=} vq_{\mathrm{in}(\tau_i)} = \mathrm{H}_\infty[X_i],$$

verifying (7.14) when $j=i$. Thus f^i is an entropically feasible $[1:i]$-flow.

By induction, f^n is an entropically feasible $[1:n]$-flow, giving the desired (X,c)-feasible multicommodity flow (7.5)–(7.7).

7.5.2 *Sufficiency of Multicommodity Flows*

The converse part of Theorem 7.2 is easier to establish, since it is not difficult to see that the existence of a feasible multicommodity flow implies achievability. Thus only the key steps are provided below.

Suppose f is an (X,c)-feasible multicommodity flow (7.5)–(7.7) on an n-pair network structure Σ, with X stationary, and further suppose that (7.5) is satisfied strictly. In the decomposition (7.10) for each i-flow $f_{\mathbf{A},i}$, no cycle flow can enter any sink, since it has no departing arcs. Consequently, the cycle flows may be taken

to be zero in (7.10) without violating (7.5)–(7.7), yielding

$$f_{\alpha,i} = \sum_{1\le k\le p_i:\pi_{k,i}\ni\alpha} u_{k,i}, \tag{7.33}$$

where $\pi_{1,i},\dots,\pi_{p_i,i}$ are the i-paths and $u_{1,i},\dots,u_{p_i,i}\ge 0$, the i-path flows. In particular,

$$\mathrm{H}_\infty[X_i] \overset{(7.6)}{=} f_{\mathrm{out}(\sigma_i)} = f_{\mathrm{in}(\tau_i),i} = \sum_{k=1}^{p_i} u_{k,i}. \tag{7.34}$$

For an arbitrary $\varepsilon > 0$, divide the time axis $\mathbb{W}$ into epochs of sufficiently long duration $m \in \mathbb{N}$ such that $\forall\, j \in \mathbb{N}, i \in [1:n]$,

$$\frac{\mathrm{H}[X_i((j-1)m : jm-1)]}{m} = \frac{\mathrm{H}[X_i(0:m-1)]}{m} \le \mathrm{H}_\infty[X_i] + \varepsilon, \tag{7.35}$$

where the first equality arises from stationarity. Next use Huffman coding [6] to losslessly encode each source-block $X_i((j-1)m : jm-1)$, $j \in \mathbb{N}$, into binary codewords $Z_{i,j}$ of variable length $L_{i,j}$, where

$$\mathrm{E}[L_{i,j}] \le \mathrm{H}\big[X_i(0:m-1)\big] + 1 \overset{(7.35)}{\le} m\mathrm{H}_\infty[X_i] + m\varepsilon + 1. \tag{7.36}$$

Then partition the bits of $Z_{i,j}$ into p consecutive sub-blocks $Z_{i,j,k}$, $k \in [1:p_i]$, of length $L_{i,j,k} := \lceil \frac{u_{k,i}}{\mathrm{H}_\infty[X_i]} L_{i,j} \rceil$. This is always possible since $\sum_{k=1}^{p_i} L_{i,j,k} \overset{(7.34)}{\ge} L_{i,j}$, padding the last sub-blocks with zeros if necessary.

Transmit and route each sub-block $Z_{i,j,k}$ along the kth i-path $\pi_{k,i}$. On every arc $\alpha \in \mathbf{A}$ apart from those leaving sources, let the arc-signal be $S_\alpha(t) = 0$ when $t \pmod m \ne m-1$ and by $S_\alpha(t) = (Z_{i,j,k})_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha}$ when $t \equiv jm-1$, $j \in \mathbb{N}$.[8] The arc signals leaving sources are set to the respective source signals to satisfy (7.2). Clearly, S is setwise causal (Definition 7.1), since every arc-signal is constructed by routing blocks along acyclic paths. In addition, $\forall \alpha \in \mathbf{A}_{\mathrm{f}}$,

$$\begin{aligned}
\mathrm{H}^\infty[S_\alpha] &= \frac{1}{m}\mathrm{H}\big[(Z_{i,j,k})_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha}\big] \\
&\le \sum_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha} \frac{\mathrm{H}[Z_{i,j,k}]}{m} \qquad (7.37)\\
&\le \sum_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha} \frac{\mathrm{E}[L_{i,j,k}]}{m} \qquad (7.38)\\
&\le \sum_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha} \frac{1}{m} + \frac{u_{k,i}}{m\mathrm{H}_\infty[X_i]}\mathrm{E}[L_{i,j}]
\end{aligned}$$

[8] If an arc is not on any i-path, then its arc signal may be taken to be 0.

$$\begin{aligned}
&\leq \mathrm{O}(1/m) + \sum_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha} u_{k,i}\frac{m\mathrm{H}_\infty[X_i]+m\varepsilon+1}{m\mathrm{H}_\infty[X_i]} \\
&= \sum_{i\in[1:n],k\in[1:p_i]:\pi_{k,i}\ni\alpha} u_{k,i} + \mathrm{O}(\varepsilon) + \mathrm{O}(1/m) \\
&\overset{(7.33)}{=} \sum_{i\in[1:n]} f_{\alpha,i} + \mathrm{O}(\varepsilon) + \mathrm{O}(1/m) \\
&= f_\alpha + \mathrm{O}(\varepsilon) + \mathrm{O}(1/m) \leq c_\alpha
\end{aligned}$$

for ε sufficiently small and m sufficiently large. In the above, the bound (7.37) is due to the subadditivity of entropy, and (7.38) is due to the fact that the expected number of bits needed to uniquely specify the value of a random variable is never less than its entropy. Furthermore,

$$Y_i(jm-1) = (Z_{i,j,k})_{k=1}^{p_i} \equiv Z_{i,j} \equiv X_i\big((j-1)m : jm-1\big).$$

Consequently, S is a solution to the n-pairs information network problem (Σ, X, c), establishing achievability (Definition 7.2).

7.6 Examples

In this section, several examples are given to illustrate the applicability of Theorem 7.2. However, to begin with a well-known counterexample is discussed.

To avoid cluttering the figures in this section, the infinite-capacity arcs leading out of sources and into sinks are not explicitly depicted.

7.6.1 Butterfly Network

The first example, a 2-pairs butterfly network, is adapted from [11, 16] and depicted in Fig. 7.1. For this network, it is well-known that routing does not achieve linear coding capacity, and it is a useful exercise to verify that it is not downward dominated.

Consider the viable 2-cut having the set $\mathbf{O}^2 = \{\alpha\}$ of outgoing arcs in the 2-bundle. Clearly, both β, γ are downstream of $\mathbf{O}^2 = \mathbf{O}^2_{*2}$, so

$$\mathbf{C} = \{\alpha, \beta, \gamma\} \subseteq \mathrm{SDOM}\big(\mathbf{O}^2\big).$$

No other arcs are downstream of $\mathbf{C}$. Furthermore, the indirect 2-walk concisely represented by $(\varphi, \delta, \varepsilon)$ does not pass through α, and neither does the indirect 1-walk (δ, φ, χ). Thus $\mathbf{C}$ is the smallest set satisfying all the conditions of Definition 7.4), i.e., $\mathbf{C} = \mathrm{SDOM}(\mathbf{O}^2)$. As $\mathbf{C}$ does not include any source or sink arcs, this network is not downward dominated (Definition 7.8) and Theorem 7.2 does not apply.

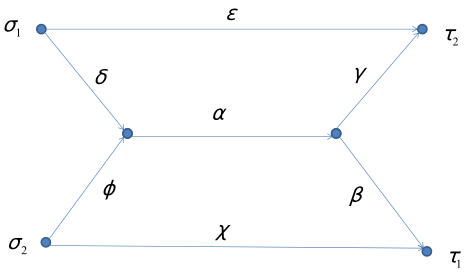


Fig. 7.1 Butterfly network

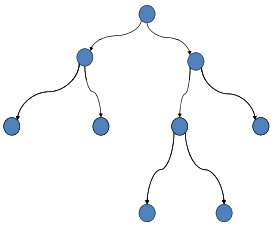


Fig. 7.2 A directed tree. Sources and sinks may be attached to any of the nodes depicted

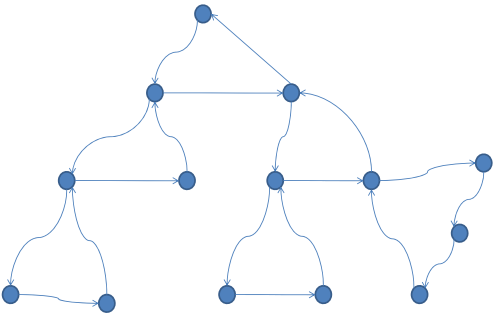


Fig. 7.3 A tree of directed cycles. Sources and sinks may be attached to any of the nodes depicted

7.6.2 Examples that Satisfy Lemma 7.4

Any network where there is at most one (directed) path from any vertex to any other automatically satisfies the conditions of Lemma 7.4, and is therefore downward dominant and structurally routable (Theorem 7.2). This includes in the first instance both directed lines and directed cycles, agreeing with results in [10, 11]. It also covers more complicated structures, for instance, directed trees (Fig. 7.2), and directed cycles arranged in a line or tree structure via one or more gateway nodes (Fig. 7.3). In all these networks, routing achieves coding capacity regardless of where sources and sinks are placed.

In networks where there are vertex pairs with two or more connecting paths, downward dominance will still hold by virtue of Lemma 7.4 if there is at most one path between each pair of source and sink vertices, or at least from each σ_s to each τ_h, where $1 \leq s \leq h \leq n$. Examples include directed versions of the undirected Okamura–Seymour network (Fig. 7.4).

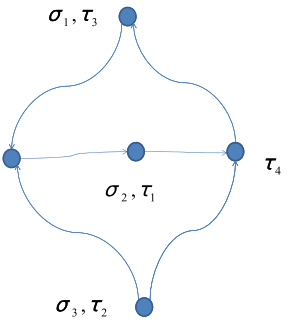


Fig. 7.4 A directed version of the Okamura–Seymour Network. Only one path exists from any source to any sink

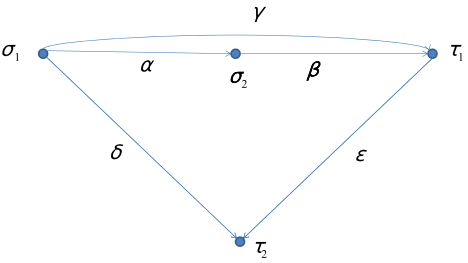


Fig. 7.5 An acyclic network covered by Lemma 7.3

7.6.3 Examples that Satisfy Lemma 7.3

Now consider the acyclic 2-pairs network in Fig. 7.5. Observe that there is one 1-path, concisely represented by the arc-sequence $\beta\varepsilon$, but two 2-paths, $\alpha\beta$ and γ. Hence Lemma 7.4 cannot be applied. Neither would it become applicable if the indices 1 and 2 were relabeled $2'$ and $1'$, respectively. To see this, consider the viable $2'$-cut with $\mathbf{O}^{2'} = \{\beta\}$. Clearly, $\mathbf{A}^{2'} \cap \mathbf{O}^{2'}$ and $\mathbf{A}^{1'} \cap \mathbf{O}^{2'} \neq \emptyset$, since the $1'$-path $\alpha\beta$ and $2'$-path $\beta\varepsilon$ both pass through $\mathbf{O}^{2'}$. However, the path γ from σ_1' to τ_1' does not.

In this instance, Lemma 7.3 can be applied. The possible viable 2-cuts have sets $\mathbf{O}^2$ of outgoing arcs in the 2-bundle equal to either $\{\alpha, \gamma\}$ or $\{\beta, \gamma\}$. In the first case, $\mathbf{O}^2$ has no intersection with any arcs in the 1-bundle, and all 2-paths obviously pass through it. In the second case, all paths from σ_s to σ_h, $1 \leq s \leq h \leq 2$, pass through $\mathbf{O}^2$. This the requirements of the lemma are met and the network is downward dominant.

Another example of the use of Lemma 7.3 is the cyclic 2-pairs network of Fig. 7.6. Observe that there is one 1-path, $\varepsilon\beta$ and two 2-paths, φ and $\beta\gamma$. The possible viable 2-cuts have sets $\mathbf{O}^2$ of outgoing arcs in the 2-bundle equal to either $\{\varphi, \beta\}$ or $\{\varphi, \gamma\}$. In the second case, $\mathbf{O}^2$ has no intersection with any arc in the 1-bundle, and all 2-paths obviously pass through it. In the first case, $\mathbf{O}^2$ intersects all 2-paths and a 1-path, $\varepsilon\beta$, and it can be seen that all from σ_s to σ_h, $1 \leq s \leq h \leq 2$, pass through $\mathbf{O}^2$. This the requirements of the lemma are met and the network is downward dominant.

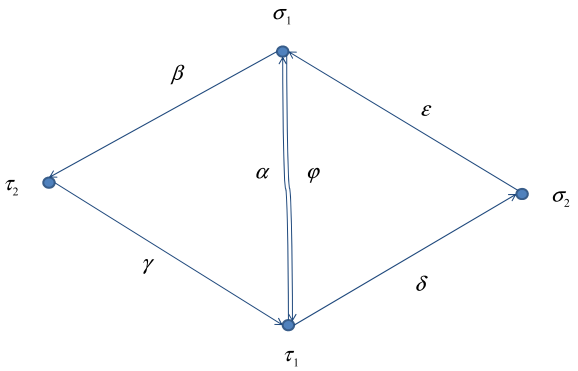


Fig. 7.6 A cyclic network covered by Lemma 7.3

7.7 Conclusion

This chapter examined the routability of possibly cyclic n-pairs information networks from a structural perspective. The concepts of downward dominance was introduced, and it was shown that for downward dominated networks, achievability always implies the existence of a feasible multicommodity flow.

Downward dominance is a conservative structural condition, and future work will focus on trying to relax it. One refinement is to partition the set of source–sinks paths into subsets having no arcs in common; it is anticipated that this approach will allow bidirectional cycles [15] and other networks to be handled.

From [2, 16], it is known that routing does not achieve linear network coding capacity, and in more recent work [7], it has been shown that linear coding does not generally achieve network coding capacity. Thus, two other important lines of research are to investigate whether the inductive approach of this chapter can be adapted to find structural conditions under which routing achieves linear network coding capacity, and under which linear coding achieves network coding capacity. It is anticipated that techniques from linear systems and control theory would be needed, especially if loops are present.

Acknowledgements The author is indebted to the reviewers for numerous helpful suggestions. He also acknowledges discussions on the decentralized control version of this problem with Prof. Rob Evans at the University of Melbourne.

This work was partially supported by Australian Research Council grant DP110102401.

This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Adler, M., Harvey, N.J.A., Jain, K., Kleinberg, R., Lehman, A.R.: On the capacity of information networks. In: Proc. 17th Annual ACM-SIAM Symp. Discrete Algorithms, Miami, USA, pp. 241–250 (2006)
2. Ahlswede, R., Cai, N., Li, S.W.R., Yeung, R.W.: Network information flow. IEEE Trans. Inf. Theory **46**(4), 1204–1216 (2000)

3. Antsaklis, P., Baillieul, J. (eds.): Special Issue on the Technology of Networked Control Systems. Proc. IEEE, vol. 95. IEEE Press, New York (2007)
4. Bang-Jensen, J., Gutin, G.: Digraphs Theory, Algorithms and Applications. Springer, Berlin (2007)
5. Cannons, J., Dougherty, R., Freiling, C., Zeger, K.: Network routing capacity. IEEE Trans. Inf. Theory **52**(3), 777–788 (2006)
6. Cover, T.M., Thomas, J.A.: Elements of Information Theory. Wiley, New York (1991)
7. Dougherty, R., Freiling, C., Zeger, K.: Insufficiency of linear coding in network information flow. IEEE Trans. Inf. Theory **51**(8), 2745–2759 (2005)
8. Fragouli, C., Soljanin, E.: Network Coding Applications. Now Publishers, Hanover (2008)
9. Han, T.S.: Multicasting multiple correlated sources to multiple sinks over a noisy channel network. IEEE Trans. Inf. Theory **57**(1), 4–13 (2011)
10. Harvey, N.J.A., Kleinberg, R., Lehman, A.R.: On the capacity of information networks. IEEE Trans. Inf. Theory **52**(6), 2345–2364 (2006)
11. Kramer, G., Savari, S.A.: Edge-cut bounds on network coding rates. J. Netw. Syst. Manag. **14**(1), 49–67 (2006)
12. Leighton, T., Rao, S.: Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. J. ACM **46**(6), 787–832 (1999)
13. Li, S.Y.R., Yeung, R.W., Cai, N.: Linear network coding. IEEE Trans. Inf. Theory **49**(2), 371–381 (2003)
14. Nair, G.N.: When is n-pairs information a multicommodity flow? In: IEEE Int. Symp. Info. Theory Proc., St. Petersburg, Russia, pp. 169–173 (2011)
15. Savari, S.A., Kramer, G.: The multimessage unicast capacity region for bidirectional ring networks. In: Proc. Int. Symp. Information Theory, pp. 763–767 (2006)
16. Yeung, R.W.: Multilevel diversity coding with distortion. IEEE Trans. Inf. Theory **41**(2), 412–422 (1995)
17. Zhou, K., Doyle, J., Glover, K.: Robust and Optimal Control. Prentice Hall, New York (1996)

Chapter 8
Computing over Unreliable Communication Networks

Nicola Elia, Jing Wang, and Andalam Satya Mohan Vamsi

8.1 Introduction

In this chapter, we present analysis and design tools for networked systems. Networked systems often need to perform cooperatively control, estimation, optimization, or computation with limited computational, power, and communication resources. The distributed nature of the system and the limited capability of the individual subsystems (nodes or agents) force them to cooperate by exchanging information on the network, through alternative physical sensing and through their interaction with the environment. The information exchange is limited by the structure of the network, the nature of the interactions, and by various sources of uncertainty (e.g., noisy channels, packet drops, delays), but also sensor and actuators limitations and environment unpredicted changes.

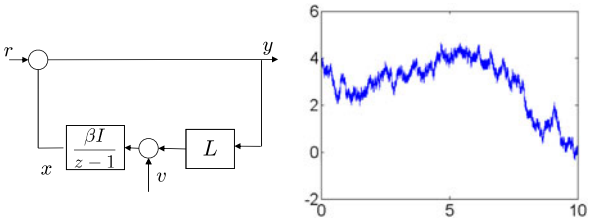
As a starting point, let us consider the problem of computing the average of the private quantities of n agents/nodes in a distributed fashion. The agents, connected over ideal communication network, exchange information with their nearest neighbors. The network is represented by its Laplacian, L, and is assumed to be strongly connected and undirected, but not necessarily complete. This is a well known and studied problem [14, 19, 38, 49, 50]. An interesting solution would have each agent exchange relative errors with their neighbors, as described by the following algo-

N. Elia (✉) · A.S.M. Vamsi
Dept. of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, USA
e-mail: nelia@iastate.edu

A.S.M. Vamsi
e-mail: vamsi68@iastate.edu

J. Wang
Cummins Inc., Columbus, IN, USA
e-mail: jingwang@alumni.iastate.edu

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_8,



(a) (b)

Fig. 8.1 (**a**) Distributed averaging system of [47]; (**b**) Its fragility to noise

rithm in terms of dynamic equations:

$$x_i(k+1) = x_i(k) + \beta \sum_{j \in N_i} \big(y_j(k) - y_i(k)\big),$$

$$y_i(k) = r_i - x_i(k).$$

The agent private quantity is r_i and N_i is the set of neighbors of node i. $\beta > 0$ is a small enough update gain. This version of the algorithm is due to [47]. Figure 8.1(a) shows the block-diagram of the computational system. Under some mild conditions, all agent's output variables y_i converge to the average of r, i.e., $\frac{1}{n}\sum_{i=1}^{n} r_i$. It is interesting to go over the implicit assumptions behind the convergence of the algorithm. (i) The communication network is ideal, each link is perfect, namely it can communicate perfectly any number of bits between time epochs. (ii) Each node can compute sums with arbitrarily high accuracy within time epochs. (iii) r_i does not change with time and is known with arbitrarily high accuracy. These idealized assumptions remove noise and physical time from the problem. However, there are some lingering issues related to the presence of uncertainty not accounted for in the idealized model.

What happens to the algorithm if r_i is measured and is therefore uncertain with some limits? Does it still make sense to use virtually infinite communication bandwidth and computation precision for computing an uncertain quantity? Similarly, when r_i is re-measured at every epoch, the uncertainty in the measure will make r_i different over time. What happens if noise enters the system? Like when the communication links are not ideal and transmit messages with a little additive noise? It is not difficult to show that the simple algorithm is fragile to additive noise, and in this case, it completely loses its ability to compute averages even with infinite computational accuracy and speed, as shown in Fig. 8.1(b).

Finally, assume the agents' state and output correspond to desired physical locations (for example, in a formation control problem) where the computational abstraction no longer applies, and we would need a way to guarantee that the mobile agents can actually move to the desired locations, with arbitrary accuracy, within epochs.

This example embodies some critical issues of a large class of systems where *the agents need to compute timely and approximately correct decisions to pursue a collective/private goal, based on partial and unreliable information about the environment and the other agents*.

Motivated by the above questions, in this chapter we merge the dynamical system interpretation of distributed computational systems with the integrated theory of communication and control and provide a unified view of computational systems over unreliable communication networks.

While a great deal of fundamental understanding has been drawn from characterizing the minimum transmission rate necessary for stabilization for different channel models [3, 34, 48, 63], most of the research has focused on stabilization of single loop systems closed over various kind of unreliable or uncertain communication links. In these settings, the feedback system becomes a stochastic system and the main finding is that different notions of stochastic stability require different notions of reliable delivery of information through the loop. In particular, the bounded moment stability requires the anytime notion of reliable communication, which is stronger than Shannon's notion, but weaker than the zero-error in [43, 44]. Sahai [43] provides the explicit characterization of the anytime capacity limitations for a single state unstable system with a binary erasure channel. More often, the anytime limitations have been indirectly captured in terms of critical channel dropout probability for bounded second moment stability in a variety of settings mostly involving a single channel on the sensor side. Limitations and coding for finite-rate packet-drop channel [30, 33] in the presence of bounded and unbounded stochastic disturbance. Analog intermittent channels are studied in [12, 13, 16, 17, 23, 24, 45, 46]. Recently, [10] has provided the explicit link between anytime capacity and dropout probability limitations for several channels models used in the literature.

The results points out that channel fading or intermittency rather than quantization is a greater limiting factor in moment stabilization. This allows us to focus on analog intermittent communications/interactions among networked dynamical systems and to neglect quantization and coding aspects. In this chapter, we provide a common framework that efficiently allows considering the mean square stability limitations of networked systems with multiple intermittent channels and additive noise. We focus on how unreliable channels can be efficiently used in real-time networked systems and how they limit the networked system performance and behavior. The framework allows studying complex interconnections with multiple loops and channels that are typical of complex systems organizations. We consider networked systems mostly made of simple agents with limited computational and communication capabilities, and study how such systems can compute and communicate at a physical level without sophisticated coding or computational intelligence, and how the intermittent unreliable interactions affects the overall networked system behavior and capabilities.

The chapter summarizes our recent results in a coherent way and a tutorial style, and it is organized as follows: after some basic graph theory and some basic notation in Sect. 8.2, in Sect. 8.3 we introduce a large class of networked systems. We first describe a class of systems over noiseless and delay free network. This is the basis for studying limitation of networked systems only due to the network topology. We then extend the models to the case where the networks are unreliable. In Sect. 8.4, we summarize the main stability analysis result for systems over fading networks. An application of the mean square stability analysis to a classical distributed averaging system when the agents' interactions happen over intermittent

noisy networks is presented in Sect. 8.5. We point out that the otherwise benign and well behaved networked system can exhibit complex behavior for certain parameter values when the communication is unreliable. We then present algorithms that are resilient to additive noise and channel dropouts in Sect. 8.6, which are designed using multi-variable feedback system properties. The systems ideas used in developing the new distributed averaging algorithms can be extended to develop dynamical systems solving distributed convex optimization problems. We introduce optimization systems in Sect. 8.7 and show how simple agents can cooperatively solve complex optimization problems using simple gradient sensing capabilities and by communicating over the network. The dynamical system perspective leads to many new research directions. In particular, least squares optimization can be solved by LTI systems. Thus, we can use the large set of analysis and design tools to develop improved and ad-hoc versions of such systems. In particular, we can design networked controllers to improve stability and performance of a network distributed least squares system. In order to do so, we rely on a newly developed networked controller design method, which provides controller implementable over the given networks. We review such methodology in Sect. 8.8, and apply it to a simple distributed least squares system. Finally, in Sect. 8.9, we provide some concluding remarks on the proposed unified view of networked computational systems put forward in this chapter.

8.2 Notation

8.2.1 Elements of Graph Theory

A *directed* graph is represented as $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, n\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. An edge of $\mathcal{G}$ is denoted by (i, j) implying there exists a directed link from node j to node i.

Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, define the adjacency matrix $\mathcal{A}(\mathcal{G})$ to be a binary matrix such that

$$\left[\mathcal{A}(\mathcal{G})\right]_{ij} = 1 \quad \text{if } (j,i) \in \mathcal{E}, \qquad = 0 \quad \text{otherwise.} \tag{8.1}$$

Define directed neighborhoods around each node i, the in-neighbors $\mathcal{N}_i^- = \{j | (j,i) \in \mathcal{E}\}$ and the out-neighbors $\mathcal{N}_i^+ = \{j | (i,j) \in \mathcal{E}\}$, which are the sets of nodes that have edges to and from node i. The in-degree of each node is defined as $d_{\text{in}}^i = |\mathcal{N}_i^-|$ and the out-degree of each node is defined as $d_{\text{out}}^i = |\mathcal{N}_i^+|$. We denote the graph Laplacian matrix of $\mathcal{G}$ as $\mathcal{L} = -\mathcal{D} + \mathcal{A}(\mathcal{G})$, where $\mathcal{D} = \text{Diag}[d_{\text{in}}^i]$. Thus $\mathcal{L}\mathbf{1} = 0$, where $\mathbf{1}$ is the vector consisting of all ones. The Laplacian is said to be *balanced* if $d_{\text{in}}^i = d_{\text{out}}^i$ for all $i \in \mathcal{V}$. Therefore, the balanced Laplacian has the property that $\mathbf{1}'\mathcal{L} = 0$. The incidence matrix $\mathcal{I}n$ of a directed graph $\mathcal{G}$ is an $n \times q$ matrix where n and q are the number of vertices and edges, respectively, such that $\mathcal{I}n_{ik} = 1$ if the edge $k = (*, i)$ leaves vertex i, $\mathcal{I}n_{ik} = -1$ if edge $k = (i, *)$ it enters vertex

i and 0 otherwise. A graph is undirected if for any $(i, j) \in \mathcal{E}$, the edge $(j, i) \in \mathcal{E}$. The graph is said to be *strongly connected* if every node can reach every other node through a sequence of elements in $\mathcal{E}$. If the graph is strongly connected, $\mathcal{L}$ has only one eigenvalue at zero and all other eigenvalues are in open right half plane, and the left eigenvector associated with the zero eigenvalue has its components all positive.

8.2.2 Vectors and Partitions

To make representations compact, we use the notation $\mathbf{vec}[x_i]_{i\in\mathcal{I}}$ for vertical arrangement of vectors $\{x_i\}_{i\in\mathcal{I}}$, of appropriate dimension, where $\mathcal{I}$ is an index set. Let $[x_{ij}]_{i,j\in\mathcal{I}}$ represent a matrix formed by arranging the sub-matrices $\{x_{ij}\}_{i,j}$ according to $\mathcal{I}$. Also, let $\mathrm{diag}[x_i]_{i\in\mathcal{I}}$ denote the matrix formed by arranging the vectors or matrices $\{x_i\}_{i\in\mathcal{I}}$ in a block diagonal fashion and the remaining entries being zeros. Sometimes, if the index set $\mathcal{I}$ equals $\{1, \ldots, n\}$, then we will not explicitly mention the index set.

8.2.3 Structured Systems

In this section, we define finite Dimensional Linear Time Invariant Discrete-Time (FDLTI-DT) Systems that have structures in their state-space or input–output representations consistent with the sparsity structure induced by a graph defined earlier. These structures are amenable to efficient searches and capture the representations of systems over networks. These representations become relevant when we want to design networked systems, and networked controllers in particular, as we will see in later sections.

Definition 8.1 We say a block matrix $A = [A_{ij}]_{i,j\in\{1,\ldots,n\}}$ is *structured according to* an $n \times n$ binary matrix J if the sub-matrix A_{ij} is a zero matrix whenever $J_{ij} = 0$. The dimensions of the sub-matrices $\{A_{ij}\}_{i,j}$ are described using two integer-valued vectors as follows. Let $\mathcal{P}_a = (a_1, \ldots, a_n)$ and $\mathcal{P}_b = (b_1, \ldots, b_n)$ be two n-tuples with a_i and b_i being integers for all $i \in \{1, \ldots, n\}$. Then, matrix A is said to be *partitioned according to* $(\mathcal{P}_a, \mathcal{P}_b)$ if the sub-matrix A_{ij} has dimensions $a_i \times b_j$ $\forall i, j$.

This definition of partitioning easily extends to the case of vectors. A vector x is said to be *partitioned according to* $\mathcal{P}_a$ if it can be written as $\mathbf{vec}[x_i]_{i\in\{1,\ldots,n\}}$ where x_i is a real vector of size a_i for all $i \in \{1, \ldots, n\}$.

Definition 8.2 Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and the partitions $\mathcal{P}_x$, $\mathcal{P}_u$ and $\mathcal{P}_y$, let $\mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$ denote the set of state-space realizations (A, B_u, C_y, D_{yu}) where A, C_y are structured according to the adjacency matrix $\mathcal{A}(\mathcal{G}) + I$, while B_u, D_{yu} are block-diagonal and the state-space matrices are partitioned as follows: A $(\mathcal{P}_x, \mathcal{P}_x)$, B_u $(\mathcal{P}_x, \mathcal{P}_u)$, C_y $(\mathcal{P}_y, \mathcal{P}_x)$, and D_{yu} $(\mathcal{P}_y, \mathcal{P}_u)$.

Let $\mathcal{R}_p$ denote the set of real-rational proper transfer function matrices, $\mathcal{R}_{sp}$ denote the set of real-rational strictly proper transfer function matrices and $\mathcal{RH}_\infty$ denote the set of real-rational proper stable transfer function matrices.

Definition 8.3 Given a graph $\mathcal{G}$ and the input and output partitions, $\mathcal{P}_u$ and $\mathcal{P}_y$, let $\mathfrak{T}(\mathcal{G}, \mathcal{P}_u, \mathcal{P}_y)$ denote the set of transfer function matrices $P(z) = [P_{ij}(z)]_{i,j}$ that are partitioned according to $(\mathcal{P}_y, \mathcal{P}_u)$ where $P_{ij}(z)$ is of the form:

$$P_{ij}(z) = \begin{cases} H_{ii}(z) & \text{if } i = j, \\ z^{-l(j,i)} H_{ij}(z) & \text{if } l(j,i) \geq 1, \\ 0 & \text{otherwise,} \end{cases} \tag{8.2}$$

where $l(j,i)$ is the length of a shortest path from node j to node i on the graph $\mathcal{G}$ and $H_{ij}(z) \in \mathcal{R}_p$ for all i, j.

We want to point out that, if a system P has a state-space representation in $(A, B_u, C_y, D_{yu}) \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$, then its transfer function $P(z) \in \mathfrak{T}(\mathcal{G}, \mathcal{P}_u, \mathcal{P}_y)$ [51].

8.3 Networked Systems

In this section, we describe systems built over communication networks and discuss some of the properties of their state-space and input–output descriptions.

Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n nodes, we associate a sub-system $\{P_i\}_{i \in \{1,\dots,n\}}$ to each node.

Each sub-system P_i is a DT system with local inputs $u_i(k)$, local outputs $y_i(k)$, network inputs $v_i(k)$ and network outputs $\eta_i(k)$. Each P_i has the following state-space description (assumed to be minimal)

$$P_i: \quad \begin{bmatrix} x_i(k+1) \\ y_i(k) \\ \eta_i(k) \end{bmatrix} = \begin{bmatrix} A_i & B_i^u & B_i^v \\ C_i^y & D_i^{yu} & D_i^{yv} \\ C_i^\eta & 0 & 0 \end{bmatrix} \begin{bmatrix} x_i(k) \\ u_i(k) \\ v_i(k) \end{bmatrix}, \tag{8.3}$$

where $x_i(k)$ is the local state. The subsystems interact over communication links corresponding to the edges of the graph. The vectors $\eta_i(k)$ and $v_i(k)$ are the stacking of the corresponding vectors indexed according to the out-neighbors and in-neighbors of node i, namely, $\eta_i(k) = \mathbf{vec}[\eta_{ij}(k)]_{j \in \mathcal{N}_i^+}$ and $v_i(k) = \mathbf{vec}[v_{ij}(k)]_{j \in \mathcal{N}_i^-}$ $\forall i$. They correspond to the overall set of messages transmitted and received by P_i, where $\eta_{ij}(k)$ is the message vector transmitted from plant P_i to P_j at the time instant k and $v_{ij}(k)$ is the message received by P_i from P_j at time instant k.

For convenience and without loss of generality, we often allow each node to explicitly send and receive messages from itself. In these cases, η_i and v_i include self-messages.

Define $P = \text{diag}[P_i]_i$ as a system with a state-space representation given by

$$\begin{bmatrix} x(k+1) \\ y(k) \\ \eta(k) \end{bmatrix} = \begin{bmatrix} \hat{A} & \hat{B}_u & \hat{B}_v \\ \hat{C}_y & \hat{D}_{yu} & \hat{D}_{yv} \\ \hat{C}_\eta & 0 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \\ v(k) \end{bmatrix}, \tag{8.4}$$

where all the matrices are block diagonal consistent with the dimension of the corresponding state-space matrices of $P_i's$. The state, input and output vectors are given by $x(k) = \mathbf{vec}[x_i(k)]_i$, $u(k) = \mathbf{vec}[u_i(k)]_i$, $y(k) = \mathbf{vec}[y_i(k)]_i$, $\eta(k) = \mathbf{vec}[\eta_i(k)]_i$, and $v(k) = \mathbf{vec}[v_i(k)]_i$. Let the corresponding partitions of $x(k)$, $u(k)$, $y(k)$, $\eta(k)$, $v(k)$ be $\mathcal{P}_x$, $\mathcal{P}_u$, $\mathcal{P}_y$, $\mathcal{P}_\eta$, and $\mathcal{P}_v$, respectively. Based on these partitions, we can see that the matrices in (8.4) are partitioned accordingly.

8.3.1 Systems over Noiseless Delay-Free Networks

The case of noiseless delay-free network is important to study limitations of networked systems induced solely by the network topology.

Definition 8.4 A noiseless and delay-free network, N, is a static relation between η and v,

$$v = N\eta, \tag{8.5}$$

partitioned consistently with P and structured according to $\mathcal{A}(\mathcal{G}) + I$.

From the above definition, a networked system G is obtained by the feedback interconnection of P with N, by substituting Eq. (8.5) in Eq. (8.4), or equivalently, by $G = \text{F}_l(P, N)$, where F_l stands for lower Linear Fractional Transformation (LFT)[1] as defined in [65].

$$\begin{aligned} G &= \text{F}_l(P, N) \\ &= \left[\begin{array}{c|c} \hat{A} + \hat{B}_v N \hat{C}_\eta & \hat{B}_u \\ \hline \hat{C}_y + \hat{D}_{yv} N \hat{C}_\eta & \hat{D}_{yu} \end{array}\right] \\ &:= \left[\begin{array}{c|c} A & B_u \\ \hline C_y & D_{yu} \end{array}\right]. \end{aligned} \tag{8.6}$$

Note that A, C_y are structured according to $\mathcal{A}(\mathcal{G}) + I$, while B_u, D_{yu} have a block diagonal structure. Thus, $G \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$.

Figure 8.2 describes the basic block diagram of such networked systems.

[1] Note that having $D_i^{\eta u} = 0$, and $D_i^{\eta v} = 0 \, \forall i$ assures that the feedback interconnection of P and N is well-posed.

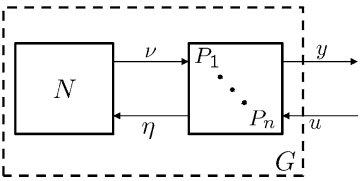


Fig. 8.2 The networked system over a noiseless and delay-free network

8.3.2 *Fading Networks*

More generally, the network interconnections are unreliable, subject to noise and intermittency or fading. To model systems over such networks, we introduce the following definitions adapted from [12]. The idea is to describe the random variables, which are parts of the link model as the source of uncertainty in an otherwise purely deterministic model, the Mean Network. In the next definition, we abuse the notation and define the Mean Network by N as it is a generalization of N in (8.5).

Definition 8.5 An analog Fading Network is composed of two parts:

1. The Mean Network, N.
2. The stochastic perturbation, Δ.

The Mean Network is, for this chapter,[2] a deterministic static system described by the following static map:

$$N: \quad \begin{bmatrix} \nu \\ z_\Delta \end{bmatrix} = \left[\begin{array}{c|c|c} L_{\nu\eta} & L_{\nu\upsilon} & L_{\nu w_\Delta} \\ \hline L_{z_\Delta\eta} & L_{z_\Delta\upsilon} & L_{z_\Delta w_\Delta} \end{array}\right] \begin{bmatrix} \eta \\ \upsilon \\ w_\Delta \end{bmatrix}, \tag{8.7}$$

where η and ν are partitioned according to P and are the Network input vector coming from P and output vector going to P, υ is a vector representing additive noise in the links, $w_\Delta \in \mathbb{R}^p$ and $z_\Delta \in \mathbb{R}^p$. The stochastic perturbation Δ maps $z_\Delta \to w_\Delta$ and is defined as $\Delta = \text{diag}[\Delta_i]_{i=1:p}$. For each $i = 1, \ldots, p$, $\Delta_i(0), \Delta_i(1), \ldots, \Delta_i(k), \ldots$ are IID random variables with

$$\mathbf{E}\{\Delta_i(k)\} = 0, \quad \text{and} \quad \mathbf{E}\{(\Delta_i(k))^2\} = \sigma_i^2 \quad \forall k \geq 0.$$

Moreover, $\Delta_1(k), \ldots, \Delta_p(k)$ are independent for each k, although not necessarily identically distributed. Δ acts as multiplication operator on z to provide w, i.e., $w_{\Delta,i}(k) = \Delta_i(k) z_{\Delta,i}(k)$ for $i = 1, \ldots, p$, $\forall k \geq 0$. Finally, υ is a vector of white noise signals independent from each other and independent from Δ.

Several fading channel models can be represented in the framework. In particular, a simple model of a packet-drop link can be represented as a Bernoulli switch $\xi(k)$ (neglecting quantization effects). $\xi(k)$ can in turn be re-parameterized as $\mu + \Delta(k)$

[2] More general dynamical models can be studied [12].

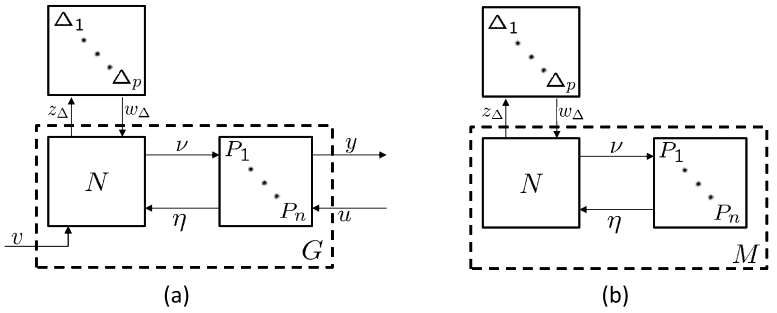


Fig. 8.3 (**a**) Networked system over a Fading Network; (**b**) Subsystem relevant to MS stability analysis

where μ is the mean of ξ and $\Delta(k)$ is a zero mean random variable with variance equal to the variance of $\xi(k)$.

When there is no uncertainty and additive noise, then the fading network model reduces to the noiseless and delay-free network of the previous section, with $N = L_{\nu\eta}$ partitioned consistently with P and structured according to $\mathcal{A}(\mathcal{G}) + I$.

An example of fading network is presented in Sect. 8.5.

8.4 Systems over Fading Networks

The interconnection of the plant and the Fading Network can be put into the robust control framework described in Fig. 8.3(a). In particular, let

$$G = \mathrm{F}_l(P, N) : \begin{bmatrix} y \\ z_\Delta \end{bmatrix} = \begin{bmatrix} G_{yu} & G_{y\upsilon} & G_{yw_\Delta} \\ G_{z_\Delta u} & G_{z_\Delta \upsilon} & G_{z_\Delta w_\Delta} \end{bmatrix} \begin{bmatrix} u \\ \upsilon \\ w_\Delta \end{bmatrix}.$$

Note that G_{yu} corresponds to G in (8.6) when the network is noiseless and delay-free. Note also that networked system is given by $\mathrm{F}_l(G, \Delta)$. To study the MS stability of the networked system we can assume that the inputs u and υ are zero as they enter linearly in the equations [12], and concentrate on the MS stability of

$$H = \mathrm{F}_l(G_{z_\Delta w_\Delta}, \Delta).$$

Let $M = G_{z_\Delta w_\Delta}$ be the system from w_Δ and to z_Δ, which is in feedback with random uncertainty Δ. Let the state space equations of M be the following.

$$M: \quad \begin{aligned} \chi^+ &= \mathcal{A}\chi + \mathcal{B}w_\Delta, \\ z_\Delta &= \mathcal{C}\chi + \mathcal{D}w_\Delta. \end{aligned} \tag{8.8}$$

Following the setup described in Fig. 8.3(b), the linear time-invariant discrete-time system M has p-inputs and p-outputs and is in feedback with the diagonal uncertainty Δ described in Definition 8.5.

8.4.1 Mean Square Closed-Loop Stability

The state-space representation of H is given by

$$H: \quad \begin{aligned} \chi^+ &= \mathcal{A}\chi + \mathcal{B}w_\Delta, \\ z_\Delta &= \mathcal{C}\chi + \mathcal{D}w_\Delta, \\ w_\Delta &= \Delta z_\Delta. \end{aligned} \tag{8.9}$$

Assume that $\chi_0 = \chi(0)$ is independent of $\Delta(k)$ for all k and that the feedback interconnection of Δ and M is well-posed, namely that the solution to system (8.9) exists for any realization of Δ.

Definition 8.6 System (8.9) is Mean Square stable if $\lim_{k\to\infty} \mathbf{E}\{\chi(k)\chi(k)'\} = 0$.

Theorem 8.1 [12] *Assume that $M = (\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D})$ is stable and that $\mathcal{D}$ is either strictly lower triangular or strictly upper triangular. Let $\Sigma^2 = \mathrm{diag}[\sigma_i^2]_{i=1:p}$. The feedback interconnection of M and Δ is Mean Square Stable iff $\rho(\Sigma^2\hat{M}) < 1$ where $\rho(\cdot)$ denotes the spectral radius and*

$$\hat{M} = \begin{bmatrix} \|M_{11}\|_2^2 & \dots & \|M_{1p}\|_2^2 \\ \vdots & \ddots & \vdots \\ \|M_{p1}\|_2^2 & \dots & \|M_{pp}\|_2^2 \end{bmatrix}.$$

See [12] for a more extended version of the theorem and for references to related work. Although rooted into the rich literature on stochastic systems [1, 8, 9, 22, 32], the approach provides new insights on the role of the channel generated uncertainty in affecting the stability of the closed loop.

8.5 Distributed Averaging over Unreliable Channels

The fading network framework applies to the MS stability analysis of a simple but classical distributed averaging algorithm when the communication links are unreliable. In [55, 59], we have shown that the classical average consensus algorithm is a prototype system for the emergence of Levy Flights, a type of complex behavior ubiquitous in complex systems. We have considered a classical consensus algorithm with the addition of noise, and communication intermittency:

$$x_i(k+1) = x_i(k) + \beta \sum_{j\in N_i} \xi_{ij}(k)\big[x_j(k) - x_i(k)\big] + \upsilon_i(k), \quad i = 1, \dots, n, \tag{8.10}$$

where $\xi_{ij}(k)$s are IID Bernoulli random variables with probability μ_{ij} of being 1 characterizing the fading property of the channel at time k. $\upsilon_i(k)$s (the communication noise) are Gaussian random variables independently identically distributed across both k and i with zero mean and unit variance.

One important characteristic of many distributed computation systems, including those in this chapter, is their lack of minimality. This is due to the presence of marginally stable poles corresponding to eigenvalues on the boundary of the stability region that are unobservable and/or uncontrollable. For example, consider the system in (8.10). Assume that all $\xi_{ij}(k)$ are independent but with the same mean μ and variance σ^2. Then

$$\begin{aligned}
x_i(k+1) &= x_i(k) + \beta \sum_{j\in N_i} \big(\mu + \Delta_{ij}(k)\big)\big[x_j(k) - x_i(k)\big] + \upsilon_i(k), \quad i = 1,\dots,n,\\
x_i(k+1) &= x_i(k) + \beta \sum_{j\in N_i} \mu\big[x_j(k) - x_i(k)\big] + w_{\Delta_{ij}}(k) + \upsilon_i(k), \quad i = 1,\dots,n,\\
z_{\Delta_{ij}}(k) &= x_j(k) - x_i(k), \quad j \in N_i,\\
w_{\Delta_{ij}}(k) &= \Delta_{ij}(k) z_{\Delta_{ij}}(k).
\end{aligned} \tag{8.11}$$

The agents are discrete-time integrators (adders), $P_i = P_A$ with $P_A(z) = \frac{\beta}{z-1}$ and the Mean Network is given by

$$N: \quad \begin{bmatrix} \nu \\ z_\Delta \end{bmatrix} = \left[\begin{array}{c|c|c} \mu L & I & B \\ \hline C & 0 & 0 \end{array}\right] \begin{bmatrix} \eta \\ \upsilon \\ w_\Delta \end{bmatrix}, \tag{8.12}$$

where $L = \mathcal{L}$ is the graph Laplacian with 0–1 weights, C is the transpose of the incidence matrix of the graph, and B is such that $L = BC$.

It follows that $M = \mathrm{F}_l(\mathrm{diag}[P_A]_{1:n}, N)$ has state-space

$$\begin{aligned}
x^+ &= (I + \beta\mu L)x + Bw_\Delta + \upsilon,\\
z_\Delta &= Cx.
\end{aligned}$$

The Laplacian of a strongly connected graph has one eigenvalue at zero with eigenvector $\mathbf{1}$, i.e., $L\mathbf{1} = 0$. One the other hand, $C\mathbf{1} = 0$, too. Therefore, M has one eigenvalue at 1 (marginally stable) with eigenvector $\mathbf{1}$, but this mode is not observable from C. Thus, the feedback interconnection of P and N is not detectable as it has one eigenvalues on the unit circle (not strictly stable). The impulse response of M, and therefore its transfer function matrix $M(z)$ will not show the marginal internal instability. We see that the distributed averaging system cannot be (strictly) stable, thus Theorem 8.1 will not directly apply and the system is indeed not MS stable.

However, since the mode at 1 in the direction $\mathbf{1}$ is not observable from Δ, we could study how Δ affects the MS stability of the rest of the modes which can be made stable by β small enough. A simple approach would be to obtain a minimal state-space representation for M and to study that. However, this will destroy the networked system structure and reduce our insights. Alternatively, we use a natural decomposition of the state. Let γ be the left eigenvector of $A = I + \beta\mu L$ associated with this eigenvalue, and by normalization such that $\gamma^T A = \gamma^T$, $\gamma^T\mathbf{1} = 1$. It is

natural to define operators

$$\mathcal{P} := I - \mathbf{1}\gamma', \qquad \mathcal{F} := \mathbf{1}\gamma'. \tag{8.13}$$

Let the *deviation state* and *conserved state* of the system be defined as $x_d(k) := \mathcal{P}x(k)$ and $x_c(k) := \mathcal{F}x(k)$, respectively, where $\mathcal{P}$ and $\mathcal{F}$ are defined in Eq. (8.13). Then, their evolutions are governed by

$$x_d(k+1) = \mathcal{P}A\mathcal{P}x_d(k) + \mathcal{P}B\Delta(k)Cx_d(k) + \mathcal{P}\upsilon(k), \tag{8.14}$$

$$x_c(k+1) = x_c(k) + \mathcal{F}B\Delta(k)Cx_d(k) + \mathcal{F}\upsilon(k), \tag{8.15}$$

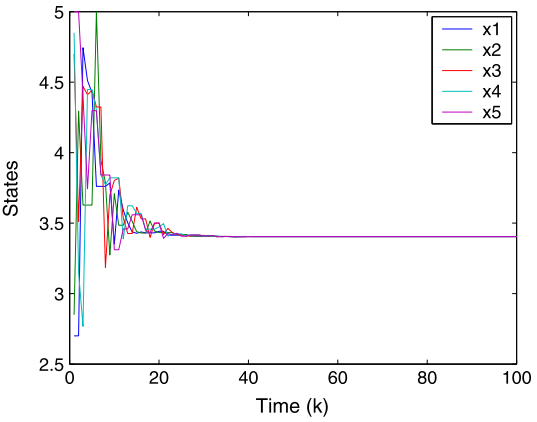
where we have used the facts that $\mathcal{P}A = A\mathcal{P}$ and $\mathcal{P}^2 = \mathcal{P}$, which are easy to verify. In the sequel, we refer to (8.14) as the *deviation system* and (8.15) as the *conserved system*. Although x_c is in $\mathbb{R}^n$, it evolves on the subspace spanned by $\mathbf{1}$. Thus, essentially x_c is one-dimensional. It is convenient to define any component of x_c as $\hat{x}_c$, or equivalently,

$$\hat{x}_c(k) = \gamma' x_c(k). \tag{8.16}$$

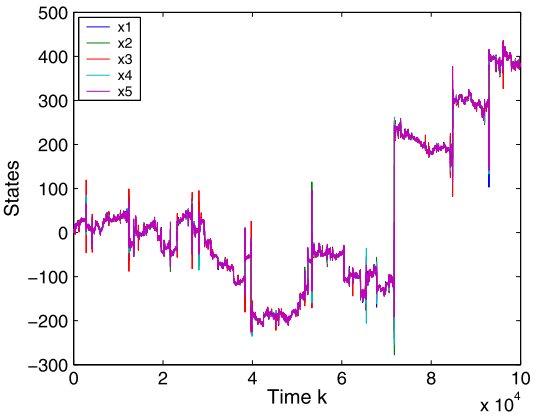
From (8.15), $\hat{x}_c$ is the result of the integration (adder) of a process that depends on x_d through the coupling provided by $\Delta(k)$ and $\mathcal{F}B$. Instead of reducing the system M, with this decomposition we are augmenting it by one dimension preserving most of the structure of the system. In the Deviation system, the eigenvalue at 1 is replaced with one at 0. This eigenvalue at 0 is still unobservable and does not impact the behavior of the networked system but makes the Deviation System stabilizable and stable for small enough β. The marginal internal instability is now captured in x_c by the conserved system. A similar state decomposition can be done for other algorithms that are more general. Although some of the details may be omitted, we will continue to use the terminology of conserved and deviation system to refer to such decompositions.

8.5.1 Emergence of Complex Behavior

We can now apply Theorem 8.1 to the deviation system and study how its MS (in)-stability impacts the behavior of the networked system. In [55, 59], we have shown that the MS instability of the deviation system is responsible for the emergence of a complex behavior of the networked system when additive noise is present. In particular, we have characterized $\hat{x}_c$ as a hype-jump diffusion process and as an uncorrelated Levy flight in a special two-agents case. We refer the reader to the cited references for details and in depth analysis. Thus, the MS instability of the deviation system is an indication of a fragility of the system to the combined effect of additive and multiplicative noises, which are typically both present in networked systems. Figure 8.4(a) shows that the state of system (8.10) converges to a common value when noise in not present although the deviation system is MS unstable for the



(a) Convergence without additive noise



(b) Levy flights in the presence of additive noise

Fig. 8.4 Hypersensitivity to noise that can happen when the channels are fading

chosen parameters. However, the system is in a critical state and is hypersensitive to noise. This fragility emerges when noise is added to the system. Figure 8.4(b) shows a remarkable new collective behavior of the system when a tiny noise is added to the communications. There are abrupt jumps in the agents' states, and between jumps, periods where the agents' states are reasonably close to each other. Levy flights have been observed in many natural sciences as well as economics and many other fields. Examples include biological searching patterns [4, 41], the distribution of human travel [7], financial series of stock markets [28, 29] and photons in hot atomic vapors [31]. This is the first example of a simple multi-agent model, inspired by natural behaviors, like swarming and flocking, that exhibits such a collective diffusion behavior [59] and for which such behavior can be predicted from the MS instability condition.

Besides the fascinating connection with the physics and mathematics of complex systems, the result above shows that limited communication can have serious effects on the behavior of computing networked systems and motivates further de-

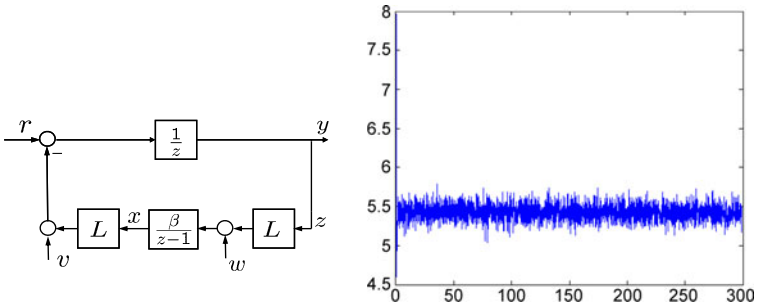
velopments geared toward prediction, prevention, or mitigation of the emergence of critical phenomena when we consider systems over unreliable networks. It also points toward the study of networked systems organizations that may be more resilient to such behavior [55, 62].

8.6 Distributed Computing of Averages Resilient to Noise

Even without channel dropouts, we have already pointed out that the simple averaging system (8.10) is fragile to additive noise. $\hat{x}_c$ becomes a random walk and the algorithm loses its ability to compute averages (Fig. 8.1(b)). This behavior is due to the integration (accumulation) of the network noise in the states of the agents (8.15). From a dynamical system viewpoint, there is a marginally stable MIMO pole-zero cancellation in the system. The Laplacian provides a MIMO zero at 1, which cancels the pole at 1 in the direction $\mathbf{1}$. The marginal instability is internal and not manifested without noise, but its effect emerges when noise is present. At the same time, this internal instability is embedded in the inner working of the algorithm and is not immediately removable. To resolve this problem, we have proposed a new algorithm that places another MIMO zero at 1 after the bank of integrators, preventing the random walk to be visible at the output. The block diagram is shown in Fig. 8.5(a), and uses the network twice [56]. Note that now not all Laplacians guarantee stability of the system. However, symmetric Laplacians $L = L'$ (corresponding to undirected graphs) are feasible:

$$
\begin{aligned}
x_i(k+1) &= x_i(k) + \beta \sum_{j \in N_i} \big(z_j(k) - z_i(k)\big), \\
z_i(k+1) &= r_i - \sum_{j \in N_i} \big(x_j(k) - x_i(k)\big), \\
y(k) &= z(k).
\end{aligned}
$$

The algorithm is a modification of the dynamic consensus scheme of [47]. Each node now uses the network twice sending two states, x_i and z_i. Under the assumption of symmetric L, it can be shown that $y(k)$ converges under quite general conditions to the average of r. In this case, though, the algorithm is resilient to noise coming into the system from various sources. Figure 8.5(b) shows the simulation when the additive noise comes from the use of the communication links. Note that the algorithm is a linear dynamical system and can be analyzed with many available tools. When the noise is Gaussian, the result of the computation $y(k)$ is a Gaussian random variable with mean equal to the average of r. Its variance can be pre-computed based on the assumed variance of the noise. We see that the dynamical system approximately computes averages under uncertainty. Moreover, if the average of $r(k)$ changes with time, $y(k)$ will be able to detect it and track it.



(a) (b)

Fig. 8.5 (**a**) New averaging system; (**b**) Its resilience to noise

8.6.1 Resilience to Channel Intermittency

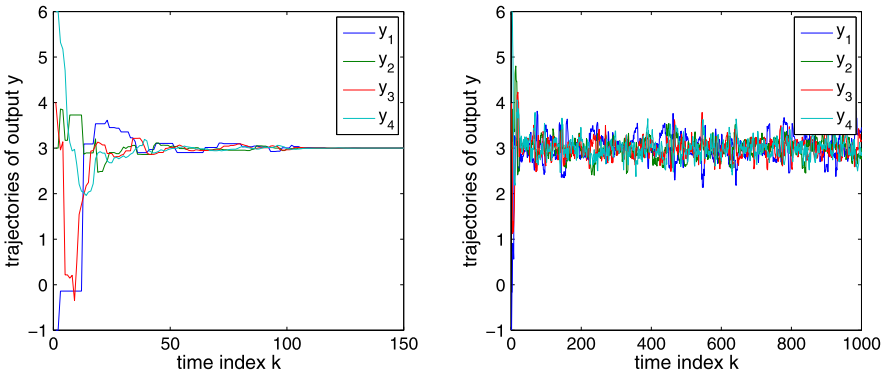
Although the computation of averages is now robust to additive noise, it is still fragile to unreliable intermittent communication. To mitigate this limitation, the agents need to be more intelligent and use the channel-state information from their neighbors. In the following algorithm [56], each agent holds and uses the last good received message from its neighbors:

$$\begin{aligned} h(k) &= \Lambda_2(k)\big(Cx(k) + v(k)\big) + \big(I - \Lambda_2(k)\big)h(k-1), \\ y(k) &= u(k) - Bh(k), \\ x(k+1) &= x(k) + \beta B \Lambda_1(k)\big(Cy(k) + w(k)\big). \end{aligned} \tag{8.17}$$

$\Lambda_1(k)$ and $\Lambda_2(k)$ are diagonal matrices of IID Bernoulli random variables, of dimension equal to the number of edges, representing the unreliable link among nodes. As before, $L = BC$ where C is the transpose of the incidence matrix of the graph. h is the vector of holding states. When a component of $\Lambda_2(k)$ equal 0, the corresponding component of h is held to the previous value. Otherwise, the component of h updated with the current channel output. v and w are additive networks noises. The MS stability of this system can be accessed using Theorem 8.1. Figure 8.6 shows the resilience of the algorithm to noise and channel intermittency with appropriate choice of parameters. Details can be found in [56, 61]. We see that the double Laplacian scheme is a robust network architecture that allows approximate computing in the presence of unreliable communication under certain conditions. It is the basis for extending the approach to distributed systems that compute solutions of unconstrained optimization problems as we will see next.

8.7 Optimization Systems

Although computing averages in the presence of noise has a variety of applications, our approach would still be limited in scope. A main development is given by the



(a) Robustness to packet drops using LSI. (b) Robustness to both uncertainties.

Fig. 8.6 Simulation examples to illustrate robustness of our algorithms to additive and multiplicative uncertainties

extension of the dynamical systems view to solving distributed optimization problems.

Although extremely advanced, current optimization methods (including distributed ones [5]) have several limitations when the optimization algorithms need to run real-time, over unreliable networks, distributed on physically separated (mobile) agents immersed in a noisy environment. On the other hand, biological social and other natural systems could be solving optimization problems, and therefore being able to compute, without being "digital" or algorithmically based. This motivates us to explore the dynamical system view, which could provide a natural implementation of these computational systems integrated with the dynamics of the physical systems.

We consider the following constrained optimization problem:

$$\begin{aligned} p^* &= \min_{x \in \mathbb{R}^m} f(x) \\ \text{s.t.} \quad & Ax = b, \end{aligned} \tag{8.18}$$

where $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^n$. To provide the main idea, we assume that f is strictly convex and differentiable, and that problem (8.18) is feasible and has a finite optimal cost $-\infty < p^* < \infty$. We use x^* to denote the optimal solution to problem (8.18). The Lagrangian function is given as

$$F(x, \lambda) = f(x) + \lambda^T (Ax - b).$$

We consider the following dynamic system to solve the optimization problem above:

$$\begin{aligned} \dot{x} &= -\nabla_x F(x, \lambda), \\ \dot{\lambda} &= \nabla_\lambda F(x, \lambda), \end{aligned} \tag{8.19}$$

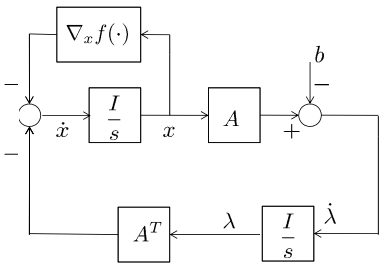


Fig. 8.7 The block diagram of system (8.20)

where ∇_x and ∇_λ represent the gradients with respect to x and λ, respectively.

Substituting $F(x, \lambda)$ in the above equations, we have

$$\begin{aligned} \dot{x} &= -\nabla_x f(x) - A^T \lambda, \\ \dot{\lambda} &= Ax - b. \end{aligned} \tag{8.20}$$

It is important to note that, under the current assumptions, the above dynamical system always converges to the optimal solution x^*.

Theorem 8.2 *Consider system* (8.20) *with* f *strictly convex and differentiable. For any initial values of* x *and* λ, *we have* $\lim_{t\to\infty} x(t) = x^*$.

The proof (see [2, 58]) links the global under-estimator property of the gradient of a convex function with dissipation. This view is consistent with that one emerged in the study of control of networks [11, 15, 20, 21, 27, 39].

8.7.1 Control Perspective

It is instructive to consider the block diagram of system (8.20) shown in Fig. 8.7. We note several control features, which now apply to the optimization system. First, the optimization system is subject to the fundamental limitations of feedback, eventually determining the limitations of disturbance and noise rejection. Second, note that the vector b, which is part of the constraints, is a command or a disturbance. From the classical control theory, we know the integrators (of the dual variable) guarantee zero steady state tracking to constant b. This ensures the satisfaction of the constraints. Moreover, we see that b can now change overtime. Thus, the system can adapt the convergence point, leading to the possibility of real-time adaptive optimization. Finally, we may design advanced controllers to improve the convergence speed and the performance of the optimization system.

8.7.2 Distributed Optimization Systems

We apply the main idea described in the previous section to derive distributed optimization systems. The new architectures allow for simple interconnected agents to solve complex optimization problems collectively with only local gradient sensing means. Consider the following unconstrained problem [57, 58]:

$$p^* = \min_{\theta \in \mathbb{R}^m} \sum_{j=1}^{n} f_j(\theta). \tag{8.21}$$

We assume there are n agents. Each has its private cost function f_j, strictly convex. The agents connected over an ideal communication network exchange information with their nearest neighbors. The network is represented by its Laplacian, $\mathcal{L}$ and is assumed strongly connected and undirected, but not necessarily complete. We propose the following equivalent optimization problem, where each agent has its individual estimate of θ, called x^j:

$$\begin{aligned} p_d^* &= \min_{x \in \mathbb{R}^{nm}} \sum_{j=1}^{n} f_j\big(x^j\big), \\ Lx &= 0, \end{aligned} \tag{8.22}$$

where $x = \mathbf{vec}[x^j]_{j=1:n}$, $x^j \in \mathbb{R}^m$ and $L = \mathcal{L} \otimes I_m$. The constraint $Lx = 0$ can be implemented over the network and eventually enforces $x^j = x^i = y$ for all $i, j \in \{1, \ldots, n\}$, due to the property of the Laplacian.

Following (8.19), the associated optimization system is governed by the following equations:

$$\begin{aligned} \dot{x} &= \mathbf{vec}\big[\nabla_{x^j} f_j\big(x^j\big)\big] - L\lambda, \\ \dot{\lambda} &= Lx, \end{aligned} \tag{8.23}$$

where we have used $L = L^T$. Figure 8.8(a) shows the resulting block diagram. We note how the double Laplacian emerges naturally in the dynamic of the optimization system. System (8.23) can be seen as a generalization of the popular average consensus system of [38]. Expanding the equations reveals their distributed nature more clearly. Each agent has the following dynamics:

$$\begin{aligned} \dot{x}^j &= -\nabla_{x^j} f_j\big(x^j(t)\big) + \sum_{i \in N_j} \big(\lambda^i(t) - \lambda^j(t)\big), \\ \dot{\lambda}^j &= -\sum_{i \in N_j} \big(x^i(t) - x^j(t)\big), \end{aligned} \tag{8.24}$$

where N_j is the set of neighbors of agent j; x^j, $\lambda^j \in \mathbb{R}^m$ are the agent's states; the sum of relative errors represents a row of the Laplacian, and $\nabla_x f_j(x^j)$ is the (sub-) gradient of $f_j(y)$ at x^j.

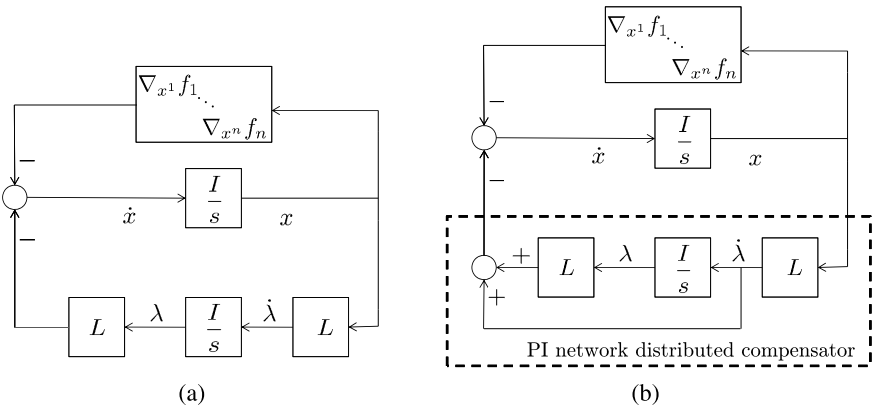


Fig. 8.8 Adapted from [58]. (**a**) The networked dynamical system solving the optimization (8.21) distributively; (**b**) A variant based on the augmented Lagrangian method. Note the presence of a networked PI (instead of an Integrator) feedback controller

8.7.3 Augmented Lagrangian and Control Interpretation

Consider the following optimization problem:

$$\begin{aligned} p^*_{au} &= \min_{x\in\mathbb{R}^m} \sum_{j=1}^{n} f_j\left(x^j\right) + \frac{1}{2}x^T Lx, \\ Lx &= 0. \end{aligned} \tag{8.25}$$

Clearly, $p^*_{au} = p^*_d$ since the constraint force $Lx = 0$, thus the added extra cost is zero if the problem is feasible. This technique is known as augmented Lagrangian method, see, e.g., [18]. It is easy to verify that the dynamical system for solving the above problem is (8.26):

$$\begin{aligned} \dot{x} &= \mathbf{vec}\left[\nabla_{x^j} f_j\left(x^j\right)\right] - Lx - L\lambda, \\ \dot{\lambda} &= Lx. \end{aligned} \tag{8.26}$$

Figure 8.8(a) shows the resulting block diagram.

It is interesting to note that the extra term, namely Lx in the first equation, is equal to $\dot{\lambda}$ from the second equation. Therefore, the dynamical system (8.26) includes a proportional action. Figure 8.8(b) shows the resulting scheme where the Proportional Integral compensation is highlighted.

Having a PI compensator instead of just an integrator in the loop leads to improved stability/convergence properties. The augmented Lagrangian method is known for its better convergence properties. We now see that these properties are justified from the feedback control system perspective.

This leads to the question of finding more general and powerful controllers and points to a new research direction of controller design for (distributed) optimization systems, which we will address later.

8.7.3.1 Comparison with Existing Approaches

Problem (8.21), with or without more general constraints, has been considered in [25, 26, 35, 36, 40, 66]. However, most of the proposed algorithms in the literature adopt a local convex mixing and vanishing step size on local gradient searching.

The literature offers various other approaches, based on the Method of Multipliers (MoM) [5], to impose the agreement among the x^js. Our approach, although related to MoM, is more natural when a network interconnection is already in place. Compared to alternating direction method of multipliers, which relies on the decomposability of the augmented Lagrangian to provide distributed solutions, the proposed approach allows one to explore the problem structure more directly. For example, in the previous distributed optimization problem, our approach does not require a global network collector as proposed in [6]. Note that we do not require each node to solve an optimization problem at each step, as done in primal–dual algorithms (e.g., [27]), and classical alternating direction method of multipliers (see [6]) but only to move along a favorable direction. This last point may have important implications that link to physical interconnected computing systems, as we see that no much intelligence or computational capabilities are required at each node.

8.7.4 Distributed Least Squares

Least Squares problems have an important role in science and engineering. Thus, advances in the distributed version of these problems are bounded to have a significant impact in many cooperative sensing, learning, and decision-making applications.

Consider the problem of estimating a common vector $\theta \in \mathbb{R}^m$, using a network of n distributed sensors. Each sensor has the measurement

$$b_i = A_i\theta + v_i,$$

where $b_i \in \mathbb{R}^{n_i}$, A_i is the known matrix to sensor i and v_i is assumed to be a Gaussian random variable with mean zero and covariance I. The most interesting case is when $n_i = 1$ and $n > 1$, which we consider next, and thus the sensors by themselves do not have enough data to estimate θ. When a fusion center is present, all sensor measurements can be aggregated as

$$b = A\theta + v,$$

where b, A and v are concatenations of b_i, A_i and v_i, respectively. The fusion center can then compute the best maximum-likelihood estimation of θ, by the least squares

solution $\hat{\theta} = (A'A)^{-1}A'b$ (assuming A has full column rank), which solves

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^m} \|A\theta - b\|_2^2.$$

Beside the scalability problem, it is desirable to depart from the concept of a fusion center, and understand if and how each sensor could obtain the above centralized maximum likelihood estimate using scalable computations and only local information exchange with its neighbors. With this objective in mind, note that the Least Squares optimization can be simply rewritten in terms of sum of cost functions local to the agents:

$$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^m} \sum_{i=1}^{n} \|A_i\theta - b_i\|_2^2 = \arg\min_{\theta \in \mathbb{R}^m} \sum_{i=1}^{n} f_i(\theta). \tag{8.27}$$

Let $\mathcal{L}$ be the Laplacian associated with the undirected strongly connected communication network graph. Associated with $\mathcal{L}$, let $L = \mathcal{L} \otimes I_m$. We assume that each agent will need to compute the optimal and unique θ^*. Each agent then must update a copy of θ, which we denote by x^i. Let $x = \mathbf{vec}[x^i]_{i=1:n}$.

Then, the following optimization system converges to the optimal solution of problem (8.27) for any initial condition:

$$\begin{aligned} \dot{x} &= -\nabla x - L\lambda - q, \\ \dot{\lambda} &= Lx, \end{aligned} \tag{8.28}$$

where $\nabla = \mathrm{diag}[2A_i^T A_i]_{i=1:n}$, and $q = -\mathbf{vec}[2b_i^T A_i]_{i=1:n}$. The dynamical system (8.28) is clearly Linear Time Invariant and has q as input.

The basic challenge with the above problem is that each node does not know the objective function of other nodes. Yet, they need to find the solution cooperatively. However, the problem is now in the form of (8.21) and can be solved by system (8.23) or (8.26).

The important feature of Least Squares, or convex Quadratic Programming problems, is that they lead to LTI optimization systems, as shown in the block diagram of Fig. 8.9. Thus, a vast array of advanced analysis and design methodologies from control theory apply to them. This may require us to move away from the general-purpose optimization algorithms, which work for most problems, and to design special-purpose optimization systems for specific applications and cost functions. Another important feature of the system is that the vector q is an exogenous input to the system. Thus, if q_i changes over time as function of the problem data, $b_j(k)$, the system can track and adapt the optimal point, as shown in Fig. 8.9, leading to real-time adaptive optimization. We want to point out that our approach is different from approaches based on the distributed implementation of the known optimal solution with assistance of consensus algorithms, as done in [64] or [37].

Finally, we want to point out that the approach presented in Sect. 8.6.1 for making the distributed averaging system more resilient to channel dropouts also applies to the distributed least squares systems. We refer the reader to [60] for details.

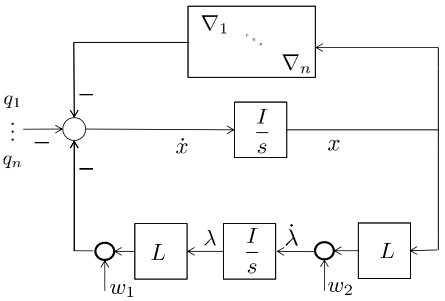


Fig. 8.9 Block diagram of the LTI optimization system solving least squares. w_1 and w_2 represent additive channel noise

8.8 Networked Controller Design for Networked Systems

Until now, we have focused on the analysis of networked systems. In this case, besides the network topology, the nodes' dynamics are known and the signals exchanged over the network predefined. In the case of distributed optimization problems, we have encountered simple examples of networked system design. We have seen how the agents' dynamics and the signals exchanged are derivable from the optimization problem setup or from simple equivalent transformations. In this last part of the chapter, we consider the problem of designing networked systems in more general terms. Given the network topology, the problem is to define the nodes' dynamics, as well as the signals that the nodes need to exchange over the network, in order to accomplish a certain task. In particular, we consider the problem of designing networked distributed optimal controllers for a given networked system under the assumption that the controller's network topology is the same as that of the plant. Motivated by the derivation of the previous sections, we will consider the design of optimal networked controllers to improve the performance of a distributed least squares system. Because we view distributed computational systems as dynamical systems, we can now look at them as plants for which to design networked controllers. To our knowledge, this is the first time that advanced controller design methods are applied to distributed computation systems. To do so, we need to rely on new results on the design of optimal networked controllers for networked plants [52, 53].

Because the size and nature of the signals to be exchanged over the network as well as the dynamics of the nodes are not a priori specified, we need to identify systems' structures that are easily searchable and from which the dynamics of the nodes with the necessary network signals can be recovered.

We have seen in Sect. 8.3 that networked systems over noiseless and delay-free networks have structured state-space representations in $\mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$ and structured transfer functions in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_u, \mathcal{P}_y)$ defined in Sect. 8.2.3. A key question to resolve is to find out if systems in these classes can be transformed back into networked systems with specified nodes dynamics and network signals.

8.8.1 Networked Controllers

The setting of Sect. 8.2.3 applies to networked controllers. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with n nodes, we associate a sub-system $\{P_i^K\}_{i\in\{1,\dots,n\}}$ to each node. P_i^K maps $\left[\begin{smallmatrix} y_i \\ v_i \end{smallmatrix}\right] \to \left[\begin{smallmatrix} u_i \\ \eta_i \end{smallmatrix}\right]$, i.e., each sub-system P_i^K is a minimal DT system with local inputs $y_i(k)$, local outputs $u_i(k)$, network inputs $v_i(k)$, and network outputs $\eta_i(k)$. Define

$$P_K = \text{diag}\left[P_i^K\right]_i \tag{8.29}$$

which, through rearrangement of input and outputs, maps $\left[\begin{smallmatrix} y \\ v \end{smallmatrix}\right] \to \left[\begin{smallmatrix} u \\ \eta \end{smallmatrix}\right]$ with partitions $\mathcal{P}_u$, $\mathcal{P}_y$, $\mathcal{P}_\eta$, and $\mathcal{P}_v$, and $\mathcal{P}_x^K$ if represented in state-space.

Definition 8.7 For given $\mathcal{G}$ and $\mathcal{P}_y, \mathcal{P}_u$, we denote by $\mathcal{N}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ the set of networked controllers (systems) $K = \text{F}_l(P_K, N)$ obtained by the feedback interconnection of some block diagonal P_K defined as in (8.29) and a network interconnection matrix N structured according to $\mathcal{A}(\mathcal{G}) + I$ for appropriate $\mathcal{P}_\eta, \mathcal{P}_v$.

It is not difficult to see that any networked controller $K \in \mathcal{N}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ over a noiseless and delay-free network $\mathcal{G}$ has a state-space representation that belongs to the set $\mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$, for some state partition $\mathcal{P}_x^K$, and has a transfer function matrix $K(z)$ that belongs to $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$.

If a controller (or a system more generally) is built over a network, it is important that the feedback interconnection of P_K and N is detectable from u and stabilizable from y. This excludes the presence of unstable internal modes in the networked system.

Definition 8.8 Let $\mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u) \subset \mathcal{N}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ be the subset of networked controllers that are detectable from u and stabilizable from y.

What are the systems in $\mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$ or in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ for which we can find a networked controller (i.e., P_K and N) in $\mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$?

8.8.2 Implementing and Realizing Systems over the Given Network

Definition 8.9 A system $K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$ is said to be network implementable if there exist P_K, defined as in (8.29) with partitions $\mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u$, and a network interconnection matrix N structured according to $\mathcal{A}(\mathcal{G}) + I$ for appropriate $\mathcal{P}_\eta, \mathcal{P}_v$ such that $K = F_l(P_K, N) \in \mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$.

Excluding non-stabilizable and/or non-detectable systems from those, we consider the property of being implementable over the network as important since it excludes the possibility that the networked system has hidden unstable modes. These hidden unstable modes will make the networked system useless in practice as noise

or uncertainty, always present in the network interconnections, will excite them and manifest the "internal" instability. Thus, although we assume ideal network interconnections, we do not want our results to be fragile to this assumption.

Lemma 8.1 [52, 53] *Any stabilizable and detectable system* $K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$ *is implementable over the network* $\mathcal{G}$, *i.e.*, $K \in \mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$.

This lemma says that it is possible to find η and ν, P_K and N with appropriate partitions, from the detectable and stabilizable structured state-space $K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$, so that the networked controller is stabilizable and detectable.

For a given $\mathcal{G}$ and input and output partitions $\mathcal{P}_y$ and $\mathcal{P}_u$, respectively, we denoted by $\mathfrak{S}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u) = \bigcup_{\mathcal{P}_x \in \mathbb{N}^n} \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$.

Definition 8.10 A system K with transfer function matrix $K(z) \in \mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ is said to be *realizable over* $\mathcal{G}$ if there exists a system $\tilde{K}$ with partitions $\mathcal{P}_y$, $\mathcal{P}_u$ implementable over $\mathcal{G}$ (for some $\mathcal{P}_x$), i.e., $\tilde{K} \in \mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, such that $K(z) = \tilde{K}(z)$.

Note that the definition of network realizability does not impose for the state-space realization of a transfer function to be minimal, but only detectable and stabilizable. Thus, the least requirement for acceptable realizations over the network is that they cannot have unstable modes unobservable and/or uncontrollable while they could have stable unobservable and/or uncontrollable modes. This allows more flexibility in the realizing networked systems. We note that it is not known how to obtain a *minimal* networked realization of a stable system in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, in general.

It turns out that any stable system in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ is realizable over the network $\mathcal{G}$ based on Definition 8.10. Unfortunately, we do not know yet if any system in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ is network realizable in general, and if is, how to realize it. Nevertheless, the network realizability of stable systems plays an important role in the parametrization of all stabilizing network realizable controllers.

Theorem 8.3 *Given a network represented by a directed graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ *and the input and output partitions,* $\mathcal{P}_y$ *and* $\mathcal{P}_u$, *any bounded-input bounded-output* (*BIBO*) *stable system* $Q(z) \in \mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ *is realizable over the given network.*

The proof is constructive and given in [51, 53].

Remark 8.1 The lack of a complete theory limits the applicability of input–output approaches to design optimal network realizable distributed controllers that may be unstable. These approaches search directly over the impulse responses and thus provide optimal structured controllers in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. Realizing such unstable controllers over the network would be a problem, in general.

We denote the set of all stable real-rational proper transfer function matrices in $\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ by $\mathfrak{T}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. Note that, if $Q(z) = [Q_{ij}(z)]_{i,j} \in \mathfrak{T}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, then $Q_{ij}(z) \in \mathcal{RH}_\infty$ for all i, j.

8.8.3 Generalized Networked Systems

We begin by extending the definition of the agents to allow local exogenous inputs and local regulated outputs. A networked plant P is defined as in (8.4), but with each sub-system now including local exogenous inputs $w_i(k)$ and local regulated outputs $z_i(k)$. Thus, the state-space description of the sub-system P_i is written as

$$P_i: \quad \begin{bmatrix} x_i(k+1) \\ z_i(k) \\ y_i(k) \\ \eta_i(k) \end{bmatrix} = \begin{bmatrix} A_i & B_i^w & B_i^u & B_i^v \\ C_i^z & D_i^{zw} & D_i^{zu} & D_i^{zv} \\ C_i^y & D_i^{yw} & D_i^{yu} & D_i^{yv} \\ C_i^\eta & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_i(k) \\ w_i(k) \\ u_i(k) \\ v_i(k) \end{bmatrix}, \tag{8.30}$$

where B_i^w, C_i^z, D_i^{zw}, D_i^{yw}, D_i^{zu}, and D_i^{zv} have dimensions compatible with the local exogenous inputs $w_i(k)$, local regulated outputs $z_i(k)$, local control inputs $u_i(k)$, the local measurement outputs $y_i(k)$, the local network outputs η_i, and the local network inputs v_i:

$$\begin{aligned} G &= \mathrm{F}_l(P, N) \\ &= \left[\begin{array}{c|c|c} \hat{A} + \hat{B}_v N \hat{C}_\eta & \hat{B}_w & \hat{B}_u \\ \hline \hat{C}_z + \hat{D}_{zv} N \hat{C}_\eta & \hat{D}_{zw} & \hat{D}_{zu} \\ \hline \hat{C}_y + \hat{D}_{yv} N \hat{C}_\eta & \hat{D}_{yw} & \hat{D}_{yu} \end{array}\right] \\ &:= \begin{bmatrix} A & B_w & B_u \\ C_z & D_{zw} & D_{zu} \\ C_y & D_{yw} & D_{yu} \end{bmatrix}. \end{aligned} \tag{8.31}$$

Note that A, C_z, C_y are structured according to $\mathcal{A}(\mathcal{G}) + I$, while B_w, B_u, D_{zw}, D_{zu}, D_{yw}, D_{yu} have a block diagonal structure.

We would like to find optimal $\mathcal{H}_2$ stabilizing networked controllers over the same network as the plant. Let $T_{zw} = \mathrm{F}_l(G, K)$ be the networked closed loop system. We consider the following optimization problems:

$$\begin{aligned} \min \quad & \|T_{zw}\|_2 \\ \text{subject to} \quad & K \in \mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_y), \\ & T_{zw} \text{ is internally stable.} \end{aligned} \tag{8.32}$$

To solve this problem, it is natural to search over the structures in the state-space, $K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$, or, as is often done, in the input–output $K(z) \in$

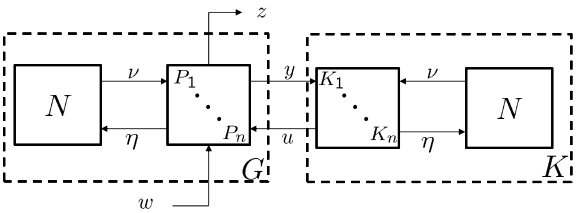


Fig. 8.10 The Networked Control Framework for Networked Plants

$\mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. (See Fig. 8.10.) However, for any stabilizing $K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x^K, \mathcal{P}_y, \mathcal{P}_u)$, or $\in \mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, we need to find a corresponding networked system $K \in \mathcal{N}^I(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. In particular, this last step is not usually addressed in the literature, which implicitly assumes that searching over input–output structures is equivalent to searching over networked systems.

8.8.4 All Stabilizing Network Implementable Controllers

Summarizing, we have seen that networked systems, over noiseless and delay free networks, have specific state-space and transfer function matrix structures. We have also seen that the state space structures can be implemented back as networked systems. The implementation is constructive and allows identifying the agents dynamics and the signals they need to exchange on the given network topology. On the other hand, obtaining networked systems out of structured transfer functions' matrices is possible for stable systems; it is not known if this is possible in general. It is then natural to work with the structured state-spaces. In this section, we characterize all the stabilizing controllers for a networked system, which are implementable over the given network, provided that one networked stabilizing controller can be found. This is always possible if the networked plant is stable. Sufficient conditions based on LMI for finding a stabilizing controller in the general case are given in [51, 53].

Theorem 8.4 *Given a networked plant G assume that $G_{22} := (A, B_u, C_y, D_{yu}) \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$ is stabilizable and detectable with $D_{yu} = 0$. If there exist matrices $\mathbf{F}$ partitioned according to $(\mathcal{P}_u, \mathcal{P}_x)$ structured according to $\mathcal{A}(\mathcal{G}) + I$ and $\mathbf{L}$ partitioned according to $(\mathcal{P}_x, \mathcal{P}_y)$ block-diagonal such that $A + B_u\mathbf{F}$ and $A + \mathbf{L}C_y$ are stable, then the set of all stabilizing FDLTI controllers for G, which are implementable over $\mathcal{G}$, is parametrized by*

$$K = \mathrm{F}_l(J, Q), \tag{8.33}$$

where

$$J = \left[\begin{array}{c|cc} A + B_u\mathbf{F} + \mathbf{L}C_y & -\mathbf{L} & B_u \\ \hline \mathbf{F} & 0 & I \\ -C_y & I & 0 \end{array}\right] \tag{8.34}$$

is implementable over the network $\mathcal{G}$, Q is any FDLTI system stable and implementable over the network $\mathcal{G}$.

The structured state-space of the networked plant is compatible with the structured state-space of J, the nominal stabilizing controller, provided suitable structured $\mathbf{F}$ and $\mathbf{L}$ can be found. Then all the implementable stabilizing controllers can be found by the feedback interconnection of J with any stable Q either in state-space or in transfer function matrix form. In particular, we can now search over stable Q in the frequency domain.

8.8.5 Optimal Solution for the Distributed H_2-Problem

A distributed controller implementable over the network $\mathcal{G}$ can be seen as a stabilizable and detectable system in $\mathfrak{S}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$.[3] Thus, given a networked plant G over a network $\mathcal{G}$, the network distributed H_2 control problem can be written as

$$\begin{aligned} \min \quad & \|T_{zw}\|_2 \\ \text{subject to} \quad & K \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u), \\ & T_{zw} \text{ is internally stable,} \end{aligned} \tag{8.35}$$

where $T_{zw} = \mathrm{F}_l(G, K)$ denotes the closed-loop mapping from $w(k)$ to $z(k)$. Since all the stabilizing controllers for T_{zw} are the stabilizing controllers for G_{22}, we have from Theorem 8.4 that

$$T_{zw} = F_l\big(G, F_l(J, Q)\big),$$

where J is given by (8.34) and $Q \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. If there exist matrices $\mathbf{F}$ and $\mathbf{L}$ with the properties described in the hypothesis of Theorem 8.4, then the set of all closed-loop transfer matrices from $w(k)$ to $z(k)$ can be obtained using Theorem 8.4 and the results from [65] by an internally stabilizing proper controller implementable over the network $\mathcal{G}$ as

$$T_{zw} = \mathrm{F}_l(T, Q) = \big\{T_{11} + T_{12} Q T_{21} \colon Q \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)\big\}, \tag{8.36}$$

where T is given by

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

[3] Note that the input and output partitions of the controller have to match the output and input partitions of the interconnected plant, while the state partition is not fixed.

$$= \left[\begin{array}{cc|cc} A + B_u\mathbf{F} & -B_u\mathbf{F} & B_w & B_u \\ 0 & A + \mathbf{L}C_y & B_w + \mathbf{L}D_{yw} & 0 \\ \hline C_z + D_{zu}\mathbf{F} & -D_{zu}\mathbf{F} & D_{zw} & D_{zu} \\ 0 & C_y & D_{yw} & 0 \end{array}\right]. \tag{8.37}$$

Since the closed-loop transfer matrix is simply an affine function of the controller parameter matrix Q, we can rewrite the distributed H_2-problem in (8.35) as a convex optimization problem

$$\begin{aligned} &\min \quad \|T_{11} + T_{12}QT_{21}\|_2 \\ &\text{subject to} \quad Q \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u). \end{aligned} \tag{8.38}$$

It is convenient to solve the above problem in the frequency domain. This is possible since $Q \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ is equivalent to $Q(z) \in \mathfrak{T}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$. We can maintain the implementability of the controller due to the results of Theorem 8.3 which guarantees the realizability of $Q(z)$ with a state-space realization $\tilde{Q} \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ which is implementable over the network $\mathcal{G}$. Thus, the optimization problem in (8.38) can equivalently be expressed as

$$\begin{aligned} &\min \quad \left\|T_{11}(z) + T_{12}(z)Q(z)T_{21}(z)\right\|_2 \\ &\text{subject to} \quad Q(z) \in \mathfrak{T}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u). \end{aligned} \tag{8.39}$$

The problem is now reduced to a standard convex optimization form, where $Q(z)$ is structured (see [42, 54]). Once the optimal $Q(z)$ is found, we obtain a network implementable state-space realization $\tilde{Q} \in \mathfrak{S}^s(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$ for $Q(z)$ using Theorem 8.3, and then a network implementable state-space realization for the optimal K from Theorem 8.4, even if K is unstable. On the other hand, obtaining $K = F_l(J, Q(z))$ directly from $Q(z)$ would lead to K having the right input–output structure, i.e., $K \in \mathfrak{T}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, but its state-space structure would be destroyed, i.e., $K \notin \mathfrak{S}(\mathcal{G}, \mathcal{P}_y, \mathcal{P}_u)$, leaving us with a structured but centralized solution, in general. We finally stress that, in order to guarantee the resulting controller is implementable over the network, we need a networked plant model in the state-space.

8.8.6 Networked Controllers for Distributed Least Squares Systems

We are now ready to apply the previous development to a least squares system. We consider the Euler discretization of (8.28) with sampling time β for the least squares system

$$\begin{aligned} x(k+1) &= x(k) - \beta\nabla x(k) - \beta L y(k) - \beta q, \\ y(k+1) &= y(k) + \beta L x(k). \end{aligned} \tag{8.40}$$

We propose designing a networked controller to improve the performance of the discrete-time approximation of the optimization system and extending the approach

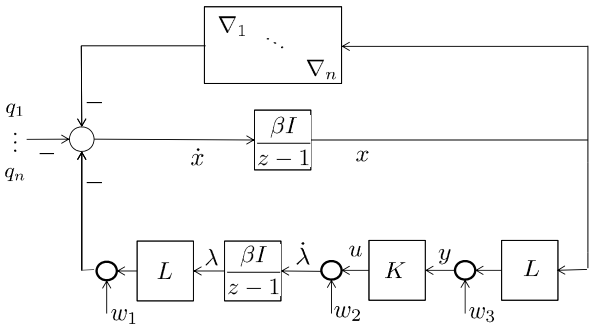


Fig. 8.11 Controller for ad-hoc least squares optimization

to directed graphs. Note that when the Laplacian is not symmetric as when the graph is undirected, the dynamical system (8.28) (or (8.40)) does not converge, in general.

We look for more advanced controllers that still maintain the property of the system to converge to the optimal solution. Although this problem can be posed in grater generality, here we consider the setup shown in Fig. 8.11.

We consider the following generalized model for the networked plant:

$$\begin{aligned} x(k+1) &= x(k) - \beta\nabla x(k) - \beta L\lambda(k) - \beta q + \beta w_1(k), \\ \lambda(k+1) &= \lambda(k) + \beta u(k) + w_2(k), \\ z_1(k) &= x(k), \\ z_2(k) &= \alpha u(k), \\ y(k) &= Lx(k) + w_3(k), \end{aligned} \tag{8.41}$$

where w_1, w_2, and w_3 represent noise sources to which we want to be resilient, and z_1 z_2 are the regulated variables. β is a scalar weight. w_1 may model noisy values of q. w_3 models measurement noise, and w_2 models noise generated within the networked controller, possibly due to noisy communication links. We assume for simplicity that all the noises are independent Gaussian distributed among components and time, with zero mean and variances $\sigma_1^2 I$, $\sigma_2^2 I$, and $\sigma_3^2 I$, respectively. More general correlated noise models are possible, but avoided here for simplicity. The generalized plant has the following state-space structure from (8.41):

$$G = \left[\begin{array}{cc|ccc|c} I-\beta\nabla & -\beta L & -\beta\sigma_1^2 I & 0 & 0 & 0 \\ 0 & I & 0 & -\beta\sigma_2^2 I & 0 & -\beta I \\ \hline I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha I \\ \hline L & 0 & 0 & 0 & \sigma_3^2 I & 0 \end{array}\right]$$

with

$$G_{22} = \left[\begin{array}{cc|c} I - \beta\nabla & -\beta L & 0 \\ 0 & I & -\beta I \\ \hline L & 0 & 0 \end{array}\right].$$

Note that $G_{22} \in \mathfrak{S}(\mathcal{G}, \mathcal{P}_x, \mathcal{P}_u, \mathcal{P}_y)$, by appropriate permutation of the states, is structured according to the network, but it is not detectable since it has unobservable modes at 1. It is not known yet how to obtain a minimal realization guaranteed to be network implementable, as standard approaches will destroy the network structure of the state-space matrices.

To maintain the networked structure, we need to work with the non-minimal G_{22}. Theorem 8.4 can be extended in this setting (details omitted), however, to obtain implementable controllers which stabilize the detectable and stabilizable modes of G_{22} using the setup of Theorem 8.4, we now need matrices $\mathbf{F}$ structured according to $\mathcal{A}(\mathcal{G})$ and $\mathbf{L}$ block-diagonal such that $A + B\mathbf{F}$ and $A + \mathbf{L}C$ have all the eigenvalues strictly inside the unit disc besides those that are not detectable and/or not stabilizable. This allows obtaining J in (8.34), which is crucial for network realizability. When we pass to the frequency domain obtaining T in (8.37) and searching for Q, the undetectable unstabilizable modes will automatically disappear from the transfer functions, which are always minimal.

With these modifications, we next apply the design method of the previous section to specific examples. We use distributed least squares to solve the distributed averaging problem. A final remark is to clarify that the $\mathcal{H}_2$-design is to be intended in the input–output sense, as the system has uncontrollable and/or unobservable simple eigenvalues on the unit circle, which prevent computing the $\mathcal{H}_2$-norm using state-space methods.

8.8.7 Examples

Consider the following least squares problem:

$$\arg \min_{x_i = x_j, \forall i,j} \sum_{i=1}^{n} \frac{1}{2}(x_i - q_i)^2.$$

where $x_i \in \mathbb{R}$, and $q_i \in \mathbb{R}$. This is a special case of the least squares problem which has as optimal solution $x^* = \frac{1}{n}\sum_{i=1}^{n} q_i$, i.e., the average of the agent private quantities. It provides a different algorithm for computing distributed averages base on optimization [67].

In this case, $\nabla = I$. We consider two network topologies of 3 agents. The first has the following graph Laplacian:

$$L_1 = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}.$$

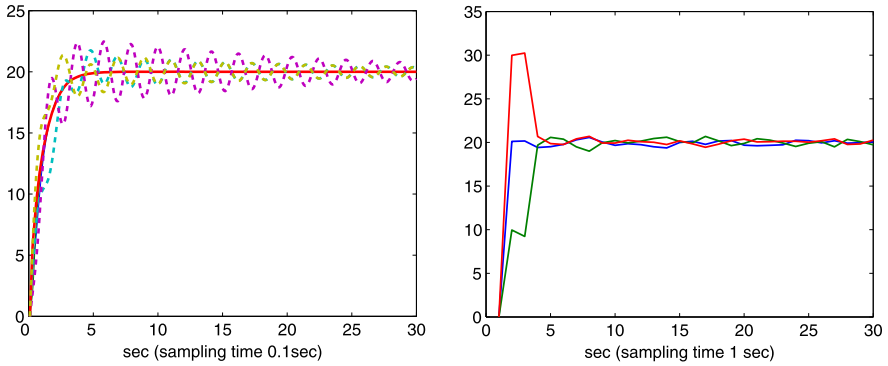


Fig. 8.12 Convergence to average with different β: (**a**) $\beta = 0.1$ s, (**b**) $\beta = 1$ s

L_1 is symmetric, so Theorem 8.2 applies. However, since we are using discretization, there is a limit on the largest sampling time $\beta < 0.12$ that guarantees convergence of (8.40) using the topology associated with L_1.

Figure 8.12(a) shows the response of the controlled new algorithm with respect to lightly damped response of (8.40); both algorithms have sampling time $\beta = 0.1$ s: $\mathbf{F} = 0.01[L_1, 0]$, $\mathbf{L} = 0.01\left[\begin{smallmatrix} 0 \\ I \end{smallmatrix}\right]$. The order of each agent's controller is 8, 11, 8, respectively, and the overall networked controller is unstable with a pole at -1.1334 although networked.

Figure 8.12(b) shows the response of the controlled new algorithm designed for sampling time of 1 s; (8.40) is unstable for such β. A small noise is added to show the resilience of the algorithm to additive noise. In this case, each sub-controller has order 6, 8, 6, respectively.

The second network is directed and with the following Laplacian:

$$L_2 = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}.$$

Theorem 8.2 does not apply, and system (8.40) does not converge without a networked controller.

Figure 8.13 shows the tracking response of the controlled new algorithm designed for sampling time of 1 s and the topology associated with L_2. Here $\mathbf{F} = -[0, I]$ and $\mathbf{L} = 0.1\left[\begin{smallmatrix} I \\ I \end{smallmatrix}\right]$. Each agent has a time-varying private input, $q_1(k) = 10\cos(1/600\pi k)$, $q_2(k) = 20\cos(3/600\pi k)$, $q_3(k) = 30\cos(6/600\pi k)$. Then the average of the inputs is $r(k) = \frac{1}{3}\sum_{i=1}^{2} q_i(k)$, which is a periodic signal composed of the three harmonics. The figure shows how each agent's state tracks the average over time and in the presence of additive noise.

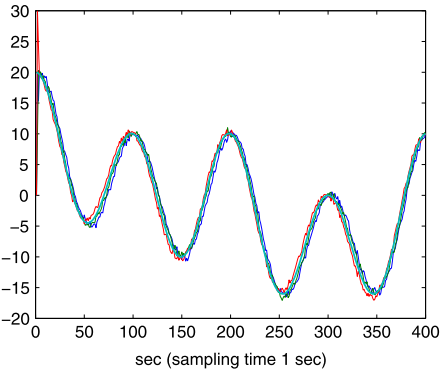


Fig. 8.13 Tracking time-varying average $q(t)$ with noise

8.9 Conclusions

In this chapter, we have put forward an integrated view of systems interacting over communication networks as distributed computing systems and proposed to study them as networked control systems. We pointed out that, when the network interconnections are unreliable, they can lead to the emergence of complex behavior in the networked system trying to compute distributed averages of the nodes private quantities. We believe that these results are relevant in studying and understanding a variety of aspects in biological and social networks, as averaging is central in many activities. We have also discussed how to mitigate the effects of complex behavior by reviewing network distributed averaging schemes that can approximately compute averages in the presence of noise and intermittent communication. The main ideas put forward for averaging systems can be extended to more general computing systems that solve convex distributed computations over the network. These results help understanding how systems with simple dynamics and limited resources can collectively compute solutions to complex problems outside the capabilities of an individual. The dynamical system view of optimization systems opens up many interesting directions of research and allows studying the effect of noise and unreliable communications, as well as tracking and adaptation properties of the optimization systems. In particular, designing controllers to improve the performance of optimization systems is now conceivable. In the last part of the chapter, we have presented an approach to improve the stability and convergence properties of a distributed least squares system, based on a recently developed networked controller design methodology.

Acknowledgements This research has been supported by NSF grants ECS-0093950, ECS-0524689 ECS-0901846, ECS-1239319.

This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Arnold, L.: Stochastic Differential Equations: Theory and Applications. Krieger, Malabar (1992)
2. Arrow, K.J., Hurwicz, L., Uzawa, H.: Studies in Linear and Nonlinear Programming. Stanford University Press, Stanford (1958)
3. Baillieul, J.: Feedback coding for information-based control: operating near the data-rate limit. In: Proceedings of the 2002 IEEE Conference on Decision and Control, Las Vegas, NV, pp. 3229–3236 (2002)
4. Bartumeus, F., Daluz, M.G.E., Viswanathan, G.M., Catalan, J.: Animal search strategies: a quantitative random-walk analysis. Ecology **86**(11), 3078–3087 (2005)
5. Bertsekas, D.P., Tsitsiklis, J.N.: Parallel and Distributed Computation: Numerical Methods. Prentice Hall, New York (1989)
6. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**(1), 1–122 (2011)
7. Brockmann, D., Hufnagel, L., Geisel, T.: The scaling laws of human travel. Nature **439**, 462–465 (2006)
8. Costa, O.L.V., Fragoso, M.D.: Stability results for discrete-time linear systems with Markovian jumping parameters. J. Math. Anal. Appl. **179**, 154–178 (1993)
9. Costa, O.L.V., Fragoso, M.D., Marques, R.P.: Discrete-Time Markov Jump Linear Systems. Springer, Berlin (2005)
10. Coviello, L., Minero, P., Franceschetti, M.: Stabilization over Markov feedback channels: the general case. IEEE Trans. Autom. Control. (2012). doi:10.1109/TAC.2012.2212616
11. Debb, S., Srikant, R.: Global stability of congestion controllers for the Internet. IEEE Trans. Autom. Control **48**(6) (2003)
12. Elia, N.: Remote stabilization over fading channels. In: Systems and Control Letters, vol. 54, pp. 237–249 (2005)
13. Elia, N., Eisenbeis, J.N.: Limitations of linear control over packet drop networks. IEEE Trans. Autom. Control **56**(4), 826–841 (2011)
14. Fax, J.A., Murray, R.M.: Information flow and cooperative control of vehicle formations. In: IEEE Transactions on Automatic Control Special Issue on Networked Control, vol. 49, pp. 1465–1476 (2004)
15. Feijer, D., Paganini, F.: Stability of primal–dual gradient dynamics and applications to network optimization. Automatica **46**(12), 1974–1981 (2010)
16. Gupta, V., Hassibi, B., Murray, R.M.: Optimal LQG control across packet-dropping links. Syst. Control Lett. **56**(6), 439–446 (2007)
17. Hadjicostis, C.N., Turi, R.: Feedback control utilizing packet dropping networks links. In: IEEE CDC (2002)
18. Hestenes, M.R.: Multiplier and gradient methods. J. Optim. Appl. **4**, 302–320 (1969)
19. Jadbabaie, A., Lin, J., Morse, A.S.: Coordination of groups of mobile autonomous agents using nearest neighbor rules. IEEE Trans. Autom. Control **48**(6), 988–1001 (2003)
20. Kelly, F.P., Maulloo, A., Tan, D.: Rate control in communication networks: shadow prices, proportional fairness and stability. J. Oper. Res. Soc. **49**, 237–252 (1998)
21. Kunniyur, S., Srikant, R.: End-to-end congestion control schemes: utility functions, random losses and ECN marks. IEEE/ACM Trans. Netw. **11**(5), 689–702 (2003)
22. Kushner, H.J.: Stochastic Stability and Control. Academic Press, New York (1967)
23. Ling, Q., Lemmon, M.: Robust performance of soft real-time networked control systems with data dropouts. In: Proceedings of the 41st IEEE Conference on Decision and Control (2002)
24. Ling, Q., Lemmon, M.: Optimal dropout compensation in networked control systems. In: Proceedings of the 42nd IEEE Conference on Decision and Control (2003)
25. Lobel, I., Ozdaglar, A.: Distributed subgradient methods over random networks. Tech. report 2800, MIT LIDS (2008)

26. Lobel, I., Ozdaglar, A., Feijer, D.: Distributed multi-agent optimization with state-dependent communication. LIDS report 2834 (2010). doi:10.1007/s10107-011-0467-x
27. Low, S.H., Lapsley, D.E.: Optimal flow control, I: basic algorithm and convergence. IEEE/ACM Trans. Netw. **7**(6), 861–874 (1999)
28. Mandelbrot, B.: The variation of certain speculative prices. J. Bus. **36**, 394–419 (1963)
29. Mantegna, R., Stanley, H.: Scaling behavior in the dynamics of an economic index. Nature **376**, 46–49 (1995)
30. 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)
31. Mercadier, N., Guerin, W., Chevrollier, M., Kaiser, R.: Lévy flights of photons in hot atomic vapours. Nat. Phys. **5**, 602–605 (2009)
32. Meyn, S.P., Tweedie, R.L.: Markov Chains and Stochastic Stability. Springer, London (1993)
33. Minero, P., Francescetti, M., Day, S., Nair, G.N.: Data rate theorem for stabilization over time-varying feedback channels. Trans. Autom. Control **52**(2), 243–255 (2009)
34. Nair, G.N., Evans, R.J.: Communication-limited stabilization of linear systems. In: Proceedings of the 39th IEEE Conference on Decision and Control, vol. 1, pp. 1005–1010 (2000)
35. Nedic, A., Ozdaglar, A.: Distributed subgradient method for multi-agent optimization. IEEE Trans. Autom. Control **54**(1), 48–61 (2009)
36. Nedic, A., Ozdaglar, A., Parrilo, A.P.: Constrained consensus and optimization in multi-agent networks. IEEE Trans. Autom. Control **55**(4), 922–938 (2010)
37. Olfati-Saber, R.: Distributed Kalman filtering for sensor networks. In: Proc. of 46th IEEE Conf. on Decision and Control, pp. 5492–5498 (2007)
38. Olfati-Saber, R., Murray, R.M.: Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans. Autom. Control **49**(9), 1520–1533 (2004)
39. Paganini, F.: Global stability result in network flow control. Syst. Control Lett. **46**(3), 153–163 (2002)
40. Ram, S.S., Nedic, A., Veeravalli, V.V.: 'Distributed stochastic subgradient projection algorithms for convex optimization. J. Optim. Theory Appl. (2010). doi:10.1007/s10957-010-9737-7
41. Reynolds, A.M., Smith, A.D., Menzel, R., Greggers, U., Reynolds, D.R., Riley, J.R.: Displaced honey bees perform optimal scale-free search flights. Ecology **88**(8), 1955–1961 (2007)
42. Rotkowitz, M., Lall, S.: A characterization of convex problems in decentralized control. IEEE Trans. Autom. Control **51**(2), 1984–1996 (2006)
43. Sahai, A.: Anytime information theory. Ph.D. thesis M.I.T. (Feb. 2001)
44. Sahai, A., Mitter, S.K.: The necessity and sufficiency of anytime capacity for control over a noisy communication link. Part I: scalar systems. IEEE Trans. Inf. Theory **52**(8), 3369–3395 (2006)
45. Schenato, S., Poolla, F., Sastry, J.: Kalman filtering with intermittent observations. In: Proceedings of the 42nd IEEE Conference on Decision and Control (2003)
46. Seiler, P., Sengupta, R.: Analysis of communication losses in vehicle control problems. In: Proceedings of IEEE ACC, pp. 1491–1496 (2001)
47. Spanos, D.P., Olfati-Saber, R., Murray, R.M.: Dynamic consensus on mobile networks, presented at the 16th IFAC World Congr., Prague, Czech Republic, July 2005
48. Tatikonda, S., Mitter, S.K.: Control under communication constraints. IEEE Trans. Autom. Control **49**(7), 1056–1068 (2004)
49. Tsitsiklis, J.N.: Problems in decentralized decision making and computation. Ph.D. thesis, Massachusetts Institute of Technology (1984)
50. Tsitsiklis, J.N., Bertsekas, D.P., Athans, M.: Distributed asynchronous deterministic and stochastic gradient optimization algorithms. IEEE Trans. Autom. Control **31**(9), 803–812 (1986)
51. Vamsi, A.S.M.: Optimal networked controllers for networked plants. Ph.D. thesis, Department of Electrical and Computer Engineering, Iowa State University (Aug. 2012)

52. Vamsi, A.S.M., Elia, N.: Design of distributed controllers realizable over arbitrary directed networks. In: 49th IEEE Conference on Decision and Control, Dec. 2010, pp. 4795–4800 (2010)
53. Vamsi, A.S.M., Elia, N.: Optimal Distributed Controllers Realizable over Arbitrary Networks. IEEE Trans. Autom. Control. (2011, under review)
54. 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 (2001)
55. Wang, J., Elia, N.: Mean square stability of consensus over fading networks with nonhomogeneous communication delays. In: Proc. IEEE CDC08, pp. 4614–4619 (2008)
56. Wang, J., Elia, N.: Dynamic average consensus over random networks with additive noise. In: IEEE Conference on Decision and Control, pp. 4789–4794 (2010)
57. Wang, J., Elia, N.: Control approach to distributed optimization. In: Allerton Conference on Communication, Control and Computing, 29 Sep.–1 Oct. 2010, 557–561 (2010)
58. Wang, J., Elia, N.: A control perspective to centralized and distributed optimization. Finalist for best student paper award. In: Proc. 50th IEEE CDC-ECC (2011)
59. Wang, J., Elia, N.: Distributed averaging under constraints on information exchange: emergence of Lévy flights. IEEE Trans. Autom. Control **57**(10), 2435–2449 (2012)
60. Wang, J., Elia, N.: Distributed least squares with intermittent communications. In: Proc. 2012 American Control Conference, pp. 6479–6484 (2012)
61. Wang, J., Elia, N.: Distributed averaging algorithms resilient to communication noise and dropouts. IEEE Trans. Signal. Process. (2013). doi:10.1109/TSP.2013.2243438
62. Wang, J., Elia, N.: Mitigation of complex behavior over networked systems: analysis of spatially invariant structures. Automatica (2013). doi:10.1016/j.automatica.2013.02.042
63. Wong, W.S., Brockett, R.W.: Systems with finite communication bandwidth constraints. II Stabilization with limited information feedback. IEEE Trans. Autom. Control **44**, 1049–1053 (1999)
64. Xiao, L., Boyd, S., Lall, S.: A scheme for robust distributed sensor fusion based on average consensus. In: Proc. Int. Conf. Information Processing in Sensor Networks, Los Angeles, CA, pp. 63–70 (2005)
65. Zhou, K., Doyle, J.C., Glover, K.: Robust and Optimal Control. Prentice Hall, New Jersey (1996)
66. Zhu, M., Martinez, S.: On distributed optimization under inequality and equality constraints via penalty primal–dual methods. In: Proc. of 2010 American Contr. Conf., 30 June–2 July 2010, pp. 2434–2439 (2010)
67. Zhu, H., Giannakis, G.B., Cano, A.: Distributed in-network channel decoding. IEEE Trans. Signal Process. **57**(10), 3970–3983 (2009)

Chapter 9
On the Conditional Mutual Information in the Gaussian–Markov Structured Grids

Hanie Sedghi and Edmond Jonckheere

9.1 Introduction

9.1.1 The Smart Grid and Its Possibly Malicious Events

We are concerned with fast and reliable detection of threats in the power grid. This extra capability of the grid to detect a malicious event, even when it is triggered by a sophisticated antagonistic player, is among the attributes that make it *smart*.

Traditionally, the term *grid* is used to refer to an electricity system that supports the following four operations: electricity generation, electricity transmission, electricity distribution, and voltage stability control. In the early days, generation was co-located with distribution in what we would now call a *micro-grid* and the connections among the micro-grids were meant to transmit energy in case of such contingencies as shift in the supply/demand balance. After deregulation, however, a large-scale generation–transmission–distribution *network* became the substitute for the traditional generation–distribution *co-location*. The new network allows consumers to purchase electricity at the cheapest price across the country, as opposed to the former concept in which consumers were forced to purchase electricity from local utility companies. Other considerations calling for an overhaul of the electricity system include the reduction of carbon emission, an objective that cannot be achieved without a significant contribution from the electricity sector. This calls for a bigger share of the renewable energy resources in the generation mix and a supply/demand that must be managed more effectively. Management and control of the grid made increasingly complex by its response to electricity market conditions are,

H. Sedghi (✉) · E. Jonckheere
Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA
e-mail: hsedghi@usc.edu

E. Jonckheere
e-mail: jonckhee@usc.edu

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8_9,

next to its ability to detect contingencies, the most fundamental attributes that make it *smart*.

Automated large-scale management requires considerable exchange of information, so that the smart grid has become a two-commodity flow—electricity and information. By utilizing modern information technologies, the smart grid is capable of delivering power in a more efficient way and responding to wider ranging conditions.

Massive amount of measurements and their transmission across the grid by modern information technology, however, make the grid prone to attacks. Next to malicious events, the potential for fault events with cascading impact on the overall stability of the power grid remains. Today's power systems are not adequately equipped with fault diagnosis mechanisms against various attacks and non-malicious events such as lines sagging in trees, as it had happened right before the 2003 blackout. Thus, there is an urgent need for quick assessment of fault events so that corrective feedback control actions can be taken promptly to avoid cascading events. Fast and accurate detection of possibly malicious events is of paramount importance not only for preventing faults that may lead to blackouts, but also for routine monitoring and control tasks of the smart grid, including state estimation and optimal power flow. Fault localization in the nation's power grid network is known to be challenging, due to the massive scale and inherent complexity.

9.1.2 State Estimator (SE) Versus Phasor Measurement Units (PMUs)

Traditionally, the State Estimator (SE) processes the measurement data from the power meters to reconstruct the state (bus voltages and phase angles). More recently, however, synchronous Phasor Measurement Units (PMUs) with GPS time stamps have been deployed across the grid and are considered the most reliable sensing information to monitor the state of operation of the grid and, if necessary, to respond to contingencies. Even though PMUs are more reliable than SEs, for economical reasons, some parts of the grid will still use state estimators in a foreseeable future. Therefore, any attack—either tampering with the power measurement to the SE or compromising the PMU data, as shown in Fig. 9.6—can harm the power grid.

9.1.3 Outline of Method

In a nutshell, the conceptual foundation of our method is the reconstruction of the graphical model of the phase data.

We use the Conditional Covariance Test for this goal. CCT algorithm can be summarized as follows: Given $i, j \in \mathcal{V}$, given a separator S, that is, a subset of $\mathcal{V} \setminus \{i, j\}$, find the correlation between X_i and $X_j \in \mathbf{X}$ given the separator. If for

all reasonably chosen separators this conditional covariance remains above a certain threshold, then (i, j) is declared an edge. Under some conditions, the resulting $(\mathcal{V}, \mathcal{E})$ is the Markov graph of $\mathbf{X}$.

Next, it is shown that, under normal grid operation, and because of the grid graph structure, the Markov graph of phasors should match the power grid graph; otherwise, a discrepancy should trigger the alarm.

It turns out that our method can detect the most recently contrived attack on the smart grid, which specifically fools the State Estimator, and against which no counter-measures have been suggested thus far [23]. The attack is deemed "sophisticated" in the sense that it assumes knowledge of the bus–branch model of the grid.

9.1.4 Related Work and Exclusivity of Approach

The line fault detection method of [8] is also based on a GMRF model of the PMU data. Besides the fault versus attack detection discrepancy in motivation, the difference between [8] and the present work is two-fold. First, probably the most important contribution of this chapter is to show that the 1-neighbor property is just an approximation, and issue that was not addressed in [8]. Secondly, the fault detection method of [8] utilizes PMUs, whereas here we utilize both PMUs and State Estimator, as shown in Fig. 9.6. For economical reasons, future grids will still contain state estimators in some parts; therefore, [8] and any other monitoring system that does not have a method to check for data integrity can be deluded by such an attack as false data injection.

9.2 Gaussian Markov Random Field (GMRF): General Concept

9.2.1 Graphical Models

Probabilistic graphical models provide diagrammatic representation of probability distributions. This way they set up a simple way to visualize the structure of a probabilistic model and provide insight into properties of the model including conditional independence properties [4].

A graph consists of *nodes* $\mathcal{V}$ connected by *links* $\mathcal{E}$. In a probabilistic graphical model, each node represents a random variable or a group of random variables and the links express the probabilistic dependence relationship among random variables. The graph represents how the joint probability distribution can be decomposed into factors that depend only on a subset of the variables [4].

There are two major classes of graphical models: *Bayesian Networks*, also known as *directed graphical models* where links are directed, and *Markov Random Fields*, also known as *undirected graphical models* where links are not directed [4].

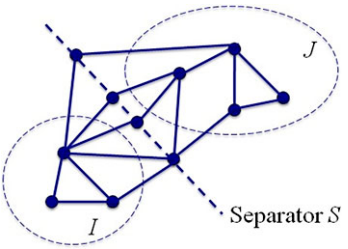


Fig. 9.1 Global Markov property: $X_I \perp X_J | X_S$

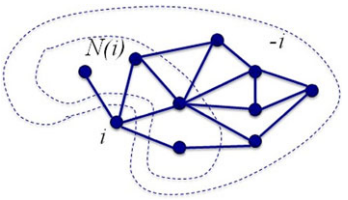


Fig. 9.2 Local Markov property: $E[X_i|X_{N(i)}] = E[X_i|X_{-i}]$

9.2.2 Gaussian Markov Random Field (GMRF)

A probability distribution is said to have *global Markov property* with respect to a graph if, for any disjoint subsets of nodes I, J, S such that S separates I and J on the graph, the distribution satisfies $X_I \perp X_J | X_S$, i.e., X_I is independent of X_J conditioned upon X_S. This is represented in Fig. 9.1.

A distribution is *pairwise Markov* with respect to a given graph if, for any two nodes i and j in the graph such that there is no direct link in the graph between i and j, X_i is independent of X_j given the states of all of the remaining nodes, i.e., $X_i \perp X_j | X_{\mathcal{V}\setminus\{i,j\}}$.

A set of random variables is said to have *local Markov property* corresponding to a graph [13] if any variable X_i is conditionally independent of all other variables X_{-i} given its neighbors $X_{N(i)}$, where $-i := \{j \in \mathcal{V} : j \neq i\}$ and $N(i) := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$. The local Markov property can be seen in Fig. 9.2.

Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a set of random variables $X = (X_v)_{v\in\mathcal{V}}$ form a *Markov Random Field* with respect to $\mathcal{G}$ if they have the global Markov property. It should be noted that the local Markov property and pairwise Markov property are equivalent and they are a special case of global Markov property. For a strictly positive probability distribution, the properties are equivalent and it can be shown that the probability distribution can be factorized with respect to the graph [13].

One instance of this positivity condition happens in case of jointly Gaussian distributions.

A *Gaussian* Markov Random Field (GMRF) is a family of jointly Gaussian distributions that factor in accordance with a given graph. Given a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \{1, \ldots, p\}$, consider a vector of Gaussian random variables $\vec{X} = [X_1, X_2, \ldots, X_p]^T$, where each node $i \in \mathcal{V}$ is associated with a scalar Gaussian random variable X_i. A *Gaussian Markov Random Field (GMRF)* on $\mathcal{G}$ has a probability

density function (pdf) that may be parametrized as

$$f_X(x) \propto \exp\left[-\frac{1}{2}x^T Jx + h^T x\right], \tag{9.1}$$

where J is a positive-definite symmetric matrix whose sparsity pattern corresponds to that of the graph $\mathcal{G}$. More precisely,

$$J(i,j) = 0 \quad \Longleftrightarrow \quad (i,j) \notin \mathcal{E}. \tag{9.2}$$

The matrix $J = \Sigma^{-1}$ is known as the *potential* or *information* matrix, the nonzero entries $J(i,j)$ as the edge potentials, and the vector $\vec{h}$ as the vertex potential vector [2].

In general, the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called the *Markov graph* (graphical model) underlying the joint probability distribution $f_X(x)$, where the node set $\mathcal{V}$ represents each random variable X_i, if the edge set $\mathcal{E}$ is defined in order to satisfy local Markov property. For a Markov Random Field, the local Markov property states that $X_i|\mathbf{X}_{-i} = X_i|X_{N(i)}$, where $\mathbf{X}_{-i}$ denotes all variables except for X_i, and $X_{N(i)}$ denotes all random variables associated with the neighbors of i. Define

$$r_{ij} \triangleq \frac{\Sigma(i,j|\mathcal{V}\setminus\{i,j\})}{\sqrt{\Sigma(i,i|\mathcal{V}\setminus\{i,j\})\Sigma(j,j|\mathcal{V}\setminus\{i,j\})}} \tag{9.3}$$

as the *partial correlation coefficient* between variables X_i and X_j for $i \neq j$ measuring their conditional covariance given all other variables. The joint distribution of the GMRF $\mathbf{X}$ follows $N(\mu, (I-R)^{-1})$, with $\Sigma = (I-R)^{-1}$ being the covariance matrix and $R \triangleq [r_{ij}]$ the matrix consisting of partial correlation coefficients off the diagonal and zeros on the diagonal entries [21] , i.e.,

$$r_{ij} = -\frac{J(i,j)}{\sqrt{J(i,i)J(j,j)}}. \tag{9.4}$$

Therefore,

$$X_i|\mathbf{X}_{-i} \sim N\left(\mu_i + \sum_{j\neq i} r_{ij}(X_j - \mu_j), 1\right), \tag{9.5}$$

where the distribution is normalized to highlight the partial correlations r_{ij}.

Setting $X \sim N(\mu, J^{-1})$, the pairwise Markov property of GMRF implies that $(i,j) \notin \mathcal{E} \Leftrightarrow r_{ij} = 0$.

9.3 Bus Phase Angles as Gaussian Markov Random Field (GMRF)

This section is the "hub" of this whole chapter. Specifically in this section, we examine the extent to which the bus phase angles of the power grid satisfy the conditions

for them to qualify as a GMRF. We discuss the approximation in neighboring property between bus phase angles.

Further, in Sect. 9.4.4, we explain the Conditional Covariance Test [2] as the method we have chosen for finding out the Markov graph of bus phase angles. Next, we explain why CCT method best describes this approximation and why this approximation is, in fact true, for a grid graph.

Finally, in Sect. 9.5, we argue that a discrepancy between the output of CCT and the grid graph structure means that the system is under stealthy deception attack.

9.3.1 AC Power Flow: Review

The AC power flow states that the real power and the reactive power flowing from bus i to bus j are, respectively,

$$P_{ij} = G_{ij}V_i^2 - G_{ij}V_iV_j\cos(\theta_i - \theta_j) + b_{ij}V_iV_j\sin(\theta_i - \theta_j), \tag{9.6}$$

$$Q_{ij} = b_{ij}V_i^2 - b_{ij}V_iV_j\cos(\theta_i - \theta_j) - G_{ij}V_iV_j\sin(\theta_i - \theta_j), \tag{9.7}$$

where V_i and θ_i are the voltage magnitude and phase angle, resp., at bus i and G_{ij} and b_{ij} are the conductance and susceptance, resp., of line ij. From [3], we obtain the following approximation of the AC *fluctuating* power flow:

$$\widetilde{P}_{ij} = (b_{ij}\overline{V}_i\overline{V}_j\cos\overline{\theta}_{ij})(\widetilde{\theta}_i - \widetilde{\theta}_i), \tag{9.8}$$

$$\widetilde{Q}_{ij} = (2b_{ij}\overline{V}_i - b_{ij}\overline{V}_j\cos\overline{\theta}_{ij})\tilde{V}_i - (b_{ij}\overline{V}_i\cos\overline{\theta}_{ij})\widetilde{V}_j, \tag{9.9}$$

where bar means steady-state value, tilde means fluctuation around the steady-state value, and $\overline{\theta}_{ij} = \overline{\theta}_i - \overline{\theta}_j$. These fluctuating values due to renewables and variable loads justify the utilization of probabilistic methods in power grid problems.

Now, assuming that for the steady-state values of voltages we have $\overline{V}_i = \overline{V}_j \simeq 1p.u.$ (per unit), and the fluctuations in angles are about the same such that $\cos\overline{\theta}_{ij} = 1$, we have

$$\widetilde{P}_{ij} = b_{ij}(\widetilde{\theta}_i - \widetilde{\theta}_j), \tag{9.10}$$

and

$$\widetilde{Q}_{ij} = b_{ij}(\widetilde{V}_i - \widetilde{V}_j). \tag{9.11}$$

9.3.2 Gaussian Distribution Assumption: Transmission Versus Distribution Network

The power flow equations can be written, conceptually, as $z = h(x)$, where $z = (P^T, Q^T)^T$ is the vector of (active and reactive) powers injected at the various buses

and $x = (\theta^T, V^T)^T$ is the state, that is, the vector of voltage phase angles and voltage magnitudes at the buses.

Whether the Gaussian distribution assumption on θ is justified depends on two considerations:

1. The nature of the injected power, which could be deterministic or stochastic.
2. The linearized approximation of $z = h(x)$, the *DC power flow equations.*

Regarding the first item, in a high-voltage transmission grid, the aggregate property of the demand justifies the Gaussian distribution assumption. On the other hand, the loads in a low-voltage power distribution network do not correspond to aggregate loads but single consumers. Hence the Gaussian distribution assumption cannot be justified on the ground of the demand. The Gaussian distribution assumption can, however, be justified by the aggregation of such renewables as wind turbines and solar panels, the power output of which is inherently random. It is suggested in [18] that as few as 5 wind turbines would suffice to see the Central Limit Theorem in action, meaning that the power generation would behave like a Gaussian random variable.

We note that the capability of detecting false data is also interesting in the distribution grid, where the high number of inexpensive sensors deployed in the grid could hardly be managed via secure communication channels.[1] For our method—which relies on the Gaussian distribution assumption—to be applicable to the distribution network, it is hence imperative to invoke the aggregation of the renewables. In the general setting where renewables need not be present, our work more realistically applies to the transmission grid, where aggregate demand is present.

Regarding the second item, if aggregate power at buses follows Gaussian distribution, by linearity of DC power flow we can reach the same conclusion for bus phase angles.

9.3.3 DC Power Flow: Active Power Versus Phase Angle

We now apply the preceding to bus phase angles. We would like to show that bus phase angles form a GMRF and then discuss the Markov graph associated with it.

The DC power flow model [1] is often used for analysis of power systems in normal steady-state operations. When the system is stable, the phase angle differences are small. In addition, DC power flow assumes that lines are highly inductive. Therefore, $\sin(\theta_i - \theta_j) \sim \theta_i - \theta_j$. Thus, the power flow on the transmission line connecting bus i to bus j is given by

$$P_{ij} = b_{ij}(X_i - X_j), \tag{9.12}$$

[1] This was brought to our attention by an anonymous referee.

where X_i and X_j denote the phasor angles at bus i and j, respectively, and b_{ij} denotes the inverse of the line inductive reactance. The power injected at bus i equals the algebraic sum of the powers flowing away from bus i:

$$P_i = \sum_{j \neq i} P_{ij} = \sum_{j \neq i} b_{ij}(X_i - X_j). \tag{9.13}$$

In the above formulation, the summation holds since $b_{ij} = 0$ is implied whenever buses i and j are not connected. Thus, it follows that the phasor angle at bus i could be represented as

$$X_i = \sum_{j \neq i} c_{ij} X_j + \frac{1}{\sum_{j \neq i} b_{ij}} P_i, \tag{9.14}$$

where $c_{ij} = \frac{b_{ij}}{\sum_{i \neq j} b_{ij}}$.

Because of load uncertainty in the transmission network, the injected power can be modeled as a random variable [15] and since injected power models the superposition of many independent factors (e.g., loads), it can be modeled as a Gaussian random variable, as already argued in Sect. 9.3.2. Thus, the linear relationship in (9.12) implies that the difference of phasor angles across a bus could be approximated by a Gaussian random variable truncated within $[0, 2\pi)$. Considering the fixed phasor at the slack bus, it is assumed that under steady-state, phasor angle measurements can be considered as Gaussian variables [8].

The next step is to find the correct neighboring relationship between the X_i's.

9.3.4 Local Markov Property: Neighboring Relationship

Here we investigate the extent to which the θ's are in a 1-neighbor relationship, by which we mean the local Markov property, $E(X_i|X_{-i}) = E(X_i|X_{N(i)})$. We look at two idealized cases: an infinite chain-structured bus system and a 2-dimensional lattice-structured bus system.

In Sect. 9.3.4.1 dealing with the idealized chain bus, we consider (9.13) along with independently injected powers P_i's and demonstrate a 2-neighbor relationship between the X_i's, i.e., the X_i's are related to their first and second degree neighbors in the grid graph; precisely, $E(X_i|X_{-i}) = E(X_i|X_{N(i)} \cup X_{N(N(i))\setminus\{i\}})$. This implies that the $J(i, j)$ matrix entry in Eq. (9.1) is nonvanishing if and only if i and j are 1-neighbors or 2-neighbors in the grid graph; in other words, $J(i, j) = 0$, $\forall d_{\text{hop}}(i, j) \geq 3$, where $d_{\text{hop}}(\cdot, \cdot)$ denotes the hop metric defined as the distance on the graph when the link weights are normalized to 1. Furthermore, using the Toeplitz structure of the coefficient matrix of the system of Eq. (9.13) and Fourier transform techniques, Sect. 9.3.4.1 shows that $J(i, j') < J(i, j)$, for $2 = d_{\text{hop}}(i, j') > d_{\text{hop}}(i, j) = 1$, that is, the J matrix entry in Eq. (9.1) for the second-neighbor is smaller than the J matrix entry for the first-neighbor. It is shown

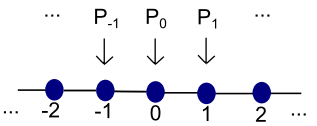


Fig. 9.3 Infinite Line Network

in Sect. 9.3.4.1 that this approximation falls under the generic fact of the tapering off of Fourier coefficients.

In Sect. 9.3.4.2, a similar result is demonstrated to hold for the idealized 2-dimensional lattice-structured grid.

Thus, we can approximate the neighboring relationship to be that of immediate neighbors in grid graph,

$$E[X_i|X_{-i}] \simeq E[X_i|X_{N(i)}]. \tag{9.15}$$

Therefore, we have an *approximate* local Markov property. It is conjectured that such an approximation holds whenever the grid has enough symmetry to allow for Toeplitz and related Fourier transform techniques.

9.3.4.1 Independent Power Injection to an Infinite Chain

Consider a doubly infinite homogeneous chain-structured power network with $b_{i,i+1} = 1$ as shown in Fig. 9.3.

The DC power flow equations, $P = BX$, in this specific case take the format

$$\begin{bmatrix} \vdots \\ P_{-1} \\ P_0 \\ P_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & \ddots & & & & \\ \ddots & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & 2 & \ddots & \\ & & & \ddots & \ddots & \end{bmatrix} \begin{bmatrix} \vdots \\ X_{-1} \\ X_0 \\ X_1 \\ \vdots \end{bmatrix}. \tag{9.16}$$

Because of the symmetry of the problem, B is a doubly-infinite Toeplitz matrix, also referred to as Laurent operator. By "Toeplitz matrix," we mean a matrix whose (i, j) entry depends only on the difference of indexes, $i - j$. Equivalently, a matrix with constant entries on the diagonal, constant entries on the super-diagonals, and constant entries on the sub-diagonals is a Toeplitz matrix, as can be seen from Eq. (9.16).

Besides the usefulness of chains as testbeds for networks with shift invariant properties, here, the most compelling justification is that the doubly-infinite chain structure secures $\sum_{k=-\infty}^{+\infty} P_k = 0$, as easily seen from the cancellation of the sum of the column elements of B, but subject to some convergence issues, which are now straightened out.

To remain within the Hilbert space setup, we restrict $X \in \ell^2(-\infty, +\infty)$. Next, it can be seen that the B-operator is bounded. Indeed, taking the Fourier transform of P and X,

$$\widehat{P}\big(e^{j\alpha}\big) = \sum_{k=-\infty}^{+\infty} P_k e^{jk\alpha}, \qquad \widehat{X}\big(e^{j\alpha}\big) = \sum_{k=-\infty}^{+\infty} X_k e^{jk\alpha}, \tag{9.17}$$

we get

$$\widehat{P}\big(e^{j\alpha}\big) = \big(2 - e^{j\alpha} - e^{-j\alpha}\big)\widehat{X}\big(e^{j\alpha}\big). \tag{9.18}$$

From the above, we notice that B is a multiplication operator in the Fourier domain:

$$\widehat{P}\big(e^{j\alpha}\big) = \widehat{B}\big(e^{j\alpha}\big)\widehat{X}\big(e^{j\alpha}\big), \qquad \widehat{B}\big(e^{j\alpha}\big) := 2(1 - \cos\alpha), \tag{9.19}$$

where $\widehat{B}$ is referred to as the *symbol* of the operator. Clearly, the multiplication operator $\widehat{B} : L^2[0, 2\pi) \to L^2[0, 2\pi)$ is bounded and since the Fourier transform is a Hilbert space isometry the operator $B : \ell^2(-\infty, +\infty) \to \ell^2(-\infty, +\infty)$ is bounded as well. This secures $P \in \ell^2(-\infty, +\infty)$ and hence gives sense to $\sum_{k=-\infty}^{+\infty} P_k = 0$.

In order to determine the neighboring structure of a chain-generated random phase angle vector $\vec{X}$, we assign a normal distribution to P. This results in $\mathbf{X}$ having a Gaussian distribution

$$f_X(P) \sim e^{-\frac{1}{2}P^T \Sigma_d^{-1} P} = e^{-\frac{1}{2}X^T B^T \Sigma_d^{-1} BX}, \tag{9.20}$$

where the covariance Σ_d is a trace class operator, a condition necessary to secure $\int_{\ell^2(-\infty,+\infty)} e^{-\frac{1}{2}P^T \Sigma_d^{-1} P} \Pi_{k=-\infty}^{+\infty} dp_k < \infty$ along with the Gaussian property of the projection of the infinite-dimensional distribution on a finite-dimensional space [19, Proposition 1.8], [17]. We take $\Sigma_d = \operatorname{diag}\{\sigma_{d,k}^2 : k = \cdots, -1, 0, +1, \ldots\}$ with $\sigma_{d,k} = 1$ for $|k| \le d$, and $\lim_{|k|\to\infty} \sigma_{d,k} = 0$ with $\sigma_{d,k} > 0$ for $|k| > d$, and such that $\sum_{k=-\infty}^{+\infty} \sigma_{d,k}^2 < \infty$. With this covariance,

$$\lim_{d\to\infty} e^{-\frac{1}{2}X^T B^T \Sigma_d^{-1} BX} = e^{-\frac{1}{2}X^T B^2 X}, \tag{9.21}$$

where B^2 is doubly-infinite Toeplitz as well with symbol

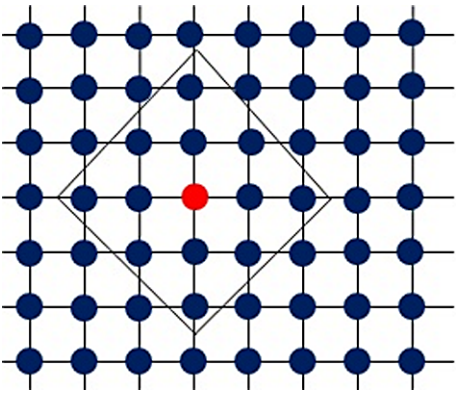
$$\widehat{B^2}\big(e^{i\alpha}\big) = \big[2(1 - \cos\alpha)\big]^2 = 6 - 8\cos\alpha + 2\cos 2\alpha. \tag{9.22}$$

From the above, it follows that

$$f_X(x) \propto \exp\left(-\frac{1}{2}X^T \begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & 1 & -4 & 6 & -4 & 1 & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} X\right). \tag{9.23}$$

According to (9.23), we can see a two-neighbor correlation between the X_i's. It can also be seen that the coefficients for the second neighbors are smaller than those

Fig. 9.4 Euclidean Lattice



of the first neighbors. It should be noted that a power grid is not infinite, hence the infinite Toeplitz structure is an idealization.

9.3.4.2 Euclidean Lattice

The preceding can be generalized to an infinite 2-dimensional Euclidean lattice. 2-dimensional Euclidean lattice is depicted in Fig. 9.4. Given a 2-dimensional lattice with vertices with integer coordinates $\{(k,l)\}_{k,l\in\mathbb{Z}}$, the neighboring relationship is $N((k,l)) = \{(k\pm 1, l), (k, l\pm 1)\}$. In other words, the susceptance $b_{(k,l),(m,n)}$ between nodes (k,l) and (m,n) is nonvanishing only if either $m = k\pm 1$ and $l = n$ or $m = k$ and $n = l\pm 1$. As in (9.17), we define the 2-dimensional Fourier transforms as

$$\widehat{P}\big(e^{j\alpha}, e^{j\beta}\big) = \sum_{k,l\in\mathbb{Z}} P_{k,l} e^{jk\alpha} e^{jl\beta}, \qquad \widehat{X}\big(e^{j\alpha}, e^{j\beta}\big) = \sum_{k,l\in\mathbb{Z}} X_{k,l} e^{jk\alpha} e^{jl\beta}.$$

As in (9.19), the DC power flow equations can be written as

$$\begin{aligned}\widehat{P}\big(e^{j\alpha}, e^{j\beta}\big) &= \widehat{B}\big(e^{j\alpha}, e^{j\beta}\big)\widehat{X}\big(e^{j\alpha}, e^{j\beta}\big),\\ \widehat{B}\big(e^{j\alpha}, e^{j\beta}\big) &= \big(4 - e^{j\alpha} - e^{-j\alpha} - e^{j\beta} - e^{-j\beta}\big),\end{aligned}$$

where $\widehat{B} : L^2([0,2\pi)^2) \to L^2([0,2\pi)^2)$ is the susceptance operator. In order to write the equivalent of (9.20) for a 2-dimensional lattice, we use Parseval's theorem as a representation of $\sum_{k,l\in\mathbb{Z}} P_{k,l}^2$ as a quadratic function of $X_{k,l}$:

$$f_X(P) \propto e^{-\frac{1}{2}\sum_{k,l\in\mathbb{Z}} P_{kl}^2} = e^{-\frac{1}{2}\frac{1}{2\pi}\oint\oint |\widehat{B}|^2 |\widehat{X}|^2 d\alpha d\beta}$$

Quadratic functions of *2-indexed* variables do not lend themselves to obvious matrix representation. The guiding idea here is to collect those $X_{k,l}$'s that are contributing

to $\sum_{k,l\in\mathbb{Z}} P_{k,l}^2$. Those $X_{k,l}$'s are the coefficients of the zeroth powers of $e^{j\alpha}$ and $e^{j\beta}$ in the integrand. Given

$$\begin{aligned}|\widehat{B}|^2 = 20 &- 8e^{j\alpha} - 8e^{-j\alpha} - 8e^{j\beta} - 8e^{-j\beta} \\ &+ 2e^{j(\alpha+\beta)} + 2e^{-j(\alpha+\beta)} + 2e^{j(\alpha-\beta)} + 2e^{-j(\alpha-\beta)} \\ &+ e^{2j\alpha} + e^{-2j\alpha} + e^{2j\beta} + e^{-2j\beta}\end{aligned}$$

and

$$|\widehat{X}|^2 = \sum_{k,l,m,n\in\mathbb{Z}} X_{k,l}X_{m,n}e^{j(k-m)\alpha}e^{j(l-n)\beta},$$

it is not hard to see that

$$\begin{aligned}\sum_{k,l\in\mathbb{Z}} P_{k,l}^2 = \sum_{k\in\mathbb{Z}} \big(&20X_{k,k}^2 - 8X_{k,k}X_{k+1,k} - 8X_{k,k}X_{k-1,k} - 8X_{k,k}X_{k,k+1} \\ &- 8X_{k,k}X_{k,k-1} + 2X_{k,k}X_{k+1,k+1} + 2X_{k,k}X_{k-1,k-1} \\ &+ 2X_{k+1,k-1} + 2X_{k,k}X_{k-1,k+1} + X_{k,k}X_{k+2,k} \\ &+ X_{k-2,k}X_{k,k} + X_{k,k}X_{k,k+2} + X_{k,k}X_{k,k-2}\big).\end{aligned}$$

Clearly, $\sum_{k,l\in\mathbb{Z}} P_{k,l}^2$ is quadratic in the $X_{k,l}$ variables, but those variables that are multiplied have their indexes within at most a 2-neighbor relationship in the lattice structure. To be somewhat more specific, what we learn over the 1-dimensional case is that the correlations decay with the ℓ^2-distance on the lattice. Indeed, for $d_{\ell^2}((k,k),(k+1,k)) = 1$, the canonical correlation $r_{(k,k),(k+1,k)} \propto 8$; for $d_{\ell^2}((k,k),(k+1,k+1)) = \sqrt{2}$, the canonical correlation $r_{(k,k),(k+1,k+1)} \propto 2$; and for $d_{\ell^2}((k,k),(k+2,k)) = 2$, the canonical correlation $r_{(k,k),(k,k+2)} \propto 1$.

As a word of technical warning, the $f_X(P)$ expression should have been written $e^{-\frac{1}{2}\sum_{k,l\in\mathbb{Z}} \Sigma_{d,(k,l)}^{-1} P_{k,l}}$, where $\sum_{k,l} \Sigma_{d,(k,l)} < \infty$ and $\Sigma_{d,(k,l)} = 1$ for $\|(k,l)\|_{\ell^2} \le d$ and $\Sigma_{d,(k,l)} \downarrow 0$ as $\|(k,l)\|_{\ell^2} \to \infty$. This brings some tempered coefficients in the correlations, which have no effect unless for $\|(k,l)\|_{\ell^2} \to \infty$. Working out this technicality explicitly would have, however, resulted in substantial clutter in the notation.

9.3.5 *Reactive Power Versus Voltage Amplitude*

It is clear from (9.10), (9.11) that we can follow the same discussion we had about real power and voltage angles with reactive power and voltage magnitudes.

It can be argued that, as a result of uncertainty, the aggregate reactive power at each bus can be approximated as a Gaussian random variable and, because of Eq. (9.11), voltage fluctuations around the steady-state value can be approximated

as Gaussian random variables. Therefore, the same path of approach as for phase angles can be followed to show the GMRF property for voltage amplitudes. Comparing (9.11) with (9.12) makes it clear that the same matrix, i.e., the B matrix developed in Sect. 9.3.4, is playing the role of correlating the voltage amplitudes; therefore, assuming that the statistics of the active and reactive power fluctuations are similar, the underlying graph is the same. This can be readily seen by comparing (9.10) and (9.11).

Therefore, voltage magnitudes provide another perspective for developing a graphical model underlying the grid structure.

The dual of our approach (linear relationship between reactive power and voltage magnitude) could be generalized to include line loss by linearizing Eq. (2) of [5] to produce a linear relationship between voltage, active and reactive power.

9.4 Model Selection

In the context of graphical models, model selection means finding the real underlying Markov graph among a group of random variables based on samples of those random variables. There are two main classes of methods for learning the structure of the underlying graphical model: convex and non-convex methods. ℓ_1-regularized maximum likelihood estimators are the main class of convex methods [7, 9, 20]. In these methods, the inverse covariance matrix is penalized with a convex ℓ_1-regularizer in order to encourage sparsity in the estimated Markov graph structure. Other types of methods are the non-convex or greedy methods [2]. As we are faced with GMRF in our problem, it would be useful to exploit one of these structure learning methods.

We have decided to use the new Gaussian Graphical Model Selection method called *Conditional Covariance Test (CCT)* [2].

It is proven in [2] that two nodes are connected in the Markov graph iff the conditional mutual information between those measurements is greater than a threshold. For Gaussian variables, testing conditional mutual information is equivalent to Conditional Covariance Test.

In order to have structural consistency, the model should satisfy two important properties:

1. α-walk-summability,
2. (γ, η)-local separation property.

9.4.1 α-Walk Summability

A Gaussian model is said to be α-*walk-summable* if $\|\bar{\mathbf{R}}\| \leq \alpha < 1$ where $\bar{\mathbf{R}} := [|r_{ij}|]$ and where $\|\cdot\|$ denotes the spectral or 2-norm of a matrix, which for symmetric matrices is given by the maximum absolute eigenvalue [2]. r_{ij} is defined in (9.3) and (9.4). The power grids that we considered satisfied this criteria.

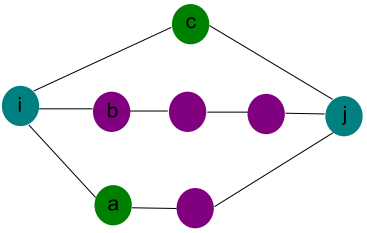


Fig. 9.5 Local Separation Property: $\gamma = 3$. $N(i) = \{a, b, c\}$ is the neighborhood of i and the γ-local separator set $S(i, j; G, \gamma) = \{a, c\}$

9.4.2 Local Separation Property

An ensemble of graphs has the (η, γ)-*local separation property* if for any $(i, j) \notin \mathcal{E}(\mathcal{G})$, the maximum number of paths between i, j of length at most γ does not exceed η [2]. The local separator concept is depicted in Fig. 9.5.

The power grid structure is an example of bounded local path graphs that satisfy the local separation property.

9.4.3 Conditional Mutual Information

Mutual information between two random variables is a quantity that measures the mutual dependence between the two random variables. In the case of continuous random variables, mutual information between random variables X and Y can be defined as

$$I(X;Y) = \int_X \int_Y f(x,y) \log\left(\frac{f(x,y)}{f(x)f(y)}\right) dx\, dy, \tag{9.24}$$

where $f(x, y)$ is the joint probability density function of X and Y, and $f(x)$ and $f(y)$ are the marginal probability density functions of X and Y, respectively.

Mutual information can be defined in terms of entropies as follows:

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X,Y) \\ &= H(X,Y) - H(X|Y) - H(Y|X), \end{aligned} \tag{9.25}$$

where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, and $H(X, Y)$ is the joint entropy of X and Y.

In the special case of Gaussian distributed random variables, for an N-dimensional Gaussian random vector $\vec{Z}$, we have

$$H(\vec{Z}) = \frac{1}{2} \log\big((2\pi e)^N |\Sigma|\big), \tag{9.26}$$

where Σ is the covariance matrix of $\vec{Z}$ [6]. This implies that the mutual information between two N-dimensional Gaussian variables is

$$I(\vec{X};\vec{Y}) = \frac{1}{2}\log\left(\frac{|\Sigma_{\vec{X}\vec{X}}||\Sigma_{\vec{Y}\vec{Y}}|}{|\Sigma_{\vec{X}\vec{Y}}|}\right), \tag{9.27}$$

where

$$\Sigma = \begin{bmatrix} \Sigma_{\vec{X}\vec{X}} & \Sigma_{\vec{X}\vec{Y}} \\ \Sigma_{\vec{Y}\vec{X}} & \Sigma_{\vec{Y}\vec{Y}} \end{bmatrix}. \tag{9.28}$$

The *conditional mutual information* is, in its most basic form, the expected value of the mutual information of two random variables given the value of a third, that is,

$$I(X;Y|Z) = E_z\big(I(X;Y)|Z\big). \tag{9.29}$$

This can be rewritten as [16]

$$I(X;Y|Z) = H(X,Z) + H(Y,Z) - H(X,Y,Z) - H(Z). \tag{9.30}$$

Conditional mutual information can also be written in terms of conditional entropies:

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z). \tag{9.31}$$

Therefore, considering two Gaussian random variables X_i, X_j, conditional mutual information between these two random variables conditioned on a set of random variables X_S is given by (see [2])

$$I(X_i;X_j|X_S) = -\frac{1}{2}\log\big[1-\rho^2(i,j|S)\big], \tag{9.32}$$

where $\rho(i,j|S)$ is the conditional correlation coefficient, given by

$$\rho(i,j|S) := \frac{\Sigma(i,j|S)}{\sqrt{\Sigma(i,i|S)\Sigma(j,j|S)}}. \tag{9.33}$$

As a result, for Gaussian random variables, for testing conditional independence, testing conditional mutual information is equivalent to testing conditional covariances [2].

If the distributions deviate from Gaussian, the conditional mutual information can still be derived from (9.32), (9.33), provided $\Sigma(i,j|S)$ is interpreted as the conditional correlation of $g_i(X_i)$ and $g_j(X_j)$, where the g_i's are nonlinear processing functions aimed at maximizing the correlation.

To be more precise, let $\rho_{g_i,g_j,g}(i,j|S)$ be the correlation coefficient between $g_i(X_i)$ and $g_j(X_j)$ conditioned upon $g(X_S)$, where g_i, g_j, and g are measurable functions. Then by nonlinear processing of X_i and X_j with the distortion functions g_i and g_j, the canonical correlation $\rho_{g_i,g_j,g}(i,j|S)$ can be made to increase towards

the mutual information:

$$\sup_{g_i, g_j, g} \left(-\frac{1}{2} \log\left(1 - \rho^2_{g_i, g_j, g}(i, j|S)\right) \right) \leq I(X_i; X_j|X_S)$$

(see [10, Corollary 1]). Furthermore, the supremum can be achieved if $g_i(X_i)$ and $g_j(X_j)$ can be made jointly Gaussian conditioned upon $g(X_S)$ (see [10, Theorem 3]). A computational procedure that precisely implements this idea is available in [10, Sect. 6]. A related computational implementation is the sequential selection [11, 12]. A simplified numerical procedure based on the canonical correlation between the powers of X_i and the powers of X_j is available in [22].

9.4.4 Conditional Covariance Test (CCT)

The Conditional Covariance Test is introduced in [2]. Using CCT method, the conditional covariance is computed for each node pair $(i, j) \in \mathcal{V}^2$ and the conditioning set that achieves the minimum, over all subsets of other nodes of cardinality at most η, is found. If the minimum value exceeds the threshold $\xi_{n,p}$, then the node pair is declared as an edge.

It is shown in [2] that under walk-summability the effect of faraway nodes on covariance decays with the distance and the error in approximating the covariance by local neighboring relationship decays exponentially with the distance. Thus by correct tuning of threshold and having sufficiently many samples, we expect the output of CCT method to follow the grid structure.

It has been shown that this method is superior to the ℓ_1-method [7, 20], as CCT distributes edges fairly uniformly across the nodes, while the ℓ_1-method tends to cluster all the edges together between the "dominant" variables, leading to a densely connected component and several isolated points [2]. Therefore, CCT is more suitable for constructing the structure of the power grid from measurements.

9.5 Stealthy Deception Attack

The most recent false data injection attack on the power grid has recently been introduced in [23]. For a p-bus electric power network, the $l = 2p - 1$ dimensional state vector $\vec{x}$ is $(\theta^T, V^T)^T$, where $\vec{V} = (V_1, \ldots, V_p)$ is the vector of voltage bus magnitudes and $\vec{\theta} = (\theta_2, \ldots, \theta_p)$ the vector of phase angles disregarding the slack bus for which $\theta_1 = 0$. It is assumed that the nonlinear measurement model for the state estimation is defined by

$$\mathbf{z} = h(\mathbf{x}) + \varepsilon, \tag{9.34}$$

where $h(\cdot)$ is the nonlinear measurement-valued function and $\vec{z}$ is the m-dimensional measurement vector consisting of active and reactive power measurements.

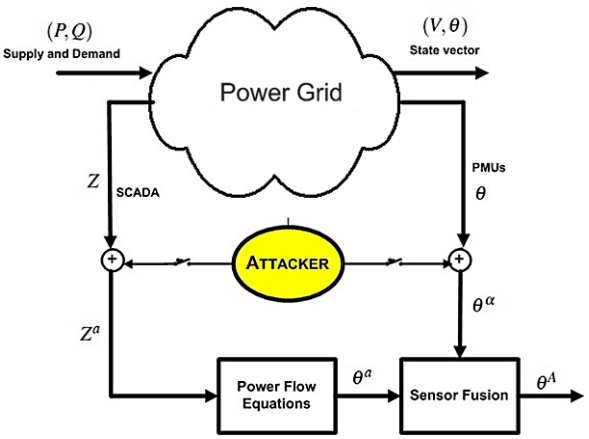


Fig. 9.6 Power grid under a cyber-attack

$H(x^k) := \{\frac{\partial h_i(\vec{x})}{\partial x_j}|_{\vec{x}=x^k}\}_{1\leq i\leq m;\, 1\leq j\leq l}$ denotes the *Jacobian matrix* of the measurement model $h(\vec{x})$ at x^k.

According to [23], the goal of a stealthy deception attacker is to compromise the measurements available to the State Estimator (SE) as

$$\vec{z}^a = \vec{z} + \vec{a}, \tag{9.35}$$

where $\vec{z}^a$ is the corrupted measurement and $\vec{a}$ is the attack vector. Vector $\vec{a}$ is designed such that the SE algorithm converges and the attack $\vec{a}$ is undetected by the Bad Data Detection (BDD) scheme. That is, the difference between $\vec{z}^a$ and the $h(x^k)$ is less than the BDD threshold. In addition, for the targeted set of measurements, the estimated values at convergence are closest to the ones compromised by the attack. The goal of attacker is to inject some data into the state estimator such that the system does not recognize that the data is manipulated and acts upon that. Then it is shown that, subject to some limitations, such an attack can be performed with $\vec{a} \in \text{Im}(H)$. The attack vector $\vec{a}$ is designed in such a way that the difference between $\vec{z}^a$ and $\vec{z}$ is the desired value. Figure 9.6 represents the attack.

It is also stated that the introduced attack is only valid if performed locally. The attack is performed under the DC flow assumption. Because of this assumption, only the $H_{P\theta}$ block of the H matrix is considered in the attack calculation and the state vector introduced in [23] reduces to the vector of voltage angles, $\vec{X}$. Since $\vec{a} \in \text{Im}(H)$,

$$\vec{z}^a = \vec{z} + \vec{a} = H(\vec{X} + d). \tag{9.36}$$

Thus, we have

$$Hd = \vec{z}^a - H\vec{X} = \vec{a}, \tag{9.37}$$

where $H = H_{P\theta}$, $\vec{z}^a$ is the attacker's goal and $\vec{X}$ is the phasor angle vector. Considering (9.13), we have $H_{ij} = -b_{ij}$ for $i \neq j$ and $H_{ii} = \sum_{i\neq j} b_{ij}$, where b_{ij} denotes the inverse of the line inductive reactance. Clearly, H is structured as a weighted graph

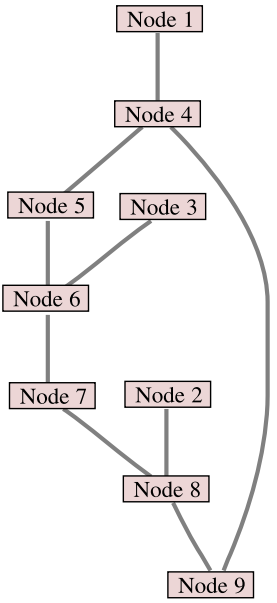


Fig. 9.7 Evaluated 9-node network

Laplacian. By "weighted graph Laplacian structure" we mean a symmetric matrix with its (i, j) entry that can be interpreted as the negative of the "conductance" of the (i, j) link and its (i, i) diagonal element equal to minus the sum of the other elements in row i or column i. This is clearly a generalization of the combinatorial graph Laplacian, where the "conductances" are normalized to 1.

Analysis of (9.36) and (9.37) shows that *the Markov graph of an attacked system changes from the grid graph.* We use this to trigger the alarm.

It should be emphasized that the attack considered here assumes the knowledge of the system's bus–branch model. Hence under this scheme the attacker is equipped with a wealth of information. Yet, we can detect such a *strong* attack with our method.

9.6 Simulation

We considered a 9-node grid suggested by Zimmerman et al. [25]. The structure is shown in Fig. 9.7. First, we fed the system with Gaussian demand and simulated the power grid. We used MATPOWER [25] for solving the DC power flow equations for various demand and used the resulting angle measurements as the input to CCT algorithm. We used YALMIP [14] and SDPT3 [24] to perform CCT.

With the right choice of parameters and threshold, and enough *uncompromised* measurements, the Markov graph follows the grid structure. Table 9.1 shows the edit distance between the Markov graph and the grid graph that is used to lead us to the correct threshold.

Table 9.1 Normalized edit distance under CCT for Fig. 9.7; measurement size = 400

Threshold	No. of links of Markov graph	Edit distance
0.0037	10	1
0.0038	**9**	**0**
0.0039	7	2

Table 9.2 Stealthy deception attack on the grid shown in Fig. 9.7

No. of attacked nodes	Detection ratio
2	100
3	100
4	100

Next, we introduced the stealthy deception attack to the system. We considered the cases where 2, 3 or 4 nodes are under attack. For each case, we simulated all possible attack combinations. In all attack scenarios, the Markov graph of tampered measurements lacked at least one link that was present in the grid graph, a discrepancy that triggered the alarm. Thus we successfully detected the attack. Table 9.2 summarizes different attack scenarios and the corresponding detection ratios.

It should be noted that Table 9.1 shows the required sample size for tuning the method to a specific network structure. So, it shows the initialization step that is enough to be performed once at the beginning of every network analysis as long as the network topology remains the same. Simulation results in the 9-bus network show that even if only 1 of the samples is corrupted it is enough to secure 100 % detection rate.

It should be noted that since we have made connections between phase angle measurements Markov graph and power grid graph, the method can be performed in a decentralized manner. In addition, as stated in [2], the complexity of CCT method is polynomial.

9.7 Conclusion

We have shown that such statistical learning techniques as the Conditional Covariance Test—which is equivalent to the conditional mutual information test in the Gaussian case—allows us to reconstruct the topology of the power grid as the Markov graph of the phase angle measurements (or the voltage magnitudes) at the buses. One of the main points of this chapter is that phase angle data only approximately satisfies the Markov property relative to the grid, a fact that was overlooked in [8]. As shown in Sect. 9.3.4, correlations indeed extend beyond the 1-neighbor relationship. Nevertheless, since the farther away neighboring relationship is weaker and less significant than the 1-neighbor relationship, as shown here on the 9-bus system, this difficulty can be overcome by correctly choosing the threshold.

Finally, if the phase angle data is compromised, the reconstructed Markov graph will be different from the grid interconnection, even when the attack is "stealthy" and launched with the knowledge of the bus–branch model.

In further work, we would like to demonstrate the same concepts and results on more realistic bus systems.

Acknowledgement This research was supported by LCCC—Linnaeus Grant VR 2007-8646, Swedish Research Council.

References

1. Abur, A., Exposito, A.: Power System State Estimation, Theory and Implementation. Dekker, New York (2004)
2. Anandkumar, A., Tan, V., Huang, F., Willsky, A.: High-dimensional Gaussian graphical model selection: walk summability and local separation criterion. J. Mach. Learn. Res. **13**, 2293–2337 (2012)
3. Banirazi, R., Jonckheere, E.: Geometry of power flow in negatively curved power grid. In: 49th IEEE Conference on Decision and Control, pp. 6259–6264 (2010)
4. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
5. Bolognani, S., Zampieri, S.: A distributed control strategy for reactive power compensation in smart microgrids. Preprint arXiv:1106.5626, Oct. 2012
6. Borga, M.: Learning multidimensional signal processing. Tech. Rep. SE-581 83, Linköping University, Sweden (1998)
7. Friedman, J., Hastie, T., Tibshirani, R.: Sparse inverse covariance estimation with the graphical lasso. Biostatistics (2007)
8. He, M., Zhang, J.: A dependency graph approach for fault detection and localization towards secure smart grid. IEEE Trans. Smart Grid **2**, 342–351 (2011)
9. Janzamin, M., Anandkumar, A.: High-dimensional covariance decomposition into sparse Markov and independence models. Preprint arXiv:1211.0919, Nov. 2012
10. Jonckheere, E., Wu, B.F.: Mutual Kolmogorov–Sinai entropy approach to nonlinear estimation. In: IEEE Conference on Decision and Control, Tucson, Arizona, pp. 2226–2232 (1992). Available at http://eudoxus2.usc.edu/CHAOS/Wu_Jonckheere_Mutual Kolmogorov-Sinai.pdf
11. Larimore, W.E.: Identification and filtering of nonlinear systems using canonical variate analysis. In: Casdagli, M., Eubank, S. (eds.) Nonlinear Modeling and Forecasting. SFI Studies in the Sciences of Complexity, vol. 12. Addison–Wesley, Reading (1992)
12. Larimore, W.E., Baillieul, J.: Identification and filtering of nonlinear systems using canonical variate analysis. In: 29th IEEE Conference on Decision and Control, Honolulu, HI, pp. 635–640 (1990)
13. Lauritzen, S.: Graphical Models. Clarendon, Oxford (1996)
14. Lofberg, J.: YALMIP: a toolbox for modeling and optimization in MATLAB. In: IEEE International Symposium on Computer Aided Control Systems Design (CACSD) (2004). Available from http://users.isy.liu.se/johanl/yalmip/
15. Luettgen, M., Karl, W., Willsky, A., Tenney, R.: Multiscale representations of Markov random fields. IEEE Trans. Signal Process. **41**, 3377–3396 (1993)
16. Makarychev, K., Makarychev, Y., Romashchenko, A., Vereshchagin, N.: A new class of non-Shannon-type inequalities for entropies. Commun. Inf. Syst. **2**(2), 147–166 (2002)
17. Maniglia, S., Rhandi, A.: Gaussian measures on separable Hilbert spaces and applications. Tech. rep., Lecture Notes of the University of Lecce, Italy, Quaderno 1/2004 (2004). ISBN: 88-8305-010-X

18. Mur-Amada, J., Sallán-Arasanz, J.: From turbine to wind farms—technical requirements and spin-off products. In: Krause, G. (ed.) Phase Transitions and Critical Phenomena, vol. 18, pp. 101–132. InTech, Rijeka (2011)
19. Prato, G.D.: An Introduction to Infinite-Dimensional Analysis. Springer, Berlin (2000)
20. Ravikumar, P., Wainwright, M., Raskutti, G., Yu, B.: High-dimensional covariance estimation by minimizing ℓ_1-penalized log-determinant divergence. Electron. J. Stat. **4**, 935–980 (2011)
21. Rue, H., Held, L.: Gaussian Markov Random Fields: Theory and Applications. CRC Press, Boca Raton (2005)
22. Shah, K., Jonckheere, E., Bohacek, S.: Dynamic modeling of Internet traffic for intrusion detection. EURASIP J. Adv. Signal Process. **2007**, 90312 (2007). 14 pp. doi:10.1155/2007/90312
23. Teixeira, A., Dan, G., Sandberg, H., Johansson, K.H.: A cyber security study of a SCADA energy management system: stealthy deception attacks on the state estimator. In: IFAC World Congress (2011)
24. Toh, K.C., Todd, M., Tutuncu, R.H.: SDPT3—a MATLAB software package for semidefinite programming. Optim. Methods Softw. **11**, 545–581 (1999)
25. Zimmerman, R.D., Murillo-Sánchez, C.E., Thomas, R.J.: Matpower steady-state operations, planning and analysis tools for power systems research and education. IEEE Trans. Power Syst. **26**(1), 12–19 (2011)

Author Index

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8, © Springer International Publishing Switzerland 2014

Subject Index

G. Como et al. (eds.), *Information and Control in Networks*,
Lecture Notes in Control and Information Sciences 450,
DOI 10.1007/978-3-319-02150-8, © Springer International Publishing Switzerland 2014