UNDERSTANDING COMPLEX SYSTEMS

Springer: COMPLEXITY

Andreas Pyka
Andrea Scharnhorst
Editors

Innovation Networks

New Approaches in Modelling and Analyzing

Springer

Springer Complexity

Springer Complexity is an interdisciplinary program publishing the best research and academic-level teaching on both fundamental and applied aspects of complex systems – cutting across all traditional disciplines of the natural and life sciences, engineering, economics, medicine, neuroscience, social and computer science.

Complex Systems are systems that comprise many interacting parts with the ability to generate a new quality of macroscopic collective behavior the manifestations of which are the spontaneous formation of distinctive temporal, spatial or functional structures. Models of such systems can be successfully mapped onto quite diverse "real-life" situations like the climate, the coherent emission of light from lasers, chemical reaction-diffusion systems, biological cellular networks, the dynamics of stock markets and of the internet, earthquake statistics and prediction, freeway traffic, the human brain, or the formation of opinions in social systems, to name just some of the popular applications.

Although their scope and methodologies overlap somewhat, one can distinguish the following main concepts and tools: self-organization, nonlinear dynamics, synergetics, turbulence, dynamical systems, catastrophes, instabilities, stochastic processes, chaos, graphs and networks, cellular automata, adaptive systems, genetic algorithms and computational intelligence.

The two major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, and the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Understanding Complex Systems

Founding Editor: J.A. Scott Kelso

Future scientific and technological developments in many fields will necessarily depend upon coming to grips with complex systems. Such systems are complex in both their composition – typically many different kinds of components interacting simultaneously and nonlinearly with each other and their environments on multiple levels – and in the rich diversity of behavior of which they are capable.

The Springer Series in Understanding Complex Systems series (UCS) promotes new strategies and paradigms for understanding and realizing applications of complex systems research in a wide variety of fields and endeavors. UCS is explicitly transdisciplinary. It has three main goals: First, to elaborate the concepts, methods and tools of complex systems at all levels of description and in all scientific fields, especially newly emerging areas within the life, social, behavioral, economic, neuro- and cognitive sciences (and derivatives thereof); second, to encourage novel applications of these ideas in various fields of engineering and computation such as robotics, nano-technology and informatics; third, to provide a single forum within which commonalities and differences in the workings of complex systems may be discerned, hence leading to deeper insight and understanding.

UCS will publish monographs, lecture notes and selected edited contributions aimed at communicating new findings to a large multidisciplinary audience.

Andreas Pyka · Andrea Scharnhorst
Editors

Innovation Networks

New Approaches in Modelling and Analyzing

Springer

Editors
Prof. Dr. Andreas Pyka
Universität Bremen
FB Wirtschaftswissenschaften
LS für Wirtschaftstheorie
Hochschulring 4
28359 Bremen
Germany
pyka@uni-bremen.de
pyka@grenoble.inra.fr

Dr. Andrea Scharnhorst
Royal Netherlands Academy
of Arts & Science
Virtual Knowledge Studio
Joan Muyskenweg 25
1090 HC Amsterdam
Netherlands
andrea.scharnhorst@vks.knaw.nl

ISSN 1860-0832 e-ISSN 1860-0840
ISBN 978-3-540-92266-7 e-ISBN 978-3-540-92267-4
DOI 10.1007/978-3-540-92267-4
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009920221

Cover design: WMXDesign GmbH, Heidelberg

Printed on acid-free paper

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

Acknowledgments

The authors wish to thank Wouter de Nooy, Loet Leydesdorff and Krzysztof Suchecki for their helpful comments. Part of this book is an outcome of the Critical Events in Evolving Networks (CREEN) project, funded by the EU under its 6th Framework, NEST-2003-Path-1, 012684.

Contents

Contributors

Petra Ahrweiler National Institute of Technology Management University College Dublin, Carysfort Avenue Blackrock Co., Dublin, Ireland, petra.ahrweiler@ucd.ie

Alex Arenas Departament de Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Campus Sescelades, Tarragona, Spain

Stefano Battiston Chair of Systems Design, ETH Zurich, Kreuzplatz 5, 8092 Zurich, Switzerland, sbattiston@ethz.ch

Frank Beckenbach Department of Economics, University of Kassel, Nora-Platiel-Str. 4, 34109 Kassel, Germany, beckenbach@wirtschaft.uni-kassel.de

Stefan Bornholdt Institute for Theoretical Physics, University of Bremen, Otto-Hahn-Allee, 28359 Bremen, Germany, bornholdt@itp.uni-bremen.de

Ramón Briegel Department of Economics, University of Kassel, Mönchebergstraße 19, 34109 Kassel, Germany

Robin Cowan BETA, Université Louis Pasteur, 61 Avenue de la Forêt Noire, 67000 Strasbourg, France; UNU-MERIT, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands, r.cowan@merit.unimaas.nl

Maria Daskalakis Department of Economics, University of Kassel, Mönchebergstraße 19, 34109 Kassel, Germany

Albert Diaz-Guilera Departament de Física Fonamental, Universitat de Barcelona, Marti i Franques 1, 08028 Barcelona, Spain

Werner Ebeling Institute für Physik, Humbold-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany, ebeling@physik.hu-berlin.de

Koen Frenken Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands, k.frenken@geo.uu.nl

Nigel Gilbert School of Human Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK, n.gilbert@surrey.ac.uk

Ingrid Hartmann-Sonntag Institute für Physik, Humbold-Universität zu Berlin, Newtonstr. 15, 12489 Berlin, Germany, ebeling@physik.hu-berlin.de

Jarno Hoekman Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Nicolas Jonard Université de Luxembourg,162a Avenue de la Faiencerie, L-1511 Luxembourg, Luxembourg, Nicolas.jonard@uni.lu

Michael D. König Chair of Systems Design, ETH Zurich, Kreuzplatz 5, 8092 Zurich, Switzerland, mkoenig@ethz.ch

Suzanne Kok Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Sergio Lozano Departament de Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Campus Sescelades, Tarragona, Spain

Frank van Oort Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Roderik Ponds Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Andreas Pyka Professor in Economic Theory, Economics Department, University of Bremen, Hochschulring 4, D-28359 Bremen, Germany, pyka@uni-bremen.de Visiting Professor at the TU Delft

Jörg Reichardt Institute for Theoretical Physics, University of Bremen, Otto-Hahn-Allee, 28359 Bremen, Germany, reichardt@physik.uni-wuerzburg.de

Pier Paolo Saviotti Université Pierre Medès, BP 47, 38040 GRENOBLE CEDEX 9, France, ppsavio@grenoble.inra.fr

Andrea Scharnhorst Virtual Knowledge Studio for the Humanities and Social Sciences, Royal Netherlands Academy of Arts and Sciences, Cruquiusweg 31, 1019 AT Amsterdam, The Netherlands, andrea.scharnhorst@vks.knaw.nl

Frank Schweitzer Chair of Systems Design, ETH Zurich, Kreuzplatz 5, 8092 Zurich, Switzerland, fschweitzer@ethz.ch

Joep van Vliet Faculty of Geosciences, Urban and Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

Chapter 1
Introduction: Network Perspectives on Innovations: Innovative Networks – Network Innovation

Andreas Pyka and Andrea Scharnhorst

The idea for this book started when we organized a topical workshop entitled "Innovation Networks – New Approaches in Modeling and Analyzing" (held in Augsburg, Germany in October 2005), under the auspices of Exystence, a network of excellence funded in the European Union's Fifth Framework Program. Unlike other conferences on innovation and networks, however, this workshop brought together scientists from economics, sociology, communication science, science and technology studies, and physics. With this book we aim to build further on a bridge connecting the bodies of knowledge on networks in economics, the social sciences and, more recently, statistical physics.

"Network" is the underlying topic in all chapters. "Innovation" as the emergence of *something new* can be described as an outcome of network activities or as part of the formation (evolution) of networks. Networks and innovation are common "objects" or "targets" in different disciplines. Sometimes the meaning of both notions can differ significantly. One of the aims of this book is to bring these different perspectives together. In this introduction we highlight some of the differences. But, let us first describe the research environment in which this book emerged. This environment is characterized currently by what one could call "network hype."

1.1 Network Hype

In the last decade, networks – understood as "complex networks" – have gained a lot of attention. As part of complexity theory, the network approach seems to be particularly well-suited to describing natural and social phenomena. In a society, sometimes labeled as "the information society" (Castells 1996), where infrastructures and workflows are vital to social life, a network approach seems more than appropriate for exploring and analyzing the complex social, technical, and natural environment of which we are a part. Visualizing the internet structure, traffic, or growth (Dodge and Kitchin 2001), or mapping science and knowledge flows

A. Pyka (✉)
Economics Department, University of Bremen, Hochschulring 4, D-28359 Bremen, Germany
e-mail: pyka@uni-bremen.de

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_1,

(Börner et al. 2007) as networks are just a few of the positive examples of possible applications. Recently, a study was conducted to test the ground for "network science" – social and natural. This idea was triggered by the possible military implications of network theories (Committee on Network Science 2005). A certain "network hype" becomes visible, if we carry out a simple topical search in the "Web of Science."

Figure 1.1 shows the result of a topical search in the Web of Science. This option (TS field) searches for documents containing the search string in the title, abstract, and keywords. The search strings were "graph" OR "graphs" OR "network∗" and " 'complex network' OR 'complex networks.' " We restricted the search to the document type "articles." Because of the overall growth of the database, we plotted the shares, that is, the number of documents for the search string divided by all documents in the database for a particular year. The search results are sensitive towards using "complex network(s)" as a phrase or as combined words.

Searching for the notions "graph(s)" and "network(s)" in the online version of the Web of Science reveals a growth of articles which is over-proportional to the overall growth of documents in this database (Fig. 1.1). A more specific search for " 'complex network' OR 'complex networks,' " shows a rapid growth process, a kind of "phase transition," starting after 2000. According to Scharnhorst (2003), this rapid growth is the result of the emergence of a new area of specialty in statistical physics, namely, complex networks theory. A number of recent review articles and books have also described this new paradigm (cf. Barabási 2002). Networks are gaining more attention and not only in physics. In sociology, for instance,

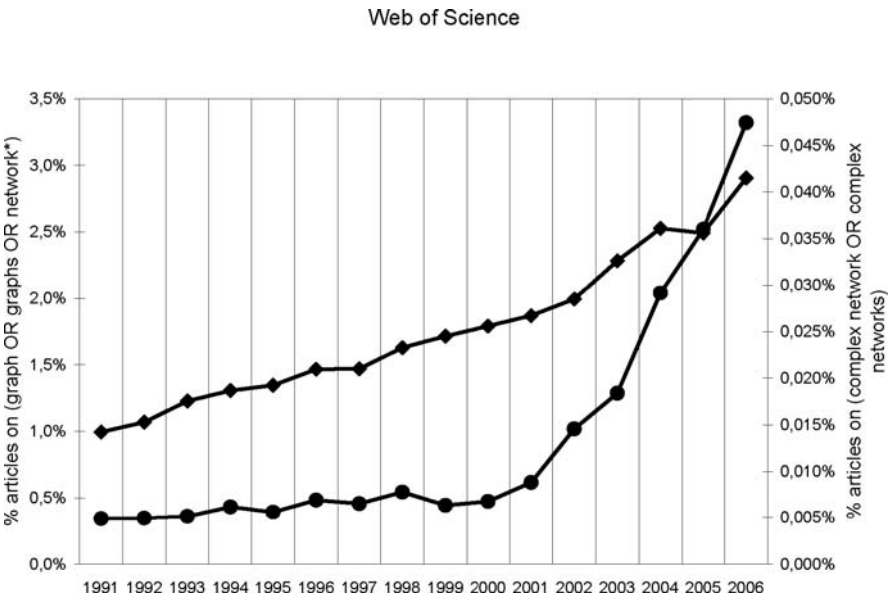


Fig. 1.1 Growth curves for articles on "graph(s)/network(s) ♦" and "complex network(s) •" (Web of Science, November 2007). Please note the two *y* axes

social network analysis has been an established field since the 1950s; in computer and information sciences, in biology, and of course in mathematics (graph theory) networks are central representations of objects and methods (De Nooy, forthcoming). More detailed bibliometric studies have examined the individual, cognitive, and institutional composition of complex network theory (Morris and Yen 2004), and social network theory (Otte and Rousseau 2002). Among the more important pieces of literature are Börner et al. (2007), Bornholdt and Schuster (2003), Buchanan (2002), Dorogovtsev and Mendes (2003), Otte and Rousseau (2002), Newman (2003), and Watts (1999, 2004). Of these, Börner et al. (2007) stand out because they have most recently re-examined network science, considering it as a possible innovation in information science. All the reviews mentioned include efforts to build bridges between different scientific disciplines and specialties. In this book we draw particular attention to the link between evolutionary economics and statistical physics.

Despite this impressive development, claims that an entirely new science has been created (Barabási 2002) have nevertheless been the subject of criticism. In-depth analyses of a subset of "complex networks" contributions (1991–2003) have shown that the notion of "complex networks" was already prevalent in a number of different fields before it became practically a "brand name" or the popular label for a new specialty area in physics, or a new cross-disciplinary paradigm. We do not point to this fact with the aim to further antagonize the battle for scientific priorities; the fields of science history, sociology of science, and philosophy of science have analyzed countless examples of *prematurity* (Hook 2002; Läsker and Plikat 1985), duplicate inventions (Merton 1961; Zuckerman 1967), and reoccurrence of ideas. Independent of the issue of priority, it is also important to take into account that a variety of meanings can be attached to the use of a single phrase. With this work, our aim is to encourage the reader to scrutinize more carefully the different connotations of the network concept. A shared phrase does not necessarily indicate commonalities. However, some hidden commonalities could be present, which would need to be first exposed and then explored systematically. Although false associations can sometimes be inferred from an apparently known concept, real commonalities have the potential to create new links between schools of scientific thought, and to contribute to the discovery and analysis of unexplored areas in the scientific landscape (Scharnhorst 2001). This book intends to make an initial contribution to defining the commonalities and dissimilarities by presenting different approaches in parallel. By creating a platform for them to coexist and where their respective proponents can, in some sense, speak to one another, we hope to have devised a kind of "trading zone" (Galison 1997) for network concepts.

In this introduction we try to identify some dimensions along which different network conceptualizations can be ordered. First, we consider the possibilities of quantitative analysis. To automatically visualize meaning through context is an old and new challenge of communication and information science (Leydesdorff 2001). What is obvious to participants in a discourse, who can identify slight differences in scientific concepts just by perusing various texts, remains almost a "mission impossible" for algorithms. This is the difference to what has also been termed "embodied

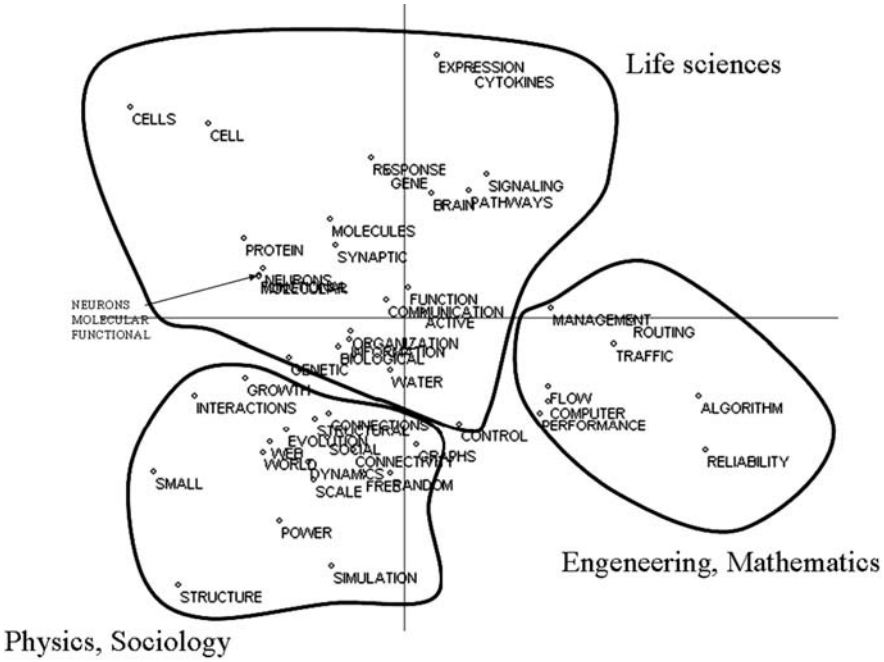


Fig. 1.2 Semantic map of the words used in the titles of complex network articles. (The 1991–2003 dataset was retrieved from the Web of Science.)

and codified knowledge."[1] Figure 1.2 presents an example in the form of a map of words occurring in titles and abstracts of complex network articles. Based on the 1991–2003 sample of articles on complex networks (Scharnhorst 2003), CATPAC[2] was used to produce and analyze a matrix of word co-occurrence and to present the results in a two-dimensional map based on multidimensional scaling. Words close to each other on the map have a tendency to occur more often together in texts.

For a reader familiar with the subject, the new concepts are "scale-free" (networks), "small world," "evolution," and "dynamic" in the fields of physics and sociology. One can also recognize the boundaries of scientific disciplines within the articles themselves. These have been made visible on our map by drawing in the lines afterwards. However, in this map, as Howard D. White, aptly noted, "we do not see the new ideas and paradigms occurring."[3] Nevertheless, these maps

[1] Personal communication Loet Leydesdorff.

[2] CATPAC is a software product marketed by Gallileo. According to their website promotion, "CATPAC is an intelligent program that can read any text and summarize its main ideas." See also: http://www.galileoco.com/N_catpac.asp (accessed 31 May 2008).

[3] Remark by Howard D. White made during the 11th International Conference of the Society for Informetrics and Scientometrics (ISSI) on network visualization, held in Madrid, Spain, 25–27 June 2007.

function well as heuristic tools to disclose hidden information, to reveal undeveloped perspectives, or to push for further clarification (Börner et al. 2003; Börner forthcoming; Thelwall 2008; Thelwall et al. forthcoming). In our case, the bibliometric analysis shows that, despite the tremendous growth of articles on networks in physics since 2000, the notion of "complex networks" (and, even more so, "networks" in general) is a shared object/subject among different disciplines.

1.2 Network Space

In trying to create a visual idea of the different connotations of "network," we borrow the notion of a "characteristics space" from innovation studies. Characteristics space in innovation studies was introduced by Saviotti and Bowman (see Saviotti and Bowman 1984) as analytic tool and to give a visual character to the development of technological output indicators. When the characteristics space is interpreted as space of genotypic or phenotypic properties, it is possible to link it to concepts from evolutionary theory, which portray or visualize evolution as search processes in fitness or adaptive landscape (Pigliucci and Kaplan 2006). A specific formulation of this approach as a generic tool was first developed in physics (Feistel and Ebeling 1989; Ebeling and Scharnhorst 2000; Ebeling et al. 2001) and later applied by Scharnhorst (2001) to different "knowledge landscapes." To be able to identify different dimensions of a conceptual space of networks, let us first examine some perspectives on networks.

Figure 1.3 presents one of the steps towards an ordering of the different network aspects. One relevant aspect is the micro versus the macro dimension. The relationship between laws or rules of behavior, on the micro-level, and emergent properties or laws, on the macro-level, has been analyzed for many different systems in self-organization and complexity theory. From the perspective of networks, individual behavior can be described in terms of networked interactions while, at the same time, emergent networks form a point of reference but sometimes also constitute a constraint on individual actions.

One of the traditional approaches in social network analysis is to measure, analyze, and interpret "ego networks." These are star-like networks (radiating from a single node) in which the social relationships of an individual are followed. Social network analysis "constructs" networks from social relations and their functions in society (Wasserman and Faust 1994). The meaning of ties or links is central to the analysis. The different roles of actors within such a network can be expressed by a link structure. The dominant role that ego networks had during a certain period of social network analysis may possibly be explained by the limitations of data sampling based on interview techniques. Nevertheless, interviewing all members of even a limited group will also yield information about the entire network. Sociometric tradition also employs this type of interview – that is, all members of a group, for instance all of the pupils in a given class.

The advantage of sampling based on interviews is that it delivers detailed insights into the nature of social relationships and the different roles an individual plays

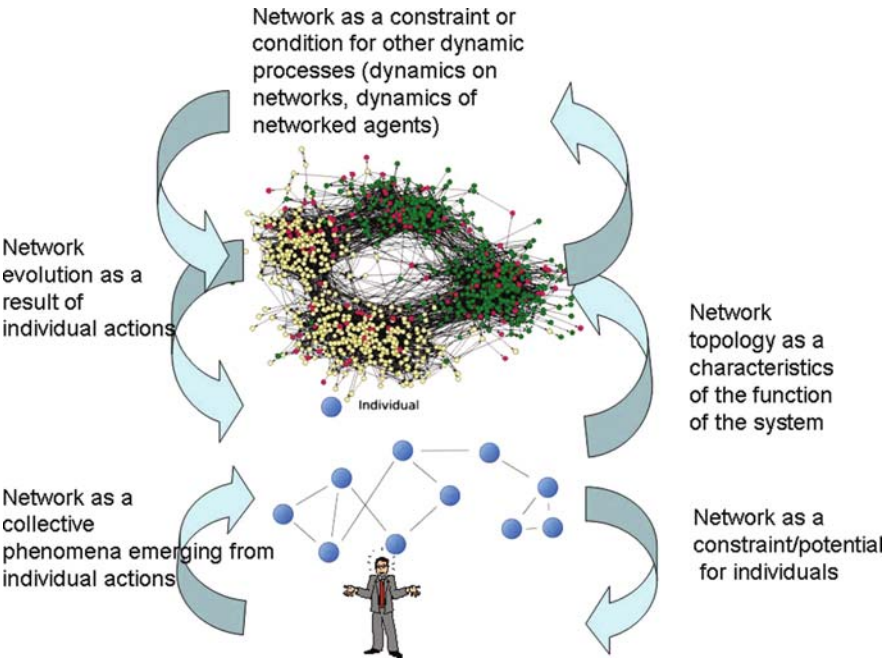


Fig. 1.3 Different aspects of network research

with respect to different communities or different processes. This is information which cannot be directly retrieved using Web 2.0 data[4] in conjunction with social science analysis techniques. This problem has also been addressed in the information sciences (Thelwall 2004) and the political sciences (Foot and Schneider 2006). Large-scale measurements can provide a more complete network, but they leave the question unanswered, as to what kind of "friendship" or other social activity is traced.

A reliable statistical analysis of network topologies requires large amounts of data. The Web and massive web-based, online retrievable data triggered these kinds of large-scale analyses in social science and interdisciplinary fields such as complexity theory (or complex system theory). From statistical physics, "complex networks theory" emerged (Huberman 2001; Scharnhorst 2003). In general, the appearance of this new field (rooted in physics), its universal claims, and a parallel increase in interest in networks in other fields of scientific inquiry can be interpreted as a specific expression of *Zeitgeist*, which reflects both the capabilities and the vulnerability of a globally networked economy and society.

From a physical perspective, the network can be seen as a space[5] with specific characteristics (a certain topology such as scale-free or small world) which influence

[4] Such as that retrievable from MySpace or LinkIn or Facebook.
[5] Personal communication from Krzyzstof Suchecki.

the dynamic processes in that space. When we speak of dynamics in this context, we are referring to the dynamic processes that influence attributes of the nodes and/or the links in the network. Sometimes dynamic processes with known outcomes, which operate on network topologies, can be used to explore these topologies further. At the same time, the evolution of networks is considered to be the determinant for this space – and this is yet another dynamic aspect. Here, the central question is how rules defined locally lead to the emergence of a certain overall topology. If we depend on automatic data mining, then the interpretation of the meaning of the links and nodes fades more into the background. Statistical physics seeks universal laws for mathematical expressions of the dynamic processes going on in networks.

Related to the micro–macro issue is the "emergence" problem, or "the relationship between the whole and its parts," as a further dimension. In this context, graph-theoretical stochastic models (as applied in social network analysis) and physics models meet, although the mathematical language is sometimes different. The common question is how the emergence of a network can be understood from the bottom up, and how we can work backwards from an existing network structure to infer the parameters of individual actions or the conditions under which such actions will occur.

To give an example, by linking back to the micro-level mechanisms of network formation, we can identify critical parameters for structural phase transitions in network topologies (Fronczak et al. 2007), although the question of social meaning for real networks still remains open. A possible link between in-depth, content-specific, systemic structural analysis emerges when networks are analyzed as the outcome of a dynamic process of node and link creation. By building on seemingly free individual actions, self-organization theories have shown that, once a structure begins to emerge from these actions, this structure begins to act as a parameter or limiting condition to constrain individual actions. Recently, in social network analysis, a new network tool, SIENA (Snijders et al. 2007), was developed, which incorporates modeling and simulation into the statistical analysis of real network data. The coupling of generic node and link introduction rules as preferential attachments to social attributes offers us a new way to combine mathematical models with social theories. It gives us the possibility both to test universal assumptions about the basic mechanisms of collective behavior and to refine our descriptions of dynamic mechanisms. It also allows us to further test, empirically and structurally, the social theories of human behavior.

In physics, network topologies are used to predict other dynamic processes such as the spreading of diseases, the movement of traffic, or the dissemination of information. Often two processes co-evolve: the growth and change of an underlying network structure and the dynamics of the processes operating on this structure. The latter are frequently described as the changing status of a node or link. Networks are seen more often as heterogeneous, multilevel, and coupled or combined, which adds another dimension of complexity (cf. Suchecki and Hołyst 2006). In the evolution of networks, characteristics such as connectivity, betweenness, or the diameter change of time point to different functions of networks during different phases of their life cycle (Barabási et al. 2002; Leydesdorff and Schank forthcoming).

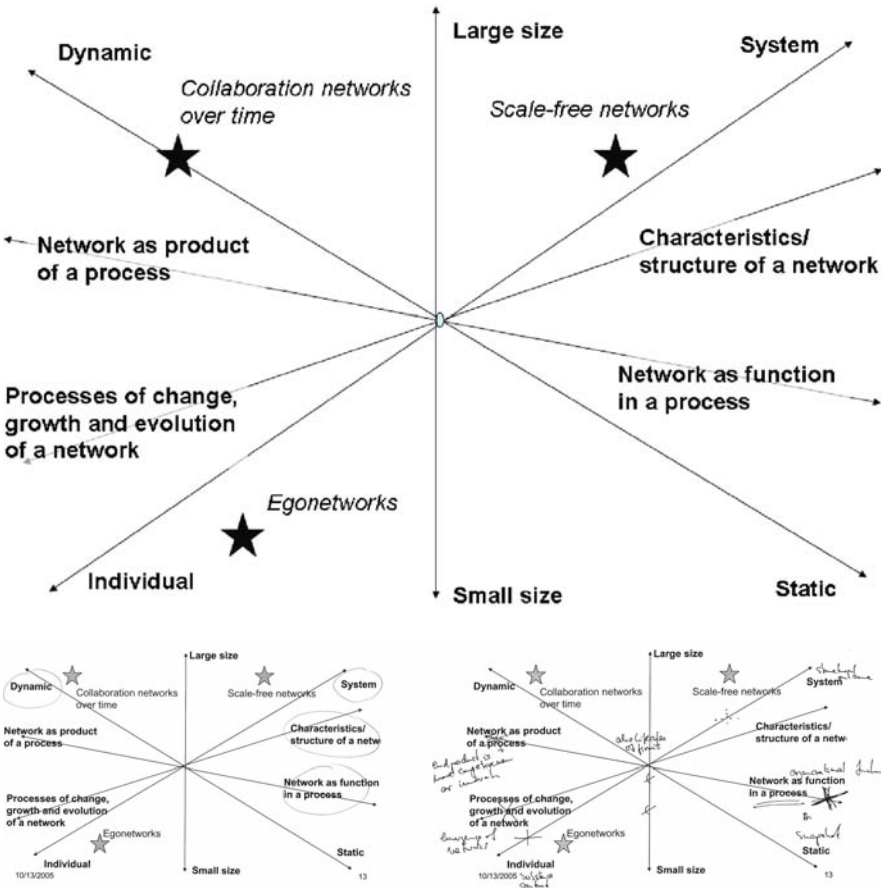


Fig. 1.4 The network space. The chart above was presented to the "Innovation Networks" workshop participants. As shown by the two small figures at the *bottom*, participants located their research by hand drawings (*left*) or even proposed other names for the dimensions (*right*)

In an initial attempt to make the different dimensions of network conceptualizations visible, we draw the following picture (Fig. 1.4) of a multidimensional space of aspects or dimensions (Scharnhorst and Hellsten 2005). We are aware that there are methodologically diverse ways to make the differences in connotations in networks visible. Among them are the above mentioned semantic webs constructed from the co-occurrence of words (Leydesdorff 2001; Leydesdorff and Hellsten 2006), ethnographic field work,[6] interviews, and literature reviews. For our purposes – on the

[6] In the course of discussion about the "space chart," Katie Vann, Virtual Knowledge Studio, Amsterdam, Netherlands, made a critical observation, which the authors of this introduction found pertinent to the present discourse. Her argument was that an *a priori* definition of network space might pre-shape or influence the outcome of our "ethnographic exercise" too much. Moreover,

occasion of the "Innovation Networks" workshop – with the aim to initiate a discussion about networks beyond disciplinary boundaries, this rough attempt worked well. Workshop participants placed their research in the frame presented, and commented critically on the proposed dimensions. This does to take away that it might be possible to use such a frame to make the "occupation" of different use of the network concept visible over time in a more sophisticated way (Chen 2003; Börner et al. 2003).

So far, we have described the research environment which occasioned this anthology. The ultimate aim of building up a new network science constitutes its background. In the previous section of this introduction, we considered the explanatory power of network perspectives, highlighting some dimensions of network research; the subsequent and final section will introduce the reader to the individual chapters of this volume and consider the various attempts by their authors to bridge economic theory and network approaches.

In economics, the concept of network is used in many different contexts, for instance, network externalities, innovation networks, networks of trade and exchange, and corporate social networks. "Interaction and the networks through which it operates are important in determining aggregate economic phenomena and that this allows us to start from more *plausible models of individuals.* . . . If this is accepted, then we must first understand how **networks** *influence aggregate outcomes*. The next step is to understand how these networks form and if, and why, they persist" (Kirman 2003, p. 274). We now turn to the different chapters of the book and consider how they address innovation networks, the methods they use to analyze them, and model networks developed in sociology, economics, and physics.

1.3 Innovation and Networks

In economics, innovation is the implementation of a new or significantly improved idea, good, service, process, or practice, which is intended to be useful or practical in the sense that either efficiency gains or new returns are generated. Further notions which play a role are invention, radical and incremental innovations, process and

---

by presenting a specific set of dimensions we could also conceivably offend other researchers who hold quite different views of networks. Indeed, in our discussion so far and in the chart we propose, we do conceptualize networks graph-theoretically, that is, as objects containing nodes and links. Other meanings are, of course, possible, by which, for instance, the term "network" is used to signify common membership or acquaintance, or to denote an object for identification without intending to further specify it as a graph of relations (Johan Heilbron, personal communication). However, we do not present this chart with the intention of providing a "complete description;" nor do we intend it as starting point for a mapping exercise. At the "Innovation Networks" workshop (Augsburg, Germany, October 2005) this chart was presented to the participants as an input to reflect the positioning of their own research. The presentation provoked neither controversial nor extended discussion. It was perceived as just an interesting "artifact" for stimulating discussion. This "neutral" perception certainly has to do with the variety of disciplinary cultures and the role of visualizations as "helping hands" in natural sciences (Paton and Neilson 1999).

product innovation, competition and logistic growth. In a more abstract systems theoretical approach, innovation can be understood as a critical event which destabilizes the current state of the system, and opens a new process of self-organization leading to a new stable state. In this latter case, the definition of what constitutes an innovation becomes dependent on the definition of the system itself. For instance, what is an innovation for a single firm must not necessarily be an innovation for a market. If the firm is considered as reference system any technology which is new to the firm is a systemic innovation. If, however, the market is the system in which the firm operates, a technology new to the firm does not need to be a technology new on the market and will therefore need not be a systemic innovation. Further, the emergence of new scientific specialties, new streams of communication, and new modes of behavior thus become innovations in a systemic sense.

The contributions to this anthology address different aspects of the relationship between innovation and networks. The chapters incorporate approaches in evolutionary economics, agent-based modeling, social network analysis, and econophysics. Although, all of the chapters belong in some sense to complexity science, the individual contributions nevertheless reflect clearly different epistemic values (theories and methods), depending on their disciplinary roots.[7] These contributions represent a spectrum of approaches between two poles, namely, *new economic theories* (Part I), on the one hand, and *mathematical models of complexity* (Part II) on the other, and therefore are supposed to give a comprehensive view on the present scientific frontier in innovation network analysis.

The authors develop different strategies to handle the epistemic tension between insights into economic processes and insights into new forms of complex dynamics. These approaches range from a conceptual inspiration from network research for empirical measurements of the regional knowledge bases (Frenken et al., this volume) to using qualitative cases of regional modes of innovation as illustration for the explanatory power of a specific model approach (Diaz-Guilera et al., this volume). Although the authors in this anthology come from different research traditions, each chapter represents a unique attempt to create explicit links between the general issues of complexity theory – network research in particular – and new branches in economic theory. This results in a comprehensive approach to innovation networks, spanning a broad range of disciplines and methodologies. We as editors of this anthology attempt in the following to address some of the intersecting links bridging the various chapters.

To be able to understand why innovation networks and why specific approaches to innovation networks have become such a hot topic in economics, we need to compare certain economic theories with actual phenomena in a global, knowledge-based economy. In the chapter by Nigel Gilbert, Petra Ahrweiler, and Andreas Pyka on “Agent-based modeling of innovation networks – the fairytale of technological

[7] Ratto (2006) discussed how epistemic commitments function as barriers to the introduction of new research technologies in scientific fields. He examined the use of new simulation tools in the humanities, in particular, archeology. A similar argument can be made concerning cross-disciplinary research and the co-opting or re-appropriation of concepts across fields.

spillovers" expectations vis-à-vis the role of innovation networks in industrial R&D, as dealt with in economic and organization theory, are compared to observations of collaborative structures between firms in innovative sectors such as biotechnology. The authors justify the introduction of "new models" of innovation behavior on the basis of explanatory gaps and faulty analogies, which can give only incomplete or unsatisfactory account of the relationship between macroeconomic impacts and microeconomic sources of observed phenomena. In particular, technological spillovers find their explanation on the microeconomic level, for instance, through the voluntary exchange of knowledge and information in innovation networks.

In the chapter by Pier Paolo Saviotti on "Knowledge networks: structure and dynamics" the author focuses not on explicit interaction networks but on the representation of knowledge as a network. Saviotti discusses the processes by which scientists have attempted to understand innovation in economic systems. These processes are embedded in a longer tradition of the exchange of ideas between physics and economics. Saviotti draws a line between the traditional approaches to self-organization theory in physics and the approaches of modern network theory. His question is what principal forms of explanation for the emergence of something new can be given by complexity theory. What kind of stylized facts (such as emergence, transition, or growth) can be explained when economics and physics meet. This question is also addressed on a general level in the chapter by Ingrid Hartmann-Sonntag, Andrea Scharnhorst, and Werner Ebeling on "Sensitive Networks – modeling self-organization and innovation processes in networks." More precisely, Saviotti describes the knowledge base of a firm which introduces an innovation as a network of different knowledge resources. Part of these resources is knowledge produced in publicly funded research laboratories (e.g. university laboratories). This research is codified in publications, and can be at least partly retrieved from scientific literature databases such as the Web of Science. Collaboration networks in science form the underlying structure from which new knowledge emerges, which, in turn, might diffuse into industrial applications and there generates another kind of network with other actors and other channels of information exchange.

The various chapters of our anthology collectively add a network perspective as a further dimension to the long research tradition of modeling innovation diffusion. Such models reveal questions for empirical measurement concerning both the development of growth distributions and information flows in an economy. Networks of interaction create a specific locality between firms. From an economic point of view, the question arises, how "neighborhoods" in networks and "neighborhood" in a geographical sense are related to each other, and how this influences or determines collaboration structure. The chapter by Koen Frenken, Jarno Hoekman, Suzanne Kok, Roderik Ponds, Frank van Oort, and Joep van Vliet on "Death of distance science? A gravity approach to research collaboration" shows the influence of geography on collaboration networks in science and by this thematizes again interaction networks. In this contribution, the authors also discuss how the "nature" of a node – in this case the type of organization of an institution – and the reference system (scientific field) exerts influence, and they consider the extent to which "distance" still matters. In line with a recent trend in economics, Frenken et al. concentrate

on geographic regions rather than countries. Pure academic collaboration seems to be less sensitive to distance, perhaps because of the higher degree of codification. In a similar vein, the agent-based model proposed by Frank Beckenbach, Ramón Briegel, and Maria Daskalakis in their chapter on "Evolution and dynamics of networks in 'regional innovation systems' (RIS)" shows the behavioral foundation of local innovation networks embedded in regional innovation systems. Beckenbach et al.'s contribution shows that regional innovation networks indeed form the backbone of the regional innovation systems which do play such a promising role in modern innovation economics (cf., e.g., Leydesdorff 2006; Cooke and Schall 2007).

The knowledge base of an innovation system or an industry or a regional cluster of firms can be measured and made visible in semantic maps of technology classes or patents, or it can be replicated as a reservoir of behavioral variants in a multi-agent model (cf., e.g., Gilbert, Ahrweiler, and Pyka, this volume). The first part of this book introduces networks on different levels of aggregation in the economy (research institutions, firms, and a single enterprise). These diverse networks form the necessary conditions for the emergence of innovation. Gilbert et al. also demonstrate the variety of ways in which networks can be represented in innovation theory, ranging from the network's description as a knowledge base to its depiction as a mode of industrial organization of research and development. From this it becomes clear that innovation itself can also be understood as a result of an interaction process between different firms.

These interaction processes are the roots of networks and their dynamics. If the interaction is modeled as knowledge exchange (as in by Gilbert et al., this volume, and König et al., this volume), then networks of firms and their structures can be depicted as adjacent matrices changing in time. According to Albert Diaz-Guilera and Sergio Lozano in their chapter on "Propagation of innovations in complex patterns of interaction," a visualization of patterns of collaboration and influence is useful if one is seeking structural patterns in different innovation networks and plausible models or explanations for their emergence. Without a rich and detailed qualitative description, formal analogies between different network types remain arbitrary and difficult to understand (see, e.g., the industrial case studies in Pyka and Küppers 2002).

The contributions to this volume demonstrate that different approaches to the modeling of innovation are possible. One such approach would be to understand innovation as the result of competing firms in differently structured interaction networks. The model proposed by Michael König, Stefano Battiston, and Frank Schweitzer in their chapter on "Modeling evolving innovation networks" links the dynamics of a network (evolution of a network by changing nodes and links) to dynamics which operate on the network (the increase of knowledge as attribute of an agent). Unlike many other econophysics applications, the rules of the model are derived from economic theories – costs and different sorts of sharing behaviors are used. Network externalities (often referred to in economic theory as positive externalities) can be traced back to concrete sources. They are modeled in terms of a cross-catalytic support term. Costs are considered the central mechanism determining the "destiny" of a collaborative link and the innovation network as a whole.

König et al. counter other findings in the literature, which assume an almost natural growth of innovation networks in a sector; they point, instead, to the vulnerability of innovation networks. It would be interesting to compare their findings to the growth of the number of networks predicted in the agent-based model and observation data. In general, König et al. point to the fact that in order to obtain a positive balance from the circulation of knowledge, one must find a certain optimum between the benefits from network externalities and the costs of maintaining the network. This positive balance also depends on the nature of knowledge, the effects of which are observable: Is the value of knowledge increased when it is shared, or does mutual knowledge sharing decrease its value? What makes their contribution so important is the demonstration of the heuristic power of models by linking certain basic rules for the behavior of agents to specific phenomena on the macro-level. At the same time, the authors critically discuss their choice of rules, in the light of the realistic behavior of firms on an economic market. König et al. make a specific connection between a population-dynamic approach (for the attributes of the agents) and changing network structure. The adjacency matrix determines when interaction takes place. Each agent interacts with other agents according to the network structure. This is different from an approach in which types of agents are grouped together, depending on the type of knowledge they are using, and where their interactions form a graph or network which usually will not change over time (see Hartmann-Sonntag et al., this volume). Innovation in the model proposed by König et al. is the increase in knowledge of a particular agent and eventually throughout the whole network. Innovation in the Sonntag et al. model is the emergence of a new type of knowledge. Innovation in the Gilbert et al. model is the result of new combinations of already existing knowledge (so-called cross-fertilization effects).

If innovation is understood as a dynamic process running *on* a certain network structure, one can ask about the influence of different typologies on this process. This issue was dealt with by Nicolas Jonard and Robin Cowan in their chapter on "Innovation networks and the distribution of knowledge." If innovation is understood as a critical event triggering diffusion processes, models of self-organized criticality can be used to describe single and overlapping avalanches (see Diaz-Guilera and Lozano, this volume). The question thus arises as to whether the mechanisms assumed in the model building approaches are reasonable to account for the social interaction patterns in economics or science. Where multi-agent approaches (Gilbert et al., this volume) try to incorporate the complexity of decision-making processes for innovations and show how this relates to collaboration, network models from physics try to find regularities in decision-making processes (for instance, considering the board of directors of various enterprises) by looking at mutual membership in different networks (König et al., this volume).

Another important aspect to the statistical analysis of network innovation data is the identification of patterns. This relates to the more general problem of community detection in networks, dealt with by Stefan Bornholt and Jörg Reichardt in their chapter on "Social network analysis – a physics approach." This chapter presents a specific algorithm for clustering. The authors link their findings to approaches in social network analysis. In particular, they show that random graphs (of finite

size) can also emulate "structures." For an analysis of data for innovation networks between firms, such a methodological foundation is absolutely necessary, because these networks are usually small, sparse, and nodes have only low average degrees. Conversely, the use of web data (such as that which one could gather from ebay, for instance) by physicists could initiate a search in economics, sociology, and science and technology studies for new sources of data.

The aim of this anthology is to create a negotiation platform for notions and research intentions of the different groups analyzing networks, and innovation networks in particular. As can be seen from this introduction, the borderline between approaches stemming from economics and from complexity theories becomes extremely fuzzy. Scholars from both disciplines are increasingly willing to step over this borderline. Today's research on innovation networks clearly shows in both camps the application orientation stemming from social sciences and the application of complexity theories stemming from theoretical physics. Our anthology wants to contribute to this promising cross-fertilization and the creation of an interdisciplinary dialogue in network research.

References

Barabási AL (2002) Linked: The New Science of Networks. Perseus Publishing, Cambridge, MA.

Barabási AL, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. Physica A 311(3–4), pp. 590–614.

Börner K, Chen C, Boyack K (2003) Visualizing knowledge domains. In: Cronin B (ed.), Annual Review of Information Science and Technology. Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, vol. 37, pp. 179–255, chapter 5.

Börner K, Sanyal S, Vespignani A (2007) Network science. In: Cronin B (ed.), Annual Review of Information Science and Technology. Information Today, Inc./American Society for Information Science and Technology, Medford, NJ, vol. 41, pp. 537–607.

Börner K (forthcoming) Atlas of Science: Guiding the Navigation and Management of Scholarly Knowledge. ESRI Press, Redlands, CA.

Bornholdt S, Schuster HG (eds.) (2003) Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH, Weinheim.

Buchanan M (2002) Nexus: Small Worlds and the Groundbreaking Science of Networks. Norton, New York, NY.

Castells M (1996) The Rise of the Network Society. Blackwell Publishers, Cambridge, MA.

Chen C (2003) Mapping Scientific Frontiers: The Quest for Knowledge Visualization. Springer, London.

Committee on Network Science for Future Army Applications, National Research Council (2005) Network Science. The National Academies Press, Washington, DC.

Cooke P, Schall N (2007) Schumpeter and varieties of innovations: Lessons from the rise of regional innovation system research. In: Hanusch H and Pyka A (eds.), The Elgar Companion to Neo-Schumpeterian Economics. Edward Elgar Publisher, Cheltenham.

De Nooy W (forthcoming) Social network analysis, graph theoretical approaches to. In: Meyers, Robert A (ed.), Encyclopedia of Complexity and System Science. Springer-Verlag, Berlin/Heidelberg, (in print).

Dodge M, Kitchin R (2001) Atlas of Cyberspace. Addison-Wesley, Harlow.

Dorogovtsev SN, Mendes JFF (2003) Evolution of Networks: From Biological Nets to the Internet and WWW. Oxford University Press, Oxford.
Galison P (1997) Image and Logic: A Material Culture of Microphysics. The University of Chicago Press, Chicago.
Hook EB (ed.) (2002) Prematurity in Scientific Discovery: On Resistance and Neglect. University of California Press, Berkeley.
Huberman BA (2001) The Laws of the Web. Patterns in the Ecology of Information. The MIT Press, Cambridge, MA.
Ebeling W, Scharnhorst A (2000) Evolutionary models of innovation dynamics. In: Helbing D, Herrmann HJ, Schreckenberg M, Wolf DE (eds.), Traffic and Granular Flow '99 – Social, Traffic, and Granular Dynamics. Springer, Berlin, pp. 43–56.
Ebeling W, Karmeshu, Scharnhorst A (2001) Dynamics of Economic and Technological Search Processes in Complex Adaptive Landscapes. Advances in Complex Systems 4(1), 71–88.
Feistel R, Ebeling W (1989) Evolution of Complex Systems. Deutscher Verlag der Wissenschaften, Berlin.
Foot K, Schneider S (2006) Web Campaigning. MIT Press, Cambridge, MA.
Fronczak P, Fronczak A, Hołyst JA (2007) Phase transitions in social networks. Eur Phys J B 59(1), 133–139 (preprint arXiv:physics/0701182v1).
Kirman A (2003) Economic networks. In: Bornholdt S, Schuster HG (eds.), Handbook of Graphs and Networks: From the Genome to the Internet. Wiley VCH, Weinheim, pp. 273–294.
Läsker L, Plikat B (eds.) (1985) Gibt es vorzeitige Entdeckungen? ITW-Kolloquien. Issue 47. Berlin, pp. 30–43.
Leydesdorff L (2001) A Sociological Theory of Communication: The Self-Organization of the Knowledge-Based Society. Universal Publishers, Parkland, FL.
Leydesdorff L (2006) The Knowledge-Based Economy – Modeled, Measured, Simulated. Universal Publishers, Boca Raton, FL.
Leydesdorff L, Hellsten I (2006) Measuring the meanings of words in contexts: Automated analysis of 'Monarch butterflies', 'Frankenfoods' and 'stem cells'. Scientometrics 67(2), 231–258.
Leydesdorff L, Schank T (forthcoming) Dynamic Animations of Journal Maps: Indicators of Structural Change and Interdisciplinary Developments. Journal of the American Society for Information Science and Technology.
Merton RK (1961) Singletons and multiples in scientific discovery – a chapter in sociology of science. Proceedings of the American Philosophical Society 105(5), 470–486.
Morris SA, Yen GG (2004) Crossmaps: Visualization of overlapping relationships in collections of journal papers. PNAS 101(1), 5291–5296.
Newman M (2003) The structure and function of complex networks. SIAM Review 45(2), 167–256.
Otte E, Rousseau R (2002) Social network analysis: A powerful strategy, also for the information sciences. Journal of Information Science 28(6), 441.
Paton RC, Neilson I (eds.) (1999) Visual Representations and Interpretations. Springer, London.
Pigliucci M, Kaplan J (2006) Making Sense of Evolution. The University of Chicago Press, Chicago and London.
Pyka A, Küppers G (2002) Innovation Networks – Theory and Practice. Edward Elgar Publisher, Cheltenham.
Ratto M (2006) Epistemic commitments and archaeological representation. In: Oosterbeek L, Raposo J (eds.), XV Congrès de l'Union Internationale des Sciences Préhistoriques et Protohistoriques. Livre des Résumés, vol. 1, p. 60 (available on-line http://www.uispp.ipt.pt/UISPPprogfin/Livro2.pdf, access December 2007).
Saviotti PP, Bowman A (1984) Indicators of output of technology. In: Gibbons M, Gummett P, Udgaonkar B (eds.), Science and Technology Policy in the 1980s and Beyond. Longman, London, pp. 117–147.

Scharnhorst A (2001) Constructing knowledge landscapes within the framework of geometrically oriented evolutionary theories. In: Matthies M, Malchow H, Kriz J (eds.), Integrative Systems Approaches to Natural and Social Sciences – Systems Science 2000. Springer, Berlin, pp. 505–515.

Scharnhorst A (2003) Complex networks and the web: Insights from nonlinear physics. Journal of Computer-Mediated Communication 8(4) (Available on-line http://jcmc.indiana.edu/vol8/issue4/scharnhorst.html).

Scharnhorst A, Hellsten I (2005) Triangulation of network metaphors. Paper given at the EXYSTENCE topical workshop Innovation networks. New approaches in modeling and analyzing, Augsburg, October 2005. (Available at http://www.virtualknowledgestudio.nl/staff/andrea-scharnhorst/current.php).

Snijders TAB, Steglich CEG, Schweinberger M, Huisman M (2007) Manual for SIENA Version 3.1. ICS/Department of Sociology, University of Groningen; Department of Statistics, University of Oxford.

Suchecki K, Hołyst JA (2006) Ising model on two connected Barabasi-Albert networks. Physical Review E 71(011122).

Thelwall M (2004) Link Analysis – An Information Science Approach. Academic Press, San Diego, CA.

Thelwall M (2008) E-social science visualizations. Collecting and visualizing internet data. In: Jankowski NW (ed.), E-Science: Transformations in Scholarly Practice. Routledge, New York (forthcoming).

Thelwall M, Wouters P, Fry J (forthcoming) Information-centred research for large-scale analyses of new information sources. Journal of the American Society for Information Science and Technology (JASIST).

Wasserman S, Faust K (1994) Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge.

Watts DJ (1999) Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton, NJ.

Watts DJ (2004) The “new” science of networks. Annual Review of Sociology 30, 243–270.

Zuckerman H (1967) Sociology of Nobel prizes. Scientific American 217(5), 25–33.

Part I
Innovation Networks in Economics

Chapter 2
Knowledge Networks: Structure and Dynamics

Pier Paolo Saviotti

2.1 Introduction

There has been recently an upsurge of interest in networks in a literature which belongs to many different disciplines, ranging from physics to biology to the social sciences. Interestingly enough, it seems that in spite of the wide differences between the entities constituting such networks, ranging from the interactions of biological molecules in cells to the Internet to citations (Barabasi et al., 1999, Barabasi, 2002; Barabasi and Bonabeau, 2003; Cohen, 2002; Watts and Strogatz, 1998), most such studies claim some kind of common intellectual framework rooted in complexity science. Yet the literature on networks does not provide any link between networks and other parts of the science of complexity which could be considered more fundamental. This contribution will proceed first to identify some possible connections between networks and other theories of complexity; second, to describe and analyse some networks of knowledge and innovation and to interpret their properties in terms of recent studies of networks; third, to formulate some generalizations about the dynamics of these networks and about their connection to the dynamics of variety and efficiency.

2.2 Networks and Complexity

Let us begin by asking the question 'why do we need complexity science?' and try and give it an answer mostly related to economic and social development. One of the most important tasks of a theory of economic development is to be able to explain some important stylized facts that we can observe in economic development. Some of the most important such stylized facts are listed as follows:

P.P. Saviotti (✉)
Université Pierre Medès, BP 47, 38040 Grenoble Cedex 9, France
e-mail: ppsavio@grenoble.inra.fr

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_2,

(i) Economic systems do not become more random, or disordered, during their process of development.
(ii) Economic development is characterized by qualitative change, since the new entities emerging during its course are not comparable to previously existing ones.
(iii) The variety/diversity of the economic system rises during the course of economic development.
(iv) Order, or structure, emerges from the process of economic development. Changes in structure occur infrequently and are followed by long periods of more incremental variations.

As a consequence of these stylized facts, the composition, defined as the list of entities and processes required to describe the economic system, changes in the course of time. The important implication of all of this is that not all theories would predict this type of development. For example, the growing variety of the economic system could be expected to lead to a growing disorder or randomness (see for example Georgescu-Roegen, 1971). Amongst the theories of complexity, the one formulated by Prigogine (see for example Nicolis and Prigogine, 1989) has the great merit of providing a potential explanation for the growing order and for the growing variety of the economic system. Let us start by defining structure as given by the components of a system and by their interactions. We can observe here that the interactions provide constraint and reduce the number of degrees of freedom of the system with respect to those of its isolated components. It is precisely this constraint that creates order. The most important explananda in our economic system are (i) the emergence of order during the evolution of the system and (ii) the discontinuous transitions that the system seems to undergo during its evolution.

Prigogine introduced some fundamental concepts and distinctions, which are very useful in this sense. An important distinction needs to be introduced between closed and open systems. The former are closed because they do not exchange anything (matter, energy etc.) with their environment. Only closed systems can achieve an equilibrium, which corresponds to the maximum possible disorder or randomness of the components of the system, that is, to the contrary of the order we observe in economic development. Open systems exchange matter, energy, information etc. with their environment. The rates of flow through the boundaries of the system measure the distance of such system from equilibrium: the more intense such flows, the more the system moves away from equilibrium. The interesting behaviour of these systems only emerges when they are far enough away from equilibrium. In these conditions, structure may emerge first and transitions between different types of structure can occur subsequently. Transitions occur when the system, as a consequence of its previous development, becomes unstable and undergoes a transition to a different configuration. According to Prigogine, instability arises in the form of *fluctuations* in some system variables. When fluctuations become sufficiently intense they induce *bifurcations*, or transitions to different configurations of the system, which not are only different from pre-transition ones but can also be in higher numbers. As a consequence of bifurcations, the number of possible states of

the system can increase, thus providing an in-principle justification for the growing variety of the economic system.

Thus, Prigogine's theory gives a potential explanation to (i) the emergence of structure, (ii) the presence of transitions between distinguishable states of the system and (iii) the growth in the number of possible system states after the transition. Prigogine's theory provides a potential explanation for the stylized facts listed above. Incidentally, let us note that some very similar stylized facts apply to biological development. Prigogine's ideas provide an underlying theoretical framework for both biological and economic development. Of course, the existence of this underlying theoretical framework does not mean that exactly the same type of explanation will apply to biological and economic phenomena. Complexity theory is expected to be applicable to many different types of systems at a high level of generality. The complete analysis of each system involves such a general level and a more specific one requiring the use of concepts and variables specific to the system itself. Thus, the actual dynamics of biological and of economic systems cannot be expected to be identical even if at a high level of generality they share the non-equilibrium nature of evolution and the possibility of bifurcations.

What relationship can we expect to exist between Prigogine's theory of complexity and the existence of networks? First, networks are composed of nodes/vertices and of links/edges. The nodes are the components of the socio-economic systems and the links are their interactions. Thus, networks are the structure of socio-economic systems. Following the previous considerations, we can expect (i) networks to emerge away from equilibrium and (ii) transitions between different types of networks to occur, as the relevant system becomes unstable and undergoes a bifurcation.

We can imagine the development of human networks to have begun as a consequence of the most fundamental aspect of human behaviour, adaptation to the external environment (Ext.Env). From very early on, it must have turned out that collective adaptation was superior to individual adaptation. However, collective adaptation involved coordination of individuals' actions, resulting in a fall in the number of degrees of freedom that would otherwise have been available to individuals. The constraints existing in the communities that adapted collectively occurred in the form of *rules* (Dopfer, 2004) which limited and streamlined people's behaviour. Furthermore, coordination involved at its roots communication. Thus, the earliest systems of rules that could provide coordination must have been language and law. Such constraints on individual behaviour shaped inter-individual interactions and shaped the earliest networks.

On the basis of the previous considerations on Prigogine's theory of complexity, we expect the evolution of a system, represented by a network, to occur by means of a combination of the gradual adjustments of the system to its Ext.Env, combined with inter-system transitions, induced by fluctuations in the previous system configuration. It is to be pointed out that we can expect fluctuations to arise endogenously within the system as a result of its previous evolution. Fluctuations induce bifurcations, leading to discontinuous changes, giving rise to the emergence of new structures/networks. The emergence of new structures/networks occurs by

the creation of new boundaries and of new subsets of the system under consideration, with the possible disappearance of older subsets. In network terms, this amounts to the creation of new nodes and links and to disappearance of older ones.

In this context, fluctuations take an interesting meaning. Instability within an existing structure/network is likely to arise as a consequence of the instability of some existing links. This instability can be interpreted as a departure from dominant rules, which in turn can allow the exploration of new subsets of Ext.Env and the creation of innovations. This would be compatible with observations which imply that historically innovations have tended to arise more frequently in societies which were freer and less bound by tradition (see for example Landes, 1998).

According to stylized fact (iii) stated above, the variety of the economic system increases during the process of economic development. In terms of networks this could imply that the number of networks within the system increases. Another important feature of network dynamics is the evolution of their connectivity. In fact the variety of an economic system at best measures the number of its nodes but does not say anything about its links, which are an important part of its structure. Therefore, connectivity is an important part of network dynamics. We can expect innovations to be introduced by entrepreneurs in a rule-poor environment, which provides the required freedom and the scope for fluctuations. Not all fluctuations are successful. At any time most of them are likely to be selected out. Precisely for this reason, a society which is able to create more fluctuations will have a greater chance of having successful ones. However, if the innovation is successful we can expect it to be widely imitated and to diffuse gradually in society. In order to acquire its 'economic weight', an innovation also requires the co-evolution of appropriate institutions (Nelson, 1994). The creation of complementary technologies and of appropriate institutions leads to the formation of new links, thus raising the connectivity of the system. For example, the creation of a regulatory institution can be expected to lead to interactions with the firms and the other organizations responsible for the production and use of the new technology. These interactions may be impersonal and simply provide constraint, as it would happen in the case of standard-creating institutions, or be more localized and directed, as in the case of a firm producing complementary inputs to the innovation and technology concerned. Examples of these situations for automobile could be (i) the ministries responsible for issuing driving permits or driving rules and (ii) the firms producing and distributing tyres or petrol (Saviotti, 2005). In all these cases, the general meaning of links is that they reduce the number of degrees of freedom of each node and provide constraint. The behaviour of the nodes then becomes more highly correlated. This progressive increase in connectivity as an innovation and the relative technology mature on the one hand increases the potential market size of the new technology by improving the technology with respect to its initial form but, on the other hand, makes the new technology progressively more rigid, even if more coherent. In this way, an increasing connectivity allows a technology to acquire its full 'economic weight' but contributes to the process, whereby diminishing returns gradually take over and slow down the rate of improvement of maturing technologies.

Network dynamics at the industry/technology level can be expected to be characterized by low connectivity during the emergence phase and by growing connectivity as the sector matures. At the aggregate level of the whole economic system, increasing diversity/variety means an increasing number of nodes. However, this increase is likely to be unevenly distributed in time and space. The creation of new nodes cannot be expected to be followed immediately by the creation of new links. The emergence of important innovations can be expected to lower connectivity while the subsequent process of diffusion can be expected to raise connectivity. Thus, aggregate connectivity cannot be expected to grow at all times, but it could easily oscillate around a given value.

We can then expect some relationships to exist between the evolution of variety and that of networks. Variety grows by the creation of new economic species (new products, services etc.):

- First, we can expect the creation of new economic species to lead to a growing number of distinguishable networks.
- Second, we can expect the phase of emergence of new economic species to occur in an institutionally poor environment characterized by a low connectivity, but we can also expect the subsequent phases of diffusion and of maturation of new technologies to lead to a growing connectivity.

It has been recently discovered that a large class of networks possess some common properties for which they are called scale-free (Barabasi and Reka, 1999; Barabasi et al., 1999; Barabasi, 2002; Barabasi and Bonabeau, 2003; Reka et al., 2000; Reka and Barabasi, 2002). In particular, these networks have a very asymmetrical distribution of links around nodes: few nodes have many links and many nodes have few links. This distribution is very different from that predicted for previously studied networks, called exponential networks, which had a much more egalitarian distribution of links around nodes. Scale-free networks have a power law distribution while exponential networks have a Poisson distribution of links around nodes. As a consequence, scale-free networks have the interesting property of being very resistant to random attack: almost 80% of the links can be cut before a scale-free network is destroyed, while the corresponding percentage for an exponential network is less than 20%. However, a targeted attack selectively cutting links around the most central nodes (hubs) destroys the network by cutting less than 20% of the links.

Two conditions are required in order for scale-free networks to exist:

(i) growth – the number of nodes must grow;
(ii) preferential attachment – new links tend to be formed more easily with already linked nodes.

These conditions are often present in socio-economic networks. In general, the observed growth in variety of a number of economic species (technologies etc.) can be expected to lead to a growing number of nodes. Thus, growing variety could supply one of the two conditions required for the existence of scale-free networks. In

socio-economic networks, the second condition – preferential attachment – depends on sources of increasing returns to adoption. Examples of these sources are reputational structure and various types of resources. Let us take the example of an alliance between an incumbent large diversified firm (LDF) and a start-up. In its choice of partner the start-up is likely to favour the LDF with the best reputation. If the alliance leads to a further enhancement of the LDF's reputation, other start-ups will continue to favour it with respect to other LDFs. Moreover, if a growing number of alliances raise the resource base of incumbent LDFs, those already having a greater number of alliances will be better able to form further ones than other LDFs. Thus, conditions (i) and (ii) can be often found in socio-economic networks. The presence of these two conditions leads to a higher probability of creation of scale-free networks than to other types of networks, or to a higher relative rate of *variation* for this type of networks. This condition is at best necessary, but not sufficient, to justify a high concentration of scale-free networks. However, one of the most important findings about this type of networks is their resistance to attack. If we interpret attack as selection, scale-free networks are likely to have a high rate of variation and a low rate of selection whenever conditions (i) and (ii) are satisfied. On the other hand, since conditions (i) and (ii) are often present in socio-economic networks, we can expect the scale-free geometry to be quite common in this type of network.

These network properties are obviously interesting and highly relevant for socio-economic networks. However, although research in scale-free networks has concentrated on the distribution of links around nodes, other related network properties are of great importance. For example, the existence of scale-free networks implies an uneven distribution of the degree of centrality of nodes. Few nodes are highly central while others have a low centrality. Furthermore, the distribution of centrality is likely to change dynamically, for example with the distribution becoming at times more skewed or more even, the relative centrality of some nodes falling and that of others rising. Another property whose role has already been discussed is connectivity. We have already seen how we can expect connectivity to rise or to fall during the evolution of networks. The meaning of this property will be discussed in greater detail in the next section.

2.3 Examples of Knowledge Networks

In what follows three examples of knowledge-related networks will be discussed:

(i) The network of knowledge itself, which will be presented only conceptually but which will provide a good basis for the subsequent discussion.
(ii) The network representing the knowledge base of firms.
(iii) Innovation networks (INs) in biotechnology.

2.3.1 Knowledge as a Network

Two very important properties of knowledge are (Saviotti, 2004)

(i) Knowledge is a correlational structure.
(ii) Knowledge is a retrieval/interpretative structure.

According to property (i) knowledge establishes correlations between different concepts and variables (see also Loasby, 2001). It is therefore possible to represent knowledge as a network whose nodes are concepts or variables and whose links are given by the joint utilization of the concepts or variables. In this representation, we would attain a complete knowledge of our Ext.Env if we had all the nodes corresponding to all the concepts and/or variables of our Ext.Env and if the corresponding network was fully connected. However, if we examine the way knowledge develops we can realize that we are very far from complete knowledge. As the exploration of our Ext.Env proceeds, we detect new observables and create appropriate variables. This is done on a *local* basis, that is, starting from a casual observation or from the solution of a practical problem (see also Popper, 1972). The subsequent evolution of different fields of knowledge gives rise to disconnected networks. Let us take an example.

In the past astronomy and medicine developed in completely separate ways. There was no awareness that the entities to which the problems could be reduced in the two fields had anything in common. Thus, astronomy proceeded by identifying observables (the sun, the earth, planets, stars etc.) and constructed models of the movements of these entities. Medicine on the other hand proceeded by identifying organs and by trying to explain the behaviour of the whole body by means of its organs. The awareness that organs were constituted by cells, cells by molecules and molecules by protons, electrons and neutrons took centuries to come. In other words, the networks of knowledge of astronomy and medicine were for a very long time separate, and it was not realized that they could in principle be connected. The awareness of the potential connectedness of these two and of other networks of knowledge came only during the nineteenth century and gave rise to the so-called Laplacian dream (Mirowski, 1989). One could say that the research programme of molecular biology aims at connecting the networks of biology and physics. However, and in spite of the considerable successes achieved in this direction in the last 30 years, we are still very far from having identified all the possible nodes and links. Thus, we can conclude that, although the final objective of knowledge is to construct a complete network, containing all the possible variables of our Ext.Env and all the possible connections linking these variables, the present state of our knowledge is very far from that.

Some observations can help in understanding the present state of our knowledge network. First, new observables are continuously discovered, although at a different speed in different disciplines. Some disciplines (e.g. chemistry) are closer to maturity and generate few new observables, while others (e.g. biology) keep generating new observables and variables. Second, the rate of creation of new nodes by the

discovery of new observables and variables precedes in general the creation of links between the corresponding variables. In many cases we can expect the construction of links to be a much slower process than the creation of new nodes. When, as a result of new discoveries, new nodes are introduced in our network of knowledge we can expect connectivity to fall. As links are established with the newly created nodes, the connectivity of the system can start rising again. We can reinforce here some trends we had already seen at a more general level:

(i) The emergence of novelty tends to create new but poorly connected nodes, thus temporarily reducing the connectivity of the system.
(ii) The subsequent diffusion of the innovations establishes new links and raises again the connectivity of the system.
(iii) As a result of (i) and (ii), the connectivity of the system is likely to fluctuate around a given value.

It is important to realize here the role that connectivity can play in the dynamics of socio-economic networks. Connectivity is generally measured by the density of links per node in a network. Since in knowledge networks links represent the existence of correlations between nodes/variables, in a high-connectivity knowledge network variables are highly correlated. In turn, the existence of correlations leads to a high probability of predicting the values of some variables from those of other correlated variables. On the whole a high-connectivity knowledge network leads to a high probability of predicting the behaviour of some parts of the network starting from the knowledge of other parts. When our knowledge network is used to modify a subset S of Ext.Env, a higher probability of predicting some parts of the knowledge network leads to lower costs of modifying a subset S(Ext.Env) (Saviotti, 2004). Thus, connectivity is a relevant property of a knowledge network, both in a cognitive and in a technological sense.

The dynamic representation of knowledge used in this contribution (but see Saviotti, 2004 for greater details) is compatible with Kuhn's (1962) analysis of the evolution of science. New observables and variables are likely to be created when new paradigms emerge. In this early phase, we can expect new variables to be poorly connected to those existing in the previous network of knowledge. Thus, the emergence, or revolutionary, phase of a new paradigm is likely to be accompanied by a falling connectivity of the network of knowledge. On the other hand, we can expect the subsequent phase of normal science to be characterized by a growing number of links, and thus by a growing network connectivity.

In a broader sense, the representation of knowledge as a correlational and as retrieval/interpretative structure is compatible with the idea of knowledge as an organized structure, to which both Kuhn's and Lakatos' theories belong (Chalmers, 1980). In particular, the representation of knowledge described in this contribution is compatible with some recent structuralist theories of science (Balzer et al., 1987; Franck et al., 1999) according to which the collection of all empirical science forms a theoretical *holon*, composed of constellations of elementary theories, theories that would be *connected* by inter-theoretical links of different types, such as equivalence, specialization, connection.

2.3.2 The Knowledge Base of Firms and Organizations

We can define the knowledge base of a firm or organization as the collective knowledge that can be used to achieve the firm's productive objectives. The term collective is due to the fact that the process of knowledge creation in the firm is based on division of labour and on coordination. Many individuals, departments, subsidiaries etc. of the firm contribute to the creation of new knowledge, each carrying out a small subset of the whole process. The production of the resultant knowledge necessarily involves the coordination of all these activities. Clearly the organizational structure of a firm can be expected to have an impact on the process of knowledge creation.

The study of firms' (and organizations') KBs is a very important component of the creation of an economics of knowledge. In this section, two different methods to map the KB and to measure its properties will be described. In both cases we start by identifying some basic units of knowledge. In principle, we could refer to the considerations of Sect. 3.1 and attempt to find all the variables corresponding to a given piece of knowledge. This is generally impossible and we use instead more aggregate units of knowledge, such as the technological classes contained in patents or the themes contained in patents or publications. The representation of the KB that we obtain by examining, for example, the patents of a firm is a network in which the nodes are constituted by our units of knowledge and the links by the interactions of the units of knowledge. In the work described here the interactions are measured by the co-occurrence of the units of knowledge in the patents or in the other sources of information that we are using.

The two methods we use to study firms' KBs are different in that they refer to different levels of aggregation. The first method, *lexicographic analysis* (*LA*), detects the units of knowledge in the texts that we use as sources of information. LA can detect in the text of patents short phrases corresponding to technological themes, or alternatively the technological classes contained in the patent. The links of these units are determined by their frequency of co-occurrence in the patents used. This provides us with a graphic representation of the network of knowledge constituting the KB at a given time. Repeating the study at different times we can map the evolution of the KB and relate it to changes in firm strategy, firm organization etc. (see Saviotti et al., 2003, 2005). In other words, LA allows us to represent the 'brain' of the firm.

The second method we use starts by constructing a matrix of co-occurrences of technological classes and provides us with a more aggregate representation of the KB. In particular, it allows us to measure some of the properties of the KB, such as its coherence, specialization, differentiation and similarity. On the basis of these measures, it is possible to show that the KB of a firm is a determinant of the firm's performance (Nesta and Saviotti, 2005). These two methods are complementary. LA provides us with a more disaggregate representation, by means of which we can enter the firm's KB, while the method based on co-occurrence matrices gives us measures of the resultant properties of each KB. The graphic representation shown here (Figs. 2.1, 2.2, 2.3, 2.4 and 2.5) was obtained by means of LA.

Figures 2.1 and 2.2 represent the KB of Hoechst for the periods (1993–1995) and (1996–1998). Figures 2.3 and 2.4 represent the KB of Rhône Poulenc for the

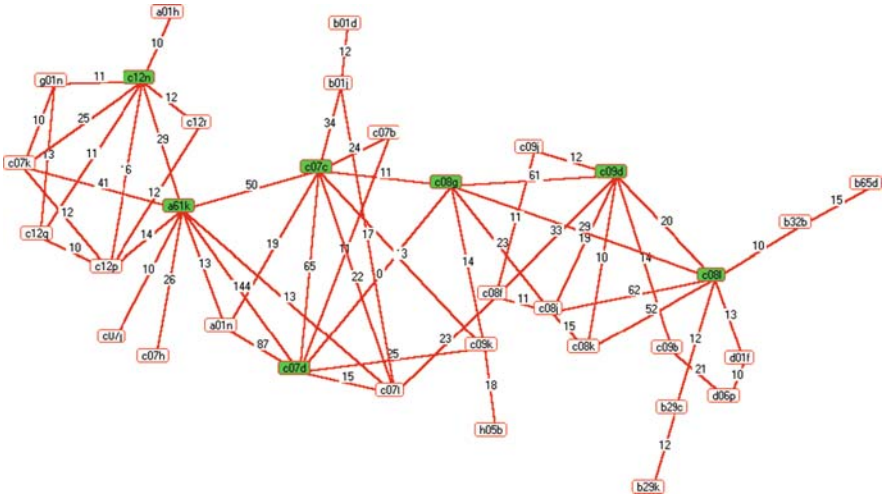


Fig. 2.1 Diagram for period 2 (1993–1995) of the co-occurrences between the main IPC classes of Hoechst (the most central classes are represented in dark grey)

periods (1993–1995) and (1996–1998). Figure 2.5 represents the KB of Aventis, the firm created by the merger of Hoechst and Rhône Poulenc. We can immediately notice some properties of this KB network:

- The distribution of the frequencies and intensities/strengths of links around nodes is highly heterogeneous. Few nodes have many links and other nodes have very few links. Also, some nodes have very strong links while others have very weak ones.
- In the KBs of both Hoechst and Rhône Poulenc we can see two parts which have a high internal density of links while being very lightly connected to each

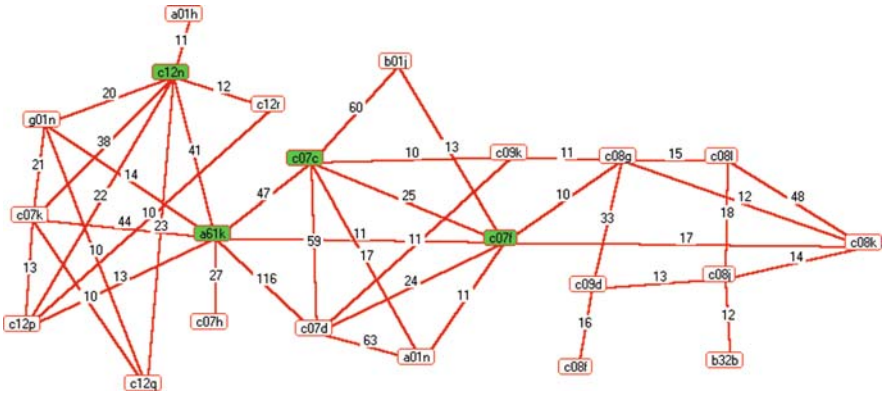


Fig. 2.2 The KB of Hoechst for period 3 (1996–1998) as represented by the co-occurrences between the main IPC classes of its patents (the most central classes are represented in dark grey)

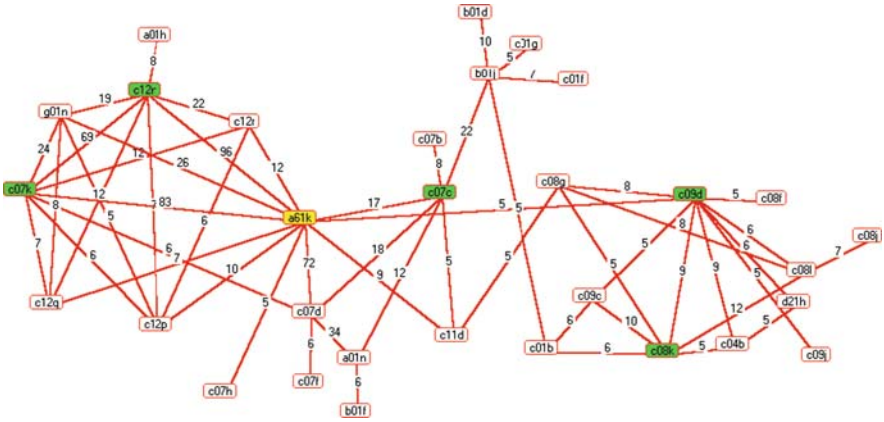


Fig. 2.3 The KB of Rhône Poulenc for period 2 (1993–1995) as represented by the co-occurrences between the main IPC classes of its patents (the most central classes are represented in dark grey)

other. Each of the two separate subsets of the KB of Hoechst and Rhône Poulenc corresponds to a different type of knowledge, chemistry and biology, respectively. Taking into account that during the period studied the two firms were attempting to move away from chemistry and towards life sciences, the two subsets correspond to the past and to the intended future of the firm, respectively. In other words, Figs. 2.1, 2.2, 2.3 and 2.4 represent the changes in the KB of both firms following from their strategic reorientation.

- Figure 2.5, representing the KB of Aventis after the merger, no longer shows the separation into two distinguishable subsets, corresponding to chemistry and

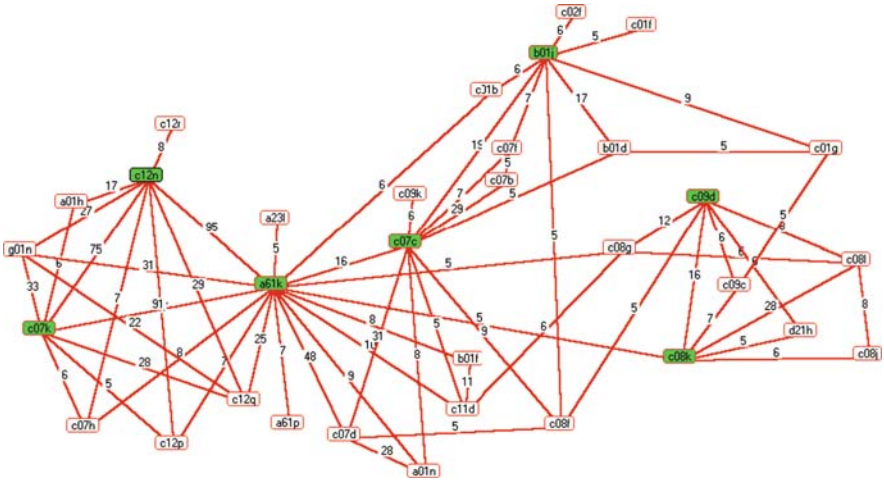


Fig. 2.4 The KB of Rhône Poulenc for period 3 (1996–1998) as represented by the co-occurrences between the main IPC classes in its patents

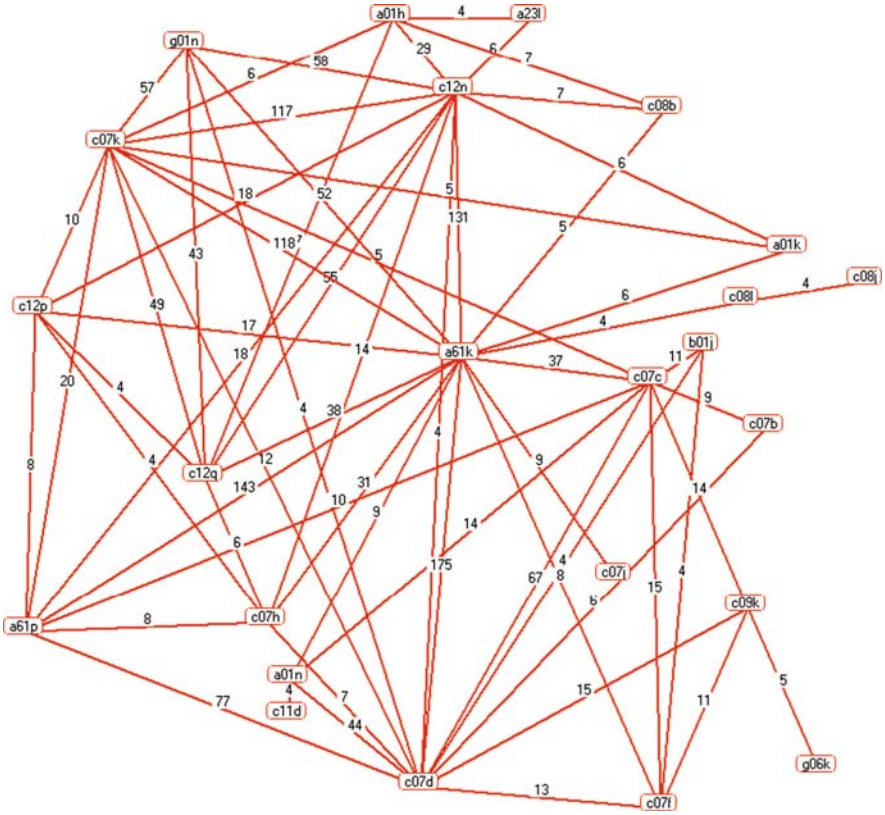


Fig. 2.5 The KB of Aventis after the merger, as represented by the co-occurrences of the technological classes in the patents of Aventis

to life sciences. The whole KB seems to be better connected, realizing a more complete integration of old and new knowledge. This better integration is confirmed by a measure of the number of links per node, which can be considered an approximate measure of network density and which is considerably higher for Aventis after the merger with respect to both Hoechst and Rhône Poulenc before the merger (Table 2.1). Furthermore, the distribution of links around nodes is highly asymmetrical.

Table 2.1 Number of nodes, links and links per node in the above representation of the KB of Hoechst, Rhône Poulenc and Aventis

	Hoechst		Rhône Poulenc		Aventis
	P2 (1993–1995)	P3 (1996–1998)	P2 (1993–1995)	P3 (1996–1998)	2002
Nodes (N)	33	22	31	32	24
Links (L)	55	36	46	54	73
L/N	1.67	1.64	1.48	1.69	3.04

2.3.3 Innovation Networks

Another type of knowledge-related network that we studied is that of INs in biotechnology. They are constituted by small dedicated biotechnology firms (DBFs), by LDFs, mostly pharmaceutical but sometimes belonging to other sectors, and by public research institutes (PRIs). These three types of actors interact by forming alliances one of whose main objectives is the creation of new knowledge. These networks were studied during the period 1973–1999 using data from the RECAP database (Catherine, 2005). INs are part of a class of inter-firm alliances which emerged starting from the early 1980s. At that time INs were considered by many economists a temporary form of industrial organization. Existing economic theories predicted that only markets and hierarchical organizations could be stable. Inter-firm alliances were considered a temporary response to shocks, a response which would have disappeared once the shocks had been absorbed by the economic system. Yet, 25 years later the number of INs keeps increasing (Fig. 2.6). INs seem to have become a new and stable form of industrial organization. However, the full answer is more subtle than that.

If we break down the whole period 1973–1999 into two sub-periods distinguished by the main technologies used in each one, the picture becomes considerably different. If we consider that modern biotechnology was created by the potential industrial applications of molecular biology, we can distinguish in its subsequent evolution two generations of biotech, linked to recombinant DNA and monoclonal antibodies and to genomics, respectively. The former begins in the mid-1970s and the latter in the mid-1980s. By classifying all the technological alliances in the data set as belonging either to the first or to the second generation, we can plot separately curves describing their numbers, represent separately their networks and measure properties of these networks, such as density, centrality (Figs. 2.15, 2.16 and Tables 2.2 and 2.3). A number of interesting results emerge:

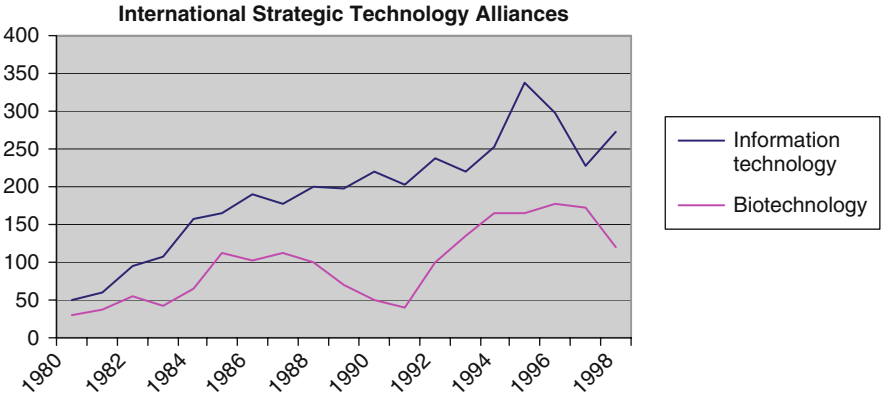


Fig. 2.6 Evolution of the number of technological alliances in ITC and in biotechnology

Table 2.2 Centrality of the different actors (DBFs = PM, LDFs = GGI, PRIs = INS) involved in the first generation of biotechnology alliances

First-generation biotechnology	Period 1976–1984			Period 1985–1992			Period 1993–1999		
Actor's type	DBFs	LDFs	PRIs	DBFs	LDFs	PRIs	DBFs	LDFs	PRIs
Average number of agreement	5.67	3.20	2.63	6.20	10.79	2.61	6.12	15.50	2.05
Average Ndegree centrality	5.84	3.30	2.71	1.71	2.97	0.72	1.41	3.58	0.47
Median Ndegree centrality	3.09	2.06	2.06	1.10	1.38	0.55	0.92	1.73	0.23
Average betweeness centrality	4.95	2.82	0.85	0.71	1.38	0.20	0.40	1.40	0.10
Median betweeness centrality	1.17	0	0	0.20	0.53	0	0.07	0.18	0

- First, the total number of INs in biotechnology keeps increasing (Fig. 2.7).
- Second, the number of INs corresponding to the first generation reaches a maximum in 1996 and then starts declining (Fig. 2.8).
- Third, if in each generation we separate the alliances linked to R&D from those linked to marketing we see that they have a different dynamics. In particular, alliances related to R&D dominate in the early phases but peak out and decline later, while marketing alliances emerge later but dominate the late phases of the life cycle of the biotech generation we are considering (Figs. 2.9 and 2.10).

One of the most interesting results of this study is that INs in each generation of biotechnology follow a life cycle, in which the number of INs rises at first, reaches a maximum and then declines. In the meantime, the type of agreement changes from R&D to marketing passing from the early to the late phases of the life cycle.

Other interesting results emerge when we represent graphically the INs corresponding to the two generations (Figs. 2.11, 2.12, 2.13 and 2.14). Here, we can see that for each generation both the number of nodes and the number of links rises during the life cycle. However, the density of links seems to be falling all

Table 2.3 Centrality of the different actors (DBFs = PM, LDFs = GGI, PRIs = INS) involved in the second generation of biotechnology alliances

Second-generation biotechnology	Period 1985–1992			Period 1993–1999		
Actor's type	DBFs	LDFs	PRIs	DBFs	LDFs	PRIs
Average number of agreement	5.98	5.90	5.48	12.50	25.13	9.27
Average Ndegree centrality	2.48	2.45	2.27	1.68	3.38	1.25
Median Ndegree centrality	2.08	1.66	1.66	1.21	1.55	0.81
Average betweeness centrality	2.70	2.21	2.69	0.41	0.97	0.13
Median betweeness centrality	1.91	1.36	2.14	0.18	0.15	0.07

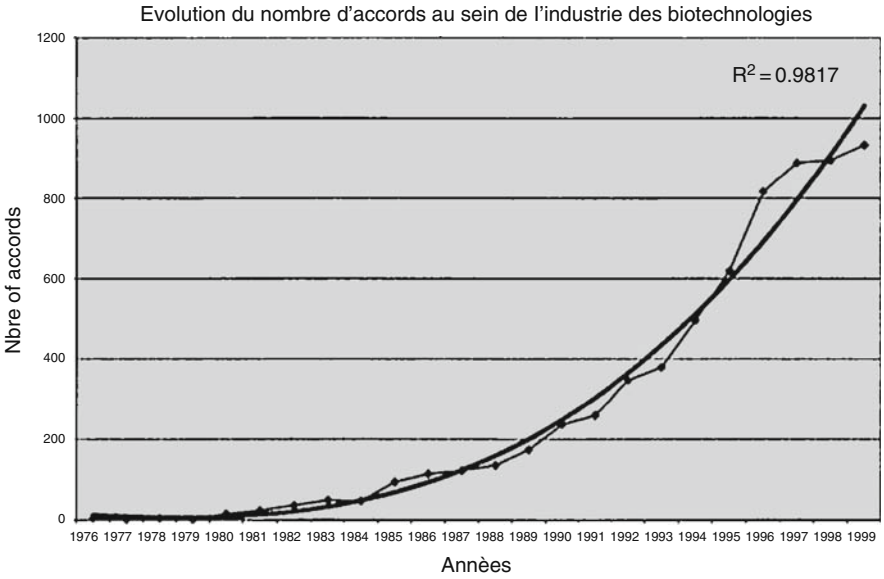


Fig. 2.7 Total number of agreements in biotechnology 1973–1999

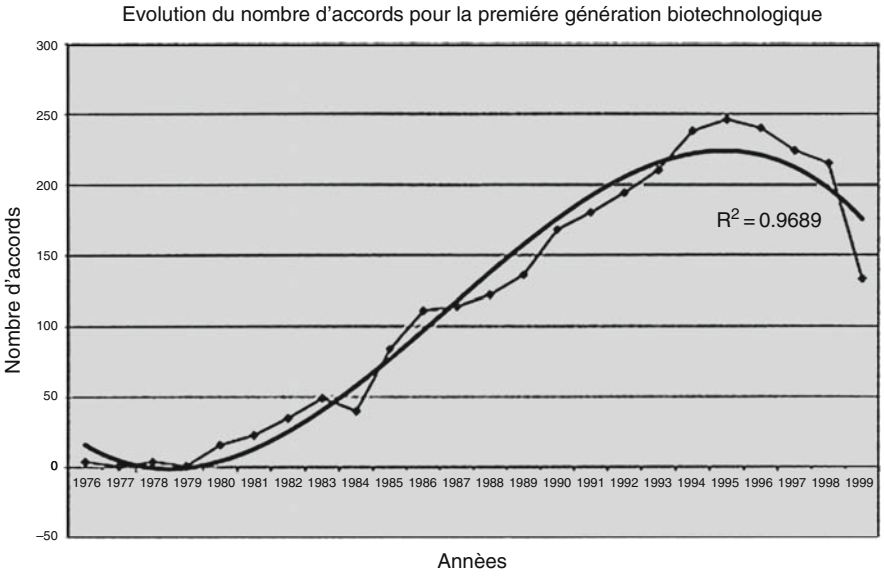


Fig. 2.8 Number of agreements in the first generation of biotechnology

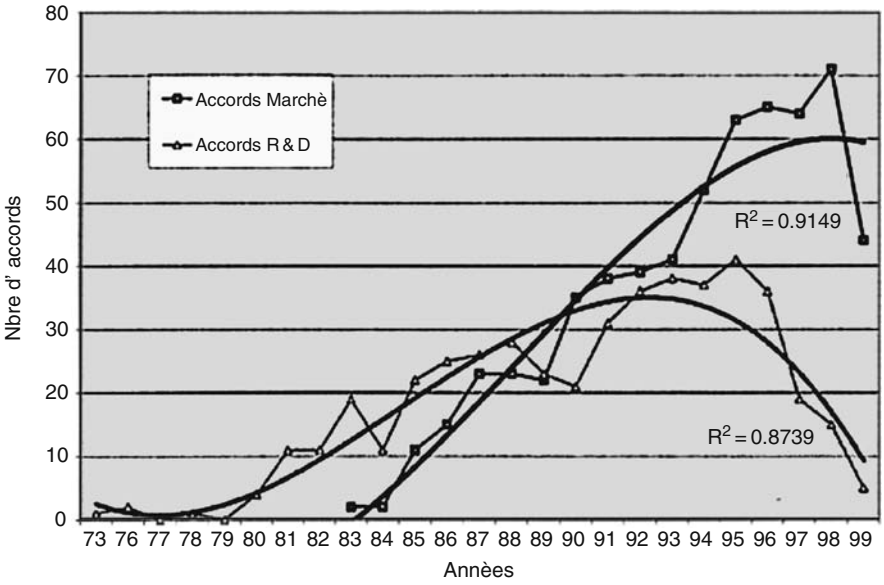


Fig. 2.9 R&D and marketing agreements in the first generation of biotech

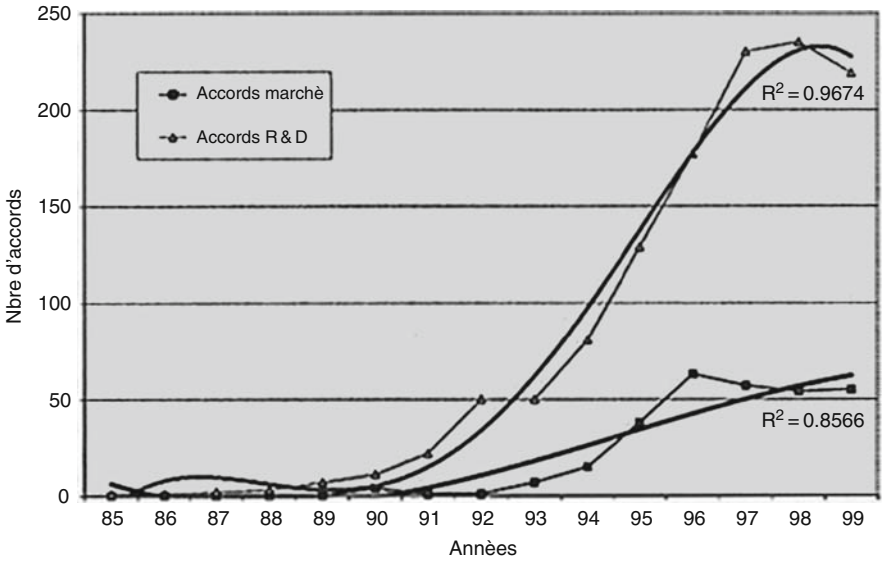


Fig. 2.10 R&D and marketing agreements in the second generation of biotech

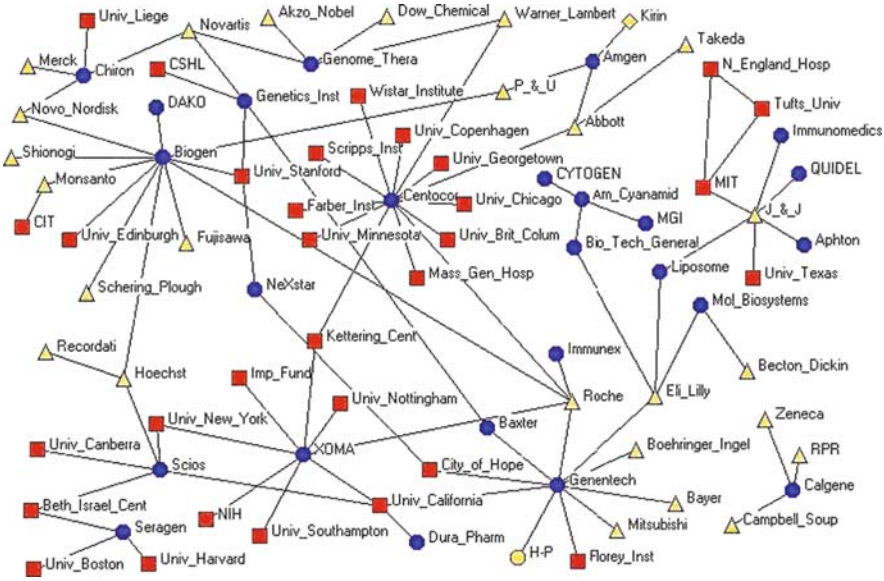


Fig. 2.11 Networks, first-generation biotech, 1973–1984

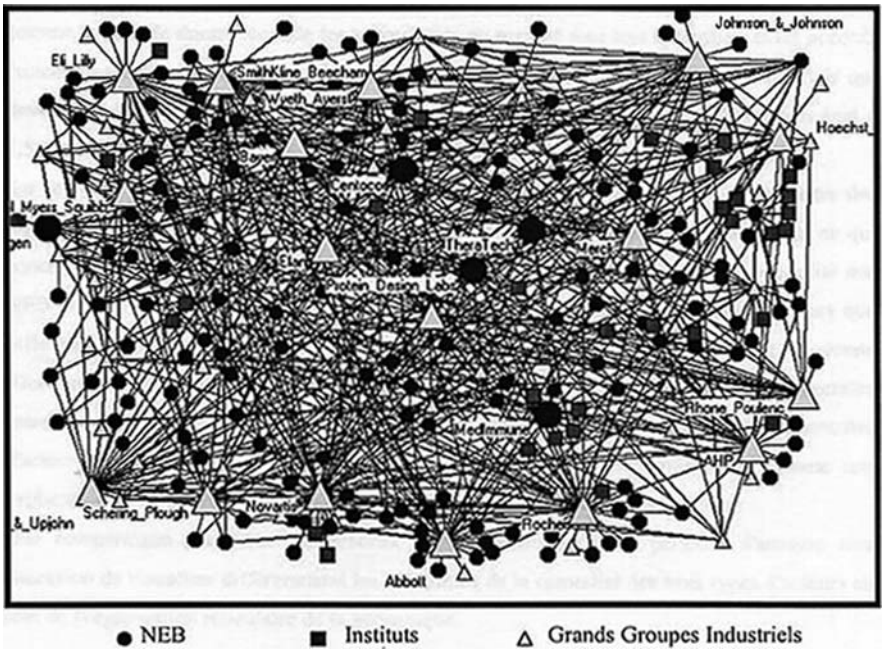


Fig. 2.12 INs, first-generation biotech, 1993–1999

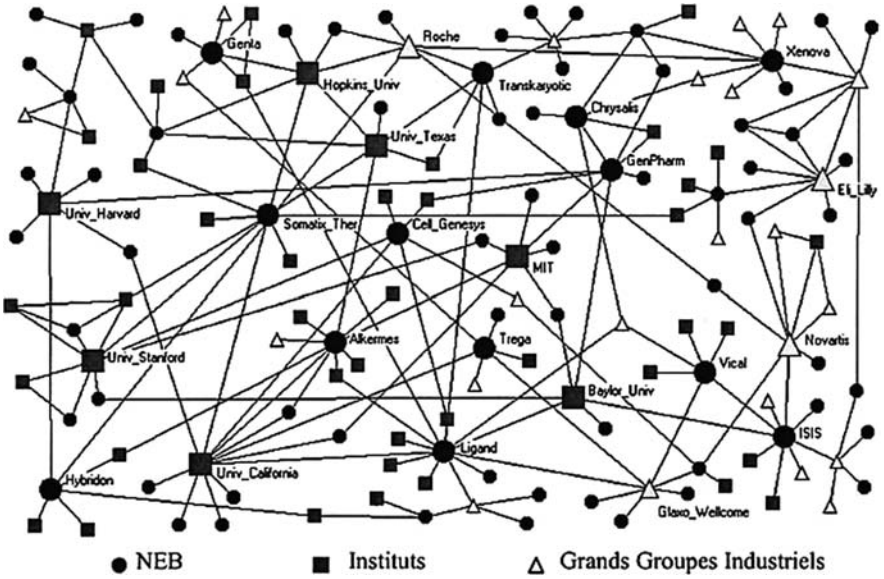


Fig. 2.13 INs, second-generation biotech, 1985–1992

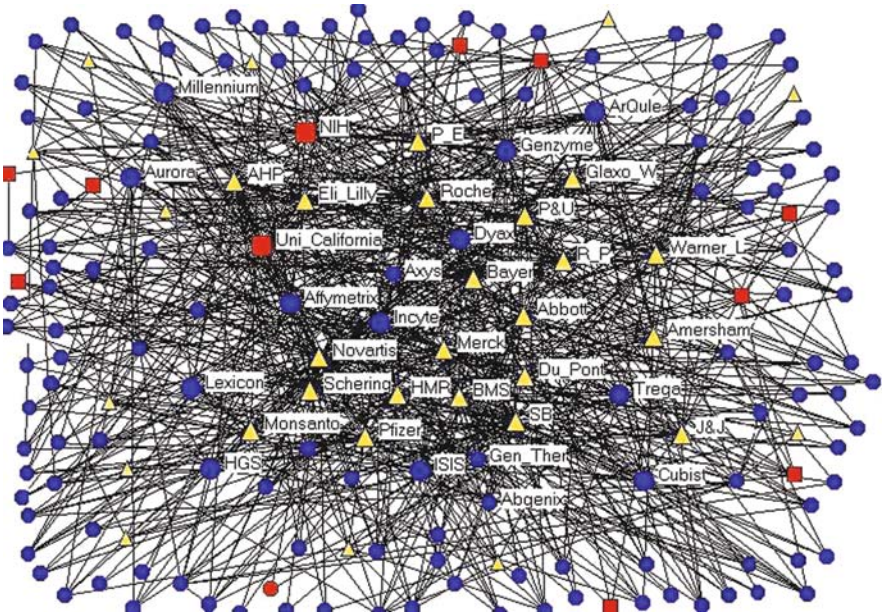


Fig. 2.14 INs, second-generation biotech, 1993–1999

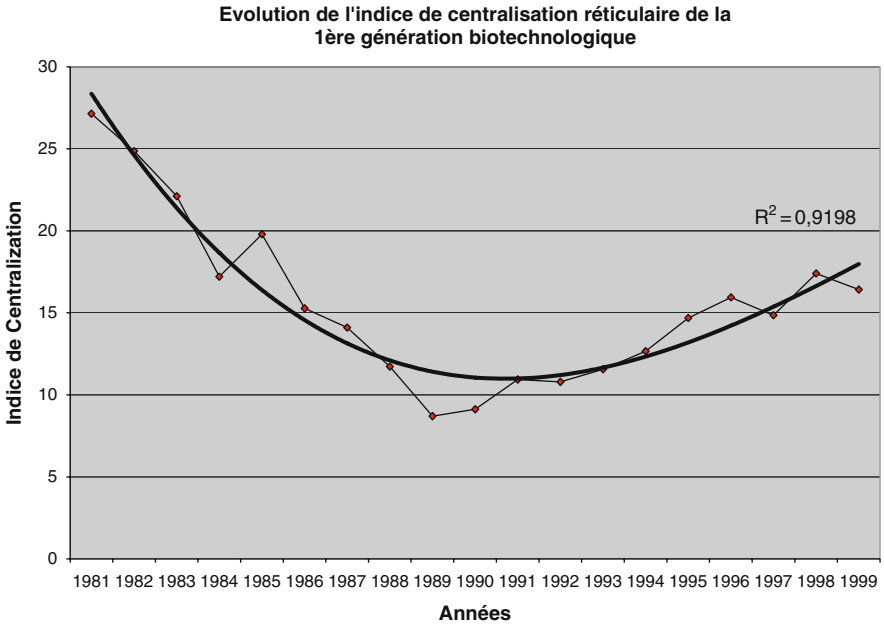


Fig. 2.15 Evolution of network density for the first-generation biotechnology

throughout the life cycle, with the possible exception of the very late phases of the first generation (Figs. 2.15, 2.16 and 2.17).

Finally, the centrality of the different actors changes during the evolution of the life cycle (Tables 2.2 and 2.3). The Ndegree centrality of DBFs (PM) is high initially and tends to fall as the life cycle tends towards maturity, that of LDFs (GGIs) falls slightly while remaining rather high, that of PRIs (INS) is initially high but falls to the lowest values of the three actors. The first and second generations show qualitatively the same type of evolution but differ for the relative extent of decline as the life cycle moves towards maturity: in the case of genomics the Ndegree centrality shows a more moderate fall, perhaps due to the shorter duration of the period studied.

Summarizing the results of this section we could say the following:

- INs in biotechnology seem to undergo a life cycle in which the number of alliances rises at first, reaches a maximum and then declines. During the same life cycle, the character of alliances changes from mostly R&D based in the early phases towards mostly marketing based in the late phases.
- This cyclical behaviour can be observed only at the level of aggregation of the generation of biotechnology (1st = Recombinant DNA + monoclonal antibodies, 2nd = genomics). The overall time profile of the number of alliances shows a continuous growth.
- The networks of both generations of biotechnology show an asymmetrical distribution around nodes. The distribution has not yet been measured.

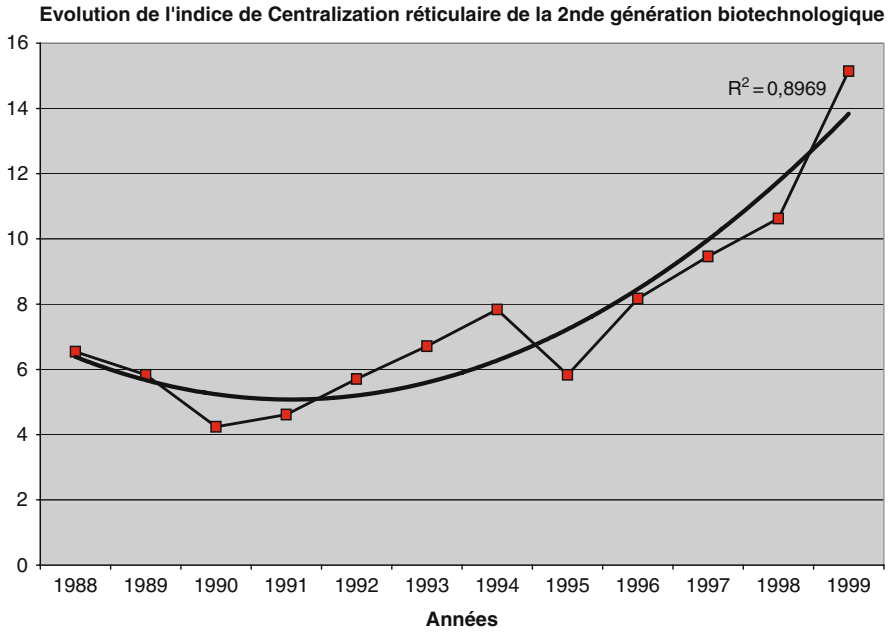


Fig. 2.16 Evolution of network density for the second generation of biotechnology

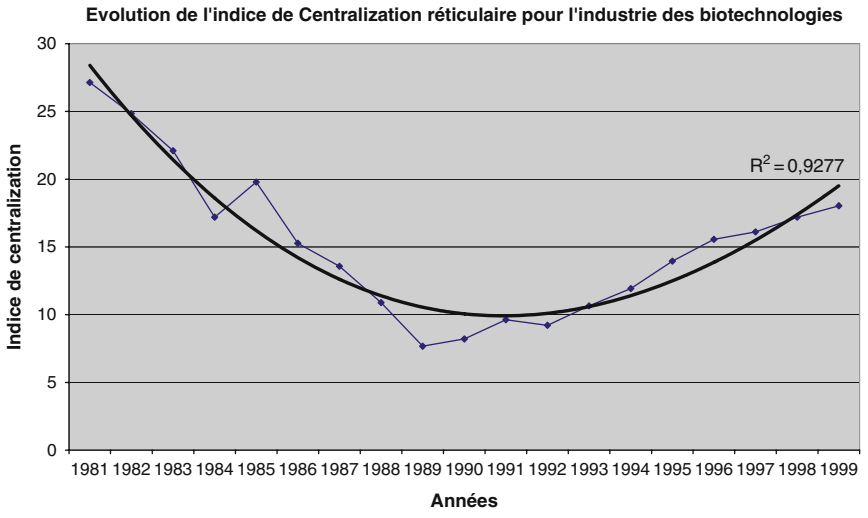


Fig. 2.17 Evolution of network density for the two biotechnology generations combined

- During the life cycle observed so far, the number of nodes grows and the number of links grows but network density seems to fall.
- The Ndegree centrality of the different actors involved in the networks (DBFs, LDFs, PRIs) tends to fall at different rates, that of LDFs remaining the highest in the long run.

A short comment is required here to interpret the meaning of network density in this case. The creation of new nodes is here due to the creation of new firms. At the beginning we can expect these firms to be relatively poorly connected to the existent economic system. Their early emergence can be expected to be accompanied by a low density of their networks, thus reducing the overall density of the networks in which they can be embedded. However, in order to grow the new firms have to establish interactions with dealers, regulators, customers, suppliers etc., thus leading to a growing density of their networks. The same situation is likely to apply to all firms in new, emerging sectors. Thus, we can expect new sectors to start with low-density networks and to undergo a growth of network density as they mature. In this sense, we can expect new sectors to benefit from the growth in network density occurring during their maturation, although this greater density will also imply a greater difficulty of introducing into the sector further innovations, especially if they are challenging the existence of the links already formed in the network. Of course, the same reasoning can be expected to apply to the technologies which give rise to the creation of new sectors or sub-sectors, as in this case. In fact, one of the main reasons for which network density can be expected to behave in the way described above is the role played by knowledge in the creation and in the dynamics of industrial sectors.

2.4 Conclusions

In this contribution, we studied the properties of a number of knowledge-related networks and tried to interpret them in the light of some recent literature on complexity and on networks. The existence and properties of networks can find their roots in more general theories of complexity, such as that of Prigogine. According to these theories the emergence of structure, or order, in a system requires the system to be open and therefore away from equilibrium. Furthermore, transitions involving changes of structure in the system occur as the open system moves further away from equilibrium. Networks, constituted by nodes and links, can be considered the structure of socio-economic systems. Thus, both their existence and their transitions can in principle find an explanation if they are open systems away from equilibrium.

In general, we can expect the number of distinguishable networks in an economic system to grow as the variety of the same system grows. Furthermore, we can expect the connectivity/density of a network to vary in meaningful ways, for example to fall during the emergence phase of new structures and to rise during the subsequent phases of diffusion and maturation. The connectivity/density of a network can be expected to involve both advantages and disadvantages. Advantages can be

the greater ease of connecting different variables and the greater economic scope acquired by new firms when they improve their links to the rest of the economic system. Disadvantages are the growing rigidity of highly connected networks and the consequent difficulty in introducing into them further innovations. In Schumpeterian terms one could expect the entrepreneurial activity leading to innovations to be accompanied by a low network density and the subsequent routinization of innovations to be accompanied by a growing network density.

The actual networks of knowledge we studied tend to confirm a number of these expectations:

- In general, they show not only a heterogeneous distribution of links around nodes but also a highly asymmetrical distribution of strengths or intensities of the links.
- The conditions required for the existence of scale-free networks are generally present. In particular, during the emergence phase of new structures or of novelty, the number of nodes tends to increase rapidly.
- As a consequence, the connectivity/density of the network tends to fall during the emergence phase and to fall during the subsequent phases of diffusion and maturation.
- The expected time path of connectivity/density is more likely to fluctuate than to rise uniformly.
- We can expect a meaningful relationship to exist between variety growth and network dynamics. As pointed out above, the emergence of new economic species, giving rise to a growing variety, is likely to be accompanied by a fall in connectivity/density while the subsequent phases of diffusion and maturation are likely to see a rise in connectivity/density.
- Since variety only measures the number of nodes, the addition of connectivity, based on the interactions of the components of the system, provides us with a more complete analysis of the evolution of the complexity of a socio-economic system.

This contribution is much more speculative than definitive, concentrating on possible generalizations of the role and dynamics of socio-economic systems. A considerable amount of further work will be required to test and articulate the propositions suggested in this contribution. However, it seems that the themes treated are of central importance to understand the complex dynamics of socio-economic systems, in particular as we move towards a knowledge-based society.

References

Balzer, W., Moulines U., Sneed J., *An Architectonic for Science, The Structuralist Program*, Dordrecth, D. Reidel Publishing Co., (1987).

Barabasi A., Reka A., Emergence of scaling in random networks, *Science*, 286, 509–412, (1999).

Barabasi A., Reka A., Jeong H., Mean field theory for scale free random networks, *Physica* A, 272, 173–187, (1999).

Barabasi A.-L., *Linked: the New Science of Networks*, Cambridge, Perseus Publishing (2002).

Barabasi A., Bonabeau E., Scale free networks, *Scientific American*, 52–59, (May 2003).

Catherine D. Changement Technologique et Réseaux d'Innovation en Biotechnologie, Ph. D. Thesis, Grenoble (2005).
Chalmers A.F., *What is this thing called science*? Milton Keynes, Open University Press (1980).
Cohen D. All the world's a net, *New Scientist*, 174(2338), 24–29, April 13 (2002).
Dopfer K. *The evolutionary foundations of economics*. Cambridge, Cambridge University Press (2004).
Franck R., La pluralité des disciplines, l'unité du savoir et les connaissances ordinaires, *Sociologie et Société*, XXXI, 129–142, (1999).
Georgescu-Roegen N., *The Entropy Law and the Economic Process*, Cambridge, Mass, Harvard University Press (1971).
Kuhn T.S., *The Structure of Scientific Revolutions*, Chicago, The University of Chicago Press, (1962).
Landes D., *The Wealth and Poverty of Nations*, London, abacus, 1999, first published (1998).
Loasby B., Time, knowledge and evolutionary dynamics: why connections matter, *Journal of Evolutionary Economics*, 11(4), 393–412, (2001).
Mirowski P., *More Heat than Light*, Cambridge, Cambridge University Press (1989).
Nelson R.R., The co-evolution of technology, industrial structure and supporting institutions, *Industrial and Corporate Change*, 3(1), 47–63, (1994).
Nesta L.J.J., Saviotti P.P., Coherence of the knowledge base and the firm's innovative performance: evidence from the US pharmaceutical industry, *Journal of Industrial Economics*, 53, 105–124, (2005).
Nicolis G., Prigogine I. *Exploring Complexity*, New York, Freeman, (1989).
Popper K.R., *Objective Knowledge, An Evolutionary Approach*, Oxford, Oxford University Press, (1972).
Reka A., Jeong H., Barabasi A., Error and attack tolerance of complex networks, *Nature*, 406, 378–382, (2000).
Reka A., Barabasi A., Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74, 47–97, (2002).
Saviotti P.P., Considerations about knowledge production and strategies, *Journal of Institutional and Theoretical Economics* 160, 100–121, (2004).
Saviotti P.P., On the co-evolution of technologies and Institutions, in Weber M., Hemmelskamp J., *Towards Environmental Innovation Systems*, Berlin, Heidelberg, New York, Springer (2005).
Saviotti P.P., de Looze M.A, Maupertuis, M.A., Nesta L., Knowledge dynamics and the mergers of firms in the biotechnology based sectors, *International Journal of Biotechnology*, 5, 371–401, (2003).
Saviotti P.P., de Looze M.A, Maupertuis, M.A., Knowledge dynamics and the mergers of firms in the biotechnology based sectors, *Economics of Innovation and New Technology*, 14, 103–124, (2005).
Watts D.J., Strogatz S.H., Collective dynamics of 'small-world' networks, *Nature*, 393, 440–442, (June, 1998).

Chapter 3
Death of Distance in Science? A Gravity Approach to Research Collaboration

Koen Frenken, Jarno Hoekman, Suzanne Kok, Roderik Ponds, Frank van Oort, and Joep van Vliet

3.1 Introduction

One of the major transitions in recent scientific research is the rise of network theory motivating a variety of new research programmes in and across various disciplines. Economic geography has been no exception. The work on networks in economic geography can be divided into two types of research. First, there are studies on inter-firm networks and their impact on firm performance. For a large part, such studies have been carried out in the context of geographical clusters, which are often characterised by strong network relations (Uzzi, 1997). A second approach, an example of which is presented below, concerns the study of inter-regional networks and their impact on regional growth. Here, the unit of analysis are territories, typically sub-national regions. The interest in this topic stems from Castells (1996) and others who have argued that regional growth increasingly depends on a region's position in global networks rather than its specific local characteristics such as institutions, endowments and amenities ('space of flows' versus the 'space of places').

The reorientation in economic geography from the study of the 'space of places' to the 'space of flows' has lead some to argue that a new 'relational economic geography' paradigm is emerging. In such a paradigm, territories are not to be seen as meaningful unit of analysis with certain objective characteristics, but as 'socially constructed' in the ongoing interactions between social actors (Bathelt and Glückler, 2003). Such a conception fits well with the concept of the knowledge-based society where economic development is increasingly dependent on intangibles.

The study of inter-regional networks and the Castells thesis also relate to evolutionary economics and its application to economic geography (Boschma and Frenken, 2006). In recent evolutionary models of network formation, network evolution is understood as an entry process of new nodes connecting with certain probability to existing nodes depending on the latter connectivity (Barabasi and Albert, 1999). This logic of 'preferential attachment' explains the emergence of

K. Frenken (✉)
Regional Research Centre Utrecht (URU), Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands
e-mail: k.frenken@geo.uu.nl

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_3,

spatial core-periphery structure among regions as a process of network growth. Such networks concern, for example, transportation systems, multinational corporations and labour mobility flows. More recent models have explicitly incorporated geography in evolutionary models of network evolution by having the connection probability of a new node to an existing node also depend on the geographical proximity between two nodes (Guimerà and Amaral, 2004; Barrat et al., 2005). In this way, a parameter can be introduced to reflect transportation costs such that different network structures can be explained by differences in transportation costs.

In this study, we are interested in inter-regional networks of scientific knowledge production as a specific example of spatial networks. Popular belief holds that geography no longer matters in scientific collaboration. With the arrival of cheap air travel, English as a global language and the Internet, science has become truly global – at least according to common wisdom. What is more, increasing funding opportunities to engage in international partnerships have further facilitated long-distance collaboration. A growing number of studies on international collaboration seem to evidence this trend (Narin et al., 1991; Luukkonen et al., 1993; Frenken, 2002; Wagner and Leydesdorff, 2005; Maggioni and Uberti, 2007).

Yet, without disputing the secular trend of the internationalisation of scientific research in recent history, the 'death of distance' hypothesis has not been proven in this particular field. One may wonder whether the forces that are 'flattening' the world indeed removed the geographical barriers to collaborate in science. Earlier studies looking at inter-regional collaboration found that geography is still relevant in facilitating scientific collaboration within countries. Studies on collaboration within the UK (Katz, 1994), China (Liang and Zhu, 2002) and The Netherlands (Ponds et al., 2007) show that geographical distance reduces the probability of researchers to collaborate.

We test the 'death of distance in science' hypothesis by focusing on both international and inter-regional research collaboration based on scientific publications with multiple addresses. Our data set consist of three distinct subsets that cover geographical areas at several spatial levels of aggregation. We explain the collaboration intensity between 36 countries in the world, 1316 regions in Europe and 40 regions in the Netherlands from their respective scientific output and geographical distance using gravity equations. In addition to possible barriers stemming from geographical distance, we also analyse barriers stemming from 'institutional distance' in the form of national borders and in the form of dissimilarities between organisational backgrounds, respectively.

3.2 Science and Proximity

If anything has characterised knowledge production in science during the twentieth century, it is its increased collaborative nature (Meyer and Bhattacharya, 2004). Co-authorships accounted for less than 10% of all publications at the start of the twentieth century, while co-authorships account for over 50% of all publications

at the end of the twentieth century (Wagner-Doebler, 2001). The share of international collaboration has also been increasing (Narin et al., 1991; Wagner and Leydesdorff, 2005).

The general facilitator of this trend has been technological advance in transport and in information and communication technology. Yet, there are also specific reasons that explain the increasing tendency to collaborate. With the universe of scientific knowledge ever expanding, researchers need to specialise to remain able to contribute to state-of-the-art knowledge production. Specialisation in turn necessitates to collaborate with relevant partners, which may only be found over longer distances. As the costs of training and research infrastructures are increasing, collaboration also provides opportunities to pool resources and to realise savings by avoiding duplication of research efforts (Katz and Martin, 1997).

Collaboration is expected to bring intellectual benefits from the cross-fertilisation of ideas that previously were unconnected. One way to indicate these benefits is by comparing citation rates. Co-authored papers receive more papers than single-authored ones, and internationally co-authored papers receive more citations than nationally co-authored ones (Narin et al., 1991; Katz and Martin, 1997; Frenken et al., 2005).

At the national and European level, particular funding schemes provide economic incentives to promote collaborative knowledge production. For instance, the particular aim of the European Union is to create an integrated pan-European research system (i.e. European Research Area). Hence, their funding schemes are explicitly focused on funding international research projects and on removing barriers that currently hinder researchers. The financial efforts of the European Union for collaboration in science and technology have once again been increased substantially in the seventh framework programme (2007–2013).

While the internationalisation trend in research collaboration has received a lot of attention in recent times, only a few scholars have focused on the specific role of geography in scientific knowledge production. Yet, we hypothesise that geography is still important for research collaboration for reasons related to the background of the scholar as well as to the context in which he/she operates. With regard to the latter aspect, transportation costs are still present and in view of this, costs of collaboration are expected to increase as a function of geographical distance. Hence, two researchers that are geographically proximate are more inclined to collaborate as compared to two researchers that are geographically distant. Furthermore, many barriers to collaboration still have to be overcome when crossing national borders as most of the relevant institutions such as property right regimes, labour markets, university regulations and funding schemes are still organised predominantly at the national level (Edquist and Johnson, 1997). Accordingly, two researchers operating in the same country are more inclined to collaborate as compared to two researchers operating in two different countries.

With respect to the background of the scholar, barriers exist when researchers from different organisations are collaborating due to differing goals and underlying incentive structures. For instance, academic scholars want to maximise the diffusion of their knowledge, while industrial agents want to minimise such diffusion.

The complexity of these collaborations renders it generally impossible to encode all contingencies in a contract and consequently, these collaborations have to rely, at least partially, on less formal institutions thereby reducing the risk of opportunism. Therefore, it is expected that in the case of collaboration between academic and non-academic organisations geographical proximity may be supportive in establishing successful partnerships. Geographical proximity may help to overcome problems related to differing goals and incentives, because of a common interest in exchanging labour, accessing local funds and mutual trust induced by informal contacts and interaction. Thus, two researchers from organisations with similar backgrounds are more inclined to collaborate over longer distances than two researchers with different backgrounds.

In short, despite the decreased costs associated with long-distance collaboration due to cheap air travel and the Internet, and despite the efforts of the European Commission to further promote international collaboration with the European Union, we expect 'proximity' – both in the form of geographical and in the form of institutional proximity – to remain an important determinant of research collaboration (Gertler, 2005; Boschma, 2005; Torre and Rallet, 2005).

3.3 Data and Methodology

3.3.1 Data

In this study, the quantity of scientific collaboration is measured using co-publications. Scientific publications are the most common form of output in scientific research, which implies that collaboration in scientific research will often be reflected in a co-publication. The data on these co-publications have been retrieved from *web of sciences* (wos). Web of science contains information on publications in all major journals in the world from 1988 onwards. From this database we retrieved the address information on publications and constructed three distinct data sets.

The first data set concerns international co-publications between countries for 36 countries in the world for the year 2004. The countries were selected based on their scientific output and population numbers.[1] Subsequently, the number of collaborations between countries was identified by using the AND query in the address field applied to each possible pair of countries and covering all scientific disciplines.

[1] More specifically, we selected all countries with a population of more than one million inhabitants and with a total scientific output of more than 5000 articles in 2004. We ended up with the following 36 countries: (1) Argentina, (2) Australia, (3) Austria, (4) Belgium, (5) Brazil, (6) Bulgaria, (7) Canada, (8) China, (9) Denmark, (10) Finland, (11) France, (12) Georgia, (13) Germany, (14) Greece, (15) Hungary, (16) Ireland, (17) Israel, (18) Italy, (19) Japan, (20) Korea, (21) Mexico, (22) Netherlands, (23) New Zealand, (24) Norway, (25) Poland, (26) Portugal, (27) Romania, (28) Singapore, (29) South Africa, (30) Spain, (31) Sweden, (32) Switzerland, (33) Taiwan, (34) Turkey, (35) United Kingdom, (36) United States.

The second data set (Frenken et al., 2007) was constructed for two scientific disciplines. The disciplines concern 'biochemistry and molecular biology' (BMB) and 'electrical and electronical engineering' (EEE), following the classification of Verbeek et al. (2002). This database consists of all inter-regional co-publications in the EU27 and Norway and Switzerland for the period 1988–2004. Regions are classified on the NUTS3 level, which roughly corresponds to labour market areas. We were not able to locate all addresses and also removed some remote locations from the database. The outcome is a total number of 1316 NUTS3 regions instead of 1329.[2] Hence, all addresses occurring in publications have been assigned to one of the 1316 NUTS3 regions in the aforementioned 29 countries in Europe.

The third database (Ponds et al., 2007) contains all scientific publications in the Netherlands for eight selected scientific disciplines in physical and life sciences, again following the classification of Verbeek et al. (2002).[3] More specifically, all publications with at least one address in the Netherlands have been retrieved for the period 1988–2004 and subsequently classified at the NUTS3 level. This database also distinguishes between three different types of organisations: academic organisations, firms and governmental/non-profit-making organisations. In order to do so, we used an algorithm with a list of abbreviations and words to assign each address to one of three types of organisations. For example, organisations with 'univ' in its name are assumed to be a university and therefore an academic organisation. Furthermore, specific names of Dutch research organisations were included in the algorithm. In the end, 99% of the organisations were assigned correctly to one of the three types of organisations and one of 40 NUTS3-regions.

In all the data sets, a collaboration link is represented by a co-publication with multiple addresses, either in different countries (database 1) or in different regions (databases 2 and 3). The collaboration intensity between region i and j, labelled I_{ij}, is then defined by the number of times addresses from these two countries/regions co-occur in a publication. Intensity of collaboration between two countries or regions is thus measured by 'full counting'. For example, if a publication contains three addresses in three different regions, the collaboration intensity between each pair of regions is 1. Alternatively, one can use fractional counting where a

[2] We were not able to locate the addresses within the greater urban areas of London and Manchester and as a result consolidated them into two new ones. Furthermore, we excluded some islands due to their remote locations and disproportional great geographical distances to other regions. These islands are: Guadeloupe Las Palmas (ES), Santa Cruz de Tenerife (ES), Guadeloupe (FR), Martinique (FR), Guyane (FR), Réunion (FR), Região Autónoma dos Acores (PT) and Região Autónoma da Madeira (PT). The outcome is a total number of 1316 NUTS3 regions instead of 1329.

[3] These technologies are: (1) agriculture and food chemistry, (2) biotechnology, (3) organic fine chemistry, (4) information technology, (5) optics, (6) semiconductor technology, (7) telecommunication technology, (8) analysis, measurement and control technology. The two scientific disciplines analysed in exercise 2 are subdisciplines of biotechnology and semiconductor technology, respectively.

co-occurrence of two regions in a publication is divided by the total number of collaborations. For example, if a publication contains three addresses in three different regions, the collaboration intensity between each pair of regions is 1/3.[4]

3.3.2 Gravity Model

We apply a gravity model to explain the number of co-publications between two countries or regions from the respective size of the entities and their geographical and institutional proximities.[5]

Spatial interaction, the process whereby actors at different points in physical space make contacts, can be revealed by applying an analogical model of Isaac Newton's Theory of Universal Gravitation (Tinbergen, 1962; Sen and Smith, 1995; Roy and Thill, 2004). In a gravity model, the gravitational force – in this case the collaboration intensity between two objects – is assumed to be dependent on the mass of the objects and the distance between them. The basic gravity equation is therefore as follows:

$$I_{ij} = \alpha_1 \frac{MASS_i^{\alpha_2} MASS_j^{\alpha_3}}{DISTANCE_{ij}^{-\alpha_4}}. \tag{3.1}$$

Taking logarithms on both sides of the equation and introducing general exponents, such a gravity model can be estimated using linear regression:

$$\ln I_{ij} = \ln \alpha_1 + \ln \alpha_2 MASS_i + \ln \alpha_3 MASS_j + \alpha_4 \ln DISTANCE_{ij} + \varepsilon, \tag{3.2}$$

where $MASS_i$ stands for the total number of publications in country/region i, $MASS_j$ stands for the total number of publications in country/region j, and $DISTANCE_{ij}$ for the geographical distance between two countries/regions.[6] It is important to take into account the total number of publications in a country or a region, because collaboration intensity is highly dependent on size. If collaboration would be random, most collaborations will automatically occur between the largest spatial units.

---

[4] The final matrix of inter-regional interaction strength based on full counting is very similar to the final matrix obtained by fractional counting.

[5] In this, we follow Maggioni and Uberti (2007) who applied this model to EU regions using datasets other than co-publications (co-inventorships, student mobility flows, hyperlinks and EU framework projects). A similar approach was also followed by Maurseth and Verspagen (2002) using inter-regional patent citation data.

[6] Because collaboration links are undirected by definition, we included a pair of countries/regions only once, which implies that the value of the coefficient of the two masses may slightly differ. Note also that in the second exercise we added one to all masses in order to allow for logarithmic transformations of observations without any publications.

In the first exercise we analyse collaborations between countries. Although measuring geographical distances between countries is not straightforward, we use a proxy that captures the flight distances between the capitals of the respective countries. We also include a dummy for the countries that are members of the European Union to test whether these countries are more inclined to collaborate with each other. So, we get

$$\ln I_{ij} = \ln \beta_1 + \beta_2 \ln MASS_i + \beta_3 \ln MASS_j + \beta_4 \ln DISTANCE_{ij} + \beta_5 EU_{ij} + \varepsilon. \tag{3.3}$$

In the second exercise, we analyse collaboration intensity between NUTS3 regions, where we use the logarithm of geographical distance in kilometres, where geographical distance $DISTANCE_{ij}$ is calculated between the central points of the regions (as the crow flies) using maps made available by the European Spatial Planning Observation Network (ESPON). In addition, we specify institutional proximity by a dummy $COUNTRY_{ij}$ which takes on a value of 1 for two regions belonging to the same country and 0 otherwise. So we get

$$\ln I_{ij} = \gamma_1 + \gamma_2 \ln MASS_i + \gamma_3 \ln MASS_j + \gamma_4 \ln DISTANCE_{ij} + \gamma_5 COUNTRY_{ij} + \varepsilon. \tag{3.4}$$

In the third exercise, we analyse collaborations between Dutch NUTS3 regions, where we distinguish between three different types of collaborations: collaborations between academic organisations, collaborations between academics and firms and collaborations between academic and governmental organisations. Since we were able to trace back the travel times between NUTS3-regions in the Netherlands, we used the more accurate variable $TRAVELTIME_{ij}$ instead of the physical distance between the regions.

$$\ln I_{ij} = \delta_1 + \delta_2 \ln(MASS_i MASS_j) + \delta_3 \ln TRAVELTIME_{ij} + \varepsilon. \tag{3.5}$$

Following Ponds et al. (2007), we treat masses here as the product, which is a different yet equivalent specification of the gravity equation.[7]

For collaborations between countries in the first data set and for collaborations within the Netherlands in the third data set, the gravity equation is estimated using negative binomial regression techniques. As we deal with count data and we have a conditional variance that is larger than the conditional mean (overdispersion), the negative binomial regression model seems to be most appropriate.

In the second data set (inter-regional collaboration) an excessive number of zero counts biases the results, for which we corrected by the use of a zero-inflated negative binomial regression. This method considers the existence of two (latent) groups

[7] The treatment of zeroes differs as well compared to the previous studies. For an overview of the exact data-treatment of the third exercise, we refer to Ponds et al. (2007).

Table 3.1 Specification of regression models

Model	Scale	Period	Countries	Disciplines and types of collaborations	Regression technique
1	Countries	2004	36 countries	All scientific disciplines	Negative Binomial
2–3	NUTS3	1988–2004	EU27 Norway Switzerland	Biochemistry and molecular biology (BMB) Electrical and electronical engineering (EEE)	Zero Inflated Negative Binomial
4–11	NUTS3	1988–2004	Netherlands	Eight disciplines in life sciences and physical sciences differentiating between academic, academic-firm and academic-governmental collaborations	Negative Binomial

within the population: a group having strictly zero counts and a group having a non-zero probability of counts different than zero. Correspondingly, its estimation process consists of two parts. The first part contains a logit regression of the predictor variables on the probability that there is no collaboration between two given regions at all. The second part contains a negative binomial regression on the probability of each count for the group that has a non-zero probability of count different than zero. A good technical discussion of the zero-inflated negative binomial model is provided by Long (1997). An overview of the three data sets and estimation techniques is provided in Table 3.1.

3.4 Results

Table 3.2 shows Model 1 reporting on the estimates for collaborations between countries. The alpha-statistic turns significant indicating that the estimates of the negative binomial regression model are most reliable in this case. The fit statistics of the model suggest that the added covariates adequately fit the data. Indeed size contributes positively indicating an increase in collaboration between two countries if the actors in these countries produce a larger number of publications.

The explanatory of main interest, geographical distance, shows a statistical significant effect on the knowledge collaboration between countries too. Geographical distance yields a negative and significant effect, indicating major impediments towards collaborations over longer distance. What is more, a significant effect for research collaborations within the EU is not found, suggesting that, apart from

Table 3.2 Regressions results for international collaborations between 36 countries in the world (Eq. 3.3)

	Model 1
Constant	3.711 (0.443)**
$MASS_i$ (ln)	0.535 (0.020)**
$MASS_j$ (ln)	0.754 (0.018)**
$DISTANCE_{ij}$	−0.425 (0.039)**
EU	0.076 (0.096)
Fit statistics	
Over dispersion (α)	0.571 (0.031)**
Log-likelihood	3795.263
Mc Fadden's Adj. R^2	0.119
AIC	12.068
N	630

Note: Significance levels: ** 0.99, * 0.95, Standard error in parentheses.

advantages that accrue from shorter distances between EU member states, collaborations between these states do not occur more often than collaborations between other countries in the world.

In Table 3.3 Models 2–5 are presented. The table shows the inter-regional regression models with both models presenting successively a negative binomial part, a zero-inflated part and some general fit statistics. The latter include tests checking

Table 3.3 Regressions results for inter-regional collaborations in Europe (Eq. 3.4)

	Model 2	Model 3
	BMB	EEE
Negative binomial part		
Constant	−5.401 (0.086)**	−4.064 (0.133)**
$MASS_i$ (ln)	0.649 (0.005)**	0.533 (0.009)**
$MASS_j$ (ln)	0.636 (0.005)**	0.552 (0.010)**
$DISTANCE_{ij}$ (ln)	−0.368 (0.010)**	−0.301 (0.016)**
$COUNTRY_{ij}$	1.160 (0.022)**	0.824 (0.036)**
Zero-inflated part		
Constant	7.366 (0.165)**	6.999 (0.202)**
$MASS_i$ (ln)	−0.770 (0.009)**	−0.851 (0.013)**
$MASS_j$ (ln)	−0.779 (0.009)**	−0.832 (0.014)**
$DISTANCE_{ij}$ (ln)	0.359 (0.021)**	0.423 (0.027)**
$COUNTRY_{ij}$	−1.359 (0.048)**	−1.112 (0.059)**
Fit statistics		
Over dispersion (α)	0.881 (0.014)**	1.333 (0.034)**
Vuong-statistic	27.250**	20.410**
Log-likelihood	−99774.550	−51301.529
Mc Fadden's Adj. R^2	0.458	0.439
AIC	0.231	0.119
N	865270	865270
Non-zero observations	25589	12531

Note: Significance levels: ** 0.99, *0.95. Standard error in parentheses.

whether the choice of zero-inflated negative binomial regression models is appropriate. Overall, the likelihood ratio test of over dispersion and the Vuong-statistic are significant suggesting that the zero-inflated negative binomial regression model presents the most efficient estimates.

It is essential to keep in mind that a positive sign in the zero-inflated part indicates that with a one percent positive change in the predictor, the chance of belonging to the 'strictly zero group' increases, ceteris paribus. Thus, the coefficients in the zero-inflated part should be interpreted reversely in comparison to the negative binomial part: a positive value in the negative binomial part has the same meaning as a negative value in the zero-inflated part and *vice versa*.

Again the estimations indicate that geographical proximity matters for collaborations. An increase in distance negatively affects the chance and intensity of collaboration between two regions. The dummy variable indicating institutional proximity (here, whether two regions belong to the same country) shows the expected positive signal and turns out to be significant. This indicates that apart from a general effect of geographical distance on research collaboration, there is also an extra effect of institutional distance on the chance and intensity of research collaboration.

In Tables 3.4, 3.5 and 3.6, Models 4–11 further zoom in and show the results for inter-regional collaborations between Dutch NUTS3-regions differentiated to institutional backgrounds. Within life sciences (Models 4–6), travel time has a significant and negative effect on the intensity of collaboration for all the three distinguished forms of collaboration.

The coefficient for travel time is higher for collaboration between academic and governmental organisations than for academic collaboration and collaboration between firms and academics. Differences are, however, relatively small. The higher coefficients for collaboration between academic and non-academic organisations suggest that geographical proximity is more important for these forms of collaboration.

In the case of the physical science-based technologies (Models 7–11), travel time has no significant effect on the intensity of academic collaborations at all. Collaborations between all other types of collaborations turn out to be significant with the exception of academic-company relations in the field of optics. Yet, in physical sciences differences between academic-governmental relations and academic-firm relations are absent with the exception of the fields of optics[8].

We thus obtain quite different results for academic collaborations when compared to all other types of collaborations. Even at the relatively small scale of the Netherlands, geographical proximity is still important for collaborations. Yet, the importance varies between forms of collaboration, and academic collaborations do not seem to be sensitive for geographical distance. The result suggests that university scholars are less sensitive for geographical distance in collaboration. One can

[8] A more explicit treatment of differences between scientific disciplines in the extent and reach of collaboration can be found in Ponds et al. (2007).

Table 3.4 Regression results for inter-regional collaborations in the Netherlands (Eq. 3.5)

Life sciences	Constant	MASS (ln)	TRAVELTIME_{ij}	Over dispersion	Log-likelihood	Adj. R^2	N
Model 4: Agriculture and food chemistry							
Total	−7.647 (0.186)**	0.857 (0.018)**	−0.008 (0.001)**	0.649 (0.053)**	−2481.06	0.297	1521
Acad	−8.363 (0.328)**	0.937 (0.030)**	−0.003 (0.001)**	0.411 (0.074)**	−713.76	0.284	324
Acad-com	−7.032 (0.308)**	0.957 (0.039)**	−0.004 (0.001)**	0.687 (0.100)**	−833.60	0.326	1024
Acad-gov	−8.177 (0.274)**	0.955 (0.028)**	−0.007 (0.001)**	0.855 (0.080)**	−1681.04	0.312	1444
Model 5: Biotechnology							
Total	−7.616 (0.184)**	0.851 (0.018)**	−0.009 (0.001)**	0.692 (0.054)**	−2598.28	0.292	1521
Acad	−8.122 (0.308)**	0.912 (0.028)**	−0.004 (0.001)**	0.338 (0.055)**	−713.90	0.299	324
Acad-com	−6.940 (0.323)**	0.956 (0.041)**	−0.005 (0.001)**	0.920 (0.119)**	−895.36	0.309	1024
Acad-gov	−8.258 (0.268)**	0.942 (0.027)**	−0.006 (0.001)**	0.919 (0.085)**	−1784.66	0.301	1444
Model 6: Organic fine chemistry							
Total	−7.417 (0.175)**	0.827 (0.017)**	−0.008 (0.001)**	0.710 (0.050)**	−2989.88	0.283	1600
Acad	−8.642 (0.281)**	0.943 (0.025)**	−0.004 (0.001)**	0.256 (0.044)**	−700.48	0.335	400
Acad-com	−7.663 (0.334)**	1.017 (0.043)**	−0.005 (0.001)**	1.142 (0.129)**	−1034.85	0.288	1089
Acad-gov	−8.769 (0.281)**	0.975 (0.028)**	−0.006 (0.001)**	1.100 (0.092)**	−1957.25	0.291	1600

Note: Significance levels: **0.99, *0.95. Standard error in parentheses.

Table 3.5 Regression results for inter-regional collaborations in the Netherlands (Eq. 3.5)

Physical sciences	Constant	MASS (ln)	TRAVELTIME$_{ij}$	Over dispersion	Log-likelihood	Adj. R^2	N
Model 7: Information technology							
Total	−7.047 (0.289)**	0.927 (0.034)**	−0.006 (0.001)**	0.841 (0.110)**	−991.23	0.312	1089
Acad	−6.838 (0.516)**	0.953 (0.066)**	−0.002 (0.001)	0.798 (0.123)**	−481.66	0.246	289
Acad-com	−5.874 (0.343)**	0.927 (0.049)**	−0.006 (0.001)**	0.683 (0.142)**	−434.24	0.372	784
Acad-gov	−5.381 (0.448)**	0.894 (0.071)**	−0.006 (0.001)**	0.777 (0.143)**	−494.71	0.242	441
Model 8: Optics							
Total	−7.890 (0.271)**	0.955 (0.030)**	−0.003 (0.001)**	0.581 (0.079)**	−1076.16	0.322	1024
Acad	−8.124 (0.501)**	1.043 (0.061)**	−0.000 (0.001)	0.420 (0.081)**	−447.97	0.256	225
Acad-com	−6.831 (0.335)**	0.935 (0.046)**	−0.002 (0.001)	1.030 (0.152)**	−667.01	0.325	900
Acad-gov	−5.517 (0.445)**	0.881 (0.065)**	−0.005 (0.001)**	0.686 (0.123)**	−411.03	0.290	361
Model 9: Semiconductor technology							
Total	−7.612 (0.291)**	0.954 (0.033)**	−0.005 (0.001)**	0.518 (0.077)**	−815.79	0.3522	784
Acad	−7.477 (0.475)**	0.974 (0.057)**	−0.001 (0.001)	0.312 (0.061)**	−377.47	0.310	196
Acad-com	−6.113 (0.398)**	0.902 (0.060)**	−0.004 (0.020)*	1.557 (0.248)**	−480.78	0.270	529
Acad-gov	−5.487 (0.504)**	0.882 (0.077)**	−0.005 (0.001)**	0.932 (0.156)**	−413.58	0.261	324

Note: Significance levels: **0.99, *0.95. Standard error in parentheses.

Table 3.6 Regression results for inter-regional collaborations in the Netherlands (Eq. 3.5)

Physical sciences	Constant	MASS (ln)	$TRAVELTIME_{ij}$	Over dispersion	Log-likelihood	Adj. R^2	N
Model 10: Telecommunication technology							
Total	−7.722 (0.278)**	0.964 (0.032)**	−0.004 (0.001)**	0.588 (0.083)**	−975.47	0.344	1089
Acad	−7.326 (0.469)**	0.989 (0.060)**	−0.002 (0.001)	0.558 (0.096)**	−488.21	0.267	289
Acad-com	−6.611 (0.363)**	0.960 (0.049)**	−0.004 (0.001)**	0.883 (0.159)**	−497.37	0.354	784
Acad-gov	−5.688 (0.410)**	0.908 (0.063)**	−0.004 (0.001)**	0.513 (0.111)**	−456.64	0.279	441
Model 11: Analysis, control & measurement technology							
Total	−8.107 (0.224)**	0.911 (0.022)**	−0.005 (0.001)**	0.637 (0.064)**	−1746.27	0.332	1521
Acad	−7.695 (0.420)**	0.914 (0.043)**	−0.002 (0.001)	0.583 (0.087)**	−685.66	0.236	289
Acad-com	−6.736 (0.301)**	0.932 (0.039)**	−0.006 (0.001)**	0.868 (0.103)**	−975.01	0.319	1089
Acad-gov	−6.333 (0.289)**	0.855 (0.033)**	−0.007 (0.001)**	0.734 (0.101)**	−958.93	0.340	1156

Note: Significance levels: **0.99, *0.95. Standard error in parentheses.

assume that academic collaborations involve the production of knowledge that is more codified than the knowledge produced in collaborations in which a firm or government agency participates. In this light, our results are in line with the idea that collective production of codified knowledge is less dependent on face-to-face contact, and thus less sensitive to geographical distance.

3.5 Conclusions

In this contribution, we tested the 'death of distance' hypothesis for research collaboration using data on publications with multiple addresses. In contrast with previous studies focusing on collaborations between nation states, this study also analysed inter-regional collaboration both within countries and between countries. We tested the effect of geographical proximity and institutional proximity using the gravity equation and found strong evidence that geographical distance and national borders still hamper research collaboration. However, by distinguishing between different types of collaborations for the Netherlands, we find that geographical proximity matters significantly less in establishing collaboration between academic organisations than in case of all other types of collaborations.

In light of research policy, our analysis proves that policies to enhance the propensity of collaboration over long distances and across national borders remain necessary. We found evidence that the effect of geographical proximity exists independently of national borders, suggesting that the process of integration between as well as within countries is incomplete. This means that in further integrating research systems there is a role for actors at several spatial scales. Thus, in light of European policy the efforts to create a European Research Area seem well justified but also need to be complemented by efforts of its member states.[9]

References

Barabasi AL., Albert R (1999) Emergence of scaling in random networks. Science 286(5439): 509–512

Barrat A, Barthelemy M, Vespignani A (2005) The effects of spatial constraints on the evolution of weighted complex networks. Journal of Statistical Mechanics Art. No. P05003

Bathelt H, Glückler J (2003) Toward a relational economic geography. Journal of Economic Geography 3: 117–144

Boschma RA (2005) Proximity and innovation. A critical assessment. Regional Studies 39(1): 61–74

Boschma RA., Frenken K (2006) Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. Journal of Economic Geography 6(3): 273–302

Castells M (1996) The Rise of the Network Society. Blackwell: Oxford.

Edquist C, Johnson B (1997) Institutions and organisations in systems of innovations, in: Edquist C. (ed.) Systems of Innovation. Technologies, Institutions and Organizations. Pinter: London, pp. 41–63

Frenken K (2002) A new indicator of European integration and an application to collaboration in scientific research. Economic Systems Research 14(4): 345–361

[9] For an in-depth discussion of EU research policies in this context, see Frenken et al. (2007).

Frenken K, Hoekman J, Van Oort F (2007) Towards a European Research Area (Rotterdam/The Hague: NAi Uitgevers/RPB), www.rpb.nl

Frenken K, Hölzl W, de Vor F (2005) The citation impact of research collaborations: The case of European biotechnology and applied microbiology (1988–2002). Journal of Engineering Management and Technology 22(1–2): 9–30

Gertler MS (2005) 'Being there': proximity, organization, and culture in the development and adoption of advanced manufacturing technologies. Economic Geography 71(1): 1–26

Guimerà R, Amaral LAN (2004) Modelling the world-wide airport network. European Physical Journal B 38 (2): 381–385

Katz JS (1994) Geographical proximity and scientific collaboration. Scientometrics 31: 31–43

Katz JS, Martin BR (1997) What is research collaboration? Research Policy 26 (1): 1–18

Liang LM, Zhu L (2002) Major factors affecting China's inter-regional research collaboration: Regional scientific productivity and geographical proximity. Scientometrics 55 (2): 287–316

Long JS (1997) Regression Models for Categorical and Limited Dependent Variables. Advanced Quantitative Techniques in the Social Sciences, Volume 7. Sage publications: Thousand Oaks, CA

Luukkonen T, Tijssen RJW, Persson O, Sivertsen G (1993) The measurement of international scientific collaboration. Scientometrics 28: 15–36

Maggioni MA, Uberti TE (2007) International networks of knowledge flows: an econometric analysis. In: Frenken, K. (ed), Applied Evolutionary Economics and Economic Geography. Edward Elgar: Cheltenham UK, pp. 230–255

Maurseth P, Verspagen B (2002) Knowledge spillovers in Europe: a patent citation analysis. Scandinavian Journal of Economics 104(4): 531–545

Meyer M, Bhattacharya S (2004) Commonalities and differences between scholarly and technical collaboration. An exploration of co-invention and co-authorship analyses. Scientometrics 61: 443–456

Narin F, Stevens K, Whitlow ES (1991) Scientific cooperation in Europe and the citation of multinationally authored papers. Scientometrics 21(3): 313–332

Ponds R, Van Oort FG, Frenken K (2007) The geographical and institutional proximity of scientific collaboration networks. Papers in Regional Science 86: 423–443

Roy JR, Thill J (2004) Spatial interaction modelling. Papers in Regional Science 83: 339–361

Sen A, Smith TE (1995) Gravity Modelling of Spatial Interaction Behaviour. Springer: Berlin

Tinbergen J (1962) Shaping the world economy. Suggestions for an International Economic Policy. The 20th century fund: New York

Torre A, Rallet A (2005) Proximity and localization. Regional Studies 39(1): 47–59

Uzzi B (1997) Social structure and competition in interfirm networks: the paradox of embeddedness. Administrative Science Quarterly 42(1): 35–67

Verbeek A, Debackere K, Luwel M (2003) Science cited in patents: A geographic "flow" analysis of bibliographic citation patterns in patents. Scientometrics 58(2): 241–263

Wagner C, Leydesdorff L (2005) Mapping the network of global science comparing international co-authorships from 1990 to 2000. International Journal of Technology and Globalization 1(2): 185–208

Wagner-Doebler R (2001) Continuity and discontinuity of collaboration behaviour since 1800 – from a bibliometric point of view. Scientometrics 52: 503–517

Chapter 4
Evolution and Dynamics of Networks in 'Regional Innovation Systems' (RIS)

Frank Beckenbach, Ramón Briegel and Maria Daskalakis

4.1 Introduction

In 1992, Cooke introduced the notion of 'regional innovation systems' (RIS) (Cooke, 1992). The notion synthesizes two strands of research: the discussion about 'national innovation systems' (NIS) and empirical studies on regional development dynamics.

The research about NIS was founded by Freeman (1987) and Lundvall (1988) and rapidly gained acceptance (e.g., McKelvey, 1991; Nelson, 1993; Metcalfe, 1995). A lot of studies specified the notion of NIS since then (e.g., Freeman, 1995; Lundvall et al., 2002; Lundvall, 2007). The background of the seminal work of Freeman and Lundvall was a new perspective on innovation. At that time, Kline and Rosenberg (1986) had introduced their model of a 'chain linked process of innovation.' The concept of NIS goes beyond this model in that it takes into consideration the influences of space, common norms, institutions, and cooperation on the innovation process. The characteristics of these aspects together with the ability of the firms to absorb and transform knowledge constitute an NIS.

During the 1980s seminal studies about the fundamentals of regional development have been done, including the investigation of regional innovative capacities (e.g., in the Emilia Romagna (Italy), Baden-Württemberg (Germany), and Wales (England)) (cf. Sabel et al., 1989; Cooke and Morgan, 1990, 1991). Those studies gave evidence that not only the national level is of importance for the analysis of innovation systems but also the regional level. Particularly, not only as regards common norms and culture, but also as regards regional policy institutions a more 'locally' oriented perspective was called for.

Hence, following Asheim and Isaksen (2002, p. 83) RIS can be defined as '. . .places where close inter-firm communications, social structures, and the institutional environment may stimulate socially and territorially embedded collective learning and continuous innovation.' This definition is backed by the observation

F. Beckenbach (✉)
Department of Economics, University of Kassel, Nora-Platiel-Str. 4, 34109 Kassel, Germany
e-mail: beckenbach@wirtschaft.uni-kassel.de

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_4,

that agents in a region can intensify their relations due to 'proximity' in terms of space and social as well as technological compatibility. This proximity is promoted (and in some cases even initiated) by different forms of institutions ranging from 'hard' infrastructure and regulation to 'soft' mental models and trust. In comparing RIS with regions where no such integration of different resources takes place, researchers claim that making use of regional proximity is a source of additional generation and diffusion of knowledge, a faster learning process, and an increase of innovations. Consequently, the selective formation of RIS is used as an explanation for the significant differences in the development of regions in terms of gross domestic product (GDP) and employment.

For rendering plausible these allegations about the additional yields of RIS it is not sufficient to relate the regional boundary conditions (e.g., in terms of political regulations and knowledge transfer from universities) on one side and the outcome (e.g., in terms of innovations, GDP, and employment) on the other side. Rather, it is necessary to analyze the internal processes in RIS more closely. First, innovations – representing a specific mode of action – are neither 'coming out of the blue' nor are they automatically generated by competitive market conditions. This mode of action requires agents willing to innovate and promoting the corresponding activities (thereby leaving other modes of actions). Second, attributing the binary distinction 'innovation' vs. 'not innovation' to these agents is too simplistic in the context of RIS. As regards the regional effect, mere imitations are different from innovations; incremental innovations have to be distinguished from radical innovations, and individual innovations differ from cooperative innovations. Third, multiple coordination mechanisms interact in the formation of RIS: market relations are supplemented as well as structured by institutional relations and network relations are overarching both. Hence, analyzing RIS more closely necessitates to take into account the involved agents, their different modes of action, and their embeddedness in different coordination mechanisms.

Due to the complexity of the subject matter and to the dominant aggregated perspective (meso-level) of the research about RIS, the aforementioned topics are not yet systematically taken into account. A starting point for such a research endeavor is simulation studies about the importance of proximity and networking for innovations in a regional context using an agent-based perspective (cf. Brenner, 2001; Gilbert et al., 2001; Zhang, 2003). The present elaboration follows this strand of research, but tries to be more specific in terms of integrating the agent-based perspective and regional network analysis.

Taking agents as the basic unit of a RIS, it seems necessary to pick up modern insights of cognitive psychology (cf. Anderson, 2000) for explaining the agent's repertory of the ways to act as well as the conditions for selecting a specific mode of action in a given situation (in terms of internal and external states). Even if there is a common understanding of what is going on within a given mode of action (like routine, imitation, innovation) it is crucial for a behavioral foundation of RIS to focus the triggering conditions for the different modes of action (including innovation and networking). Hence, we use a multi-mode approach to human action (cf. Camerer et al., 2005; Svenson, 1996; Louis and Sutton, 1991) and systematize it

by incorporating insights from the Ajzen school as well as the Carnegie school of behavioral research.

This agent-based perspective implies a specific approach to networks as a part of activities in RIS. Contrary to the usual top down perspective focusing global network properties (like efficiency, stability) a bottom up perspective is required taking into account the agent's motivations, capabilities, and constraints to act in a network context.[1] Due to the different coordinating mechanisms operating at the same time in a RIS, it seems appropriate not to subsume all relations in that system under a broad notion of network but to reserve this notion for a specific way to coordinate activities lying between market coordination on one side and hierarchical inner firm coordination on the other side (cf. Ménard, 2004). Taking into account the 'distance' (in spatial, social, cultural, technological, or cognitive terms) necessarily inherent in market operations (cf. Cowan, 2004) as well as the 'threat' (again in spatial, social, cultural, technological, or cognitive terms) of hierarchical relations (cf. Miller, 1992), networks are creating a mid-term proximity for which the region is the generic place.[2]

Integrating the agent-based perspective and regional network analysis then has a two-fold meaning: on one side, the micro-economic conditions for regional networking (constraints as well as possibilities) have to be specified; on the other side, the feedback from ongoing regional networks to the level of individual activity has to be included in such an analysis. In pursuing such a perspective, Sect. 4.2 addresses the required conceptual foundations: in Sect. 4.2.1 layers of RIS in terms of modes of action and institutional embeddedness are distinguished; in Sect. 4.2.2 the factors determining the behavior of agents and correspondingly their willingness to contribute to a RIS are analyzed more closely. These conceptual foundations (as well as additional empirical observations) are the background for the architecture of the simulation model which is presented in Sect. 4.3. The results of this simulation model are dealt with in Sect. 4.4 especially focusing the network dynamics being the core of a RIS. Some conclusions are discussed in Sect. 4.5.

4.2 Conceptual Foundations

4.2.1 A Multi-layer Concept of RIS

Taking into consideration that RIS have a specific environment and a differentiated internal structure, multiple layers can be distinguished according to the internal

---

[1] Hence, neither large 'scale-free' networks (Barabasi, 2003) nor the usual 'small-world' network approach (Watts, 1999) is appropriate for the perspective chosen here; for different reasons they are not related to the micro-level of economic agents.

[2] Cowan (2004) identifies three (inter-related) drivers toward a network-type of interaction and coordination: (i) 'Imperfectness' of markets as regards to product- and person-related background information; (ii) outsourcing of firm activities bringing about new coordination requirements; and (iii) transfer of knowledge as a specific part of economic interaction.

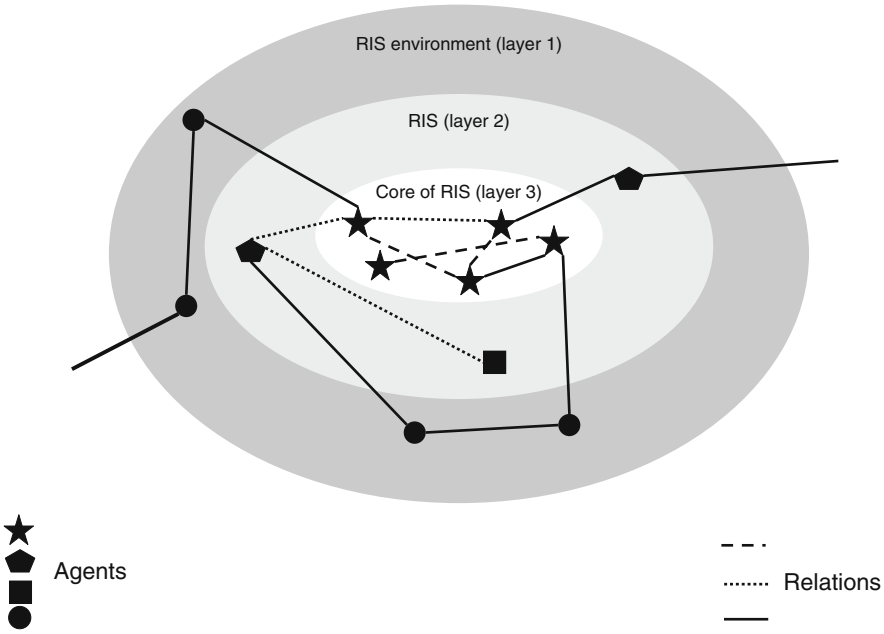


Fig. 4.1 'Onion concept' of RIS

mode of action of agents and their institutional embeddedness.[3] Every additional layer adds at least one additional type of relation to the relations already existing (cf. Fig. 4.1).[4] Switching between layers then implies a modification of the internal states of the agents being the entity promoting the different levels of activities. This modification may include a difference in the orientation toward other agents as well as a difference in mobilizing capabilities to act.

The *first* outer layer – actually the environment of RIS – is given by political (legal), geographical, and cultural attributes determining the interaction of agents. Apart from this common denominator, the agents are only loosely coupled (e.g., by regional markets and regional political processes). The agents in this layer are not (yet) engaged in a novelty-creating activity (innovation or imitation). The *second* layer – the first layer of RIS proper – is defined by additional institutional structures being relevant for those agents who are willing to create novelties (either by imitating or by innovating). The bases for these additional institutional components are declarations or commitments of a subset of agents (firms, politicians, and scientists)

[3] In the analysis of RIS the (multi-)layer concept normally is used for grasping the relations between functionally different groups and institutions like firms, universities, and political administration (cf. Braczyk et al., 2004; Etzkowitz and Leydesdorff, 2000). In the present study layers are characterized by different modes of activities of the same groups.

[4] The relations that are important for RIS are: market relations, institutional relations, knowledge relations, and – most importantly – network relations.

to promote a subpart of the region leading to some kind of supplementary structures (e.g., regional manager, regular meetings). Correspondingly, the agents on this layer are internally in a mode of action aiming to create a novelty (innovation or imitation).[5] Furthermore, having a strong willingness to innovate is a necessary (but not sufficient) condition for collaboration and cooperation in terms of innovation. Agents being on this layer have to figure out (at least potentially) the appropriate interests in the knowledge of others as well as the requirements for institutional settings dealing with transaction costs and overcoming opportunism in the context of networking. If the necessary conditions for such cooperation are given, the agents start to build up trust against each other, by reinforcing the process of collaboration and cooperation. Hence, the *third* layer – the core of RIS proper – is given by those agents who actually decide to cooperate in searching for and implementing innovations (they are internally in the cooperative innovation mode). In this layer there is an agreement about the goals pursued by innovative activities of different agents and the resources dedicated to this goal followed by processes to implement this agreement. Stabilizing trust between cooperating agents is a crucial factor for the success of these processes, i.e., its prosecution until all possible comparative advantages of cooperation (due to transferring specific knowledge, realizing economies of scale and scope, and implementing the division of tasks and parallelism in research activities) are exhausted. The existence of institutions stabilizing trust will increase the probability of success for these cooperation processes.

Taking the perspective of a single agent, he/she either remains on a given layer or switches between different layers according to the institutional embeddedness and to the mode of action he/she is pursuing. The RIS as a whole and its dynamics is specified by the frequency of the agents in the different layers over a longer time span.

The difference between the three layers of RIS can be further specified in terms of the respective knowledge dynamics as well as the institutional dynamics. The *knowledge dynamics* being a result as well as a condition for the agent's switching between the different layers can be characterized by referring to the different nature of knowledge and the various ways it comes into existence from the agent's perspective. Here, first a distinction is made between general knowledge which is more or less public and specific knowledge which is accessible at a cost. Second, a distinction is made between virtual and real knowledge, i.e., between knowing that there is something to know and actually knowing something.

On layer 1 nothing else but a constitution of virtual general knowledge about the region happens. Agents in a region are endowed with information about their legal and political boundary conditions as well as about the regional associations, public, and semi-public institutions in the region.

---

[5] As regards to the distance to the core of RIS, this layer can be further differentiated into two subparts, one consisting of agents with moderate willingness to create novelties (imitation) and another with agents having a stronger willingness to create novelties in terms of looking for innovations proper.

On layer 2 the establishment of additional regional institutions is tantamount to the (partial) transformation of virtual general knowledge into real general knowledge by way of a simple knowledge transfer. This transfer takes place on regional meetings, within regional associations and may be induced by political programs. Crucial for the knowledge transfer on layer 2 is an agreement on how to proceed in the endeavor to cooperate. Hence, to homogenize the procedural knowledge to a certain degree is another condition for success on this level. On the other side, it seems necessary to maintain some heterogeneity of the declarative knowledge as a basis for the division of research activity between different agents on level 2: this establishes one of the comparative advantages over research done by isolated agents. Correspondingly, the beginning of the communication between different regional agents (or groups of agents) leads to a (virtual) formation of specific knowledge relevant for the different domains in the regional context.[6]

Finally, on layer 3 the transfer of common knowledge is continued and accompanied by a (partial) transfer of specific knowledge. Both takes place within the cooperation formed for the sake of innovation. Private declarative knowledge is selectively revealed between cooperating partners without becoming public knowledge. The heterogeneity of the declarative knowledge between cooperating agents is an important source for the propensity to cooperate. Revealing (at least partly) this knowledge in the process of cooperation is therefore at the same time eroding (at least partly) the cooperation.

For the *institutional dynamics* on the different layers (being also – like knowledge – cause and effect of the RIS-process at the same time but more influenced by exogenous political and cultural factors) of RIS two attributes (and the respective attribute's expression) seem to be important (cf. Quéré and Ravix, 2003):

(i) the difference between 'hard' (technological, legal) and 'soft' (cognitive, social, cultural) institutions; and
(ii) the difference between a genesis of institutions by design and self-organization.

Going from level 1 to level 3 there is an increasing importance of soft institutions. On level 1 and especially on level 2 legal and political regulations as well as infrastructure are necessary initial conditions for any kind of cooperation procedures. These auxiliary instruments are required for overcoming regional inertia and routines and for establishing a critical mass for the RIS-dynamics. But for level-3 activities this is not sufficient: a self-enforcing trust building process strongly depends on soft institutions like shared visions and goals as well as mechanisms for punishing defecting agents.

Hard institutions are often designed by political actors and – due to lack of knowledge of their designers – not necessarily appropriate for the needs of their

[6] The main problem for low-developed regions in terms of knowledge is a lack of declarative knowledge (about technological fields, market domains, and other regional agents). Hence, the transfer of declarative knowledge is crucial for these regions.

respective target groups. Hence, their binding function for these groups requires a self-organized adaptation process of agents as well as of institutions (e.g., by exit and entry of agents and/or a modification of the institutional design). These adaptation processes are necessary for establishing coordination procedures in which these – adapted – hard institutions are accomplished by soft institutions.[7]

In the following, only the internal processes of the agents on the different layers are conceptualized more closely.

4.2.2 Behavioral Foundation of Agent's Mode of Action on the Different Layers of RIS

The behavioral foundation for agent's mode of action on the different layers of RIS is elaborated in two steps: (i) In the first step we are looking for a behavioral foundation for creating novelties comprising activities different from routine and choice. Here we differentiate between the behavioral elements leading to innovation and those elements leading to imitation. (ii) In the second step we are investigating more closely the behavioral elements in favor of innovation, and by drawing on additional circumstances we distinguish further between a behavioral foundation for a cooperative and an isolated way to innovate. Step (i) is backed by referring to two different concepts of action being synthesized for the given context. Step (ii) is done by specifying different conditions for the emergence and proceeding of innovation networks from a micro-perspective.

4.2.2.1 Synthesizing the Ajzen- and the Carnegie-Approach

In the given context of RIS there are two requirements for the behavioral explanation: First, the different modes of action specific for the different layers of RIS (routine, imitation and innovation) should be part of the explanandum. Second, this explanation should be empirically meaningful. Unfortunately, there is no concept in the behavioral scientific literature fulfilling simultaneously these two requirements.[8] Therefore, we synthesize two well-known behavioral approaches each of which has been applied to empirical problems and was used to explain more than one mode of action.

The focus of the *approach of Ajzen* is to explain intentional activities, i.e., activities resulting from a conscious plan to do something. According to this approach, this plan (intention) is influenced by three different cognitive factors (cf. Ajzen, 1991, p. 181): (i) the attitude of the agent toward the attributes of the

---

[7] The dynamics of RIS (i.e. their 'paths') differ according to the degree in which their institutions are dominated by hard institutions and according to the amount these institutions are simply designed or accomplished (or even dominated) by self-organized processes.

[8] Attempts in this direction are: Jager, 2000; Fagiolo and Dosi, 2003; Beckenbach, 2004; Briegel, 2006. Our suggestion is in the tradition of this research at the intersection of economics and psychology.

planned activity itself, (ii) the appropriateness of this activity for the social norms the agent is pursuing, and finally (iii) the agent's ability to control the intended activity. This implies that the agent tries to anticipate the salient attributes of the activity under consideration, that he is aware of norms he wants to follow (as well as the persons or groups representing these norms), and that he has an idea about his ability to control the intended activity. All these circumstances are valued by the agent and – in the case of multiple arguments – aggregated for each factor. Hence, for each of the factors (i–iii) there exists a subjective weight, influencing the overall intensity to pursue an intended activity.

From an economic perspective this approach can be assessed in a twofold manner: On one side, it contains the essential features of an economic approach in that it combines goals (attitudes) and constraints/endowments (abilities to control) for explaining activities. On the other side, this framework is broader than the economic standard approach to decisions in that it includes a subjective perception of social embeddedness (in terms of reference groups, norms, etc.) as a core element for intentional activities. This approach is also different from a game theoretic treatment of interaction patterns in that it is not bound to any type of common knowledge assumptions and well-defined strategies.

In the context of explaining agent's activities in RIS, the *first* shortcoming of this approach is the exclusion of dynamics. Even if Ajzen principally concedes that the experience of past behavior influences the initial conditions (in terms of subjective attitudes, subjective norms, and subjective control beliefs) for explaining present behavior (cf. Ajzen, 1991, p. 203), it is not clear at all how this influence takes place. The *second* shortcoming of the Ajzen approach in the given context is the neglect of other modes of action: in focusing the deliberative procedures leading to an intention, this approach deals exclusively with (a component of) decision processes. Neither circumventing intentions by automatic processes nor the search for new options of activity are explicitly part of this framework. Nevertheless, attitudes, norms, and control still seem to play a role when these other modes of actions are selected. So it appears to be worthwhile to integrate these modes of actions in the given framework.

Obviously there is only a narrow scope for integrating modes of action based on automaticity into the Ajzen approach. Ajzen himself (1991, p. 203) suggested that in the case of a very low level of subjective control (few elements of control with a low influence on behavior) intentions do not determine behavior. Rather, in this case there is a direct influence of the strong constraints on the resulting behavior. This kind of automaticity urged by the situation-specific circumstances is different from a 'learned' automaticity. The latter takes place if the situation remains rather similar for a longer time, and a specific way to act proves to be rather successful in this situation (cf. Verplanken and Aarts, 1999, p. 104).[9] Then the perception of situational cues seems to be sufficient for performing the well-known activity. Implementing an

[9] To act in a manner which conforms to the requirement of the situation can be a result of previously planned behavior; but it can also be derived from other types of activity such as teaching.

intention by consciously selecting attributes of action, reference persons, and assessing control abilities is dispensable in such a case. Rather, what takes place here is an unconscious activation of schemata or scripts stored in the long-term memory. This kind of learned automaticity (different from urged automaticity mentioned above) is at the core of what usually is called 'habit' or 'routine.'[10]

The *Carnegie-approach* is important in two respects: First, this approach sheds light on two usually neglected modes of decision making: routines and search. Furthermore, the decision units (especially firms) are not conceptualized as a consistent unit but rather as entities with internal conflicts being moderated in different ways (cf. Cyert and March, 1992, p. 214, p. 229). A common denominator of these features of such a decision unit is the restriction of rationality, i.e., of perceiving information as well as of the ability to transform this information into activity. Therefore, a specification of the bounded rational way to settle and use goals as well as capacities by constituting an aspiration level, by following satisficing behavior, and by varying organizational slack is the second reason for referring to this approach.

Originally, the *aspiration level* has been conceptualized in psychological field theory backed by observations about the context-dependence of the expected result of an activity and about the role of these expectations for future activities. In its economic adaptation, the core idea behind the aspiration level is to internally fix a level of goal attainment which is related to past experience and/or to the observable experience of other agents being in a similar situation (cf. Cyert and March, 1992, p. 162, p. 172).[11] Generally, the divergence between the aspiration level and the actual performance level is seen as a source for modifications in behavior. This is due to an evaluation according to which a negative discrepancy (performance level is lower than aspiration level) leads to the internal state of dissatisfaction whereas a positive discrepancy (performance level is higher than aspiration level) leads to an increase of ambition.

Orienting behavior toward such an aspiration level then defines *satisficing* instead of optimizing in goal attainment. Hence, the aspiration level gives a cue for dealing with bounded rationality: 'Actually satisficing is less a decision rule than a search rule. It specifies the conditions under which search is triggered or stopped, and it directs search to areas of failure' (cf. March, 1994, p. 27). A negative discrepancy triggers a different mode of action in terms of information gathering and risk taking (cf. Cyert and March, 1992, p. 228).[12]

---

[10] 'Habit' is the notion used in sociology and is meant to include a wide range of social and cultural explanantia; 'routine' is the notion used in economics and is more focused on explanantia inherent in the activity under consideration.

[11] In terms of the concept of Ajzen the aspiration level has to be classified as a norm because it indicates a social interaction leading to an 'appropriate' level of goal attainment.

[12] Basically the notions of aspiration level as well as satisficing are related to the individual. In organizations like firms the level of goal attainment may be group-specific and therefore be a source for a conflict. Managing this conflict is then a boundary condition for the pursuit of the original goal (cf. March and Simon, 1993, p. 132; Simon, 1997, p. 159).

The aspiration level as well as the satisficing goal attainment do not take into account the endowment of agents (in terms of knowledge, finance, and time) and their corresponding capacities to act. Given the (fast) dynamics of the endeavors to meet the aspiration level and the statics (or slow dynamics) of these endowments, there is a varying surplus in the capacity to act. This surplus is called *'slack.'* If this slack is large, it is seen as an additional source for novelty-generating procedures because then constraints are loose giving room for playful experimentation. Contrary to the 'failure-induced' search, this kind of search is 'success-induced' (cf. March, 1994, p. 31; cf. Cyert and March, 1992, p. 188; March and Simon, 1993, p. 203).[13]

This success-induced search can be specified on the level of the individual as well as on the level of organizations (firms). The corresponding individual trait is curiosity as a search for new information, new knowledge, and new experience for its own sake.[14] Recent research in this field reveals that curiosity is not simply a genetically programmed drive (activated in a crude stimulus response context)[15] but rather has cognitive sources either in searching for congruity and sense making or in practicing idle competences (cf. Loewenstein, 1994, p. 80). Hence, curiosity is in the neighborhood of creativity and intimately related to the above-mentioned phenomenon of slack.

On the level of firms, the slack is a necessary implication of the organizational 'coordination failure.' Given an organization consisting of a multitude of agents and resources, bounded rationality as well as opportunism will play a role in the coordination of these organizational elements. 'When the presence of slack relaxes coordination and control pressures, decision makers are free to pursue idiosyncratic local preferences' (cf. March, 1994, p. 31; Cyert and March, 1992, p. 41). This will release a search for novelties on the different levels of the firm because it opens up possibilities for the different organizational departments to strengthen their relative position. The harmonizing of these different (potentially) innovative activities will be a central task for the strategic management.

For the *behavioral synthesis* at stake here the architecture of the Ajzen approach is used as the basic heuristics. This is due to the broader scope of this approach and to its greater flexibility. But first, the factors used by Ajzen have to be specified for the domain of a firm in a competitive market economy and for the given context of explaining the dynamics of regional innovation. Curiosity, risk orientation, and expectations about the properties of an option are taken into account as personal

---

[13] Similar to the approach of Ajzen the endowment of the agent (and the correlating control capacity) is taken into account here. The difference to this approach is given by the assumption that the amount of control capacity does not simply determine an intention to act but rather a specific way to act, e.g., the switch to searching behavior.

[14] In the Carnegie-approach curiosity and the corresponding exploration drive is dealt with as a component of the risk attitude. Apart from this curiosity component, the risk attitude is seen to consist of a personal trait and risk taking depending on the level aspiration attainment (cf. Cyert and March, 1992, pp. 227–228; March, 1994, pp. 35–55).

[15] This notion of curiosity is related to a part of the work of Berlyne and was 'imported' into economics by Scitovsky (1976).

(or organizational) attitudes in the context of market and firms and are accomplished by including goals as an important feature of economic actors. Aspirations in terms of different economic goals are specified as the relevant 'norms.' The control component is derived from the usual elements of economic endowment (knowledge, finance, and time). Second, these factors are not used for explaining a specific activity (or intention to do that activity); rather, according to the necessary behavioral enhancement, they are used for explaining the selection of the mode of action[16] and hence, for explaining one source of the dynamics between the different layers of a RIS (cf. Fig. 4.2).

Integrating aspirations and slack in this framework allows for dealing with the behavioral dichotomy in terms of switching between search and routines. According to the context under investigation, here the search mode is further differentiated in a twofold manner: a distinction is made between a search for novelties already practiced by others (imitation) and novelties which are created by the searching agent (innovation).[17] These different modes compete for being activated by the agent due to endogenously generated forces. Activating imitation necessitates overriding the 'default mode' of routine. Practicing innovation requires higher innovation pressure to overcome the lower pressure toward mere imitation.

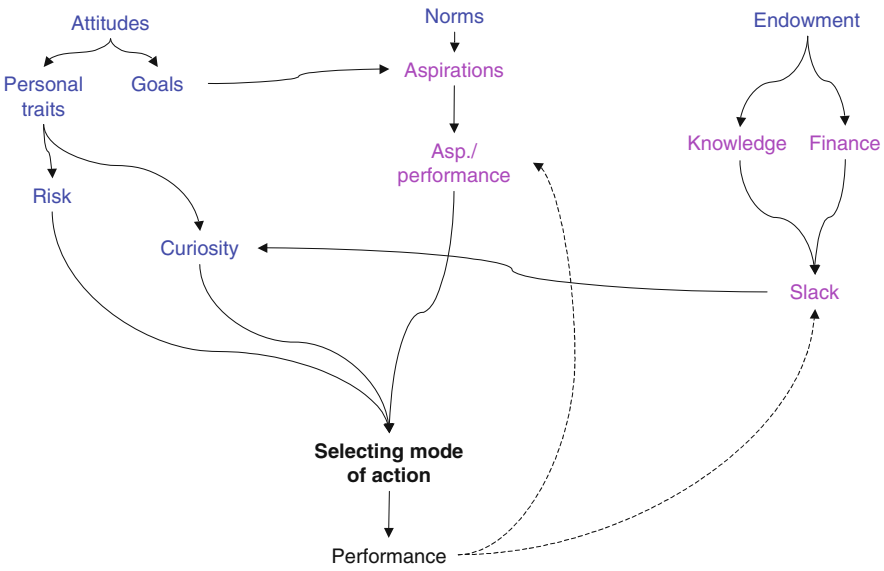


Fig. 4.2 Causal chain diagram for selecting mode of action

[16] Cf. Svenson, 1990; Louis and Sutton, 1991 and Jager, 2000 for such a multi-mode concept of human action.

[17] It is assumed that slack (and the curiosity derived therefrom) is only relevant in the case of innovations and that the expected costs of imitations are lower than the expected costs of innovation.

4.2.2.2 Micro-economic Conditions for Innovation Networks

In the previous section we have specified the behavioral conditions or the willingness of the agent to innovate. What has not dealt with is the question, whether, and if so why, agents choose to conduct their innovations in a collaborative manner. More precisely, the reasons why agents are moving to and from the core of the RIS are not explained. This moving as well as the collaborative innovation at the core constitutes the dynamics of the RIS. In the following, we will take into consideration the network aspect of RIS and distinguish two interdependent levels of analysis: the morphology of the network and the states of the agents involved in network activities.

4.2.2.3 'Social Network Analysis' (SNA)

Although networks are undoubtedly a central feature of RIS, the respective research usually does not comprise the network morphology and, accordingly, not the feedbacks between states of agents and the network as a whole. However, recently two aspects have been investigated: the morphology of patent networks (cf. Breschi and Lissoni, 2004; Cantner and Graf, 2006) on one side and the specific contents of the relationship between agents involved in an innovation network, namely the role of trust and knowledge transfer (cf. Nooteboom, 2004, 2006; Daskalakis and Monz, 2006; cf. Sect. 4.2.1 above) on the other side. Both research fields lack a unified perspective: whereas the former deals with the multiple contents of network relations, the latter is less concerned with the morphology of the network.[18] The common denominator of both fields of research is the social network analysis (SNA). Hence, it is worthwhile to take a closer look at the SNA and the options this paradigm might provide for the analysis of innovation networks.

The SNA provides a formal method to analyze the morphology of networks and the impacts it has on the network's performance. The main thesis of the SNA is that an agent is embedded in a social environment and is influenced by this. Agents (either individuals or organizations) are connected to other agents in their social environment by interaction. These connections between the agents are called 'relational ties' and the agents themselves are called 'nodes.' The main focus of the SNA lays on the analysis of the structure of these ties, especially on the pattern and/or regularities of these structures (cf. Wasserman and Faust, 2007, p. 3). Besides the structure, the SNA is also concerned with the contents of ties which might be, for example, biological elements, material resources, information, knowledge, or trust.

Within the SNA, a basic distinction is made between an analysis of individual agent networks, especially ego-centric networks, the analysis of sub-groups within networks, and the investigation of networks as a whole (e.g., Wasserman and Faust, 2007; Scott, 2005). The formal tools provided by the SNA differentiate between all three levels. Relevant measurements are, for example: centrality, cluster,

[18] An exception is given by Kauffeld-Monz and Fritsch, 2007.

cohesion, density, (mean) degree, multiplexity, path distance as well as value and direction of ties (cf. Wasserman and Faust, 2007).

There are several possibilities to apply SNA to the analysis of RIS. As regards the simulation model (Sect. 4.3), four measures are considered more closely: (i) multiplexity, (ii) values of graphs, (iii) direction of ties, and (iv) mean degree.

(i) *Multiplexity* refers to the circumstance that between two nodes (n_i, n_j) there might be more then one tie (Wasserman and Faust, 2007, p. 73). For example, if two agents at the core of the RIS are engaged in mutual knowledge transfer and if this transfer is based on mutual trust, then one tie $n_i \overset{k}{\longleftrightarrow} n_j$ reflects the knowledge (k) transferred and a second tie $n_i \overset{tr}{\longleftrightarrow} n_j$ contains trust (tr). Multiplexity can be measured by simply counting the number of different ties between two agents.

(ii) The *value* of a tie describes the frequency and/or intensity of the interaction (cf. Wasserman and Faust, 2007, p. 140). With regard to the relationship $n_i \overset{k}{\longleftrightarrow} n_j$, for example, the value ν_k is measured in terms of knowledge-units being transferred and/or in terms of frequency of interaction. Accordingly, the value of trust ties represents the level of trust. In case of multiplexity, the values of the respective ties are summed up.

(iii) For both measurements, the multiplexity and the value of ties, it is reasonable to distinguish between *undirected* and *directed ties*. The two measures introduced above are related to undirected ties. They give no information about differences in case the ties are decomposed, e.g., by analyzing the amount of knowledge given from n_i to n_j ($n_i \overset{k}{\longrightarrow} n_j$) and vice versa ($n_j \overset{k}{\longrightarrow} n_i$). Those differences are measured by directed ties. For example, if n_i is giving more knowledge to n_j than n_i receives from n_j then the value of the ties is different with $\nu_{ijk} > \nu_{jik}$. Different amounts of trust between n_i and n_j can be formalized accordingly.

(iv) The *degree* measures the number of ties with which one agent is connected to others. The minimum degree (d) is 0, if an agent has no connection; the maximum is given with $d=g-1$ (g being the number of nodes). The mean degree of a network comprises the ratio of the number of actual ties between the agents and the number of nodes.
Thus, the mean degree ($\bar{d}$) is given by

$$\bar{d} = \frac{\sum_{i=1}^{g} d(n_i)}{g} = \frac{2L}{g},$$

where L denotes the number of ties in the network. If directed ties are investigated, a difference can be made between the indegree (amount of ties n_i is receiving) and the outdegree (amount of ties leaving n_i).

If analyzing RIS from a social network perspective, it seems plausible to view RIS as network which is subdivided into several interconnected sub-networks.

Within our model, those sub-networks are expressed by the different layers of RIS. Applying the methods of the SNA to the analysis of the RIS-layers then provides useful insights to the internal components of the networks and the respective resulting output of the RIS. How many agents are connected in and between the layers and how does this share impacts the outcome of the RIS (e.g., in terms of completed innovations)? To what extent does the structure of degrees at the core influences the dynamics of the network? Are there observable negative effects of network size?

Applying SNA to the analysis of RIS, however, implies necessarily to take into account the inherent economic logic of innovation networks. First, a formal description of the relevant contents and strength of ties between agents in a RIS has to be given. Second, the dynamics of the networks within the RIS have to be considered: at least within the core of RIS and the movement from and to the core, the dynamics are supposedly much stronger than in the networks usually investigated with SNA tools (as, for example, friendship networks). This dynamic aspect is a challenge for the application of the SNA. Furthermore, if one aims to explain and analyze the movement between the layers of RIS, one has to take into account the reasons of why the agents are moving into the core (instead of conducting their innovation alone).

4.2.2.4 Network Behavior of Innovative Agents

If a combination of behavioral traits (risk attitude, curiosity) and a critical relationship between aspiration level and actual performance is given activating the agent's willingness to innovate, the question arises if such an agent is oriented toward doing the innovation alone or by cooperating with other agents in a network.

Filling up the gap in terms of micro-foundation for networking in general (cf. Sect. 4.1) necessitates to distinguish between the following:

- the triggering condition for building a network (in what situation does a firm look out for network partners?);
- the matching condition for the networking agents (what is expected from the partners in a network?);
- the exchange condition (which resources are transferred between network agents under what conditions?); and finally
- the replication condition (is there a positive or negative feedback of the network results on the continuation of the network?).

The ability of a network to overcome the 'distance' inherent in any market operation (cf. Sect. 4.1) is the common denominator of all these conditions. In economic terms this means to relate the cost for reducing the distance inherent in market relations to the yields which can be expected from this networking. The important costs in this context are the transportation cost (if distance is related to space)[19] and the transaction costs (in all other dimensions of distance). The yields of networks are

[19] To analyze these costs is the basic tenet of economic geography.

due to specificity advantages, to economies of scale and scope, to socializing risk (or uncertainty), and to solving a critical mass problem.

Cooperative innovations represent a specific type of network: (i) there are no structural asymmetries in terms of power and command between the networking agents; (ii) there is a formalized agreement about the goals of the network and the tasks of the network participants; and finally (iii) generating and exchanging knowledge is an important part of the network activities (cf. Fritsch and Franke, 2004; Antonelli, 2000; Sternberg, 2000).

An economic analysis of this type of network necessitates to relate the prospects for distance-reducing costs and yields to the nature of knowledge transformed and transferred in the network as well as to the nature of the relationships between the network partners. As regards the nature of knowledge from an economic perspective, the most important distinction is between 'public/explicit' on one side and 'private/implicit' on the other side. Between these two extreme cases, there are many possible degrees of these attributes. It seems reasonable to assume that the cost to acquire a given set of knowledge (a subspecies of transaction costs) is higher the more specific the knowledge, i.e., the more it tends to the 'private/implicit' extreme. Correspondingly, the more this is the case the higher is the specificity advantage the knowledge acquiring agent can expect. Contrary to this economic *specificity effect* of knowledge transfer, the *scale effect* is related simply to the number of knowledge components which are transferred between network partners. Here it seems appropriate to assume that the cost of transferring knowledge linearly depends on the number of knowledge components, whereas the yields are increasing with the number of knowledge components due to economies of scale and scope.[20] Hence, as regards the knowledge transfer, both the degree of knowledge specificity and the number of knowledge components are influencing the micro-economic efficiency prospects of a network.

Concerning the nature of the relationships between the network partners trust is the most important factor in economic terms. Corresponding to the literature about trust and dynamic transaction costs (cf. Nooteboom, 2002; Lorenz, 1999) it can be expected that trust operates on the side of cost and yields at the same time: On the one hand, the more often a cooperation between two agents has been successfully accomplished the lower are the transaction costs in terms of finding an appropriate partner and coordinating the activities. On the other hand, a well-functioning partnership makes it easier to find new knowledge components and to increase the economic yield (e.g., by improving the product quality).

In formal terms this means that the transaction cost (TC) of an innovation network are a function of the knowledge specificity (s), the number of knowledge components (q), and trust (tr) in terms of the frequency of past successful knowledge transfers:

$$TC = f(s, q, tr) \tag{4.1}$$

[20] Cf. the formal analysis of transaction cost and asset specificity in Williamson, 1985.

The yield which can be expected from that innovation network (Y_{cop}) is influenced by the same arguments:[21]

$$Y_{cop} = g(s, q) \tag{4.2}$$

The countervailing effects of these factors are depicted in Fig. 4.3.

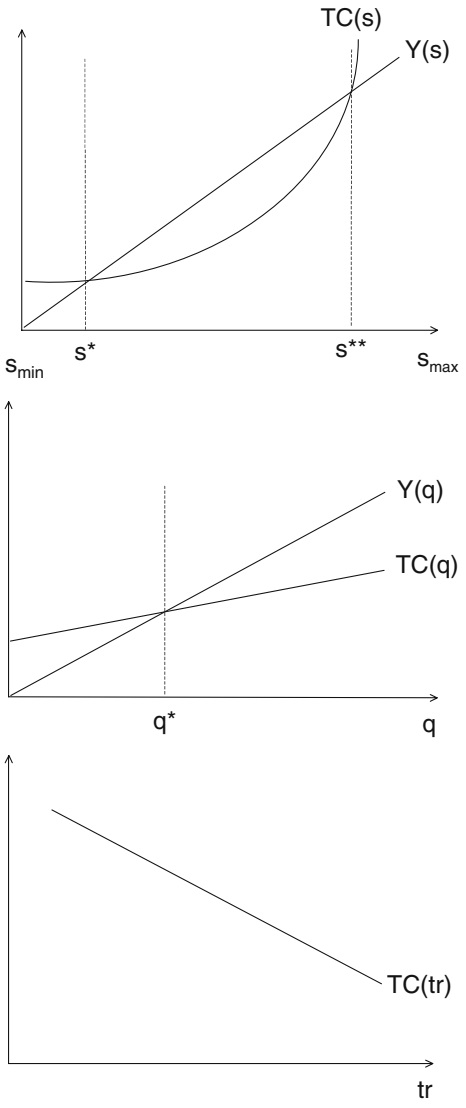


Fig. 4.3 Graphical representation of factors influencing the micro-economic network efficiency

[21] Due to strong irregularity we skip the influence of trust here.

These factors for the micro-economic efficiency of networks are not transparent for the agents: they are realized in a time-dependent manner due to constraints and path-dependencies of the agents.

Taking into account these peculiarities of innovation networks, the above-mentioned general micro-economic conditions for networking can be specified:

- the triggering condition comprises not only a willingness to innovate but more specifically a combination of individual traits and competitive environment leading to a preference for cooperative innovation;
- the matching condition implies a correspondence between the dominant innovation drive of the agent looking for network partners on one side and knowledge endowments of these partners (including that the expected yields of a network fulfilling this matching condition exceed the transaction costs) on the other side;
- the exchange condition is related to a transfer of different types of knowledge; and finally
- the replication condition is constituted by memorizing and communicating the experience of the network activity leading either to reinforcing or exiting the network.[22]

4.3 Simulation Model

4.3.1 Agents and Their State Variables

Agents are characterized by state variables and parameters in all the branches of the behavioral concept:

(i) They are endowed with financial and cognitive resources giving them more or less control over the environment they operate in. The cognitive resources comprise especially procedural and declarative knowledge. The procedural knowledge is tantamount to the ability of the agent to pursue different modes of action. These modes of action each comprise a specific way to perceive a situation and the possibilities to act, to evaluate these possibilities, and to select one way to act. The declarative knowledge is related to the domain of activity the agent is engaged in. It is composed of common and specific elements.

(ii) They are determined by attitudes, i.e., mental commitments giving them a basic orientation and mood in their way to act especially as regards selecting the different modes of action. Hence, it is not only an external situation which influences the selection of a mode of action; additionally, these attitudes play an essential role. Apart from risk acceptance and curiosity as basic personal traits, the goals of the agents are considered as an important element of attitude.

[22] In terms of the graphs depicted in Fig. 4.3, this means that the frequency of successful network in relation with a given partner is stagnating and the trust-dependent component of the transaction cost is no more diminishing.

Agents have two (monetary) goals, namely to gain profit and to achieve a high market share.

(iii) As proposed by SNA agents are considered as socially embedded entities being influenced by proximate reference groups or past interaction experience. According to the Carnegie-approach this is reflected in the notion of the aspiration level, the plimsoll line for goal achievement. This aspiration level is dynamically adjusted according to the actually reached current profit resp. market share (moving target). Furthermore, especially for the cooperative innovation, trust is an important result of interaction, being an agent-specific norm for selecting partners to cooperate with and for the decision about the continuation of a cooperation.

4.3.2 Modes of Action and Their Selection

In the given context the modes of action are routine (as the default mode), imitation, and innovation (the latter either as individual or as cooperative activity). Both novelty-creating modes (innovation and imitation) require not only cognitive effort, but also temporal and financial resources.[23] The temporal and financial cost of an imitation project is smaller than the one of an innovation project. However, as regards the possible returns, there is a disadvantage for an imitator compared to an innovator. The reason for this is that an imitator can only participate in the fraction of the total demand potential of the imitated innovative product which is not yet exhausted (in the course of diffusion of the product) at the time of accomplishment of the imitation. Generally, the expectation value for the demand potential assigned to a novel product just created is set proportionally to a certain power of the amount of declarative knowledge of the firm(s) that has/have created it, more precisely, to the number of knowledge domains where the firm(s) has/have got knowledge.[24]

Picking up the conceptual reflections in Sect. 4.2.2.1, the *innovation force* (F_2) itself consists of different components such as curiosity, the degree of satisfaction as regards profits, and the degree of satisfaction as regards market shares. These degrees of satisfaction are indicated by the relationship of the aspiration level for profits (asp) and for market shares (asm), respectively, to the corresponding actual performance level (p and m): asp/p, asm/m. The activation level of these different components is influenced by personal attitudes such as the exploration drive (w_0) as well as the weight and elasticity of the profit resp. market share aspiration (w_1 and ε_1 resp. w_2 and ε_2). Finally, the innovation force is modulated by the expected cost for the innovation endeavor (cin) and the risk acceptance (α) which maps the willingness to accept the higher risk of innovation compared to imitation. The aspiration

[23] The time to develop an innovation (resp. imitation) is set probabilistically (drawn from a uniform random distribution) for each innovation (resp. imitation) project; the minimal and maximal values of the corresponding distributions are model parameters.

[24] The background for this assumption is the positive relation between the broadness of knowledge and the firm's ability to meet the needs of (potential) customers.

levels are updated at the end of each time step according to the difference equation

$$\mathrm{asp(t+1)} = (1-\phi)\mathrm{asp(t)} + \phi\,\mathrm{p(t)} \tag{4.3}$$

(and analogously for asm) where ϕ is the flexibility of adaptation, which is another personal trait ($0 \leq \phi \leq 1$).

According to the discussion in Sect. 4.2.2.1, curiosity is strongly related to the phenomenon of 'slack,' i.e., the reserve capacities in terms of knowledge and finance. In any given time step this slack is tantamount to balancing the given state of knowledge and finance on one side and the amount needed of these resources for a given mode of action on the other side. Again, the intensity of curiosity triggered by this slack is depending on a personal trait, the exploration drive (w_0).

These considerations can be formalized as follows: We define three component forces f_i for curiosity (i=0), profit aspiration (i=1) and market share aspiration (i=2) by

$$f_0 = w_0(\mathrm{kr} + \mathrm{fr}) \tag{4.4}$$

$$f_1 = w_1 \left(\frac{\mathrm{asp}}{\mathrm{p}}\right)^{\varepsilon_1} \tag{4.5}$$

$$f_2 = w_2 \left(\frac{\mathrm{asm}}{\mathrm{m}}\right)^{\varepsilon_2} \tag{4.6}$$

Here kr and fr denote the knowledge resp. financial reserves of an agent in a given time step. The knowledge reserves are operationalized as the relation of the number of specific sharable knowledge domains (see below) where the agent possesses knowledge to the total number of sharable knowledge domains; the financial resources are operationalized as the share of the current profit with respect to the current turnover. Finally, the innovation force F_2 is defined by

$$F_2 = \alpha\,\frac{f_0 + f_1 + f_2}{\mathrm{cin}} \tag{4.7}$$

The *imitation force* (F_1) is different from the innovation force in three respects: First, curiosity (or the personal exploration drive) plays no role in it (f_0 is omitted here). Second, there is a comparative difference in expected cost: cim < cin. Third, the risk acceptance α, which was introduced to map the willingness to accept the higher risk of innovation compared to imitation (see above), is omitted here. Hence, the imitation force can be formalized as

$$F_1 = \frac{f_1 + f_2}{\mathrm{cim}} \tag{4.8}$$

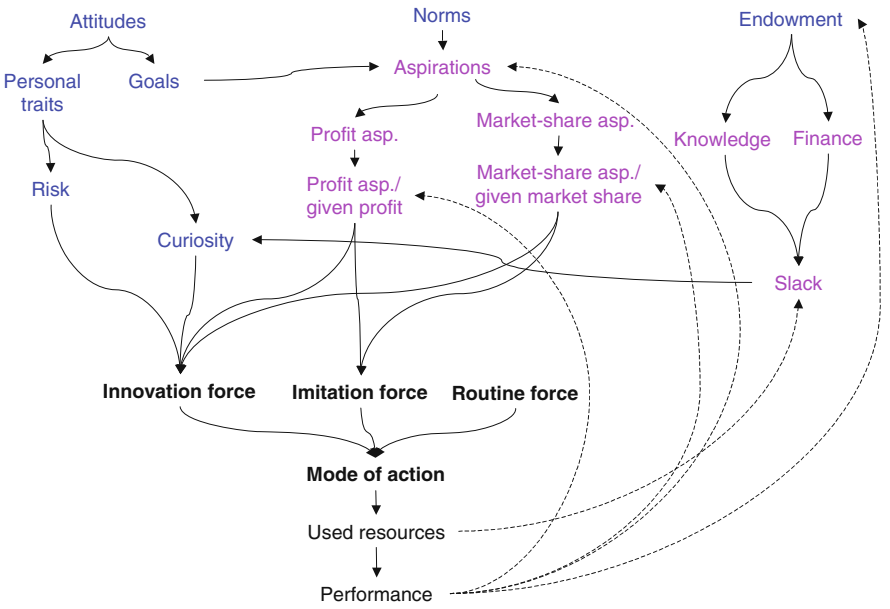


Fig. 4.4 Causal chain diagram for selecting a mode of action in the simulation model

Finally, we set the *preservation* (or routine) *force*

$$F_0 = 1 \tag{4.9}$$

as a reference value.[25] Figure 4.4 depicts the triggering conditions for the different modes of action and the feedback from the economic performance on these conditions as it is implemented in the simulation model.

A firm agent who has selected the action mode of innovation still has two options (see above): He can try to develop an innovation on his own (individual innovation)

[25] Setting the preservation force to a constant is no restriction of generality, since the absolute values of the forces F_i don't matter; it is only the ratio between them which determines the action mode. There are three special or exceptional cases in which the selection mechanism mentioned above is not applied (or even not applicable):

(a) A new firm (start-up), which at the moment of its entry has zero turnover and zero profit, is assigned the action mode of innovation and in the case of cooperation the reason is assumed to be curiosity.

(b) When a new firm has just finished the development of its first innovation, it is assigned the routine action mode for a short period (which is determined by the mean development duration of an innovation).

(c) A firm with negative or zero profit which doesn't belong to case (b) is assigned the action mode of innovation and – if it is willing to cooperate – the cooperation reason (see the next section for a formal definition of this notion) is assumed to be profit dissatisfaction.

or he can seek for cooperation in order to enable or facilitate the development of an innovation.

4.3.3 Cooperative Innovations and Their Network Implications

Two types of cooperations and, correspondingly, network relations are distinguished in the model; actually, they are two subsequent phases of a cooperation's lifespan:

(i) An *innovation development cooperation* which aims at the development of a marketable innovative product. This process involves an exchange of (sharable) specific knowledge (see below) between the member firms of the cooperation. If the development of such an innovation is successfully finished, the innovation development cooperation is transformed into a sales cooperation (see below).
(ii) A *sales cooperation* for the innovative product or service means that the cooperation partners put the product on the market together and share the production cost and the returns of the sale of this product. The production of the innovative product is stopped and the sales cooperation is dissolved when the production cost is not covered by the sales returns.[26]

Given the importance of innovative cooperation for RIS, it is necessary to specify the micro-economic conditions for this type of network activity (cf. Sect. 4.2.2.2) in the given modeling context. The *triggering conditions* for the innovation development cooperation are derived from the behavioral synthesis (cf. Sect. 4.2.2.1) and from empirical observations (cf. Section 4.3.4). According to this behavioral framework it is a combination of personal attitudes, subjective norms, and endowment conditions which influences the orientation of an agent's willingness to innovate toward a cooperative mode for this innovation. First, the three forces shaping the innovation drive (cf. above formula 4.4, 4.5, and 4.6) are basic for the agent's willingness to cooperate in terms of innovation.[27] These forces are related to the actual market performance of the agent. Second, different behavioral types of agents have a different propensity to cooperate. This takes into consideration that not only the actual market position is an important component for explaining cooperative innovations but also some 'deeper' attitudes stemming from different communication

[26] Generally it holds: If for a firm agent – be it an innovator or an imitator – the production cost is not covered by the sales returns for two subsequent time steps and this deficit is increasing from the first of these time steps to the second, the firm agent stops the production of this product. The corresponding demand for the product is then transferred to other firm agents, which produce the same product; if there are no such agents, the product disappears from the market and the corresponding demand is substituted by conventional products of the same branch.

[27] Contrary to the innovation force, this innovation drive is not discounted by any kind of aversion against risk or uncertainty; furthermore, the calculation of expected costs does not play an essential role here.

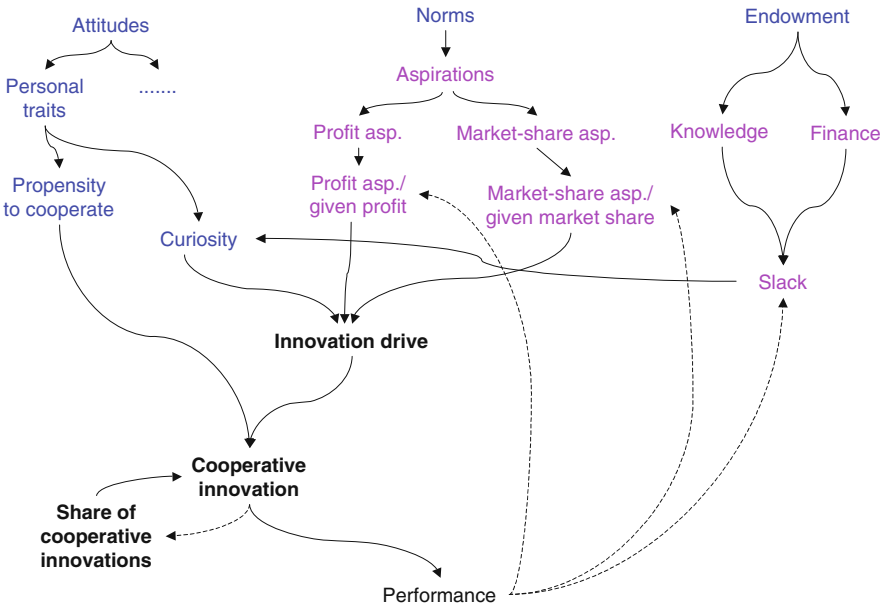


Fig. 4.5 Triggering conditions for cooperative innovation (without subsidies)

styles in different innovation milieus. Third, at least in a regional context, the frequency of cooperative innovations can be observed. Hence, they can help to reduce the uncertainty associated with this type of innovative activity in that they facilitate the search for partners and demonstrate the possibilities to overcome opportunism. Fourth, subsidies coming from political institutions and attributed to cooperating firms act as an exogenous incentive to form a cooperation. Figure 4.5 shows how this triggering of cooperative innovation is embedded in the behavioral framework (solid lines) and how the outcome of cooperative innovation in turn influences these triggering conditions (dotted lines).

Denoting the propensity to cooperate by χ, the share of cooperative innovations N_c (related to the total number of innovations N) by N_c/N, and the amount of subsidies for cooperation by sc, the cooperation force of an agent is given by

$$cp(t) = (1 - ifb - iff - ifs)\chi + ifb\frac{N_c(t)}{N(t)} + iff\sum_{i=0}^{2} f_i(t) + ifs\ sc \tag{4.10}$$

ifb, iff, and *ifs* being parameterized weights for the different triggering forces. Actually, a cooperation is pursued by an agent if this cooperation force is larger than a threshold (*ct*):

$$ct < cp(t) \tag{4.11}$$

Given the willingness of an agent to switch to the cooperative innovation mode, the matching procedure (*matching conditions*) has to be specified. *Matching condition 1* requires that the partners for a cooperation should either belong to the same branch or that they are related to each other as suppliers or customers.[28] *Matching condition 2* picks up the distinction between different kinds of knowledge the agent is equipped with (cf. Sect. 4.2.2.2):

- The first kind of knowledge is general knowledge. This is meant to be the necessary requirement for doing a business in any branch of activity and is essential for the communication between cooperation partners. This kind of knowledge is composed of generally accessible (public) explicit knowledge and basic parts of implicit knowledge.
- The second kind of knowledge is 'specific' knowledge. This kind of knowledge consists of (private) explicit knowledge which is not generally accessible and of those parts of implicit knowledge which are not basic.

It is assumed here for the sake of simplicity that only the explicit parts of the specific knowledge are sharable between the partners of a cooperation; implicit knowledge is considered as not transferable.[29] Hence, the whole (declarative) knowledge base[30] of a firm agent is composed of three elements: common knowledge that is eo ipso not subject to transfer, transferable specific (explicit) knowledge, and non-transferable specific (implicit) knowledge. This knowledge base is mapped in the model as vector components corresponding to knowledge domains in each of which the agent may or may not have knowledge.[31] Given these specifications the *matching condition 2* requires a minimum amount of common general knowledge components.[32] *Matching condition 3* relates the cooperation reason, which is defined as the dominant component in the innovation force of the agent, with the availability of certain knowledge components on the part of the cooperation partner(s):

- The cooperation reason should be the same for all partners.
- If the profit aspiration drive is the dominant innovation force, a sufficient amount of common transferable specific knowledge is required.

---

[28] At the beginning, each firm is assigned randomly a set of supplier firms, being fixed for the whole simulation.

[29] Contrary to that it is sometimes assumed that implicit knowledge can be transferred at a certain cost (cf. e.g. Nonaka et al., 2000; Cowan, 2004). This is only reasonable if there are different degrees for the specificity of knowledge (cf. Sect. 4.2.2.2).

[30] The procedural knowledge (e.g. to know what to do in a given mode of action) is assumed to be the same for all agents.

[31] This means that there is no cardinal measure for the knowledge in a certain domain; each component of the agent's knowledge vector can take only one of the two values 'available' (1) and 'not available' (0).

[32] The distinction between 'general' and 'specific' knowledge is related to the availability of knowledge; the distinction between 'common' and 'individual' refers to the distinction between knowledge actually shared with other agents and knowledge being unique for one agent.

- If market share aspiration drive or curiosity is dominating the innovation drive, it is necessary for a successful matching of partners that each of the two agents has a sufficient amount of transferable specific knowledge which is complementary for the other.

If these matching conditions are fulfilled and there is someone to cooperate with, this causes a transaction cost (cf. Sect. 4.2.2.2). In specifying Eq. 4.1 it is assumed here that the transaction cost of a cooperative innovation depends linearly on trust, i.e., it is composed of a fixed element (*tcc*) which is the same for every transaction and an element which varies according to the level of trust one agent has in the other agent to cooperate with.[33] The maximum amount for this component of the transaction cost (no trust at all) is generally given by the parameter *tct.* This maximum is the more reduced, the more trust (tr) has been built up.[34] Formally this is depicted as

$$\mathrm{TC(t)} = tcc + tct\ (1 - \mathrm{tr(t)}) \tag{4.12}$$

If there is more than one candidate for cooperation, the one for whom the trust is highest and the corresponding transaction cost is lowest is selected (cf. Sect. 4.2.2.2). This selection is repeated (and the number of partners augmented) until the given budget earmarked for this purpose is exhausted.

Once the initiator of a cooperation has selected his partner(s), a process of knowledge transfer is started (*exchange condition*). This process lasts as long as the development duration of the innovation project (cf. above). In each time step, knowledge in one domain is transferred from the agent initiating the cooperation to his partner(s) and vice versa with a certain probability[35]. This probability (kp) depends on one hand positively on the trust (tr) of the knowledge-giving agent in the knowledge-receiving agent and on the other hand, positively on the absorptive capacity (*ac*) of the receiving agent[36]. Formally, the probability for knowledge transfer is given by[37]

$$\mathrm{kp(t)} = se\,(\mathrm{tr(t)} - 1) + ac \tag{4.13}$$

where *se* denotes the sensitivity of the transferring probability with respect to trust, which is another model parameter.[38]

---

[33] Hence this variable element is not necessarily symmetric.

[34] It holds: $0 \leq \mathrm{tr} \leq 1$.

[35] This stochastic modeling reflects the conjectural nature of this process due to the bounded rationality of the agents. The corresponding random number is drawn for each partner and for each transfer direction, rendering these knowledge transfers stochastically independent.

[36] The absorptive capacity is conceptualized here simply as a given probability weight for the happening of the knowledge transfer. Sometimes the influence of absorptive capacity is modeled the other way round as a constraint for a perfect adoption of knowledge (cf. Cowan, 2004).

[37] If the right hand side of the formula is negative, the probability is set to 0. We assume $0 \leq \mathrm{tr} \leq 1$ and $0 \leq ac \leq 1$.

[38] These endeavors to exchange knowledge within the cooperation cause an additional cost in every time step the cooperative innovation is given (cf. Eq. 4.1).

Generally the *replication* or *continuation condition* for the cooperation is the expectation of the agents that in the case of a successful knowledge transfer the quality of the product can be improved and the demand for it will increase. The state variable depicting this feature of the cooperative innovation is trust. Each time this knowledge transfer doesn't happen, the trust of the 'receiving' agent in the 'giving' one is diminished by a certain decrement; on the contrary, trust is raised by a certain increment each time the transfer actually happens. If the trust of a member of the cooperation falls beyond a certain threshold, this agent is leaving the cooperation.[39] The decrement and increment of trust as well as the threshold are model parameters.[40]

Figure 4.6 shows the different micro-economic conditions for the cooperative innovation and the selection of cooperation partners as a flow diagram. Figure 4.7 depicts the dynamics of trust and knowledge transfer within the cooperation process as well as the breaking-off condition for cooperation.

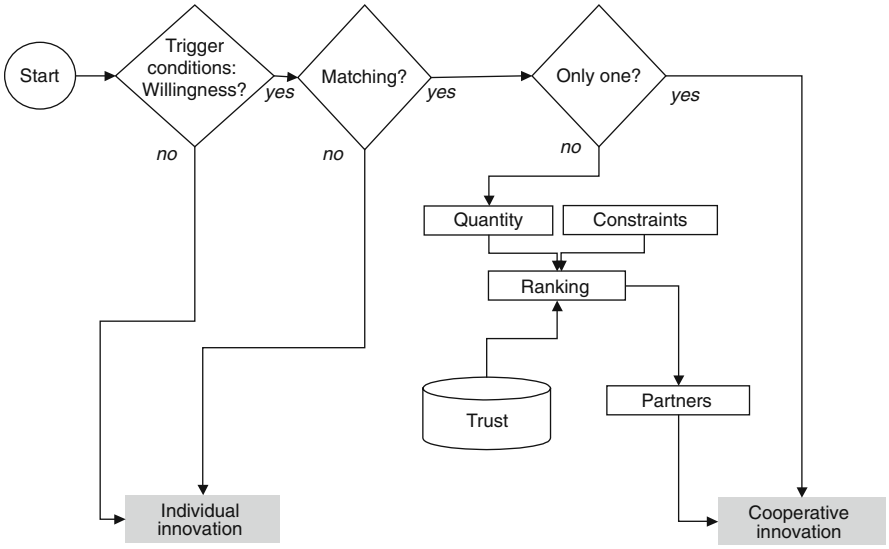


Fig. 4.6 Flow chart for selection of cooperation mode and partners for cooperative innovation

[39] This can have different consequences for the cooperation depending on the size of the cooperation and the role the leaving agent plays in it. If the leaving agent is *not* the initiator of the cooperation (i.e. he has been selected as a cooperation partner by the initiator) *and* the number of remaining members is greater than one, there is no further consequence, i.e. the remaining members continue the cooperation process. If, on the contrary, the leaving agent *is* the initiator of the cooperation *or* if there is only one member left, the cooperation is completely broken off and the joint innovation development project is cancelled.

[40] Analogously to this stochastic process for maintaining or leaving a cooperative innovation, there is a lottery in the course of the process of individual innovation: during the development of an individual innovation, in each time step, the innovation project is abandoned with a certain probability, which is a model parameter.

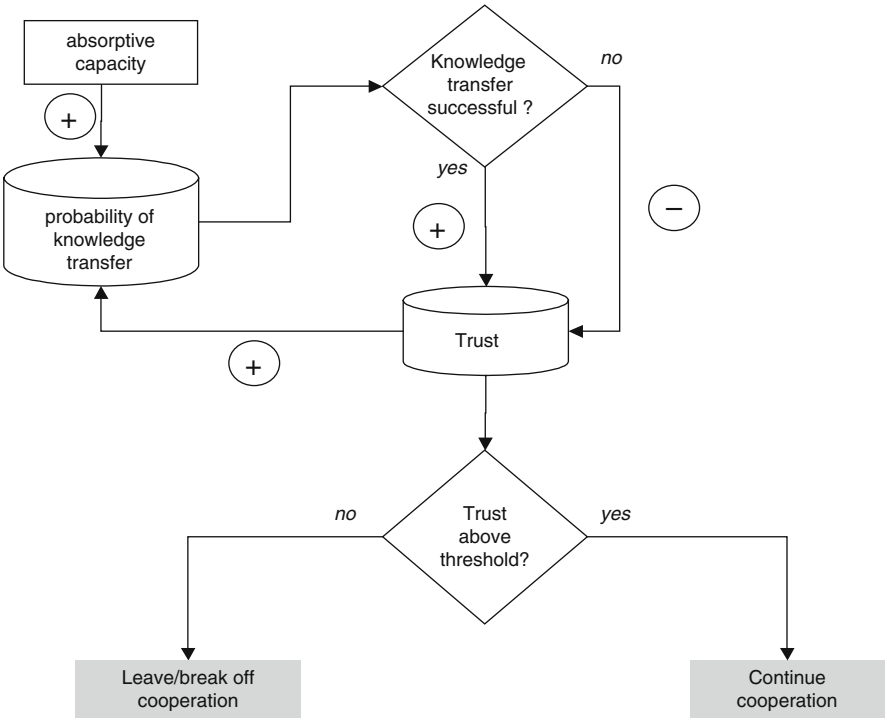


Fig. 4.7 Dynamics of trust during the innovation cooperation process

4.3.4 Empirical Calibration of Behavioral Parameters

Calibrating multi-agent systems (MAS) by using empirical data generally is a neglected topic. One reason for this might be the difficulty to gain agent-related data – especially in the economic realm where time is a scarce resource for both, the researcher and the researched. Another reason might be that there is simply no appropriate standard procedure for gathering these kinds of data. The data used for calibrating parts of the model sketched in the previous section were gained during the last year in Northern Hesse, Germany. The first data set (D1) is based on a written questionnaire which was sent out by the authors of this article to 1783 firms. This sample consisted of the whole population of firms in Northern Hesse with more than three employees. These firms all belong to the manufacturing sector and related service sectors. Altogether, 527 firms responded to the survey (response rate of 29.6%). To gain the second data set (D2), a random sample of 400 firms was drawn from the 527 firms which form the data set D1. Two hundred and seven firms responded to the second survey (51.6%). Whereas the first questionnaire dealt with innovation, cooperation, and networking more generally, the second one was especially designed to capture behavioral variables more thoroughly (cf. Daskalakis and Krömker, 2007).

The design of the survey questions is mostly based on five-point (D1) and six-point (D2) Likert scales. For calibrating the model, the means of the relevant parameters were first calculated and second linearly transformed in order to fit the model scales (cf. Table 4.1). Correlation analyses (Kendall tau-b) as well as non-parametric tests (Mann-Whitney U-test) were applied in order to get statistical proves for the calibration (see Table 4.2).

Table 4.1 Statistical calibration of behavioral model parameters

Variable	Parameter	Data set	Empirically derived means			Scale	Transformed means		
			F_IR	**F_IIM**	**F_ROUT**		**F_IR**	**F_IIM**	**F_ROUT**
Risk acceptance	α	D1	2.3	2.0	–	1–5	3	2.5	3
		D2	4.1	4.2	3.5	1–6			
Exploration drive	w_0	D1[41]	3.3	2.6	–	1–5	0.25	0.2	0.15
		D2[42]	4.6	4.3	3.7	1–6			
Profit aspiration	w_1	D2	5.1	4.7	4.7	1–6	0.04	0.04	0.04
Market share aspiration	w_2	D2	4.9	3.9	4.3	1–6	0.06	0.04	0.05
Cooperation propensity	χ	D1	(70%)	(34.5%)	–	–	1.3	1	0.7
Regional trust	tr_0[43]	D1	4.0	3.9	–	1–5	0.75	0.75	0.75

Table 4.2 Statistical tests of behavioral model parameters

Variable	Parameter	Data set	Mann-Whitney test significances (two-tailed)			Kendall's tau-b correlation coefficient		
			F_IIM and F_IR	F_ROUT and F_IIM	F_ROUT and F_IR	F_IR	F_IIM	F_ROUT
Risk acceptance	α	D1	0.051	–	–	−0.54	−0.091	–
		D2	0.724	0.075	0.092	0.082	−0.057	−0.098
Exploration drive	w_0	D1[44]	0.000	–	–	0.236**	−0.205**	–
		D2[45]	0.231	0.143	0.001	0.199**	−0.089	−0.139*
Profit aspiration	w_1	D2	0.483	0.686	0.128	0.019	0.103	−0.125
Market share aspiration	w_2	D2	0.021	0.413	0.030	0.168*	−0.025	−0.067
Cooperation propensity	X	D1	0.000	–	–	0.276**	−0.120*	–
Regional trust	tr_0	D1	0.788	–	–	0.082	0.004	–

*Significance at 5% level (two-tailed); **Significance at 1% level (two-tailed)

[41] This variable comprises the items 'developing new markets' and 'creating new needs'.
[42] This variable comprises the items 'experimental drive' and 'creativity'.
[43] This empirical value is only used to calibrate the initial value of the parameter.
[44] This variable comprises the items 'developing new markets' 'creating new need'.
[45] This variable comprises the items 'experimental drive' and 'creativity'.

To derive the relevant means it is necessary to distinguish different (behavioral) types of firms. With regard to the model, three types of firms are classified: radical innovators, imitators, and routinizing firms (cf. Table 4.3). The distribution of the type of firms (D1) is as follows: The share of Type F_IR in the product innovators was 30% and the share of the firms type F_IIM was 15%. About 20% of the firms conducted no innovation at all (firm type F_ROUT).[46] The share of the respective population of firms within the model simulation is set up accordingly.

Afterwards the relevant parameters have been identified. As regards the model parameters presented in the previous sections, two types have been distinguished:

(i) Behavioral parameters which influence the choice of the modes of action. Here the designs of the questionnaires allow for investigating the following parameters: 'risk acceptance,' 'exploration drive,' 'market share aspiration,' and 'profit aspiration.'
(ii) Behavioral parameters which affect the choice between individual and cooperative innovation as well as the course of the cooperation. Those are 'propensity to cooperate' and 'trust toward the (regional) cooperation partner.'

These variables are evaluated according to the three different types of firms. Thus, we derive the parameter means for the respective firm types. The data being used are from both data sets. D2 has the advantage that the parameters matching (i) and (ii) are available for all three types of firms whereas in D1 those questions were only answered by the innovating firms. Hence, D2 provides a better foundation for the calibration. Unfortunately, the absolute number of firms of type F_IIM in D2 is low. This might be the reason why differences between F_IR and F_IIM

Table 4.3 Mapping from empirical firm classes to agent types in the model

Firm class	Agent type
Radical innovators (F_IR)[47]	Experimental
Imitators (F_IIM)[48]	Cautious
Routinizing firms (F_ROUT)	Conservative

[46] Altogether, about 80% of the 527 firms were innovating, most of them conducted product (and service) innovations (88%), the shares for process innovations and organizational innovations were 63% and 46%, respectively. Product innovators who conducted incremental innovations and are not part of F_IR (55% of the product innovators) are excluded from our analysis. We thus analyze a very specific – yet quite relevant – section of possible innovation activities. This allows for implementing the empirical findings in the simulation model.

[47] Those firms might also have realized incremental innovation as well as imitation.

[48] Those firms solely accomplished imitation.

are quite clear cut when analyzing the respective variables of D1 but not visible when using D2.[49]

Ad (i): The means of the behavior parameters are illustrated in Table 4.1. They range from 2.3 to 4 within D1 and from 3.5 to 5.1 within D2.[50] Table 4.2 gives the evidence that the correlations between the selected items and the types of firms are significant only in the case of 'exploration drive' and 'market share aspiration.' The differences between the types of firms prove to be significant with regard to the following: in D1 imitators and radical innovators have different levels of 'risk acceptance' and 'exploration drive,' the latter accepting more risk and having more 'exploration drive.' They do not differ with regard to 'profit aspiration.' Hence the transformed means for the latter are set equally for all three types of firms. In D2 significant differences can be found, concerning the 'exploration drive' (F_ROUT vs. F_IR) as well as the 'market share aspiration' (F_IIM vs. F_IR; F_ROUT vs. F_IR).

Ad (ii): To investigate the propensity to cooperate, we only refer to data set D1. As shown above, the share of innovating firms was about 88%. Nearly 45% of these firms stated that they were involved in cooperative innovations, and 37% of the cooperative innovations were conducted with regional partners. With a cooperation rate of about 70%, the firms of type F_IR were the most cooperative firm type (cooperation rate type F_IIM: 34.5%). In more detail: about 40% of the firms of type F_IR had accomplished a cooperation during the investigated time span; about 47% still had been involved in innovation cooperation (type F_IIM: 9.1% and 29.1%).[51] Accordingly, the correlation between the items capturing the cooperation activities and both types of firms has a different level of significance and the statistics show a remarkable difference between both types of firms.

With regard to trust, our findings are ambivalent. Trust seems to be a relevant pre-condition for cooperative innovations as the means are relatively high. Yet, there are no significant correlations between the type of firms and the level of trust to be found. There are also no differences in the level of trust between the firm types. Therefore, the corresponding transformed means are set equal for all three types of firms.[52] However, the statistics show that within cooperative innovations trust and knowledge exchange are interdependent. This is approved in the data set D2 by a high correlation between the item 'we have given our cooperation partners a lot of assistance' and the cooperative firms.[53]

---

[49] The number of firm type F_IIM in D2 is 20. The respective answers of the firms concerning the behavioral foundation are thus only used to calibrate the model, if either the statistical evidences concerning those variables are sufficient and/or if statistical evidences from data set D1 support the findings.

[50] Note the different Likert scales!

[51] Most of the cooperations (respective over 90%) were aligned to product innovations.

[52] There can be significant differences in the levels of trust toward regional and national partners: the regional cooperation partners were trusted more profoundly; level of significance (Mann-Whitney test) for the two highest values: 0.001 (two-tailed).

[53] Kendall's tau: 0.202; $p < 0.01$. This is in accordance with the empirical findings in Daskalakis and Kauffeld (2006).

4.4 Simulation Results

4.4.1 Change of RIS-Layers Over Time

Analyzing the dynamics of the RIS as a whole necessitates to look at the action modes the agents pursue in the course of time and to decipher the network relations (according to the definition in Sect. 4.1) resulting therefrom. As regards the frequency of the modes of action (cf. Fig. 4.8 (above)) an inverse relationship between routine and imitation on one side and between imitation and innovation on the other side can be observed. This corresponds to the competition of behavioral forces expressed in Eqs. (4.7, 4.8 and 4.9). Contrary to that the frequency of individual and cooperative innovation is moving in about the same direction. Because the individual innovation creates the knowledge resources being exchanged in the cooperative innovation, the latter follows the former with a time lag. The only mode of action increasing over the whole time span under investigation is the sales cooperation. This is due to the difference in the entry and exit dynamics for this special mode of cooperation: the entry is determined by successful former cooperative innovations whereas the exit depends on a loose cost condition (cf. Sect. 4.3.3). Generally it can be observed that there are two phases in the development of RIS: in the first phase up to about $t = 50$ there is an increase in the modes of action related to creating novelties and a corresponding decrease of the frequency of agents being in a routine mode. In the second phase of RIS (for $t > 50$) all modes of action (except the sales cooperation) fluctuate around a rather stable level.

In terms of the multi-layer concept (onion-concept) of RIS this means that although the agents individually are continuously switching between the different layers there are patterns on the meso-level in terms of the size of subpopulations on the different layers. In Fig. 4.8 (middle) two snapshots for the RIS-onion composed of (numbered) agents on the different layers are depicted ($t = 60$: RIS (60); $t = 120$: RIS (120)). In such a graph the network relations either in an ongoing cooperative innovation or in a sales cooperation stemming from a former innovative collaboration are visible. What can be observed here is that agents with higher degrees (cf. Sect. 4.2.2.1) are more dissipated among all layers in RIS (120) compared to RIS (60) due to the persisting of sales cooperation. The possibility for an agent to earn profits or to maintain a market share by such sales cooperation might hinder the novelty creation. Figure 4.8 (below) finally shows the network of trust ties for the corresponding time steps (RIS (60) and RIS (120)). These ties indicate all experiences the agents made with each other in terms of knowledge exchange. Contrary to the relations in the actual innovation network and sales network, the trust relations are not necessarily symmetric between agents, and they differ in value.[54]

[54] In terms of SNA, this network is represented by a directed asymmetric graph with ties of different values.

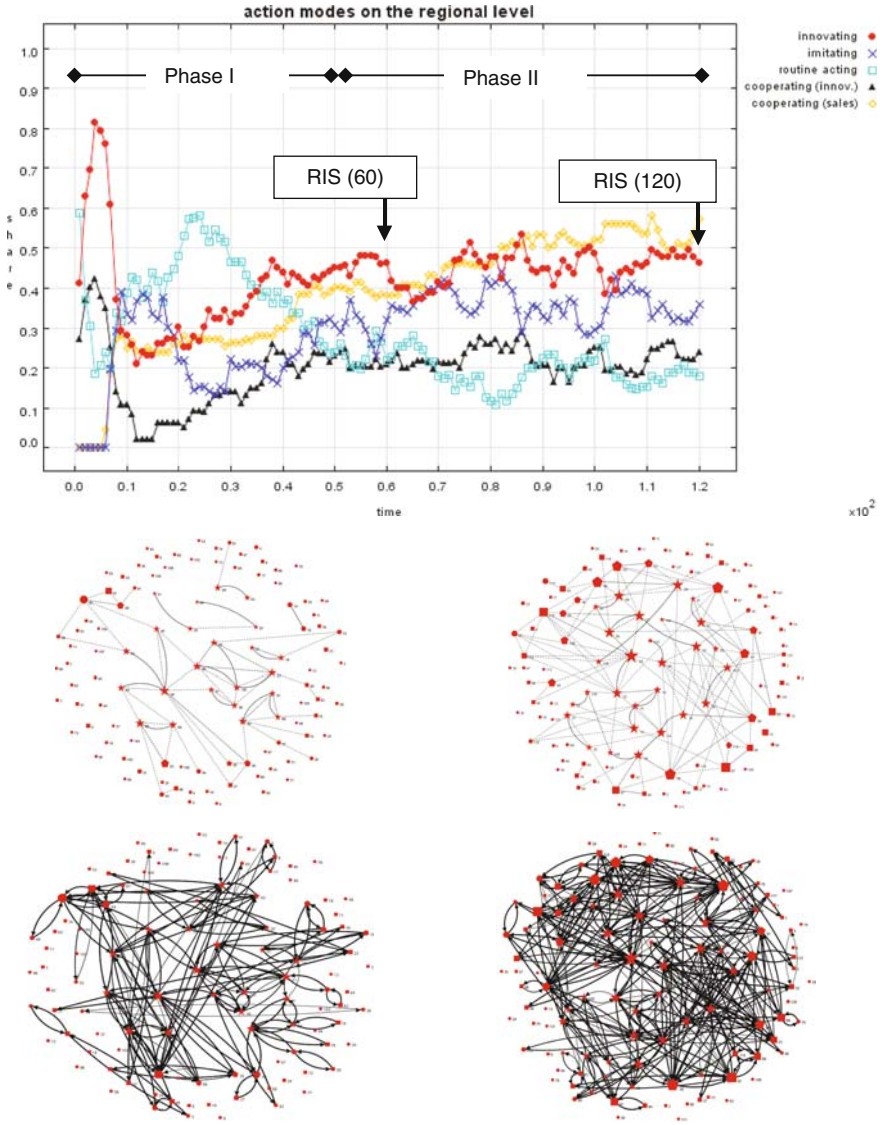


Fig. 4.8 (Above) Modes of action over time; (middle) RIS (60) (stars, pentagons, squares, and circles symbolizing agents in cooperative innovation, individual innovation, imitation, and routine respectively; solid lines and dotted lines symbolizing network relation in cooperative innovation and sales cooperation, respectively) and RIS (120); (below) Trust network in RIS (60) and RIS (120)

Due to the difference in persistence three different network components can be associated with a RIS indicating the multiplexity property (cf. Sect. 4.2.2.1): (i) a brittle and temporary innovation network, (ii) a more stable sales network on the medium range, and (iii) a long-term trust network. Whereas the sales network is a simple outcome of the innovation network, the trust network is a cause and effect for the innovation network at the same time because it is determining the replication of a given innovation network as well as the matching condition for a new innovation network.

4.4.2 Dynamic Features of Networks

According to the conceptualization of RIS given in Sect. 4.3.1, the cooperative innovations are the most important driving forces for the dynamics of RIS. In the context of the model framework proposed here, they are the only source for the network relations between the agents belonging to RIS. These network relations are either related to knowledge and trust (innovation cooperation) or based on common market operations (sales cooperation).

The implementation of a cooperative innovation requires at least the fulfilling of the triggering as well as the matching conditions specified in Sect. 4.3.3. As a first step for analyzing the background of the cooperative innovation activity in the RIS in Fig. 4.9, the triggering conditions in terms of behavioral forces relevant for this type of activity are depicted over time. What is obvious here is that the willingness to cooperate is much more spread over the new innovative agents than the actual ability to implement such a cooperation (cf. Fig. 4.8 (above)). This indicates that the knowledge- and finance-related requirements for these cooperations play an important role.[55] Furthermore, Fig. 4.9 reveals that the importance of curiosity for cooperative innovations is reduced in the second phase where the resources in terms of knowledge and finance are more efficiently used. Market share frustration and profit frustration (resulting from missing the corresponding aspiration levels) are getting more important in the second phase of RIS because competition is increasing. Nevertheless, the peaks of the fluctuating frequency of realized cooperative innovations in the second phase of RIS (e.g., between $t = 75$ and $t = 88$ or between $t = 110$ and $t = 115$; cf. Fig. 4.8(above)) seem to correlate with high levels of curiosity.

Focusing in a second step the number of realized cooperative innovations, it can be observed that the constraints implied in the exchange as well as the replication conditions are preventing an uneven participation of agents in this type of activity. As regards the frequency of being in the core of RIS, about 50% of the agents are there only up to three times; the other 50% are distributed in the range from being four times up to 16 times in the core of RIS (cf. Fig. 4.10). Hence, there

[55] Cf. the sensitivity analysis as regards to the role of transaction costs parameters (Fig. 4.15) for a specification of this argument.

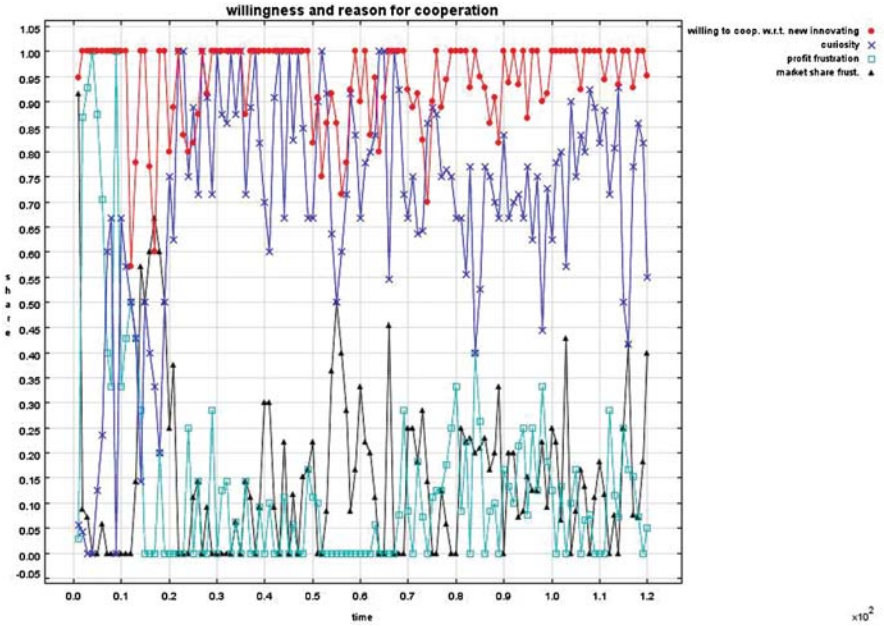


Fig. 4.9 Behavioral forces for cooperative innovation over time

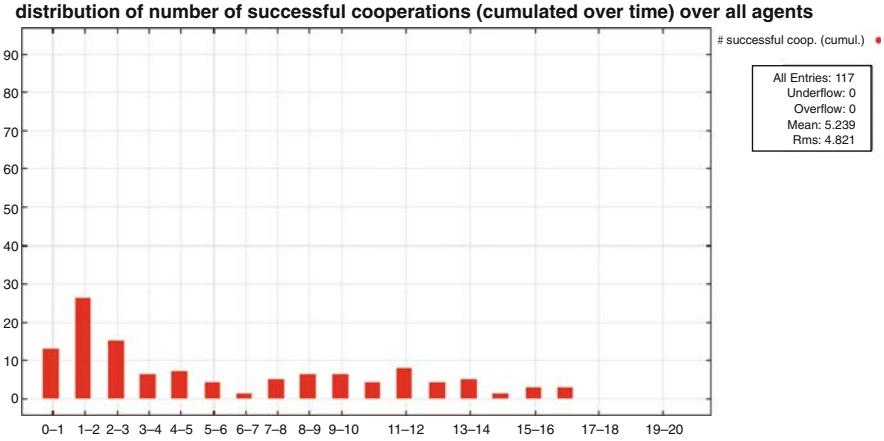


Fig. 4.10 Cumulated frequency of agents for being in the core of RIS

seems to be some 'big players' in this RIS, and also a large number of agents with moderate or small engagement in innovation cooperations. Obviously, there is no regular relationship between the number of agents and the frequency of finishing a cooperative innovation.

This assessment is confirmed if the mean degree (cf. the definition in Sect. 4.2.2.1) for the innovation network relations is plotted over time (cf. Fig. 4.11). For the first

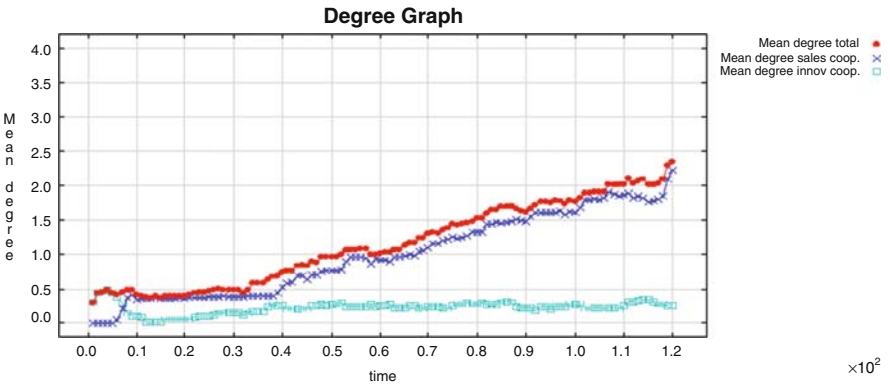


Fig. 4.11 Mean degrees for cooperation over time

phase of RIS this degree is slightly increasing. In the second phase of RIS, it is almost stable with regards to the number of cooperative innovations (cf. Fig. 4.8 (above)). A major cause for this result is certainly the limitation that a firm agent can only engage in one (cooperative or individual) innovation project at a time in our model. Hence, there is not much room for 'hubs' being established in such a network.

The picture changes if the sales cooperations (resulting from successful cooperative innovations) are included (cf. Fig. 4.11). Due to the much softer constraint for this type of cooperation, they can persist temporarily leading to an increasing number of network relations on the level of agents. This is indicated by the continuous increase in the mean degree for the network relations resulting from this type of cooperation in the second phase of RIS (after enough cooperative innovations have been finished). Against this backdrop temporary network-hubs are possible (mainly consisting of sales relations) as can be seen from Fig. 4.8 (middle, left), e.g., for agents no. 55, 99 and in Fig. 4.8 (middle, right), e.g., for agents no. 38, 55, and 110. In Fig. 4.12 the distribution of the degree for all network relations for all agents

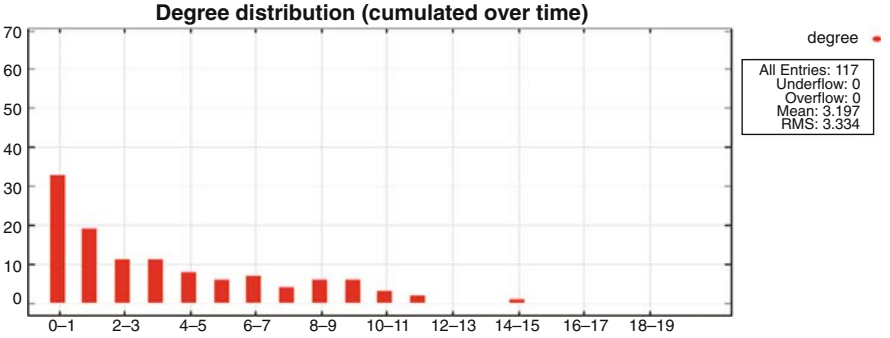


Fig. 4.12 Cumulated frequency distribution for cooperation degree

cumulated over the whole simulation span is shown. Obviously, the corresponding frequency distribution is left skewed, i.e., the frequency decreases with the level of the degree. Hence, high-level degrees do not occur very often or do not persist over a longer time period.

4.4.3 Spread of Knowledge

Individual and cooperative innovation are the two sources of new private (but sharable) knowledge in the given simulation model. For the sake of simplicity, it is assumed here that in each time step for both types of innovation, there is a chance to acquire only one additional element of sharable knowledge. This element of knowledge is related to a domain (e.g., knowledge about raw material of a certain type) and is coded in a binary manner ('1' means there is knowledge in a given domain; '0' means there is no knowledge in this domain). Hence, in every time step the number of agents getting new knowledge (switching from '0' to '1' in a given domain) can be counted for both types of innovation activities (Fig. 4.13).

For both types of innovation there is a tendency to deliver additional knowledge for an increased number of agents. In the first phase of RIS the number of agents involved in the spreading of knowledge in the case of cooperative innovation is growing faster than the corresponding number in the case of individual innovation: the former starts from a lower level and depends on the preceding of the individual innovation as a primary source. Apart from these tendencies the fluctuations of both sources for knowledge acquisition correspond to the time-dependent frequency of the respective modes of action (cf. Fig. 4.8 (above)).

It can be summarized that – except in the initial phase – the cooperative innovation plays an important role for the spreading of knowledge on the regional level. Due to the assumed possibility to transform this additional knowledge in quality improvements of commodities, the number of agents being able to increase

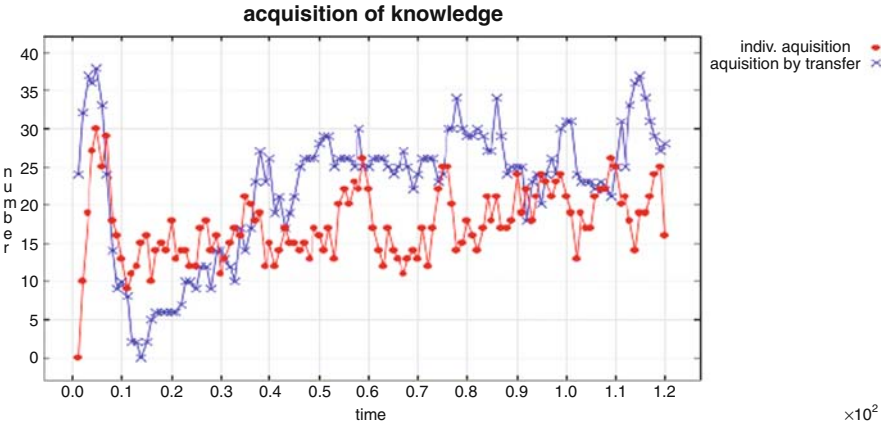


Fig. 4.13 Number of agents getting new knowledge over time

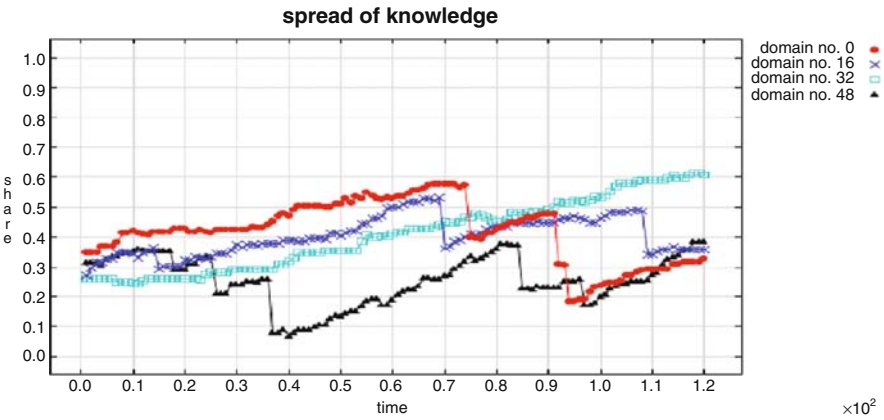


Fig. 4.14 Domain-specific spread of knowledge

their market performance increases. This additional innovation potential triggers a growing value added and employment in the region as a whole.[56]

From the perspective of agents the acquisition of knowledge is only the first step of knowledge processing. The storage and the use of knowledge are other steps in this knowledge processing of agents, which may imply that parts of the knowledge disappear due to memory constraints and to devaluation effects for unused or useless knowledge. Hence, the knowledge string of an agent varies in the course of time depending on the acquisition on one side, and forgetting and devaluation of knowledge components on the other side. Therefore, the behavioral differences of the agents do not only endow them with heterogeneous knowledge strings; furthermore, the knowledge for the region as a whole (measured by the frequency for the different knowledge domains of being a part of the agent's knowledge string) is different in every time step. Figure 4.14 shows for selected knowledge domains how their distribution over the knowledge strings of agents evolves in the course of time. This indicates that different knowledge domains are of different importance for the development of a region (for example domain no. 32 is compared with domain no. 0 in Fig. 4.14).

4.4.4 Sensitivity Analysis

According to the conceptualization of RIS in Sect. 4.2.1 and according to the findings of the simulation model presented in Sect. 4.4.3, the number of cooperative innovations and the related spread of knowledge between agents are the most important performance indicators of RIS. In the specification of the explanantia for this performance output (explanandum) of the simulation model (cf. especially

[56] Cf. Beckenbach et al. 2007 for a specification of these effects of cooperative innovation on the region as a whole.

Sects. 4.3.2 and 4.3.3) cognitive as well as economic parameters have been used. Hence, it remains to be demonstrated how the above-mentioned indicators for the core performance of RIS are influenced by these parameters.[57] In each of the following sensitivity tests the parameter space around the standard parameter configuration[58] used in Sects. 4.4.1, 4.4.2, and 4.4.3 is investigated by means of Monte Carlo simulations.

In Fig. 4.15 it is shown how the number of cooperative innovations and the average stock of sharable knowledge changes if the average aspiration flexibility (ϕ in Eq. 4.3) and cooperation threshold (*ct* in Eq. 4.11) is varied. This reveals that an increase of the aspiration flexibility leads to an increase in finished cooperative innovations only if *ct* is not too high (Fig. 4.15a). Due to the important role the cooperative innovation plays for the regional spread of knowledge basically the same critical relationship between the aspiration flexibility and the cooperation threshold holds for the average stock of sharable knowledge (although this relationship is not as strong as in the former case because individual innovation is another source for the stock of knowledge) (Fig. 4.15b).

In Fig. 4.15(c,d) both outputs indicating the performance of the RIS-core are related to the parameter *ifb*, which determines the weight of the frequency component of the cooperation force (cf. Eq. 4.10), and to the initial value for the trust level of the agents (tr(0)). The latter is important for the amount of transaction costs of a cooperative innovation (cf. Eq. 4.12) and for the probability of knowledge transfer within such a cooperation (cf. Eq. 4.13). Here it is obvious that for low and moderate values, there is almost no influence of *ifb* on the frequency of successful cooperative innovations; for high values of this parameter, there is even a negative influence on the former (cf. Fig. 4.15c). This can be explained by the constraining effect an increase of *ifb* has on the propensity to cooperate (χ) according to Eq. (4.10). Hence, only if low and medium range values for *ifb* are given an increase in the initial value for trust has a positive influence on the number of successful cooperative innovations (due to a reduction in transaction cost and an increase in the probability of knowledge transfer). Again the same delicate relationship between tr(0) and *ifb* is principally observable as regards the average stock of sharable knowledge (cf. Fig. 4.15d) (though not in the same drastic manner because private innovation is another source for knowledge generation).

Finally, the dynamics of the RIS-core have been checked for varying transaction cost parameter *tcc* (cf. Eq. 4.12) on one side and the amount of subsidies sc (cf. Eq. 4.10) on the other side (cf. Fig. 4.15e, f). In this constellation the increase of the transaction cost parameter has almost no influence on the frequency of cooperative innovations. Hence, to pursue a cooperative innovation seems not to be a purely financial problem. This constellation changes if – due to subsidies – the cooperation force (as expressed in Eq. 4.10) increases and the financial constraint matters:

---

[57] For getting a representation of the spread of knowledge by a real number, the number of domains of sharable knowledge in the knowledge strings of agents is simply summed up and divided by the number of agents (average stock of sharable knowledge).

[58] In Fig. 4.15, this configuration is marked by a black point.

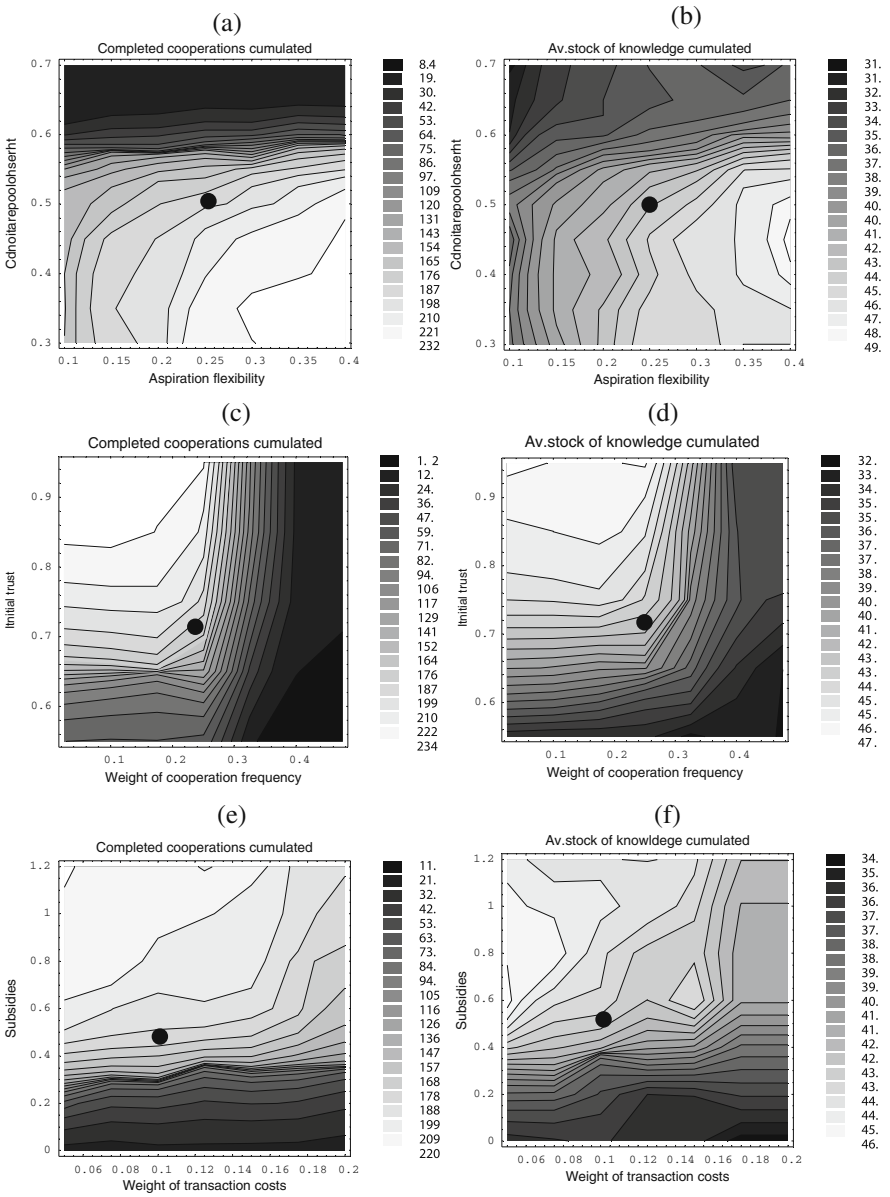


Fig. 4.15 Sensitivity analysis

now an increase in transaction costs reduces the number of cooperative innovations (cf. Fig 4.15e). Contrary to that enhancing the amount of subsidies increases the number of cooperative innovations for any given level of transaction costs. Due to the partial role of cooperative innovations for knowledge generation, these relationships do not hold in the same way for the average stock of sharable knowledge (cf. Fig. 4.15(f)). In this case a low level of transaction costs and a moderately high level of subsidies give the best results.

4.5 Conclusions

In this contribution it is shown how one can proceed if the often neglected internal dynamics of RIS shall be explained. Then the agents, their different modes of action as well as the coordination of their activities by markets, networks, and institutions have to be taken into account. It was proposed here to explain the internal dynamics of RIS by referring to different layers in terms of action modes the agents can pursue. This implies to explain the selection of these different modes of action by considering behavioral as well as situational components. In that context the core of RIS is the mode of cooperative innovation of several agents in a region. This mode of action was identified as a source for multiplex network relations lying between market relations on one side and hierarchy relations on the other side.

The simulations with the agent-based model of RIS show the emergence of patterns as regards the frequency of the different modes of action and hence the dynamics on the different layers of RIS. These patterns on the regional (meso) level 'grow' out of the ongoing behavioral dynamics of the individuals (micro-level) continuously switching between the different modes of action. These patterns on the meso-level comprise different evolution paths for the network components: a brittle development of cooperative innovations on a rather low level, a stable medium range growth of sales cooperation, and a high-level growth for trust relations.

Finally, the sensitivity analyses as regards the parameters for cooperative innovation reveal a 'network landscape' behind the observable dynamics of RIS. Broadening the perspective of a singular path to include such a network landscape sheds light on the conditions for good network performance in terms of parameter constellations. Hence, such simulations can be a starting point for a bottom up improvement of the performance of RIS and the networks included. This is in sharp contrast to the usual top down optimization perspective in network research and the corresponding perspective of a designer or even planner of a RIS.

Acknowledgments We gratefully acknowledge the funding of this research by the VW-Stiftung.

References

Ajzen, I. (1991). The Theory of Planned Behavior. Organizational Behavior and Human Decision Process 50(2): 179–211.

Anderson, J. R. (2000). Cognitive Psychology and its Implications. W.H. Freeman, New York.

Antonelli, C. (2000). Collective knowledge communication and innovation: the evidence of technological districts. Regional Studies, 34(6): 535–547.
Asheim, B. T. and A. Isaksen (2002). Regional Innovations Systems: The Integration of Local `Sticky' and Global `Ubiquitious Knowledge'. The Journal of Technological Transfer 27: 77–88.
Barabasi, A. L. (2003). How Everything is Connected to Everything Else and What it Means. Perseus Publishing. Cambridge.
Beckenbach, F. (2004). Das Handlungskonzept in der evolutorischen Mikroökonomik. Studien zur Evolutorischen Ökonomik Bd. IX. W. Kerber, Duncker and Humblot: 29–68.
Beckenbach, F., R. Briegel, and M. Daskalakis (2007). The Influence of Regional Innovation Systems on Regional Economic Growth: Linking Regional Input-Output Analysis and Agent Based Modeling. Papers on agent-based economics. Kassel.
Braczyk, H. M., P. Cooke, and M. Heidenreich (eds.) (2004). Regional Innovation Systems: The Role of Governance in a Globalized World. Routledge, London.
Brenner, T. (2001). Simulating the Evolution of Localised Industrial Clusters – An Identification of Basic Mechanisms. Journal of Artificial Societies and Social Simulation 4(3).
Breschi, S. and F. Lissoni (2004). Knowledge Networks from Patent Data: Methodological Issues and Research Targets. DRUID summer conference 2004. Helsingør, Dänemark.
Briegel, R. (2006). Nachhaltige Freizeitmobilität? Ein dynamisches akteursorientiertes Simulationsmodell. Metropolis, Marburg.
Camerer, C. F., G. Loewenstein, and D. Prelece (2005). Neuroeconomics: How Neuroscience Can Inform Economics. Journal of Economic Literature XLIII(1): 9–64.
Cantner, U. and H. Graf (2006). The Network of Innovators in Jena: An Application of Social Network Analysis. Research Policy 35(4): 463–480.
Cooke, P. (1992). Regional Innovation Systems: Competitive Regulation in the New. Geoforum, 23(2): 365–382.
Cooke, P. and K. Morgan (1990). Learning Through Networking: Regional Innovation and the Lessons of Baden-Württemburg. Regional Industrial Research.
Cooke, P. and K. Morgan (1991). The Network Paradigm: New Departures in Corporate and Regional Development. Environment and Planning D: Society and Space 11(5): 543–564.
Cowan, R. (2004). Network Models of Innovation and Knowledge Diffusion. Maastricht: MERIT, Maastricht Economic Research Institute on Innovation and Technology Research Memoranda. http://www.merit.unu.edu/publications/rmpdf/2004/rm2004-016.pdf
Cyert, R. M. and J. G. March (1992). A Behavioral Theory of the Firm. Blackwell, Oxford.
Daskalakis, M. and M. Kauffeld-Monz (2006). The Microeconomics of Innovation Networks: On the Dynamics of Knowledge Generation and Trust Building. Papers on agent-based economics. Kassel.
Daskalakis, M. and D. Krömker (2007). Investigating Innovation with Behavioral-Sciences: Empirical Evidences for the Utility of an Agent Based Approach. Mimeo.
Etzkowitz H. and L. Leydesdorff (2000). The Dynamics of Innovation: from National Systems and "Mode 2" to a Triple Helix of University-Industry-Government Relations. Research Policy 29(2): 109–123.
Fagiolo, G. and G. Dosi (2003). Exploitation, Exploration and Innovation in a Model of Endogenous Growth with Locally Interacting Agents. Structural Change and Economic Dynamics 14: 237–73.
Freeman, C. (1987). Technology Policy and Economic Performance. London: Pinter.
Freeman, C. (1995). The 'National System of Innovation' in Historical Perspective. Cambridge Journal of Economics 19: 5–24.
Fritsch, M. and G. Franke (2004). Innovation, regional knowledge spillovers and R&D cooperation. Research Policy, 33(2): 245–255.
Gilbert, N., A. Pyka, and P. Ahrweiler (2001). Innovation Networks – A Simulation Approach. Journal of Artificial Societies and Social Simulation 4(3).
Jager, W. (2000). Modelling Consumer Behaviour. Universal Press, Groningen.

Kauffeld-Monz, M. and M. Fritsch (2001). The Impact of Network Structure on Knowledge Transfer: An Empirical Application of Social Network Analysis in the Context of Regional Networks of Innovation. Mimeo.

Kline, S. J. and N. Rosenberg (1986). An Overview of Innovation. In: R. Landau and N. Rosenberg (eds.). The Positive Sum Strategy. National Academy Press, Washington, 275–305.

Lorenz, E. (1999). Trust, Contract and Economic Cooperation. Cambridge Journal of Economics, 23: 301–315.

Loewenstein, G. (1994). The Psychology of Curiosity: A Review and Reinterpretation. Psychological Bulletin 116(1): 75–98.

Louis, M. R. and R. I. Sutton (1991). Switching Cognitive Gears: From Habits to Active Thinking. Human Relations 44(1): 55–76.

Lundvall, B.-A. (1988). Innovation as an Interactive Prozess: From User-Producer-Interaction to the National System of Innovation. In G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (eds.), Technical Change and Economic Theorie. London: Pinter, 349–369.

Lundvall, B.-Â. (2007). National Innovation Systems – Analytical Concept and Development Tool. Industry and Innovation 14(1): 95–119.

Lundvall, B. -Â., B. Johnson, B. Dalum, and E. S. Andersen (2002). National Systems of Production, Innovation and Competence Building. Research Policy 31(2): 213–231.

March, J. G. (1994). A Primer on Decision Making. The Free Press, New York.

March, J. G. and H. Simon (1993). Organizations. Blackwell, Cambridge, Mass.

McKelvey, M. (1991). How do National Systems of Innovation Differ?: A Critical Analysis of Porter, Freeman, Lundvall and Nelson. In: G. Hodgson and C. Screpanti (eds.), Rethinking Economics: Markets, Technology and Economic Evolution. UK: Edward Elgar, 117–137.

Ménard, C. (2004). The Economics of Hybrid Organizations. Journal of Institutional and Theoretical Economics 160(3): 345–76.

Metcalfe, J. S. (1995). Technology Systems and Technology Politics in an Evolutionary Framework. Cambridge Journal of Economics 19(1): 25–46.

Miller, G. J. (1992). Managerial Dilemmas: The Political Economy of Hierarchy. Cambridge University Press, New York.

Nelson, R. (1993). National Innovation Systems: A Comparative Analysis. Oxford: Oxford University Press.

Nonaka, I., R. Toyama, and A. Nagata. (2000). A Firm as a Knowledge-creating Entity: A New Perspective on the Theory of the Firm. Industrial and Corporate Change 9(1): 1–20.

Nooteboom, B. (2002). Trust: Forms, Foundations, Functions, Failures and Figures, Edward Elgar, Cheltenham.

Nooteboom, B. (2004). Inter-firm Collaboration, Learning and Networks: An Integrated Approach. Routledge, London.

Nooteboom, B. (2006). *Trust and Innovation*. Essay for the Dutch Ministry of Economic Affairs. Background Document for the Innovation Lecture 2006.

Quéré, M. and L. Ravix (2003). The Austrian Theory of Institutions Applied to Science-Industry Relationships: The Relevance of Innovative Institutions. Review of Austrian Economics 16 (2–3).

Sabel, C. F., G. B. Herrigel, R. Deeg, and R. Kazis (1989). Regional Prosperities Compared: Massachusetts and Baden-Württemberg in the 1980s. Economy and Society 18(4): 374–404.

Scitovsky, T. (1976). The Joyless Economy. Oxford University Press, New York.

Scott, J. (2005). Social Network Analysis: A Handbook. 2nd edition. Sage Publication, London.

Simon, H. A. (1997). Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations. The Free Press, New York.

Sternberg, R. (2000). Innovation Networks and Regional Development–Evidence from the European Regional Innovation Survey (ERIS): Theoretical Concepts, Methodological Approach, Empirical Basis and Introduction to the Theme Issue. European Planning Studies, 8(4): 389–407.

Svenson, O. (1990). Some Propositions for the Classification of Decision Situations. In: K. Borcherding, O. I. Larichev, and D. M. Messick (eds.). Contemporary Issues in Decision Making. Amsterdam, North-Holland, 17–32.

Svenson, O. (1996). Decision Making and the Search for Fundamental Psychological Regularities: What can be Learnt from a Process Perspective? Organizational Behavior and Human Decision Processes 65(3): 252–67.

Verplanken, B. and H. Aarts (1999). Habit, Attitude, and Planned Behaviour: Is Habit an Empty Construct or an Interesting Case of Goal-Directed Automaticity? European Review of Social Psychology 10: 102–134.

Wasserman, S. and K. Faust (2007). Social Network Analysis. Methods and Applications. 4th edition. Cambridge University Press, Cambridge.

Watts, D. J. (1999). Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton.

Williamson, O. E. (1985). The Economic Institutions of Capitalism. Firms, Markets, Relational Contracting. London New York: The Free Press/Macmillan.

Zhang, J. (2003). Growing Silicon Valley on a Landscape: An Agent-Based Approach to High-Tech Industrial Clusters. Journal of Evolutionary Economics 13(5): 529–48.

Chapter 5
Agent-Based Modelling of Innovation Networks – The Fairytale of Spillover

Andreas Pyka,* **Nigel Gilbert and Petra Ahrweiler**

5.1 Introduction

Today's knowledge-based economies are more than places where goods and services are bought and sold; they are the sites where complex logistic processes are coordinated, where innovation takes place, where knowledge is generated, communicated, re-combined and exchanged. In such competitive and knowledge-intensive environments characterized by price as well as innovation competition and in which there are quickly changing global technological and economic requirements (Bahlmann, 1990; Hanusch and Pyka, 2007a) and a variety of institutional infrastructures (Amable, 2003; Hanusch and Pyka, 2007b), a firm can improve its performance only by exploiting resources more creatively and intelligently than its competitors (Lam, 2003). To keep a competitive edge, firms have to engage in continuous learning. Organizationally, cooperative research in so-called innovation networks has become a prominent alternative, which allows access to external knowledge sources. These innovation networks were almost neglected in the theory of industrial organization (IO) due to a misinterpretation of technological spillover effects. In innovation networks, which are the outcome of various horizontally and vertically bilateral/multilateral collaborations, the actors of the innovation process share and co-develop new knowledge with other actors (Pyka, 2002).

In economics knowledge transfers between different actors basically were described by technological spillovers. This concept is borrowed from macro-economics, where spillovers are the engine of economic growth due to the positive feedbacks they induce in economic development (e.g. Romer, 1990). On the level of micro- and industrial economics, the concept of technological spillovers was

A. Pyka (✉)
Economics Department, University of Bremen, Hochschulring 4, 28359 Bremen, Germany
e-mail: pyka@uni-bremen.de

*Andreas Pyka is also a Visiting Professor at the Department of Economics of Innovation at the Technical University of Delft, The Netherlands.

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_5,

carried over almost one-to-one by referring to the public good nature of new technological knowledge. Accordingly, in the incentive-based approaches of neoclassical industrial economics, technological spillovers are considered as involuntary knowledge flows, which reduce the incentives to be engaged in costly R&D. Obviously, the negative interpretation of technological spillovers on the industrial level also has had an impact on the assessment of collaborations in industrial R&D.

In the theory of IO in the first place, collaborative R&D was considered to be a temporary phenomenon only (e.g. Harrigan, 1985; Sapienza, 1989). Additionally, on the basis of well-established anti-trust principles, collaborative agreements generally were supposed to be suspicious. For example, in transaction costs theory, innovation networks were taken as a hybrid form of IO, between hierarchy and markets, which finally will disappear. Exemplarily the developments in the biotechnology-based industries were cited, where since the late 1980s well-established pharmaceutical firms were engaged frequently in cooperative research with specialized biotech start-up companies (e.g. Pyka and Saviotti, 2005). These cooperations were supposed to allow access to the upcoming biotechnology paradigm for the established pharmaceutical firms, which were specialized so far in organic chemistry. The cooperations, furthermore, were supposed to disappear soon again: either the large pharmaceutical firms will build up their own competences in biotechnologies (which include the acquisition of biotech start-ups) and therefore, no longer need collaboration partners or the biotech start-ups will eventually displace the large pharmaceutical firms and become the dominant actors in the pharmaceutical industries themselves. In this perspective, a long-lasting room for collaborations is inconceivable, and implicitly knowledge flows between firms (i.e. technological spillovers) are considered to insert negative effects on the level of R&D investment.

Until today, neither of these two extreme alternatives becomes true. In the early years of the 21st century, both populations of actors, the pharmaceutical firms and the small biotech firms, still co-exist and also innovation networks still shape the IO of R&D processes in this industry. Meanwhile, in the literature of organizational learning (OL) the term ‘liability of disconnectedness’ is describing the importance for industrial actors to get access to external knowledge sources via collaboration (Powell et al., 1996).

This lack of interest in R&D collaborations prevailing in the theory of IO until the end of the 1980s can entirely be traced back to the negative view on technological spillovers characteristic for neoclassical industrial economics. As mentioned above, technological spillovers, i.e. the involuntary and uncontrollable knowledge flow due to the supposed public good nature of new technological knowledge, were basically considered to reduce R&D incentives (e.g. Levin and Reiss, 1988) and therefore, as detrimental for economic welfare. Not before the end of the 1980s this view changed, and R&D collaborations slowly moved into the centre of interest. In 1988, D’Aspremont and Jacquemin showed that pre-competitive cooperation in R&D leads to higher R&D intensity in an industry in the case of a high degree of technological spillovers. Cohen and Levinthal (1989) introduced the concept of absorptive capacity, which firms need in order to benefit from technological spillovers. The

multi-faceted nature of new technological knowledge more and more moved in the centre of interest, challenging the idea of being a public good. Knowledge is characterized as tacit (Polany, 1962), local (Atkinson and Stiglitz, 1969), cumulative (Cowan et al., 2000) and complex (Malerba and Orsenigo, 1993). Accordingly, technological spillovers are very likely to also support industrial innovation processes by spurring the collective innovation process and allowing for the exploration of extensive technological opportunities (Coombs, 1988), whereas the incentive-reducing effects are relevant only for a small fraction of technological spillovers affecting the intensive technological opportunities (Coombs, 1988) within a single industry. Referring to this important discussion, Nelson (1989) coined the notion of knowledge being a 'latent public good' only. Appropriability conditions are far better than the traditional view on new technological knowledge suggests and accordingly, the incentives to undertake R&D are far less endangered by technological spillovers.

Of course, this switch from the 'incentive-based' to the 'knowledge-based' view on the nature of industrial R&D also has consequences for the idea to consider technological spillovers as involuntary knowledge flows. Geroski (1995, p. 85) stressed this point: 'In particular, what often appears to be an involuntary flow of knowledge between firms may be nothing more than a pair of draws from a narrow but common pool shared by a group of agents within a common set of problems.' In other words, firms may have to take positive actions to make their newly developed knowledge available to others (Nelson, 1988) in order to spur the collective innovation process which they participate. Technological spillovers are hardly conceivable without being embedded in innovation networks, which are bound by collaborative R&D projects between the different actors in an industry. In this light, the traditional discussion of technological spillovers and their impact on the behaviour of firms and on industrial dynamics is misleading without considering the innovation networks as a means for knowledge flows.

The emphasis on knowledge collaboration and innovation networks first became central in a new strand of literature, namely industrial dynamics (ID) in a Neo-Schumpeterian tradition (Hanusch and Pyka, 2007a). In ID, besides the incentive-reducing effects of spillovers, their knowledge-creating effects are heavily emphasized (see e.g. Cantner and Pyka, 1998; Eliasson, 1995). Furthermore, exactly these knowledge-creating effects have to be considered as the basic motive for industrial actors to engage themselves in collaborative R&D. To analyse the impact of these mutual knowledge flows in industrial learning processes, the traditional modelling framework with its strong assumptions, in particular the equilibrium orientation, perfect rationality and related homogeneous technologies, is not applicable. Instead, to capture the dynamics going on in industrial learning processes fed by heterogeneous actors, in ID numerical simulations are adopted (see e.g. Windrum, 2007). Within the last few years in particular, the methodology of agent-based modelling (ABM) is frequently applied (see among others Gilbert and Troitzsch, 1999; Tesfatsion, 2001).

The ABM approach offers certain advantages (Pyka, 2006; Pyka and Fagiolo, 2007) for the investigation of innovation-driven dynamics and allows for the first

time the analysis of the processes going on in innovation networks and shaping their evolution. It basically consists of a decentralized collection of agents acting autonomously in various contexts. The massively parallel and local interactions can give rise to path dependencies, dynamic returns and their interaction. In such an environment global phenomena as the development and diffusion of technologies, the emergence of networks, herd-behaviour etc. which may cause a transformation of the observed system can be modelled. This modelling approach focuses on depicting the agents, their relationships and the processes governing their transformation. Very broadly, the application of ABM offers two major advantages with respect to the knowledge and learning orientation in ID:

(i) The first advantage of ABMs is their capability to show how collective phenomena come about and how the interaction of the autonomous and heterogeneous agents leads to their genesis. Furthermore, ABMs aim at the isolation of critical behaviour in order to identify agents that more than others drive the collective result of the system. They also endeavour to single out points of time where the system exhibits qualitative rather than sheer quantitative change.
(ii) The second advantage lies in the possibility to use ABM as computational laboratories to explore various institutional arrangements, various potential paths of development so as to assist and guide, e.g. firms, policy makers in their particular decision context.

The aim of this contribution is to introduce an agent-based simulation model (SKIN – **S**imulating **K**nowledge Dynamics in **I**nnovation **N**etworks; for a detailed introduction into the model's architecture see Ahrweiler et al., 2004), which allows studying innovation networks and their developments and to overcome the shortcomings of the discussions of technological spillovers in industrial economics. The agents are designed in a way which reflects the knowledge-based view of ID, i.e. they are characterized by imperfect knowledge which they try to improve; they act in strong uncertain environments which also are continuously affected and modified by their actions. Furthermore, the actors actively search for collaboration partners in order to struggle successfully in a complex innovation process.

This chapter builds on the results of the SKIN analysis of OL (Gilbert et al., 2007), the results of the analysis of network dynamics (Pyka et al., 2007) and the results of the analysis of the relation between agency and network architectures (Ahrweiler et al., 2007) and aims at a better understanding of innovation networks without which technological spillovers hardly occur. The basic structure of the model is briefly outlined in Sect. 5.2. In Sect. 5.3, various simulation experiments are shown in order to demonstrate the model's performance. In particular, it is shown that innovation networks can be a persistent organizational form of industrial R&D (Sect. 5.3.1) and not only of transitory nature as claimed in traditional IO theory. The phenomenon of innovation networks is closely related to the emergence of the so-called knowledge-based economies since the 1980s. The innovation networks in reality show certain characteristics in their architecture and dynamics which are knowledge-driven (e.g. Powell et al., 2005) and which do also appear in our

artificial innovation networks. In Sect. 5.3.2, we analyse these features, in particular the scale-free attributes of innovation networks. Some conclusions close the chapter.

5.2 Modelling Innovation Networks

SKIN is a multi-agent model containing heterogeneous agents which act and interact in a complex and changing environment. The actors are engaged in learning processes, both individually and collectively in network organizations. The basic structure of the model is summarized in Fig. 5.1. A more detailed description can be found in Appendix 1 of this chapter.

The agents represent innovative firms who not only try to sell their innovations to other agents and end users, but also have to buy raw materials or more sophisticated inputs from other agents (or material suppliers) in order to produce their outputs. This basic model of a market is extended with a representation of the knowledge dynamics in and between the firms. Each firm tries to improve its innovation performance and its sales by improving its knowledge base through adaptation to user needs, incremental or radical learning, and cooperation and networking with other agents allowing for combinatorial innovation processes.

The core concept of the framework is knowledge, which will manifest itself in the innovative production or delivery of manufactured and service products. The approach to knowledge representation used in the model is similar to Toulmin's

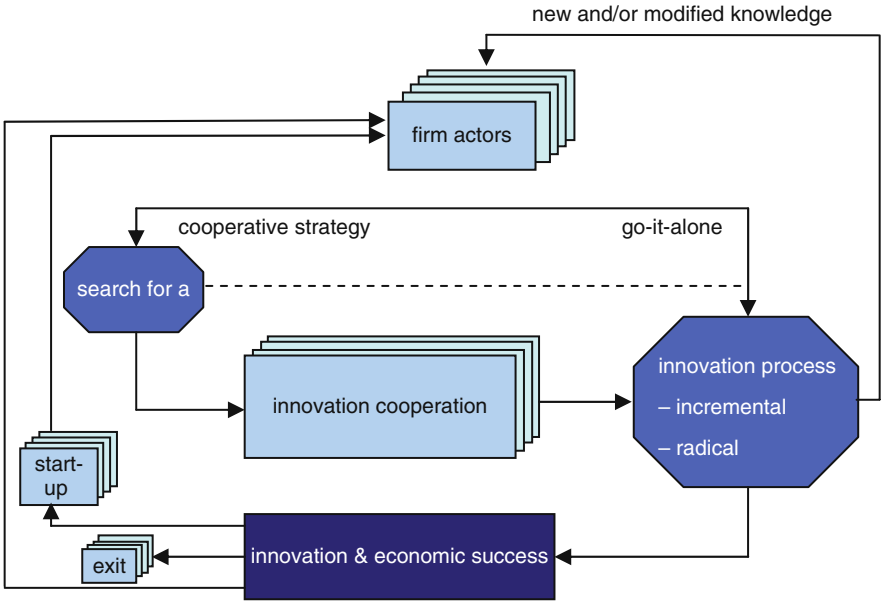


Fig. 5.1 Basic structure of the SKIN model

(1967) evolutionary model of knowledge production. He identified concepts, beliefs and interpretations as the 'genes' of scientific/technological development evolving over time in processes of selection, variation and retention. Ackermann (1970) interpreted the works of Kuhn and Popper according to this perspective allowing for different selection systems. In the SKIN model, the concept of a 'kene' is used to represent the aggregate knowledge of an organization (Gilbert, 1997).

Knowledge flows between the agents do not occur automatically but are dependent on the condition of an agreement of knowledge sharing and exchange within innovation cooperation. In this respect, our model sharply differs from traditional approaches in neoclassical industrial economics, because no (involuntary) technological spillovers exist. Of course, access to external knowledge sources is essential. However, as new technological knowledge is only a latent public good, the appropriability conditions allow excluding actors who are not participating in a collaboration. The possibilities of actors to include successfully knowledge from their collaboration partners of course depend on their absorptive capacities, i.e. the particular knowledge-make-up of their kenes.

5.3 The Emergence of Innovation Networks in an Artificial World

Following the strategies of empirically grounded and history-friendly modelling as pointed out by Malerba et al. (2001), the model draws on the theoretical frameworks and empirical insights from socio-economic innovation research.[1] For validation, namely whether the outputs for given inputs/parameters resemble empirical observations (although because the processes being modelled are stochastic and because of unmeasured factors, identical outputs are not to be expected, as discussed in detail in Gilbert and Troitzsch, 1999), we compare the model's outputs with empirical data. For this, we use our own case study on innovation networks in the UK and German bio-pharmaceutical industries (see Ahrweiler et al., 2006).

In the empirical as well as in our artificial world, we observe an initial shake-out of actors and a strong increase of collaborative activity in the early phase of industry development because of missing absorptive capacities and the inflexibility of established big firms which have to rely on specialized small high-tech enterprises. For their part, small firms need the large ones as 'commercializers' of their technological knowledge. As a result, we observe a change of the sectoral knowledge base while actors form networks in order to benefit from one another's competences and to perform combinatorial innovation processes.

In a more mature sector, the composition, attachment strategies and structural properties of the networks change – both empirically and in the simulated industry –

[1] Empirical results from five case studies of European innovation networks in different technological domains carried out in the EU project "Simulating self-organizing Innovation Networks" (see Pyka and Küppers, 2002) were initially used to set up a first version of the model.

to a stronger focus on partnerships between equals and to the growth of financing as a tie between firms, in addition to R&D links (for a description of a similar evolution of empirical network dynamics in the US biotech industries, see Powell et al., 2005). The risks and uncertainty inherent in new drug development are indicated by the high exit rate of projects and firms. Our model is able to reproduce these observed empirical features.

The following sections introduce the developments in our artificial world and show the wide possibilities to analyse network evolution there. We first show developments for a scenario which we label standard scenario and then focus on some experiments, modifying parameters and settings. For the standard scenario, the parameters of the starting distribution can be found in Appendix 2 of this chapter. The actors in the standard scenario apply the so-called conservative strategy, i.e. they were looking for actors with similar knowledge bases and the attractiveness threshold, i.e. the threshold which has to be surpassed in order to create a new partnership is set to a low value of 0.3. Figure 5.16 in appendix 3 gives the max–min-corridor of one of our observables in a Monte Carlo simulation to illustrate the robustness of our results. In Sect. 5.3.2 different scenarios, modifying the cooperation strategy, the number of large actors in the initial distribution as well as the attractiveness threshold are analysed. The purpose of this analysis is to gain a better understanding of the determinants of network evolution, their structures as well as their impact on the performance of the artificial industry.

5.3.1 Innovation Networks as a Persistent Form of the IO of R&D

The first aim of our modelling analysis is to show that innovation networks are a viable form of organization of industrial R&D and therefore are characterized by persistence. For this purpose we compile a scenario, labelled standard scenario, which should correspond closely to the conditions characteristic for knowledge-intensive industries where the phenomenon of innovation networks frequently appears in reality. The firm actors in our artificial world are engaged in innovation processes in order to survive in an environment, where besides price competition the competition for innovation plays a crucial role. Among the 500 actors in the initial distribution of the standard scenario, we find 50 large actors. The consideration of large actors and this particular relationship should reflect the situation characteristic for knowledge-intensive industries, where a few very large companies co-exist with quite a large number of small dedicated technology firms.

Figure 5.2a displays the development of the number of actors. Due to the alignment in the first iterations – products demanded by customers have to be developed as well as adequate input factors for production have to be purchased – a large share of small firm actors do not survive this period due to their only small endowment with capital. With successful market transactions appearing after around 300

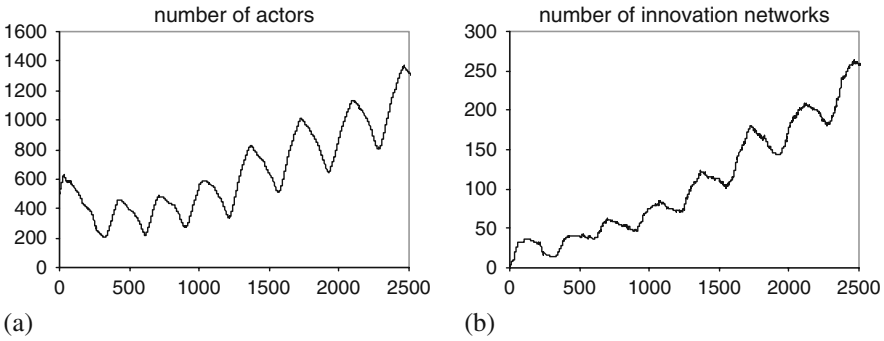


Fig. 5.2 (**a**) Number of actors. (**b**) Number of networks

iterations, however, life becomes easier for small actors and entry processes following the successful introduction of novelties are responsible for an again increasing number of actors. This situation repeats a couple of times during the simulation leading to a cyclical development, which follows the discovery of profitable regions in the innovation space, the exploitation of these possibilities and the subsequent exploration of new regions in the innovation space. These new regions are opened up by new combinations of the actors' kenes.

Figure 5.2b shows that this development is accompanied by a continuous evolution of networks measured by their number. During the shake-out periods, where the rate of successful innovations due to depleted opportunities decreases and with it the entry processes of start-ups slow down, some networks disappear temporarily. With the upswing in innovation and the related entries, the number of networks increases again. The cyclical development of cooperation activities therefore corresponds closely to the discovery of new major technologies (i.e. profitable region in the innovation space) leading to pronounced networking activities.[2]

Although the steadily increasing number of innovation networks gives a first hint on the persistence of this form of IO of R&D, only a closer look on the cooperative connections and the composition of networks allows for the conclusion that innovation networks are a viable form of R&D organization leading to a prolific industry development.

In Fig. 5.3, the frequency of the distribution of network sizes is illustrated. Rather obviously the majority of innovation networks encompass only a few actors and quite often consist of long-run bilateral cooperation. However, also a few innovation networks exist which include six and more actors. These large networks play a crucial role for knowledge flows in our artificial industry. This is underlined in Fig. 5.4, where the development of connectivity is displayed. Connectivity is a

[2] For example, Saviotti and Nesta (2006) have shown that the cyclical nature in innovation networks in the biotechnology-based industry corresponds to major technological waves (from RDNA, monoclonal antibodies, genetics to protein design and so on). Each technological shock introduces a new wave of cooperative activities in these industries.

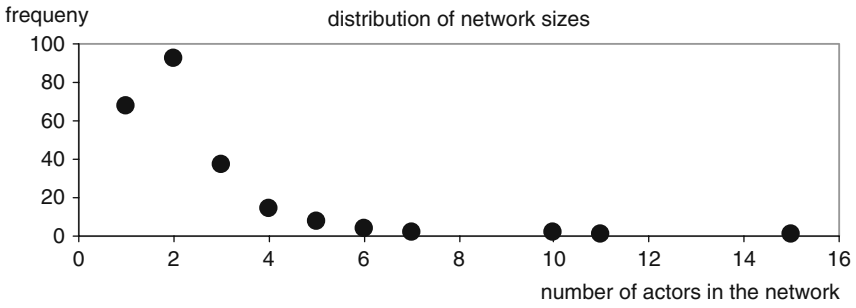


Fig. 5.3 Distribution of network sizes

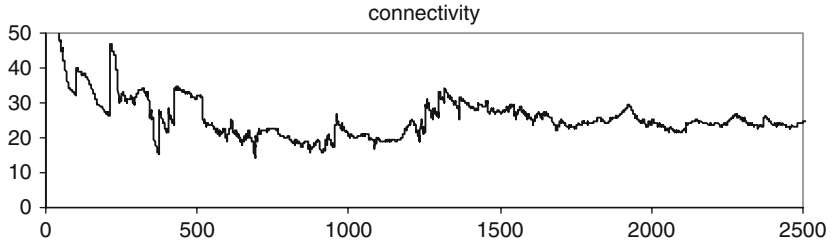


Fig. 5.4 Development of connectivity

measure in graph theory, which relates the number of actual connections to the number of potential possible connections in a network. It is usually applied as a rough indicator for the spread of information diffusion in networks. The considerable high values of connectivity over the whole time horizon of the simulation illustrates multi-channelled knowledge flows. This indicates a high degree of knowledge diffusion which has to be considered one of the major functions of innovation networks.

Following the results of our standard scenario simulations, we can show the persistent nature of innovation networks in knowledge-based industries. In the following section, we test different settings and their impact on the evolution of the innovation networks in order to develop a better intuition for the prerequisites and consequences of cooperative R&D in networks as well as of the networks' structural composition.

5.3.2 Scenario Analysis of Network Architecture

After having shown that innovation networks are a persistent organizational form of industrial R&D, we will now focus on the characteristics of networks and the design of their structures. Empirically, innovation networks in knowledge-intensive industries show features of so-called scale-free networks (Barabasi and Albert, 1999) and small worlds (Watts and Strogatz, 1998). The attribute of scale-free is related

to the dynamic growth of networks, whereas small worlds describe architectures combining a high rate of knowledge diffusion with a high speed of knowledge diffusion. The growth of these innovation networks seems to follow certain advantageous conditions concerning the stability and effectiveness of networks which leads to an unequal distribution of network connections. In these networks some actors are characterized as central actors because of their high number of collaborative relationships relative to other actors with only a few cooperative relations. Among others, Powell et al. (2005) have discovered scale-free attributes as well as small-world architectures for example in biotechnology innovation networks.

In our artificial world the question arises whether these attributes of real networks can be reproduced and which conditions are responsible for generating these particular network architectures. For this purpose, we run several experiments modifying parameters from the standard scenario and look at their impact on network evolution. The parameters modified are the initial distribution of firm sizes, the cooperation strategy as well as the attractiveness threshold.

(i) Modifying the initial distribution of firm sizes, i.e. whether we start with an equal distribution of similar (small) actors, or whether we include a number of considerably larger actors (*big firms*) in the initial distribution, can be considered as an experiment which allows the analysis of the impact large actors have on the network development. This is of particular importance in knowledge-intensive industries where in reality usually large actors, e.g. large pharmaceutical companies or the former national monopolists in telecommunication, co-exist with small technology-dedicated firms.
(ii) Modifying the cooperation strategy should give a hint on the rationale actors apply in the real world concerning the selection of cooperation partners. Obviously, the intention is to acquire knowledge which is not existent in the own firm and difficult to create alone. However, the degree of novelty of the external knowledge and its integration can vary from cases where firms try to reduce the difficulties of mutual knowledge exchange by looking for cooperation partners with rather similar knowledge (*conservative strategy*) to cases where firms try to maximize the amount of new knowledge transferred from cooperation partners (*progressive strategy*), hazarding the difficulties the integration of external knowledge might cause.
(iii) Finally, the modification of the attractiveness threshold allows analysing the general attitude of cooperation, i.e. whether the actors rather quickly decide to cooperate or not.

The different combinations of these modifications are summarized in Table 5.1, which describes the settings of the scenarios.

Figures 5.5 and 5.6 illustrate the development of actors in the different scenarios according to the applied cooperation strategy.

The scenarios with the conservative cooperation strategy are more successful with respect to the entry processes, and as entry is correlated to innovation, the actors

Table 5.1 Scenario settings

	Standard	I	II	III	IV	V	VI	VII
Number of big firms	50	50	50	50	0	0	0	0
Cooperation strategy	Conservative	Progressive	Conservative	Progressive	Conservative	Progressive	Conservative	Progressive
Attractiveness threshold	0.3	0.3	0.7	0.7	0.3	0.3	0.7	0.7

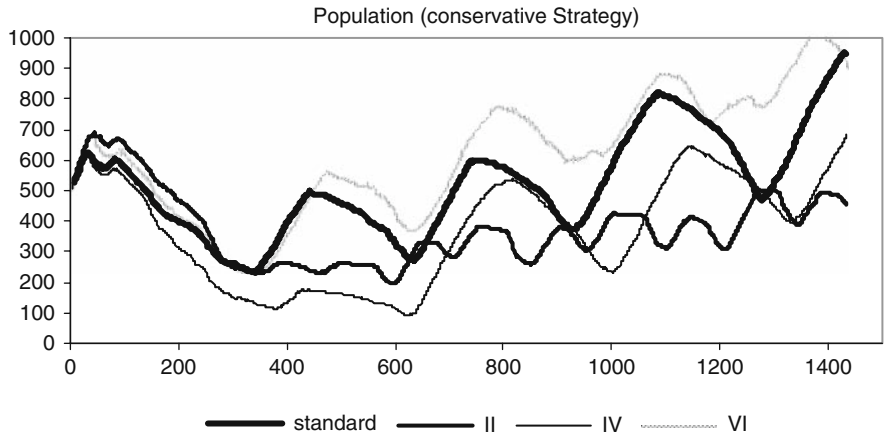


Fig. 5.5 Development of the actors' population in scenarios where the conservative strategy is applied

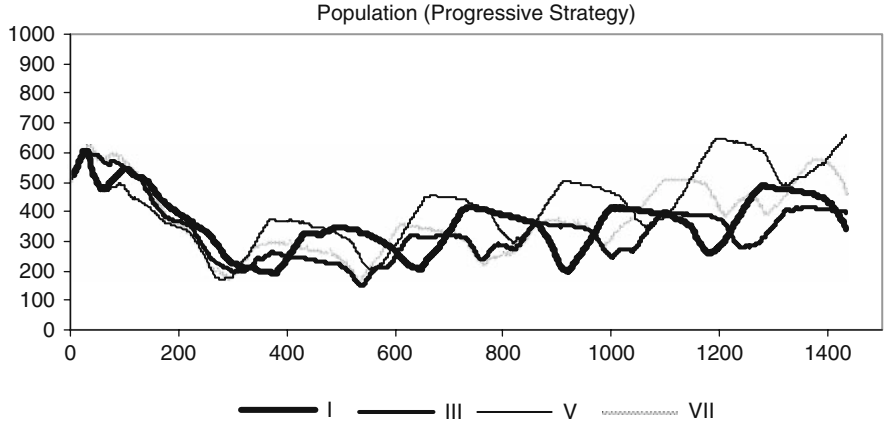


Fig. 5.6 Development of the actors' population in scenarios where the progressive strategy is applied

applying the conservative strategy are also more often introducing successful innovations compared to similar settings when only the progressive strategy is applied. Table 5.2 shows the linear growth trends[3] of the different scenarios. Only in the case of scenario II with the high attractiveness threshold, this value is (somewhat) lower compared to the corresponding scenario III with the progressive strategy.

[3] The linear growth trends are measured as the slope of a linear regression line.

Table 5.2 Linear trends in the growth of actors

	Standard	I	II	III	IV	V	VI	VII
Trend of growth in the number of firms	0.241	0.001	–0.03	–0.021	0.167	0.153	0.371	0.063

The lower attractiveness threshold is generally combined with more pronounced shake-outs or higher volatility. In these cases actors are less selective in their selection of cooperation partners accepting implicitly the potential higher risk of failures and market exits.

The diverse development in the number of actors obviously must have an impact on the development of concentration (market shares) in the industry. Market concentration is measured by the Herfindal-index[4], which is exemplarily displayed for the standard scenario and scenario IV in Fig. 5.7.

In contrast to the standard scenario, in scenario IV no large actors are in the initial distribution of firm sizes. Accordingly, we start with lower concentration values. However, the considerable shake-out of unsuccessful actors in the early periods leads to a considerable increase in concentration. The few actors remaining in the market are characterized by an extraordinary relative growth in market shares which leads to higher degrees of concentration over the whole simulated development. From this one can conclude a certain stability improvement introduced into the system by having an unequal initial distribution of firm sizes.

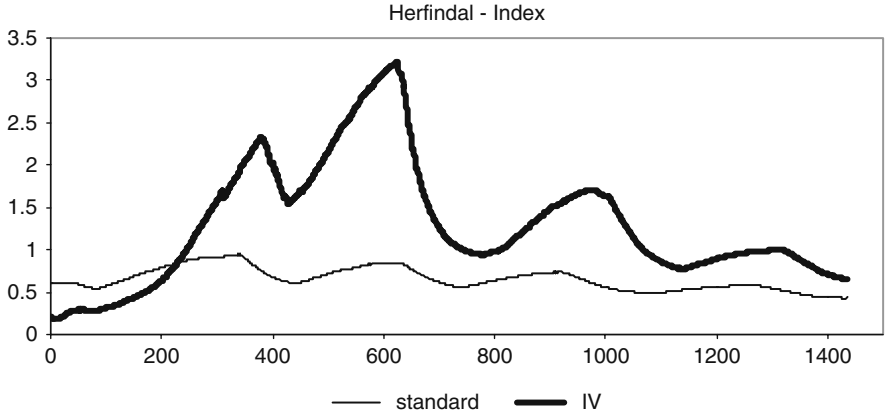


Fig. 5.7 Development of concentration

[4] The Herfindal index H is computed according to the following formula: $H \equiv \sum_{j=1}^{n} s_j^2$ where s_j stands for the market shares of the j actors, $j \in \{0, \ldots, n\}$.

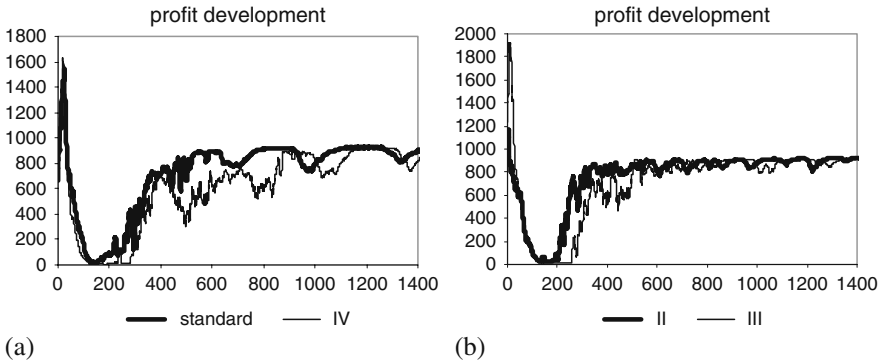


Fig. 5.8 **(a)** Comparison of profits (standard and IV). **(b)** Comparison of profits (II and III)

Figure 5.8a and b compares the development of the profit situation for different scenarios. The impact of the pronounced shake-out in scenario IV without large actors in the initial distribution of firm sizes, which has caused the higher degree of concentration, is also reflected in profit development. Compared to the standard scenario, the recovery period takes longer and during the further simulation the profit situation is almost always below to the one of the standard scenario.

Similar recovery problems characterize the profit situation of scenario III (progressive strategy) relative to scenario II (conservative strategy) both with the higher attractiveness threshold (0.7). As in both cases we have large actors in the initial distribution of firm sizes, the concentration effect cannot offer an explanation. An additional effect seems to be working, stemming from the innovation networks. Figure 5.9a and b, which shows the distribution of network sizes in the last iteration (1500) on a log–log scale, together with Table 5.3 illustrates a further commonality of the standard scenario and scenario II compared to the scenarios III and IV.

The SKIN model allows observing the structure of the emerging networks. Figure 5.9a and b shows the distribution of network sizes after 1500 iterations on a log–log scale. Besides a large number of networks encompassing only a small

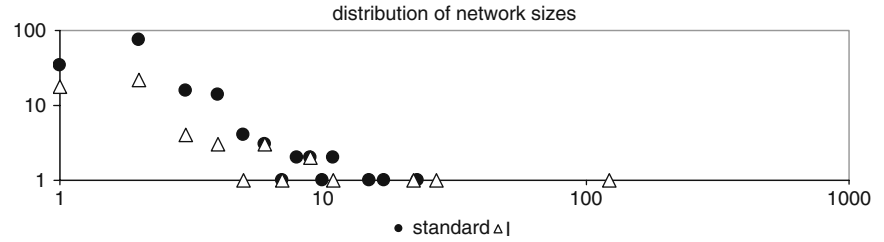


Fig. 5.9a Distribution of network sizes (conservative vs. progressive strategy, low attractivity threshold)

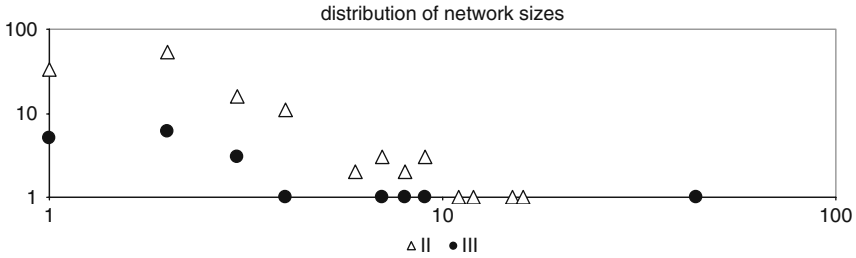


Fig. 5.9b Distribution of network sizes (conservative vs. progressive strategy, high attractivity threshold)

Table 5.3 The coefficients of power-law distributions of the networks after 1500 iterations

	Standard	I	II	III	IV	V	VI	VII
Coefficient of power-law distribution (R^2)	1.53 (0.822)	0.67 (0.582)	1.60 (0.894)	0.54 (0.603)	1.64 (0.874)	1.43 (0.697)	2.35 (0.868)	1.40 (0.521)

number of actors (up to 6), larger networks with up to 10 members and even more can be observed. This observation indicates that even if networking behaviour seems to be a widely shared strategic alternative, the actors in the various scenarios are engaged in networking to very different degrees. The particular shape also corresponds closely to the distribution of network links one might expect in so-called scale-free networks. In order to test this feature, Table 5.3 displays the coefficient of a power-law distribution for the eight scenarios under investigation.

Seyed-Allaei et al. (2006) have proved that indeed scale-free attributes of networks show up with rather small coefficients around 2. The figures in the table show that scale-free attributes of the underlying network emerge in the standard scenario as well as in scenario II (also the regression coefficient is close to one). In the cases of scenarios I and III, the coefficients are not large enough and the regressions are not significant to claim scale-free attributes for networks emerging in these scenarios. In scenarios IV and VI, we find coefficients sufficiently large (and regression coefficients sufficiently close to one) to state scale-free attributes to emerge. In scenarios V and VII, the coefficients of the power-law distribution are also not too small but the low regression coefficient in these cases does not allow to speak of scale-free attributes. From this comparison of the different scenarios one commonality of the scale-free solutions clearly stands out: only in the scenarios where the actors do apply the conservative collaboration strategy, networks with scale-free attributes emerge.

This feature of scale-free networks over longer time periods for scenarios with the conservative strategy shows that the SKIN model allows for the evolution of innovation networks, which are not only stable and a persistent organizational

form of industrial R&D but also, under certain circumstances, structurally corresponds to network architectures found in the reality of knowledge-intensive industries.

By these scenario analyses another crucial advantage of this type of model analysis can be demonstrated: with the help of our agent-based model, we can identify strategic dispositions of actors which lead to network dynamics resembling those in the real world. Empirically, these strategic dispositions are hardly to be measured but remain empirically more or less hidden.

Finally, we are interested in the possibilities of our model to generate innovation networks showing the features of small worlds. Seyed-Allaei et al. (2006) have shown that for rather extreme scale-free degree distributions with a power-law coefficient smaller than 2 automatically small-world architectures emerge. Accordingly, also in our model this particular combination of small path lengths amongst agents in the networks and a high degree of cliquishness should show up. This particular network architecture is beneficial in the organization of innovation processes because they guarantee both a fast diffusion of knowledge as well as a high spread of new knowledge (e.g. Baum, Shipilov and Rowley, 2003). Following Watts (1999) and Kogut and Walker (2001), we calculate the ratio of our artificial world's clustering coefficient and a random network with the same size concerning number of actors and ties and divided it by the ratio of the artificial networks' path length with respect to a purely random network of the same size (Fig. 5.10). Values above two are signalling small-world architectures.

Within the observed evolution of innovation networks in our standard scenario, we find almost from the beginning a strong prevalence for small worlds as an emergent property confirming the expectation by Seyed-Allaei et al. (2006).

In Fig. 5.11 we calculate in the same way the respective values for innovation networks in the biotechnology-based industries (Pyka and Saviotti, 2005). Since 1995, our real networks show features coming close to those of small worlds. These further correspondences of our artificial world with the dynamic developments of cooperation in innovation in this particular industry allow us to conclude that major mechanisms and inter-relations determining collaborative innovation processes and industrial learning are captured in our agent-based model.

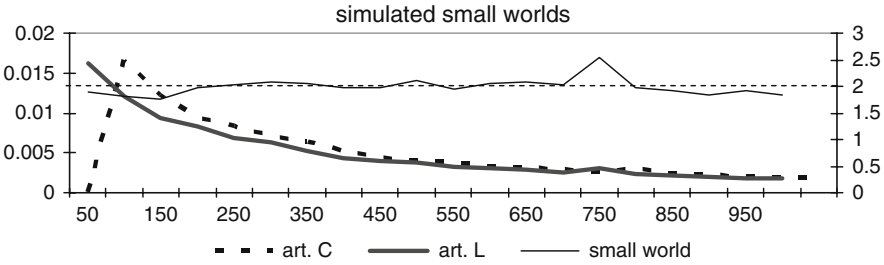


Fig. 5.10 Clustering coefficients (C) and path length (L) in the standard scenario

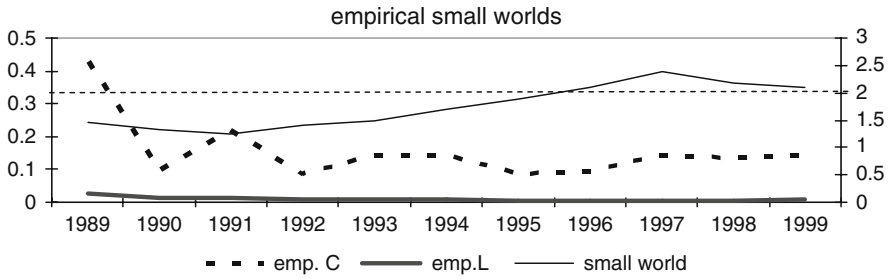


Fig. 5.11 Clustering coefficients (C) and path length (L) in biotechnology-based industries

5.4 Conclusions

The SKIN model is an attempt to improve our understanding of the complex processes going on in modern innovation processes. The model goes far beyond previous theoretical attempts in economics of analysing the IO of innovation processes. Instead of integrating strategic alliances and cooperative R&D in a standard equilibrium model of oligopolistic competition, insights coming from numerous case and industry studies are used to model the decision procedures. Technological spillovers are no longer considered to be automatically available, e.g. by correlating their size to the overall amount of industry R&D expenditures. Instead, the agents create a tight network consisting of mutual innovation collaborations which provides the channels for knowledge to flow between the actors and by this contribute to the collective innovation process. Using an agent-based simulation allows the modelling of innovation which can abstract from reality without assuming away the essentials of innovation processes (e.g. true uncertainty, historical time, heterogeneous agents learning experimentally and learning from each other in partnerships and networks) which are heavily emphasized by modern innovation economics (e.g. Nelson, 2001). In this sense, the book of fairytales of technological spillovers on the micro- and industry-level is closed and the new book on innovation networks is opened.

SKIN is designed in a way that allows the investigation of different industries. For calibration of the model's parameters different strategies can be envisaged: First, parameters can be estimated econometrically from various industry data sets describing the industry's cooperative behaviour which are increasingly available, for e.g. biotechnology-based industries. Alternatively, the second strategy focuses on the historical development of a particular industry which is reproduced in the simulation thereby identifying an adequate set of parameters.

In this chapter, we followed the second strategy and reproduced some stylized facts of real-world innovation networks in knowledge-intensive industries. It has been shown that SKIN can reproduce viable and persistent networks supporting the actors in their knowledge creation and diffusion processes. Furthermore, the networks evolving in the artificial SKIN-world show structural similarities with

real-world networks with respect to their degree distribution as well as particular network architectures. Under certain circumstances, SKIN innovation networks tend to be characterized by scale-free attributes and small worlds which are also characteristic for the architecture of innovation networks in knowledge-intensive industries.

SKIN is supposed to offer a first alternative within the broad set of possibilities of agent-based models for the analysis of the evolution of and the processes going on in innovation networks. The application of this class of models to this important phenomenon in the IO of R&D will obviously improve our understanding of the complex processes shaping industry evolution.

Appendix 1: The Agents and Their Knowledge

The individual knowledge base of a SKIN agent, its kene (Fig. 5.12), contains a number of 'units of knowledge'. Each unit is represented as a triple consisting of a firm's capability C in a scientific, technological or business domain (e.g. biochemistry), represented by an integer, its ability A to perform a certain application in this field (e.g. a synthesis procedure or filtering technique in the field of biochemistry), represented by a real number, and the expertise level E the firm has achieved with respect to this ability (represented by an integer). The firm's kene is its collection of C/A/E-triples which is of variable size and represents an artificial knowledge space.

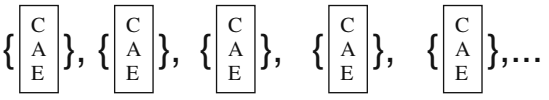


Fig. 5.12 The kene of an agent

Each firm, when it is set up, has also a stock of initial capital. It needs this capital to produce for the market and to improve its knowledge base, and it can increase its capital by selling products. The amount of capital owned by a firm is also used as a measure of its size and additionally influences the amount of knowledge that it can support, represented by the number of triples in its kene. Most firms are initially given a starting capital allocation, but in order to model differences in firm size, a few randomly chosen firms can be given extra capital. In many knowledge-intensive industries, we find a co-existence between large and small actors (e.g. the large pharmaceutical firms and the biotech start-ups, or the former national monopolists and high technology specialists in the ICT-industries). This particular distribution of firm sizes makes it necessary to discriminate between large and small actors in the simulation set-up.

Market Relationships and Innovation Processes

Firms apply their knowledge to create new products that have a chance of being successful in the market. The special focus of a firm, its potential innovation, is called an innovation hypothesis (IH) (Fig. 5.13). In the model, IH is derived from a subset of the firm's kene triples.

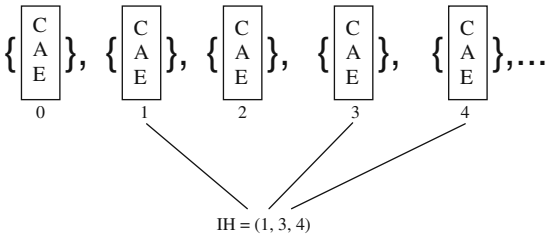


Fig. 5.13 Deriving an IH from an actor's kene

The underlying idea for an innovation, represented by IH, is the source an agent uses for its attempts to make profits. Transforming the IH into a product is a mapping procedure, where the capabilities of IH are used to compute an index number that represents the product. The particular transformation procedure applied allows similar products resulting from different kenes, which is not too far from reality where the production technologies of firms in a single industry vary considerably.

A firm's product, P, is generated from its IH as

$$P = \sum_{IH} C_i \bmod N, \tag{5.1}$$

where N is a constant.

The product has a certain quality which is also derived from the IH, by multiplying the abilities and the expertise levels for each triple in the IH and normalizing the result.

In order to realize the product, the agent needs some input factors. These can either come from outside the sector (raw materials) or from other firms, which generated them as their products. What exactly an agent needs is also determined by the underlying IH (Fig. 5.14): the kind of material required for an input is obtained by selecting subsets from the innovation hypotheses and applying the standard mapping function (see Eq. 5.1 above).

These inputs are chosen in such a way that each is different and differs from the firm's own product. In order to be able to engage in production, all the inputs need to be available on the market, i.e. provided by other agents or available as raw materials. If the inputs are not available, the agent is not able to produce and has to give up this attempt to innovate. If there is more than one supplier for a certain input, the agent will choose the one at the cheapest price and, if there are several similar offers, the one with the highest quality.

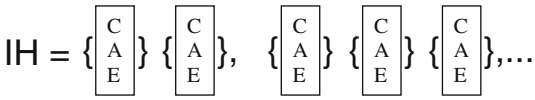


Fig. 5.14 The firm's input requirements

If the agent can start its production, it has to find a price for its own product which takes account of the input prices it is paying and a possible profit margin. While the simulation starts with product prices set at random, as the simulation proceeds, a price adjustment mechanism increases the selling price if there is high demand, and reduces it (but no further than the total cost of production) if there are no customers. A range of products are considered to be 'end-user' products and are sold to customers outside the sector: there is always a demand for such end-user products, provided that they are offered at or below a fixed end-user price. An agent will then buy the requested inputs from its suppliers using its capital, produces its output and puts it on the market for others to purchase. Using the price adjustment mechanism, agents are able to adapt their prices to demand side and in doing so learn by an adaptive feedback.

In making a product, an agent applies the knowledge in its IH and this increases its expertise in this area. This way learning by doing/using is modelled. The expertise levels of the triples in the IH are increased by 1 and the expertise levels of the other triples are decremented by 1. Unused triples in the kene eventually drop to an expertise level of 0 and are deleted from the kene; the corresponding abilities are 'forgotten' or 'dismissed' (cf. e.g. Hedberg, 1981).

Learning and Cooperation: Improving Innovation Performance

In trying to be successful on the market, the firms are dependent on their IH and thus on their kene. If a product does not meet any demand, the firm has to adapt its knowledge in order to produce something else for which there are customers (cf. e.g. Duncan, 1974). In the model, a firm has several ways of improving its performance, either alone or in cooperation, and in either an incremental or a more radical fashion. All strategies have in common that they are costly: the firm has to pay a 'tax' as the cost of applying an improvement strategy.

Incremental Innovation

If a firm's previous innovation has been successful, i.e. it has found buyers, the firm will continue selling the same product in the next round. However, if there were no sales, it considers that it is time for change (evaluating feedback). If the firm still has enough capital, it will carry out 'incremental' research (R&D in the firm's labs).

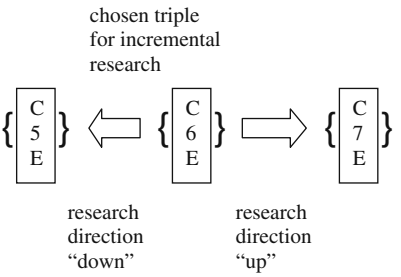


Fig. 5.15 Incremental research

Performing incremental research (Cohen and Levinthal, 1989) means that a firm tries to improve its product by altering one of the abilities chosen from the triples in its IH, while sticking to its focal capabilities. The ability in each triple is considered to be a point in the respective capability's action space. To move in the action space means to go up or down by an increment, thus allowing for two possible 'research directions' (Fig. 5.15).

Initially, the research direction of a firm is set at random. Later it learns to adjust to success or failure: if a move in the action space has been successful the firm will continue with the same research direction within the same triple; if it has been a failure, the firm will randomly select a different triple from the IH and try again with a random research direction.

Radical Innovation

A firm under serious pressure that is in danger of becoming bankrupt, i.e. using up its capital stock will turn to more radical measures, by exploring a completely different area of market opportunities. In the model, an agent under financial pressure turns to a new IH after first 'inventing' a new capability for its kene. This is done by randomly replacing a capability in the kene with a new one and then generating a new IH.

Cooperation and Networking: Combinatorial Innovation

Partnerships

An agent in the model may consider partnerships (alliances, joint ventures, etc.) in order to exploit external knowledge sources. The decision, whether and with whom to cooperate, is based on mutual observations of the firms, which estimate the chances and requirements coming from competitors, possible and past partners and clients.

The information a firm can gather about other agents is provided by a marketing feature: to advertise its product, a firm publishes the capabilities used in its IH. (Capabilities not included in its IH and thus in its product, are not visible externally and cannot be used to select the firm as a partner.) The firm's advertisement is then the basis for decisions by other firms to form or reject cooperative arrangements.

In experimenting with the model, we can choose between two different partner search strategies, both of which compare the firm's own capabilities as used in its IH and the possible partner's capabilities as seen in its advertisement. Applying the conservative strategy, a firm will be attracted by a possible partner who has similar capabilities; using a progressive strategy the attraction is based on the difference between the capability sets.

Previously good experience with former contacts generally augurs well for renewing a partnership. This is mirrored in the model: to find a partner, the firm will look at previous partners first, then at its suppliers, customers and finally at all others. If there is a firm sufficiently attractive according to the chosen search strategy (i.e. with attractiveness above the 'attractiveness threshold'), it will stop its search and offer a partnership. If the possible partner wishes to return the partnership offer, the partnership is set up.

The model assumes that partners learn only about the knowledge being actively used by the other agent. Thus, to learn from a partner, a firm will add the triples of the partner's IH to its own kene. For capabilities that are new to it, the expertise levels of the triples taken from the partner are reduced by 1 in order to mirror the difficulty of integrating external knowledge (cf. Cohen and Levinthal, 1989). For partner's capabilities that are already known to it, if the partner has a higher expertise level, the firm will drop its own triple in favour of the partner's one; if the expertise level of a similar triple is lower, the firm will stick to its own version. Once the knowledge transfer has been completed, each firm continues to produce its own product, possibly with greater expertise as a result of acquiring skills from its partner. Accordingly, cooperative innovation processes are conceived a combinatorial innovation, i.e. combing knowledge elements of the various partners.

Networks

If the firm's last innovation was successful, i.e. the amount of its profit in the previous round was above a threshold, and the firm has some partners at hand, it can initiate the formation of a network. This can increase its profits because the network will try to create innovations as an autonomous agent in addition to those created by its members and will distribute any rewards to its members who, in the meantime, can continue with their own attempts, thus providing a double chance for profits.

Networks are 'normal' agents, i.e. they get the same amount of initial capital as other firms and can engage in all the activities available to other firms. The kene of a network is the union of the triples from the innovation hypotheses of all its participants. If a network is successful it will distribute any earnings above the amount of the initial capital to its members; if it fails and becomes bankrupt, it will be dissolved.

Entry

If a sector is successful, new firms will be attracted into it representing a Schumpeterian competition by imitation. This is modelled by adding a new firm to the population when any existing firm makes a substantial profit. The new firm is a clone of the successful firm, but with its kene triples restricted to those in the successful

firm's advertisement, and an expertise level of 1. This models a new firm copying the characteristics of those seen to be successful in the market. As with all firms, the kene may also be restricted because the initial capital of a start-up is limited and may not be sufficient to support the copying of the whole of the successful firm's IH.

The code for the NetLogo model on which this paper is based is available on a request from the first author.

Appendix 2

The standard scenario uses the following parameter settings:

- Initial capital: 20,000 (big firms: 200,000)
- Initial population of firms: 500
- Number of large firms, with extra capital at the start: 50
- Range of product index numbers in the sector:]0.0 to 100.0[
- Maximum difference between product and input index numbers for them to be considered substitutable: 1.0
- All products with a product number below 5.0 are considered to be 'raw-materials' and all those with numbers above 95 are 'end-user' products.
- Price of raw materials: 1
- Maximum price of end-user products: 1000
- Profit required attracting new start-ups: 1200
- Partnering search strategy: conservative
- Attractiveness threshold to allow two firms to partner: 0.3
- Capital cut-off below which firms do radical rather than incremental research: 1000
- Taxes: per time step: 200; per incremental research attempt: 100; per radical research attempt: 100: per collaboration partner: 100

Appendix 3

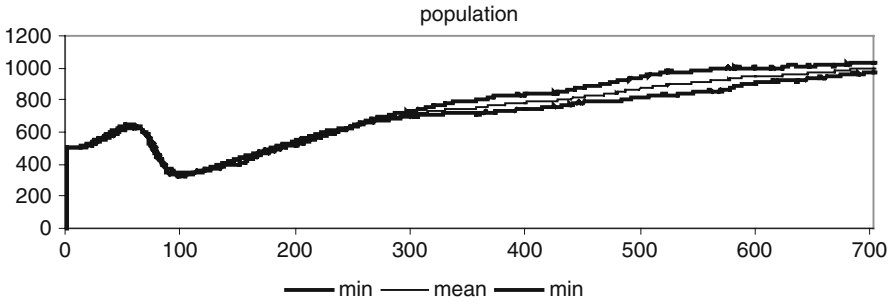


Fig. 5.16 The min–max-corridor of a Monte Carlo simulation

References

Ackermann, R. (1970). The Philosophy of Science. New York: Pegasus.

Ahrweiler, P., Pyka, A. and Gilbert, N. (2004). Simulating Knowledge Dynamics in Innovation Networks. In: Leombruni, R. and Richiardi, M. (eds.). Industry and Labor Dynamics: The Agent-based Computational Economics Approach. Singapore: World Scientific Press, pp. 284–296.

Ahrweiler, P., Gilbert, N.and Pyka, A. (2006). Institutions matter, but . . . Organisational alignment in knowledge-based industries, Science, Technology and Innovation Studies, 2(1), 3–18.

Ahrweiler, P., Gilbert, N. and Pyka, A. (2007). Agency and structure: a social simulation of European Knowledge-Intensive Industries, Mimeo.

Amable, B. (2003). The Diversity of Modern Capitalism. Oxford: Oxford University Press.

Atkinson, A.B. and Stiglitz, J.E. (1969), A new view on technological change, Economic Journal, 79, 116–131.

Bahlmann, T. (1990). The learning organization in a turbulent environment. Human Systems Management, 9, 249–256.

Barabasi, A. and Albert, R. (1999), Emergence of scaling in Random networks, Science, 286, 509–512.

Baum, J.N., Shipilov, A.V. and Rowley, T.J. (2003). Where do small Worlds come from?. Industrial and Corporate Change, 12(4), 697–725.

Cantner, U. and Pyka, A. (1998). Absorbing technological spillovers. Simulations in an evolutionary framework. Industrial and Corporate Change 7(2), 369–397.

Cohen, W.M. and Levinthal, D. (1989). Innovation and learning: the two faces of R&D. The Economic Journal, 99, 569–596.

Coombs, R. (1988). Technological opportunities and industrial organization. In: Dosi, G. et al. (eds.). Technical Change and Economic London: Theory. Pinter Publishers.

Cowan, R., David, P.A. and Foray, D. (2000). The explicit economics of knowledge: Codification and tacitness. Industrial and Corporate Change, 9(2), 211–233.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and non-cooperative R&D in duopoly with spillovers. American Economic Review, 78, 1031–1038.

Duncan, R. B. (1974). Modifications in decision structure in adapting to the environment. Some implications for organizational learning. Decision Sciences 5, 705–725.

Eliasson, G. (1995). General purpose technologies, industrial competence and economic growth. Working Paper, Stockholm: Royal Institute of Technology.

Geroski, P.A. (1995). Do spillovers undermine the incentives to innovate? In: Dowrick, S. (ed.). Economic Approaches to Innovation, Cheltenham, UK: Edward Elgar.

Gilbert, N. and Troitzsch, K. (1999) Simulation for the Social Scientist, Berkshire, UK: Open University Press.

Gilbert, N. G. (1997), A simulation of the structure of academic science. Sociological Research Online 2/2, http://www.socresonline.org.uk/socresonline/2/2/3.html.

Gilbert, N., Ahrweiler, P. and Pyka, A. (2007). Learning in innovation networks: some simulation experiments. PhysicaA, 378(1), 100–109.

Hanusch, H. and Pyka, A. (2007a), The principles of neo-schumpeterian economics. Cambridge Journal of Economics, 31(2), 275–289.

Hanusch, H. and Pyka, A. (2007b). Applying a comprehensive neo-schumpeterian approach to Europe and its Lisbon Agenda. In: Welfens, P. (Hrsg.), Europe's Innovation Performance, Heidelberg: Springer.

Harrigan, K. (1985). Strategies for Joint Ventures, Lexington: Lexington Books.

Hedberg, B. (1981). How organizations learn and unlearn. In: Nystrom, P.C. and Starbuck, W.H. (eds.). Handbook of Organizational Design. Oxford: Oxford University Press, pp. 3–27.

Kogut, B. and Walker, G. (2001). The small world of Germany and the durability of national networks. American Sociological Review, 66, 317–335.

Lam, A. (2003). Organizational learning in multinationals: R&D networks of Japanese and US MNEs in the UK. Journal of Management Studies 40(3), 673–703.
Levin, R.C. and Reiss, P.C. (1988). Cost-reducing and demand creating R&D with spillovers, Rand Journal of Economics, 19, 538–556.
Malerba, F. and Orsenigo, L. (1993), Technological regimes and firm behaviour. Industrial and Corporate Change, 2, 45–71.
Malerba, F., Nelson, R., Winter, S. and Orsenigo, L. (2001). History-friendly models – An overview of the case of the computer industry. Journal of Artificial Societies and Social Simulation, 4(3).
Nelson, R.R. (1988). Modelling the connections in the cross section between technical progress and R&D intensity. Rand Journal of Economics, 19(3), 478–485.
Nelson, R.R. (1989). What is private and what is public about technology. Science, Technology and Human Values, 14, 229–241.
Nelson, R.R. (2001). Evolutionary perspectives on economic growth. In: Dopfer, K. (ed.). Evolutionary Economics, Program and Scope. Boston, Dordrecht, London: Kluwer Academic Publishers.
Polany, M. (1962). Personal Knowledge: Towards a Post-Critical Philosophy. New York: Harper Torchbook.
Powell, W.W., Koput, K. and Smith-Doerr, L. (1996). Inter-organizational collaboration abd the locus of innovation: networks of learning in biotechnology. Administrative Science Quarterly, 41, 116–145.
Powell W.W., Douglas, R.W., Kenneth, W.K., Jason and O.-S. (2005). Network dynamics and field evolution: The growth of inter-organizational collaboration in the life sciences. American Journal of Sociology, 110(4), 1132–1205
Pyka, A. (2002). Innovation networks in economics: from the incentive-based to the knowledge-based approaches. European Journal of Innovation Management, 5, 152–63.
Pyka, A. (2006). Modelling qualitative change. In: Rennard, J.-R. (ed.). Handbook of Research on Nature-Inspired Computing for Economics and Management, Chapter XVI, Vol. 1, Idea group Reference, Hershey (USA), 211–224.
Pyka, A. and Küppers, G. (2002) (eds.). Innovation Networks – Theory and Practice. Cheltenham: Edward Elgar.
Pyka, A. and Saviotti, P.P. (2005). The evolution of R&D networking in the biotech industries. International Journal of Entrepreneurship and Innovation Management, 5, 49–68.
Pyka, A. and Hanusch, H. (eds.) (2006). Applied Evolutionary Economics and the Knowledge-Based Economy, Cheltenham, UK: Edward Elgar.
Pyka, A. and Fagiolo, G. (2007). Agent-based modelling: A methodology for neo-schumpeterian economics. In: Hanusch H. and Pyka A. (eds.). Elgar Companion to Neo-Schumpeterian Economics, Cheltenham, UK: Edward Elgar.
Pyka, A., Gilbert, N. and Ahrweiler, P. (2007), Simulating knowledge-generation and – distribution processes in innovation collaboration and networks. Cybernetics and Systems, 38(7), 667–707.
Romer, P. (1990). Endogenous technical change. Journal of Political Economy, 98(5), 71–102.
Sapienza, A. (1989), R&D collaboration as a global competitive tactic: biotechnology and the ethical pharmaceutical industry, R&D Management. 19, 285–295.
Saviotti, P.P. and Nesta, L. (2006). Firm knowledge and Market value in biotechnology. Industrial and Cooperate Change, 15, 625–652.
Seyed-Allaei, H., Bianconi, G. and Marsili, M. (2006), Scale-free networks with an exponent less than two. Physical Review E, 73, 046113.
Tesfatsion, L. (2001). Agent-based modelling of evolutionary economic systems, IEEE Transactions on Evolutionary Computation, 5, 1–6.
Toulmin, S. (1967). The Philosophy of Science. An Introduction. London: Hutchinson.
Watts, D.J. (1999). Networks, dynamics, and the small-world phenomenon, American Journal of Sociology, 105, 493–528.

Watts, D.J. and Strogatz, S.H. (June, 1998). Collective dynamics of 'small-world' networks. Nature, 393, 440–442.
Windrum, P. (2007). Neo-Schumpeterian simulation models. In: Hanusch, H. and Pyka A. (eds.), Elgar Companion to Neo-Schumpeterian Economics, Cheltenham, UK: Edward Elgar.

Chapter 6
Structural Holes, Innovation and the Distribution of Ideas

Robin Cowan and Nicolas Jonard

6.1 Introduction

This contribution examines the relationship between the architecture of an industrial R&D network and efficiency in knowledge distribution, both from the point of view of individual firm performance, and at the level of the system.

Recent technological changes have had the effect of creating multi-product firms: the knowledge base on which both production and innovation are founded has, in general, become much broader, covering more and different types of knowledge (Grandstrand and Sjolander, 1990). As a consequence, firms increasingly discover that their in-house knowledge is not sufficient for efficient production or innovation. This had driven them to seek the knowledge they need outside, in other firms. However, and precisely due to the nature of knowledge, this task is difficult to achieve through pure market interaction. Thus, firms are now forming relatively long-term alliances, formal and informal, with other, often competing, firms. This has led to the networked organization, a hybrid organizational form lying between the market and a pure hierarchy (Powell, 1990), which takes advantage of both market and non-market interactions. Networks supply firms with rapid, flexible access to resources outside their core competencies. Strong, stable contacts with other firms can provide a firm with the knowledge it needs for its immediate production or innovation without navigating the difficulties of market transactions for knowledge. In addition, contacts of this type can also provide a form of insurance – giving a firm rapid access to information about developments taking place in other firms or related industries. These observations are particularly relevant to knowledge-intensive and science-based industries.

R. Cowan (✉)
BETA, Université Louis Pasteur, 61 Avenue de la Forêt Noire, 67000 Strasbourg, France;
UNU-MERIT, Maastricht University, 6200 MD Maastricht, The Netherlands
e-mail: r.cowan@merit.unimaas.nl

This contribution was originally published as "Structural holes, innovation and the distribution of ideas", DOI 10.1007/s11403-007-0024-0 *Journal of Economic Interaction and Coordination*, Volume 2, Number 2, December, 2007, pp.93–110

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_6,

There is now a large literature, much located in the management field, examining the structural properties of innovation or R&D networks. Three properties recur. Networks tend to be sparse. That is, of the total possible connections between agents, the actual connections constitute a small proportion. Networks tend to be locally dense. Local clusters of closely interconnected agents are common. In addition, the local clusters tend to be only sparsely connected to each other. Finally, the distribution of links over agents tends to be highly skewed. It is probably too strong to say that there is common evidence of power-law distributions, but relatively heavy tails do exist. The discussion in the literature concerns how these properties arise, and how firms' performance is affected by them.

As to the relationship between network position and performance, there are roughly speaking two competing views. On the one hand, following Coleman (1988), it is possible to argue that dense sub-groups are a source of social capital.[1] A group of highly inter-connected agents generates trust, common languages and problem-solving heuristics, social disapprobation for opportunistic behaviour and so on. If firms i and j are linked, they can share information about a common partner k. This reduces significantly the incentives for k to behave opportunistically against, j, even if he will never see j again, since information about his behaviour will travel rapidly to i and to all the other members of the clique. More positively, if i is working on a problem, using information gathered from j and k, or in discussion with j and k, an ability of j and k to discuss the problem with each other, or exchange information about it, can only have a positive effect on i's ability to solve his problem. These considerations imply that structurally embedded partnerships will be important sources of value for a firm, and redundant links are privately and probably also socially valuable.[2] Thus for a firm, a useful link formation strategy is to close open triangles and create strong cliques. The value of this strategy is observed in empirical studies by Dyer and Nobeoka (2000), on the automobile industry; Gulati and Gargiulo (1999) in a study of alliance formation in several industries; Powell et al. (1996) who study the impact of network position on innovation performance in the biotechnology sector; and Rowley et al. (2000) in a study of strong and weak ties in innovation networks in the steel and semiconductor industries.

On the other hand, Burt (1992) argues that dense local links are redundant in a strong sense, that the existence of structural holes in a firm's ego network is efficient, and that locally dense networks can be a source of rigidity. A structural hole exists if two of my neighbours are not linked to each other. Through these two neighbours I am connected to different parts of the larger network, and thus have access to different sources of dispersed information. Thus if a firm is to form a new link, closing a structural hole is less valuable than finding a partner to whom none of my current partners is currently connected. This is closely related to the argument of Podolny (1993) that firms attempt to increase their betweenness centrality. If

[1] See also Walker et al. (1997) on the same subject.

[2] The third link that closes a triangle can be seen as redundant, since its effect is simply to create a path of length one between two agents where a path of length 2 already existed. Notice here that "redundant" is only strongly applicable if this reduction in path length serves no purpose, that is, if in general path lengths are not (privately) important.

many shortest paths between firms go through firm *i*, then *i* can exert considerable control over knowledge flows. Particularly in the knowledge economy, control of knowledge flows can be translated into rents. This is an argument that clique spanning ties are valuable. The value of structural holes has been examined empirically by, for example, Ahuja (2000) in the context of the international chemical industry (structural holes have a negative impact on industry performance, whereas indirect and direct ties have a positive impact on firm innovative performance); Gargiulo and Bennassi (2000) who find in a study of an Italian IT firm that dense local networks do not respond well to change (they find a trade-off associated with the safety conferred by cohesive ties (social capital) and the flexibility conferred by ties that connect different parts of a network); Baum et al. (2003) studied the sources of inter-clique link formation in the Canadian merchant banking industry.

This debate between social capital and structural holes is formalized through the notion of clustering. An agent's ego network is clustered if many of its partners are partners of each other. The structural holes argument claims that a highly clustered ego network is bad for performance; the social capital position argues the opposite. At an aggregate level, individual clustering levels can be averaged to describe an industry (or sector or economy) network. By extension, the structural holes argument implies that unclustered networks will perform well, whereas the social capital position argues that locally dense networks, which by definition are highly clustered, will perform well. One way in which these positions are sometimes reconciled (see for example Rowley et al., 2000) is that they apply to different moments in an industry life cycle. When an industry is young, technologies are being explored, and many different avenues of advance are potentially fruitful. Here, it is important to have rapid access to "distant" (both in geographical and technological space) information. Thus redundant ties are less valuable than ties that connect to different parts of the network. Structural holes are desirable. However, in a more mature industry, there are fewer technological surprises, so exploitation is more common. Here, a dense core of agents addressing similar issues creates the critical mass that is necessary to make further progress along the chosen path. Social capital becomes more valuable.

While clustering has received much attention in the literature, there is a second aspect of structure that is now considered important. This is the distribution of edges over nodes, and in particular the extent to which this distribution is skewed, with a few nodes having many links and the majority having few. Though the skewness of the degree distribution is relatively well established (see Powell et al., 2005 for example), its implications for R&D networks are not well understood. The genesis of a skewed distribution lies in some form of "preferential attachment": firms with many existing links are likely to be valuable partners, and so attract more links. On the face of it, this is intuitively appealing, since a large number of partnerships indicate that a firm both has useful knowledge and knows how to collaborate. But a skewed distribution implies the presence of stars in the network – agents through whom many (short) paths run. Consequently, while this structure is robust to random failures (since the failure of a randomly selected node affects, most likely affects, only a few other nodes), it is very fragile to specific failures (if a star fails, many other nodes, and paths between many pairs of nodes, are affected). Stars in a network

can serve as important centres of knowledge distribution, and so a highly skewed network may be conducive to very rapid diffusion of knowledge. On the other hand, though, if a star ceases to participate in the system for some reason, this can cause a serious disruption in the distribution system, and if the situation is such that agents withdraw from the system from time to time (for whatever reason) a flatter link distribution may be preferable.

While network structure will clearly have an effect on the efficiency of diffusion, structure may interact with the micro-specifics of exchange. The transmission of knowledge, particularly among competing firms, is a challenge for economists, especially if there is no market for knowledge. Two patterns of transmission have been observed empirically. Allen (1983) describes "collective invention" in which knowledge is given away as a (local) gift. In the steel industry in Cleveland, UK, in the mid 19th century, for example, steel producers met regularly under the auspices of societies like the Cleveland Institute of Engineers, the South Wales Institution of Engineers or the national Iron and Steel Institute and disclosed their own recent technological developments. As a producer made an advance in furnace height or temperature, for example, that producer would document the change – how it was accomplished, the technical effects and so on – and present this to other local firms. Knowledge was essentially given away to competitors within the local cluster, and as a consequence, the technology developed rapidly.[3] Von Hippel (1987) on the other hand documents a barter exchange. Technical managers of steel mini-mills in the US exchange technical information and explicitly help each other solve problems.[4] But here the transfer is not a gift: there is a *quid pro quo*. While the interaction is not market-based, there are social sanctions if an agent routinely receives but does not give knowledge. In essence, knowledge is bartered. In both of these cases, knowledge transmission is local, taking place in face-to-face interactions. If this represents the nature of knowledge diffusion, then the structure of local interactions will play a central role in the process through which a "piece of knowledge" moves from one geographic location to another, or more generally how it diffuses throughout an economy.[5]

These are the issues we take up in this contribution. We are interested in how the architecture of the communication network affects its performance in terms of knowledge distribution. This relationship may change, however, depending on the details of the transmission mechanism. The model we construct below permits us to examine both architecture and transmission as variables controlling knowledge diffusion.

---

[3] McGaw (1987) finds a similar pattern in paper manufacturing in New England in the early 1880s, and Lamoureaux (1999) cites other examples from the 18th and 19th centuries in the US.

[4] Later work found the same phenomenon in aerospace and waferboard industries (von Hippel, 1994). Powell et al. (1996) documented a similar phenomenon in biotech.

[5] For models examining the relationship between network structure and knowledge diffusion, see Cowan and Jonard (2003) on the gift economy, and Cowan and Jonard (2004) on a barter economy.

6.2 The Model

In general, for any agent, more knowledge is better from the point of view of producing goods, and the communication network is the infrastructure over which agents acquire the knowledge they need in order to produce.

To capture this, we model a world in which there is a fixed, finite population of agents, and a fixed, finite number of ideas relevant to production. An agent is characterized by two properties: the set of ideas he has and his production goal. Production is controlled by a Leontieff production function for which only a small number of ideas is necessary, different agents having different production functions. Production is done in isolation, but demands that an agent possesses all the ideas that are relevant to his productive activity, i.e. the ideas for which his production function has non-zero coefficients. If one or more ideas are missing, they can be acquired via an agent's acquaintances. Thus the set of ideas held by any agent evolves over time, and we shall assume this takes place through a simple process of one-to-one exchange or gift.

Ideas have the feature that agents do not lose by giving. Thus, it might be argued that agents could well give without asking for reciprocity. If agents have to compete in a second step, however, even though an agent's knowledge level does not decrease by giving away ideas, his competitive position might. Different industrial contexts will display different "terms of trade". In the present contribution, we will consider a world of knowledge traders, a world of knowledge givers and a mixed situation in which both co-exist.

6.2.1 Network Structure

Let $G(V,N)$ be the undirected graph representing the industry network, with $V = \{1, \ldots, n\}$ the set of agents and $N = \{N_i, i \in S\}$ the correspondence specifying, for each $i \in V$, the neighbourhood N_i of i. The degree of firm i is the number of direct ties of that firm $n_i = \#N_i$. Average degree in the network is then

$$n = \sum_{i \in V} n_i / n.$$

Any $j \in N_i$ is at distance 1 from i. Indirect ties connect i to individuals at a distance strictly more than 1. Define d_{ij} the distance between i and j, as the number of edges in the shortest path, or geodesic, connecting i to j. The average distance to i is then

$$d_i = \frac{1}{n-1} \sum_{j \neq i} d_{ij}$$

and average distance (characteristic path length) is

$$d = \sum_{i \in V} d_i / n.$$

An additional measure of the structure of local links is clustering. The extent to which i's neighbourhood is clustered is measured by

$$c_i = \frac{2}{n_i(n_i - 1)} \sum_{j,l \in N_i} \xi_{jl},$$

where $\xi_{jl} = 1$ if $j \in N_l$ and $\xi_{jl} = 0$ otherwise. This statistic measures the proportion of existing triangles among those which could involve i, given n_i. The clustering coefficient for the network as a whole is

$$c = \sum_{i \in V} c_i / n.$$

Though more sophisticated measures of structural position can be designed, for our purpose the distribution and organization of direct and direct ties as captured by degree, clustering and distance are sufficient.

In order to examine the relationship between structure and performance, we create an algorithm to construct a family of random graphs derived from an ordered substrate. As a first step, consider the rewiring algorithm from Watts and Strogatz (1998) that has as a control parameter the probability p that a link in a periodic lattice is randomly rewired. Start from the periodic lattice with an even number m of nearest neighbours ($n_i = m$, for all $i \in V$) and sequentially consider each edge, making a decision of uniform random rewiring (with probability p) or preservation (probability 1-p). As p increases, the regular periodic lattice ($p = 0$) is left and through intermediate states ($0 < p < 1$) a random graph with uniform degree is reached ($p = 1$). This procedure creates a small amount of variation in individual degree n_i and an average degree of m is preserved as the total number of edges is kept constant (it is exactly equal to $nm/2$, half the degree sum). We know from Watts and Strogatz (1998) that there is an interval (the small world region) over which clustering remains high while path length has fallen close to the level of a random graph of average degree m. Indeed, when the number of random links is small the removal of a few of them has a strong effect on average path length while it has only little effect on the clustering coefficient.

Besides clustering and path length, we are interested in the effect of asymmetry in the degree distribution. There has been extensive debate (see for instance Barabási and Albert, 1999) about scale-free networks (that is, networks with a power-law degree distribution having exponent between 2 and 3) and the extent to which they can be found in empirical data. Scale-free networks and power-law distribution in general are not our interest here. Rather, we will simply explore the effect of having stars in the system, in addition to the possibility of having more or less clustered random structures. We do this the following way. Assume now there are two classes

of agents: s stars with degree $d_i = D$ for $i \in S \subseteq V$, and n-s non-stars with degree $d \leq D$. The problem is now to design a procedure analogous to the one above that permits us to tune stardom (s) and randomness (p) independently, while having a constant degree sum in the graph ($sD + (n - s)d = nm$). To that end, locate the n individuals on the circle. Take an even $d \leq m - 2$, and sequentially pick s nodes at random to form S. First, the stars are taken care of: for each $i \in S$ create an even $D = (nm - (n - s)d)/s$ links. With probability 1–p, each of the D links is connected to one of the $D/2$ nearest nodes on each side of i on the circle. With probability p, it is connected at random. That takes care of the stars and creates also some links for the non-stars. Then run across all the non-stars and proceed analogously, checking that the degree constraint for non-stars is also satisfied. Because only integer numbers are handled, the procedure will in general not create exactly nm links. However, the results will be reasonably close to that target. The three-panel Fig. 6.1 summarizes a few statistics for an illustration with $n = 500$, $d = 6$ and $m = 10$, i.e. a degree sum of 5000 which the algorithm roughly preserves. The sum is displayed in the insert in the upper panel of Fig. 6.1. The concentration of links (upper panel) falls monotonically with s, across all p values. The two lower panels of Fig. 6.1 show clustering and path length versus p for $s = 21$, 107 and 500 stars. Clustering displays no significant variation with s, while displaying its usual pattern with p. Distance also behaves monotonically with degree, while the effect of s (owing to the assumption of a constant degree sum) is monotonic and quite obvious. The algorithm thus behaves well and will permit us to explore independently the effects of degree asymmetry and local disorder.

6.2.2 The Dynamics of Ideas

Over time, as ideas are exchanged, knowledge evolves. An agent's knowledge stock never shrinks, as we assume that ideas display non-rivalry (i does not lose his idea by letting j have it), thus one agent's knowledge only increases or stays constant as times passes.

6.2.2.1 Knowledge as Sets

Agents operate in a system (an industry) where there is a finite number of existing ideas indexed by $l = 1, \ldots, L$. Each agent is endowed with a subset of these ideas, and we denote $H_i \subseteq \{1, \ldots, L\}$ the set of ideas held by i. Agents use ideas for the sake of production, and in that respect agents are heterogeneous.

Production is done by agent i according to the Leontieff production function

$$\phi_i = \begin{cases} 1 \text{ if } P_i \subseteq H_i, \\ 0 \text{ otherwise }, \end{cases} \tag{6.1}$$

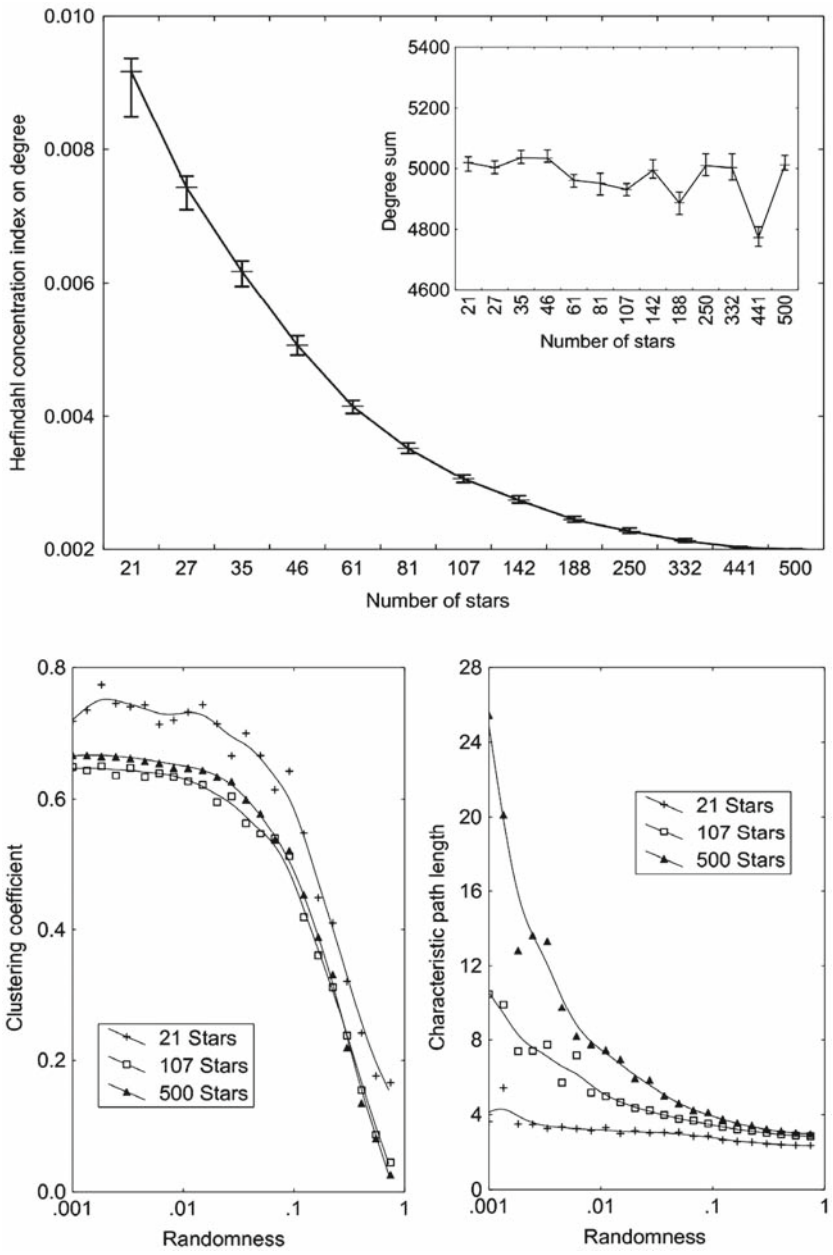


Fig. 6.1 Characteristics of the graph family: asymmetry in the degree distribution as measured by the Herfindahl concentration index (*upper panel*, pooled data across p, the whiskers are the minimum and maximum values), degree sum (insert in *upper panel*, pooled across p), clustering coefficient (*lower left panel*) and characteristic path length (*lower right panel*) for different numbers of stars

where $P_i \subseteq \{1, \ldots, L\}$ is the set of ideas agent i needs to produce. Production is impossible if even one $l \in P_i \notin H_i$. As initial endowments are random, it will typically be the case that $P_i \subseteq H_i$, so there is room for exchange.

Consider now $j \in V_i$ (equivalently $i \in V_j$, as the graph is non-directed). Denote now $N_i = \{l \in P_i - H_i\}$ the set of ideas that i uses in production but does not have, that is, i's needs. Then i is interested in j provided j has at least an idea that i uses but does not have (that is, if $N_i \cap H_j \neq \{\emptyset\}$).

6.2.2.2 Exchange

As for the "terms of trade", a number of possibilities can be simply explored in the model presented above. Suppose the sequence of events is that each time period an agent is selected and engages in knowledge exchange with one of his neighbours j. In a *gift* transaction, upon request from i, j provides him with an element of $N_i \cap H_j$ and asks nothing in return. This captures a situation in which agents tell each other freely, creating knowledge spillovers. In a *barter* transaction, upon request from i, and provided i has an idea that j needs ($N_j \cap H_i \neq \{\varnothing\}$), an idea is exchanged for another one. This is a trading situation, in which all ideas have a common price, yielding a one-to-one exchange rate. In a mixed economy, some agents give while some trade. This is controlled by $0 \leq \pi \leq 1$ the share of knowledge givers. It is important to note that agents are not interested in all the pieces of knowledge but only in those relevant to their production activity. This implies that at some point trading can stop without every agent holding every idea.

6.3 Numerical Experiment

A natural measure in the productive efficiency of the system is the average output

$$\phi = \sum_{i \in V} \phi_i / n,$$

the proportion of individuals who actually produce. As in this simple environment any agent either achieves production or does not, the ϕs consist of a collection of 0s and 1s and the variance in output is equal to $\phi(1 - \phi)$, which is largest when $\phi = 1/2$ and smallest when ϕ is either 0 or 1.

The settings of the experiment are the following: a population of $n = 600$ agents and $m = 10$ links per agents (hence a total of 6000 edges); each agent is endowed with a set of ideas that is randomly initialized by making each idea l available to i with probability $\Pr\{l \in H_i\} = q$. The production function is a 200 category one, with agents needing on average 20 ideas to produce. Which ideas an agent needs is determined randomly, independently of all other agents: $\Pr\{l \in P_i\} = \theta = 0.1$. All probabilities used to create initial knowledge endowments and production technologies are independent from each other. For each period in the simulation, one agent is

selected, he activates a connection and they trade if possible. This process continues until all possible trades have been made. To examine the space of graphs, using the algorithm described in Sect. 6.2.1 above, we vary the rewiring probability p from 0.001 to 1. For each p-value, 10 different graphs are created and on each graph a single history is run, until all exchange possibilities are exhausted.

The parameters are p, the degree of disorder; s, the number of stars (with lower s-values corresponding to more asymmetry in the link distribution, see Fig. 6.1) and π the share of agents who give rather than trade.

6.4 Results

In this section, we present the results of our experiment. As will be seen, they connect well to some of the empirical results described in the introduction.[6] We start by examining performance at the aggregate level before turning to firm performance. Results are presented under two initial conditions: one of scarcity, wherein initially any agent holds only 15% ($q = 0.15$) of all possible knowledge; and one of abundance, wherein an agent initially holds 85% ($q = 0.85$).

6.4.1 Effect of the Density of Traders

In the presentation of results, we suppress one parameter. The effect of the proportion π of agents who give rather than trade is monotonic throughout. As the number of givers increases, total knowledge levels, or equivalently the number of agents producing, increases for all levels of asymmetry in the degree distribution and for all values of randomness in the network. The cause is clear. The absence of a quid pro quo in exchange when givers are involved implies that when there are many of them more exchanges will take place, and so a larger proportion of firms will find the knowledge they need. In some of the other results which we discuss below, the number of traders changes the strength of the effects we observe, but in every case the direction of the effects remains unchanged. We thus set the proportion, π, of agents who give to 5%, which permits these agents to have a moderate influence on the diffusion dynamics, without swamping other effects.

In the figures given below, we present shaded contour plots wherein darker shades of grey indicate higher values of the variable being examined. The axes in each plot are p, the degree of randomness in network structure, and the Herfindahl concentration index, as a measure of asymmetry in the link distribution. In each figure there are two panels: the left panel shows data from the case of scarce knowledge in the initial condition; the right panel corresponds to the case of abundance.

[6] We do not explicitly test the results against empirical data. One possible route to confront the model would be that described by Werker and Brenner (2004).

6.4.2 Industry Production

A measure of industry performance is the proportion ϕ of firms that have been able to acquire the knowledge they need to produce. We refer to these as producers, and to those that have not acquired the requisite knowledge as non-producers. The relationship between aggregate efficiency, the degree of randomness p and asymmetry, is shown in Fig. 6.2a, b.

The first effect, comparing the two panels, is that not surprisingly, when knowledge is abundant more firms are able to produce. The second effect concerns the number of stars or asymmetry of the link distribution. When knowledge is abundant, as asymmetry increases (or the number of stars falls) efficiency decreases monotonically in a very clear pattern. This is explained as follows. Increasing asymmetry means that links become more concentrated on fewer individuals. If there is a trading agent who has many links, his need for a quid pro quo can block many trades, and close paths between many pairs of agents. At the same time, an agent with many links can quickly find the knowledge he needs. At this point, needing nothing, he has no reason to trade, and so withdraws, thereby closing many paths. This possibility can result in many agents not finding the knowledge they need. Clearly, this possibility recedes when there are fewer agents dominating the linked distribution. When knowledge is scarce, the pattern is not monotonic: efficiency increases and then decreases as asymmetry increases, with the peak at about 400 stars. When asymmetry is at its minimum, however (when every agent is a "star"), efficiency is still much higher than when asymmetry is at its maximum (21 stars). The overall decrease in efficiency has the same explanation as in the previous case. The non-monotonicity of the relationship is explained, however, by a second effect, namely path length. The possibility that knowledge transmission is blocked as explained above increases with the path length between sender and ultimate receiver. Thus, shorter paths reduce this effect, all else equal. Increasing asymmetry thus has

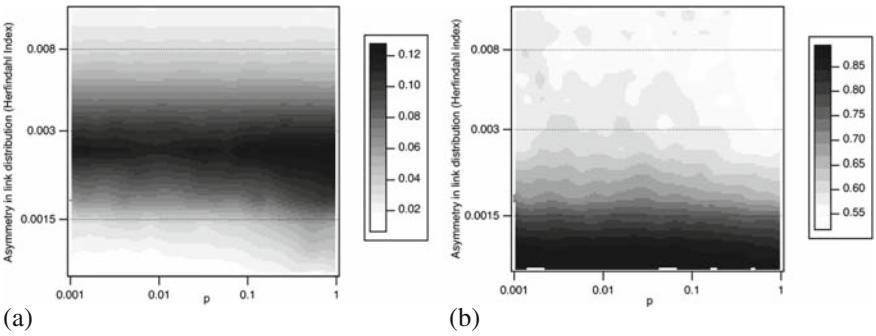


Fig. 6.2 Proportion of firms producing, with scarce knowledge (*left panel*) and abundant knowledge (*right panel*)

two effects working in opposite directions: decreasing path lengths and concentrating links in a few agents. At an intermediate degree of asymmetry, efficiency is maximized.

It is possible to observe an effect of p in Fig. 6.2, though it is of smaller magnitude than the effect of s. Recall that increasing p increases the presence of structural holes, and by the same mechanisms decreases both clustering and path length. What we observe is that when knowledge is scarce, efficiency increases with p; when knowledge is abundant, efficiency decreases with p. In the former case, at the aggregate level social capital is less valuable than are structural holes. Rapid access to distant parts of the network is highly valuable in acquiring knowledge, and at the aggregate level, short path lengths imply complete diffusion of knowledge. In the latter case, with relatively high probability the needed knowledge is close in network space, so a clustered neighbourhood will imply many paths between agents that hold reciprocally desirable knowledge.

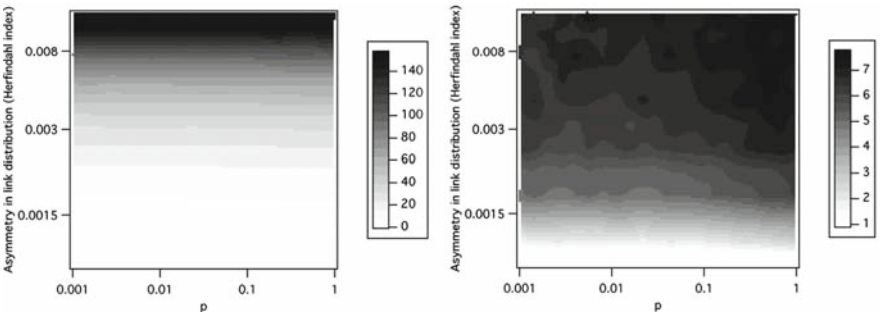
6.4.3 Individual Performance

It is less straightforward to examine performance at the micro-level, as the performance of a firm must be compared to that of other firms in the same context. We are interested here in the effect of the structure of a firm's ego network on its performance.

At the end of each run in our simulation experiments, firms can be partitioned into two groups: those who have accumulated the knowledge they need to produce, and those who have not. The obvious question is whether these two groups are different from each other along interesting network dimensions. We answer this by looking at two structural parameters of each agent's ego network: degree and clustering.

In Fig. 6.3 we show the difference in degree between producers and non-producers, averaged over runs at each point in the parameter space. The figure shows that in terms of acquiring useful knowledge, it is valuable to have many connections – the difference in degree between producers and non-producers is always positive. When knowledge is scarce, this difference is much larger than when knowledge is abundant.[7] In the former case, essentially only stars, having many connections, are able to produce; in the latter, many non-stars are also able to produce. This explains the magnitude of this difference. An agent with many connections rapidly acquires the knowledge he needs, at which point he withdraws from the system (completely if he is a trader, partially if he is a giver). In the worst case this disconnects the network, and in the best case it makes path lengths longer. This can make it impossible for other agents to acquire the knowledge they need. This effect is severe when knowledge is scarce, much less so when knowledge is abundant since many agents will be able to produce from the initial period, and those who cannot are likely to

[7] In both cases all observed differences in means are statistically significant at the 5% level or higher.



(a) (b)

Fig. 6.3 Difference in degree between producers and non-producers

need only a few pieces, which are relatively easy to find. The two structural parameters, asymmetry of the link distribution and randomness of the network both have visible effects when knowledge is common. The value of having many connections increases both as the randomness of the network increases and as asymmetry in the link distribution increases. The latter effect is driven largely by the fact that stars are more likely to produce than non-stars, and as asymmetry increases, stars have more neighbours. The effect of p is driven by clustering. In a barter economy, a clustered graph can alleviate the double coincidence of wants problem, because a transitive triple provides a short path over which two agents can make an indirect trade if they cannot trade directly. Thus, agents with few connections can take advantage of indirect connections to get the knowledge they need. When there are few transitive triples, it is very difficult to overcome a failed double coincidence of wants, so there is a large advantage to having many potential partners, since when trade becomes impossible with one of my neighbours, I can simply turn to another.

Is having a clustered ego network valuable for an agent? Figure 6.4 shows the difference in ego network clustering between those who produce and those who do not, averaged over runs at each point in the parameter space. This difference is always negative – a clustered ego network is bad for information gathering. However, the difference is roughly an order of magnitude larger when knowledge is rare than when it is common.

This is consistent with the structural holes argument given above. Access to knowledge is vital in this economy, and in a clustered neighbourhood, links increase local density rather than connect to distant parts of the network. If an agent is part of a cluster he will, in general, have long path lengths to other agents. If there is a piece of knowledge in a distant part of the network that the agent needs, the longer the path to it, the more likely that on that path is a trader who, because he has all the knowledge he needs, has effectively stopped participating in knowledge transactions. Second, as the networks become less clustered, the negative value to an individual of being in a clustered neighbourhood also decreases, becoming statistically insignificant when there are no explicit stars. As the network itself is less clustered, the average distance from a particular agent to others in the economy falls,

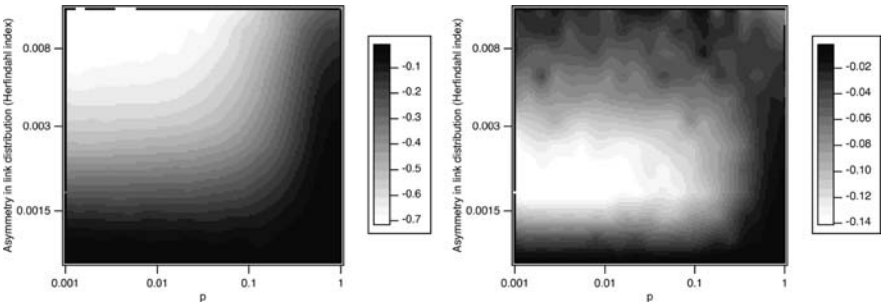


Fig. 6.4 Difference in clustering between producers and non-producers

regardless of the extent of clustering in his ego network. This provides another route by which the effect just described is attenuated. The effects seen in these figures are largely driven by the probability that there is a failure in the path between a firm and the knowledge it needs, but when knowledge is abundant in general, these failure probabilities will be lower in general, and so the effects weaker.

6.4.4 Givers and Traders

Figure 6.5 shows the difference in performance between givers and traders, measured as the difference between the proportion of traders who produce and the proportion of givers who produce. Traders always fare better, and the extent to which this is the case decreases with the asymmetry of the link distribution when knowledge is scarce, again explained by the dominance of stars among the producers. The explanation for the superior performance of traders generally lies in the nature of knowledge interactions. If a trader is involved in an interaction, he always receives desirable knowledge. If a giver is involved in a transaction, he only receives information if he is the originator of the transaction. Thus, the knowledge of traders grows faster than that of givers, so more of them will be able to produce in the long run.

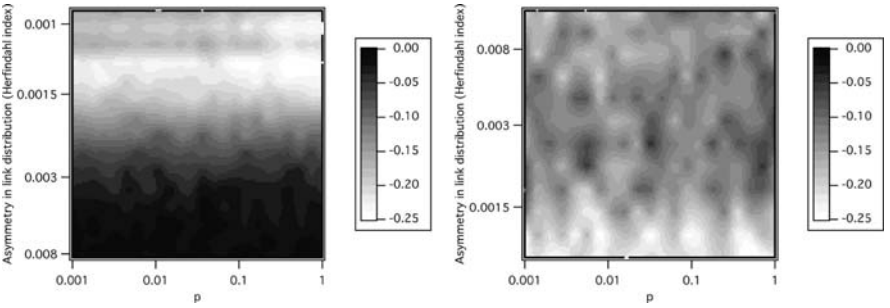


Fig. 6.5 Difference in performance between givers and traders

6.4.5 Particular Neighbours

In the structure of the model there are two types of agents that might have a significant impact on the performance of agents in their neighbourhoods: stars and givers. The effect of having a direct link to a giver is shown in Fig. 6.6. An agent has in his neighbourhood a proportion of givers. In these panels we display the data over the same range. We take the average difference in these proportions between producers and non-producers. In the left panel of Fig. 6.6, the data take values only in [−0.01, +0.01], but are not statistically different from zero, whereas in the right panel all values are above 0.02, and are statistically significant. When knowledge is scarce, having a giving neighbour has no effect; when knowledge is abundant, being close to a giver significantly improves performance.

The effect of having a direct link to a star is shown in Fig. 6.7, using the same measure as above in Fig. 6.6. In both panels of Fig. 6.7, the data are significantly negative. Having a star as a friend is never a good thing as it is almost always a star who creates the hold-up problem by exiting from the trading process. When there are few stars, each having high degree, the problem is strongest. The pattern in both panels essentially tracks the asymmetry of the degree distribution.

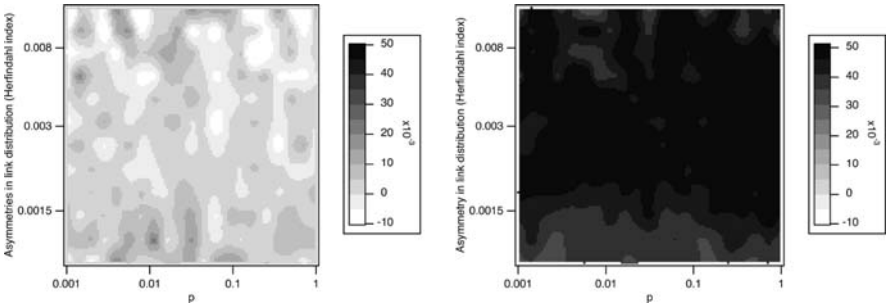


Fig. 6.6 Difference between the proportion of givers in the ego network of producers and non-producers

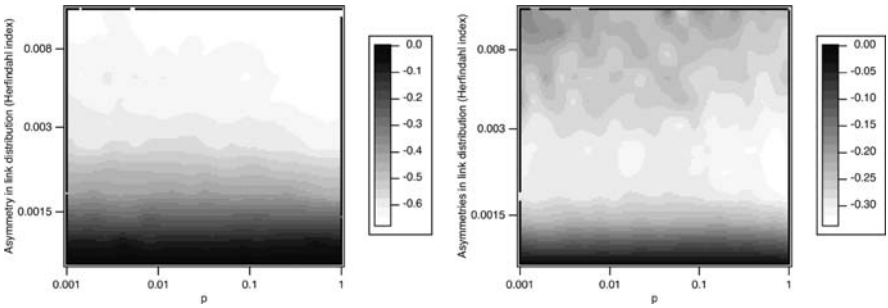


Fig. 6.7 Difference between the proportion of stars in the ego network of producers and non-producers

6.5 Conclusions

We have observed from these results that the presence of "givers" in the economy is generally a good thing. This is intuitively appealing in the first instance, as it seems natural that if agents are giving knowledge away, knowledge flows will be facilitated. Traders function differently. In a networked economy, goods and services, including knowledge flow from one agent to another through a path of intermediate agents. At each step in the path some transaction is made. If it involves a trader, then the transaction must benefit both parties. Thus, if one agent on a path is no longer interested in exchange, in our case, because a trader has all the information he needs and so no longer makes trades, this eliminates any paths which flow through that agent. Thus, the model illustrates that a network with a skewed link distribution, having a few stars, can be either good or bad. If the stars are givers, then their knowledge can flow rapidly out to their many partners, and from there to the rest of the economy. This will continue through the life of the economy. If the stars are traders, because they have many partners, they will rapidly acquire all the knowledge they need, and so stop trading. This blocks many paths between agents, and in the most extreme case, can disconnect the network.

In the introduction, we discussed briefly the debate between the structural hole position and the social capital view. On the former, a non-clustered ego network is good for a firm; on the latter, a densely interconnected neighbourhood provides a good knowledge environment. The model developed here permits some formalization of the resolution suggested by Rowley et al. (2000). In Sect. 6.4.2 we showed that when knowledge was scarce, aggregate production increased with p, whereas when it was abundant, aggregate production decreased with p. Average clustering decreases monotonically with p. Thus, when knowledge is scarce a network with structural holes performs well; when knowledge is abundant a network with high social capital performs well. A situation of scarce knowledge represents a young industry, in which technologies are new, and firms are exploring the technological space to find and create the best possible variant of their products and processes. For any firm, the necessary knowledge is difficult to find, and may reside in distant parts of the economy. Here, redundant links will be less valuable than links that create short paths to other agents. A situation of abundant knowledge may represent a more mature industry. A dominant design has emerged; most firms know what it is and have most of the knowledge needed to execute it. What is happening is that firms are exploiting their versions of the dominant design, and need details rather than new principles. Here, distant parts of the economy will have similar knowledge roughly speaking to local parts of the economy, and what is necessary is to extract the final details. Here, non-redundant links lose their advantage and we see the force of the social capital argument.

One thing that is striking in this regard is that the economy and individual firms have different responses to clustering when knowledge is abundant. At the aggregate level, to recall, when knowledge is abundant output rises as clustering increases, particularly when there are many traders. But at the individual level, being part of a cluster is almost always bad (those who produce have on average less clustered ego

networks than those who do not).[8] The explanation may lie in the fact that when a network is highly clustered on average, a firm with low clustering can connect different parts of the network. If it is the case that firms within a cluster develop similarities, then a firm between two clusters can have access to two different types of knowledge. This provides it with an advantage in terms of finding the knowledge it requires. This appears to be so in our model even though there is no strategic knowledge acquisition or control. It has been suggested by Burt (1992), and also by Baum et al. (2003), that firms that fill structural holes in the network can control information flows. In our model there is no notion of controlling flows, but we do see the first necessary condition for such control, namely access to different knowledge pools.

Efficient knowledge diffusion is the hallmark of a healthy modern economy. What the model developed here shows is that the structures necessary to promote knowledge diffusion depend to a very great extent on the details of the industry. Industries or episodes dominated by collective invention, in which knowledge is (locally) freely given, are very different from industries in which knowledge trading is the norm. But further, how they differ depends on whether knowledge is scarce, as in a newly emerging industry, or abundant, in a mature industry. When knowledge is scarce, random networks always perform well. When knowledge is abundant, and knowledge trading dominates, clustered networks perform best from the social point of view.

Is there a role for policy here? When knowledge is scarce, socially, random, unclustered networks perform well. This performance carries over to the individual firm level: firms with non-clustered ego networks also perform well. So we observe a coincidence between network structures that are socially and individually desirable. The coincidence disappears, though, when knowledge is abundant. Socially clustered networks are efficient, but any firm would prefer a non-clustered ego network – one abundant in structural holes. Here, this divergence between social and private efficiency provides scope for intervention in network formation, at least in principle. What this model has shown, though, is that policy-making in this area is a very delicate business – one size definitely does not fit all. The details about the state of knowledge and the social conventions regarding exchange matter a lot, and so in this area policy must be built on a very strong empirical foundation.

Acknowledgments We thank an anonymous referee for useful comments, and participants in the NEMO project (European Commission FP6, NEST-ADV 028875).

[8] This is obviously a very rough generalization. Exceptions exist, particularly when there are many givers in the economy, or when there are no stars.

References

Ahuja, G., 2000. "Collaboration Networks, Structural Holes and Innovation: A Longitudinal Study", Administrative Science Quarterly, 45, 425–455.

Allen, R. 1983. "Collective invention", Journal of Economic Behavior and Organization, 4,1–24.

Barabási, A. and R. Albert, 1999. "Emergence of Scaling in Random Networks", Science, 286, 509–512.

Baum, J.A.C., A.V. Shipilov, and T.J. Rowley, 2003. "Where Do Small Worlds Come from? ", Industrial and Corporate Change, 12, 697–725.

Breschi, S. and F. Lissoni, forthcoming. "Mobility and Social Networks: Localised Knowledge Spillovers Revisited", Annales d'Economie et de Statistique

Burt, R.S. 1992. Structural Holes: The Social Structure of Competition, Harvard University Press, Cambridge, Mass.

Coleman, J.S., 1988. "Social Capital in the Creation of Human Capital", American Journal of Sociology, 94(S1), S95.

Cowan, R. and N. Jonard, 2003. "The Dynamics of Collective Invention", Journal of Economic Behavior and Organization, 52, 513–532.

Cowan, R. and N. Jonard, 2004. "Network Structure and the Diffusion of Knowledge", Journal of Economic Dynamics and Control, 28, 1557–1575.

Dyer, J.H. and K. Nobeoka, 2000. "Creating and Managing a High Performance Knowledge Sharing Network: The Toyota Case", Strategic Management Journal, 21, 345–367.

Gargiulo, B. and M. Bennassi, 2000. "Trapped in Your Own Net? Network Cohesion, Structural Holes and Adaptation of Social Capital", Organization Science, 11(2), 183–196.

Grandstrand, O. and S. Sjolander, 1990. "Managing Innovation in Multi-technology Corporations", Research Policy, 19, 35–60.

Gulati, R. and M. Gargiulo, 1999. "Where Do Inter-Organizational Networks Come from?", American Journal of Sociology, 104, 1439–1493.

Lamoureaux, N.R. 1999. "Accounting for Capitalism in Early American History: Farmers, Merchants, Manufacturers, and Their Economic Worlds." Technical report, UCLA, 1999. Available at http://www.sscnet.ucla.edu/history/activities/usccpapers/lamoreaux.html.

McGaw, J.A., 1987. Most Wonderful Machine: Mechanization and social Change in Berkshire Paper Making, 1801–1885. Princeton University Press, Princeton, NJ.

Podolny, J.M. 1993. "A Status-Based Model of Market Competition", American Journal of Sociology, 98, 829–872.

Powell, W.W. 1990. "Neither Markets Nor Hierarchies: Network Forms of Organization." In B.M. Shaw and L.L. Cummings (eds.). Research in Organizational Behavior, 12, 395–336.

Powell, W.W., K.W. Koput, and L. Smith-Doerr 1996. "Inter-Organizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology", Administrative Science Quarterly, 41, 116–145.

Powell, W.W., D.R. White, K.W. Koput, and J. Owen-Smith, 2005. "Network Dynamics and Field Evolution: The Growth of Inter-organizational Collaboration in the Life Sciences", American Journal of Sociology, 110, 1132–1205.

Rowley, T., D. Behrens, and D. Krachhardt, 2000. "Redundant Governance Structures: An Analysis of Structural and Relational Embeddedness in the Steel and Semiconductor Industries", Strategic Management Journal, 21, 369–86.

von Hippel, E., 1987. "Cooperation Between Rivals: Informal Know-how Trading", Research Policy, 16, 291–302.

von Hippel, E. 1994. "The Impact of Sticky Data on Innovation and Problem Solving", Management Science, 40, 429–439.

Walker, G., B. Kogut, and W. Shan, 1997. "Social Capital, Structural Holes and the Formation of an Industry Network", Organization Science, 8, 108–125.

Watts, D.J. and S.H. Strogatz, 1998. "Collective dynamics of 'small-world' networks", Nature, 393(6684), 409–10.

Werker, C. and T. Brenner, 2004. "Empirical Calibration of Simulation Models," Max Planck Institute, Jena; Papers on Economics and Evolution number 2004–10.

Part II
Innovation Networks in Complex Theories

Chapter 7
Tools from Statistical Physics for the Analysis of Social Networks

Jörg Reichardt and Stefan Bornholdt

7.1 Introduction

This article is to introduce the social scientist concerned with social network analysis and an affinity to quantitative methods to parts of the research done by physicists in the field of complex networks. In fact, much of the research done by physicists has been inspired by examples and problems from sociology. We believe that the methods developed by natural scientists will prove to be valuable tools that allow new insights into data arising in the social sciences. We hope that these methods find their way back into social sciences and find ample application on the problems by which they were originally inspired.

After a brief revision of basic terminology, we will give a short history of physics research in complex networks that we hope is of interest for a reader not only with a social science background. The focus of this article lies on the analysis of cohesive subgroups or communities in networks, and we will show the relation of this problem with assortative mixing and introduce a quantitative measure of how modular a graph is. The community analysis which we would like to present is based on physical principles, and we introduce some basic concepts of statistical physics important for understanding this approach. We then show how the problem of community detection can be mapped onto known problems from statistical physics. It is equivalent of finding a minimal energy state of a magnetic system. From this insight we can derive a concise definition of communities and we can apply standard methods to find them. We give benchmarks on how this approach performs for both finding all communities in a network as well as finding the community a given node belongs to. Then, we show how hierarchies and overlap of communities arise in networks and how they can be discovered. We show an application of community detection in a large network of scientific co-authorship and the relation of our approach to other algorithms for community detection. Finally, we address the subject of statistical significance and theoretical limits of community detection.

J. Reichardt (✉)
Institute for Theoretical Physics, University of Bremen,
Otto-Hahn-Allee, 28359 Bremen, Germany
e-mail: reichardt@physik.uni-wverzburg.de

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_7,

7.2 Basic Notation

Before we go into the detail of our discussion, we shall introduce a number of important terms and relations to be used throughout the text. The reader already familiar with social network analysis and basic graph theory may skip this section.

Mathematically, a network is represented as a graph $\mathcal{G}(V, E)$, i.e., an object that consists of a set of nodes or vertices V representing the agents in the network and a set E of edges or links or connections representing the interactions or relations between the agents. The cardinality of these sets, i.e., the number of nodes and edges is generally denoted by N and M, respectively. We may assign different values w_{ij} to the links between nodes i and j in E, rendering an edge weighted or otherwise non-weighted ($w_{ij} = 1$ by convention), if we are only interested in the presence or absence of the relation. The number of connections of node i is denoted by its degree k_i. We can represent the set of edges conveniently in an $N \times N$ matrix A_{ij}, called the adjacency matrix. $A_{ij} = w_{ij}$ if an edge between node i and j is present and zero otherwise. Relations may be directed in which case A_{ij} is non-symmetric or undirected in which case A_{ij} is symmetric ($A_{ij} = A_{ji} \quad \forall i, j$). We shall only be concerned with networks in which self links are absent ($A_{ii} = 0 \quad \forall i \in V$). In case of a directed network, A_{ij} denotes an outgoing edge from i to j and A_{ji} an incoming link from j to i. Hence, the outgoing links of node i are found in column j of row i, while the incoming links to i are found in row j of column i. For undirected networks, it is clear that $\sum_{j=1}^{N} A_{ij} = k_i$. For directed networks, $\sum_{j=1} A_{ij} = k_i^{out}$ is the out-degree and equivalently $\sum_{j=1} A_{ji} = k_i^{in}$ is the in-degree of node i. We shall only be concerned with undirected and non-weighted networks, i.e., symmetric adjacency matrices which only have 1 or 0 entries. However, appropriate hints will be given on how to extend the analysis presented to weighted and directed networks. It is understood that, in undirected networks, the sum of degrees of all nodes in the network equals twice the number of edges $\sum_{i=1}^{N} k_i = 2M$. The distribution of the number of connections per node is called degree distribution $P(k)$ and denotes as the probability that a randomly chosen node from the network has degree k. The average degree in the network is denoted $\langle k \rangle$ and we have $N\langle k \rangle = 2M$. We can define a probability $p = 2M/N(N-1) = \langle k \rangle/(N-1)$ as the probability that an edge exists between two randomly chosen nodes from the network.

An (induced) subgraph is a subset of nodes $v \subseteq V$ with n nodes and edges $e \subseteq E$ connecting only the nodes in v. A path is a sequence of nodes, subsequent nodes in the sequence being connected by edges from E. We say, a node i is reachable from node j, if there exists a path from j to i. A subgraph is said to be connected if every node in the subgraph is reachable from every other. A network is generally not connected, but consists of several connected components. We shall restrict our analysis to connected components only, since it can be repeated on every single one of the connected components of a network.

Frequently, we will be concerned with models of random graphs, i.e., models of how a graph arises from a random process. We will need these models for comparison with real world data. When analyzing the structure of real world networks, our null hypothesis shall always be that the link structure is due to chance alone. We may

only reject this null hypothesis, if the link structure we find differs significantly from an expectation value obtained from a random model. Two such models shall be considered here. The first is due to Erdős and Rényi (ER) (1960). They consider random graphs, in which every link between nodes i and j exists with equal probability p. This leads to a Poissonian degree distribution $P(k) = \exp(-\langle k\rangle)\langle k\rangle^k/k!$ with average degree $\langle k\rangle = p(N-1)$. The second model or "configuration model" is generally attributed to Molloy and Reed (1995), who devised an algorithm for constructing actual networks, but it was first introduced by Bender and Canfield (1978). The configuration model assumes a given degree distribution $P(k)$. This means, every node i is assigned a number of stubs k_i according to its degree and then the stubs are connected randomly. For this model, the probability that two randomly chosen nodes are connected by an edge is $p_{ij} = k_i k_j/2M$ and hence proportional to the product of the degrees of the two nodes chosen. The two models make fundamentally different assumptions on the nature of the agents represented by the nodes. In the ER model, fluctuations in the number of connections of a node arise entirely due to chance. In the configuration model, they represent a quality of the node which may be interpreted as the "activity" of the agent represented by the node.

Last, we shall introduce the notion of a "scale free" degree distribution. A scale free degree distribution is characterized by a power law form of $P(k) \propto k^{-\alpha}$ with some positive exponent α. The probability of having k neighbors is inversely proportional to k^α. The reason why this special class of degree distributions is particularly interesting for physicists is that in physics, many systems that have measurable quantities which behave according to a power law fall into different "universality classes" which are characterized entirely by the exponent α. From the same α for two distinct systems, one can generally conclude the same underlying physical principle. The name "scale free" comes from the fact that there is no characteristic value of k. While in ER graphs, the characteristic k is the average degree $\langle k\rangle$, i.e., the average is also a typical k, there is no typical degree in scale free networks.

With these notations and terms in mind, we can now turn to a brief history of physicists research on networks.

7.3 A Physicist's View on Networks

When physicists started to expand their fields of research to networks, naturally, they took a very general approach, sometimes disregarding much of the detailed work done in other more specialized disciplines but always in search of underlying general principles. As more and more networked data became available, physicists soon started to apply the tools and methods from their own discipline in the analysis of the statistical properties of these networks. Sources for these empirical studies encompass networks from a wide scope of subjects and could be as diverse as the web of human sexual contacts, the world wide web, networks of scientific co-authorship, power grids, trophic webs or cellular and gene regulation networks. From these diverse starting points, the term "complex networks" emerged to subsume every system of interacting entities that can be represented as a network and

where the specific pattern of *all* interactions and not necessarily the particularities of the individual interactions themselves are thought to be crucial to the behavior of the system as a whole.

One of the first landmark papers to spark the interest in network analysis was of sociological origin. Stanley Milgram's classic paper on six degrees of separation (1967) showed that even though the agents in a personal acquaintance network had only local knowledge of the topology of the network, they were able to navigate a letter across the USA in this network, even though sender and receiver did not know each other. Furthermore, it could be shown that two randomly chosen individuals in this network of personal acquaintance are connected by a surprisingly short path of only a hand full of intermediary individuals. After this discovery, it were Duncan Watts and Steve Strogatz (1998), who provided the first model of a network that combines the high clustering characteristic for acquaintance networks (many of my friends are friends among each other) and the short average path lengths between randomly chosen individuals already known from the theory of random graphs developed by Erdős and Rényi. At the same time, it retains the fact that there is a typical number of connections or friends per node in the network. The Watts and Strogatz model came to be known as the "small world model" for complex networks.

For years, ER random graphs had been the standard model for networks and a large body of mathematical theory of random graphs had been developed (Bollobas, 2001). In a sense, this is understandable from the following mathematical observation: Drawing from an ensemble of graphs with N nodes and M links, where each graph has the same probability of being drawn, one will draw an ER graph with almost certainty. Or, stated differently, all graphs possible are already encompassed in the ensemble of graphs modeled by ER. Those that have a different degree distribution than a Poissonian are simply very unlikely to be observed. The likelihood of observation of a particular graph with N nodes and M edges is of course highly dependent on the probability of connection between the nodes. With the increasing use of the internet as a source of information and means of communication as well as the increasing availability of large online databases and repositories, more and more differences between real world networks and random graphs were discovered. Most striking was certainly the observation that many real world networks have degree distributions far from Poissonian which rather follow log-normal or power-laws. From these observations, it became clear that the assumption of equal linking probability for all pairs of nodes had to be dropped, and that specific mechanisms had to be sought which explain the link pattern of complex networks from a set of rules. Until now, many such models have been introduced which model networks to an almost arbitrary degree of detail. The starting point for this development was most likely the model by Albert-Lazlo Barabási and Réka Albert (1999). They realized that for many real world networks, two key ingredients are crucial: growth and preferential attachment. Nodes that already have a large number of links are more likely to acquire new ones when nodes are added to the network. These two simple assumptions lead to a network with a scale free degree distribution of exponent $\alpha = 3$ and provided the first model for the in-link distribution of web pages.

Apart from these topological models mainly concerned with the link structure, a large number of researchers is concerned with dynamical processes taking place on networks and the influence the network structure has on them. Among the most widely studied processes is epidemic spreading and one of the most salient results is certainly that by Cohen et al. (2000; Pastor-Satorras and Vespignani, 2001), which showed that for scale free topologies with low clustering, the epidemic threshold (the infectiousness a pathogen needs to infect a significant portion of the network) drops to zero. Liljeros showed that networks of sexual contacts have indeed such a topology (Liljeros et al. 2001). At the same time, these results brought about suggestions for new vaccination techniques such as the vaccination of acquaintances of randomly selected people which allows to vaccinate people with higher numbers of connections with higher efficiency (Pastor-Satorras and Vespignani, 2002). Consequently, a number of researchers are also studying the interplay between topology of the network and dynamic processes on networks in models that allow dynamic rewiring of connections in accordance with, for instance, games being played on the network to gain insights into the origin of cooperation.

The wealth of data from various different disciplines has triggered an enormous amount of data driven research and modeling. All these studies have in common that based on statistical observations on real world data, models are built to reproduce these data. If successful, this approach may give some insight into the underlying microscopic processes that tie, untie or rewire a network and drive the processes taking place on the network. Three excellent and well readable review papers may serve as a summary and starting point into the research of complex networks (Albert and Barabàsi, 2002; Newman, 2003; Dorogovtsev and Mendes, 2005). A more comprehensive overview can be found in (Bornholdt and Schuster, 2003).

In this article, we shall focus on one particular aspect of complex networks research that is of special interest to social network analysis: homophily or the fact that people generally have a tendency of interacting more with alike people than with different ones. This leads to the formation of cohesive subgroups or communities in social networks. The rest of this article will be concerned with the detection of such community structures in networks.

7.4 Community Detection

Experience in our own social environment tells us that often in social networks cohesive subgroups or communities form. All of us have an implicit knowledge which communities we are part of and which not. As a community, we experience groups of actors in social networks that are more densely connected among themselves than with the rest of the network. The reason for such inhomogeneity in the link density is homophily, or the fact that people who are somehow similar, be it by interest, in life style or by any other characteristic, have a greater chance to interact with alike people than with very different people. This effect, however, is not limited to social networks alone. If all the proteins a particular organism produces are tested to find

whether they bind with one another, and a protein interaction network is constructed in which two proteins are linked if they form a stable complex, then, finding groups of densely interconnected proteins may give valuable hints on biologically active protein complexes. The underlying assumption behind both of these examples is that the nodes in the network can be characterized as belonging to different classes. The probability of intra-class connections is believed to be higher than for inter-class connections. The problem is then to assign the nodes into the *a priori* unknown classes. Figure 7.1 illustrates the problem considering the adjacency matrix of a network as example. The problem consists in making the differences in the link density within and between classes visible in the adjacency matrix by reordering rows and columns, such that rows and columns corresponding to nodes from the same class are next to each other.

This presents us with an example of combinatorial data analysis (Arabie and Hubert, 1992). Since the number of possible ways to group the N nodes of the network into an unknown number of classes grows exponentially fast with N, it is generally impossible to compare all possible groupings. One then has to find a clever way to reduce the problem to make it computationally feasible. The first paper from the physics community used again an example from sociology (Girvan and Newman, 2002). Girvan and Newman (GN) took a recursive approach to community detection, dividing the network into two parts, each of which is divided into two parts again. This repeats for all parts until no parts are left to be divided anymore. The procedure results in a complete dendogram representation of the network. The way GN find the partitions in two at each step of the algorithm is inspired by the sociological notion of betweenness. GN defined an "edge betweenness" as the number of shortest paths between all pairs of nodes in the network that run across an edge. After calculating the edge betweenness for every edge in the network, they

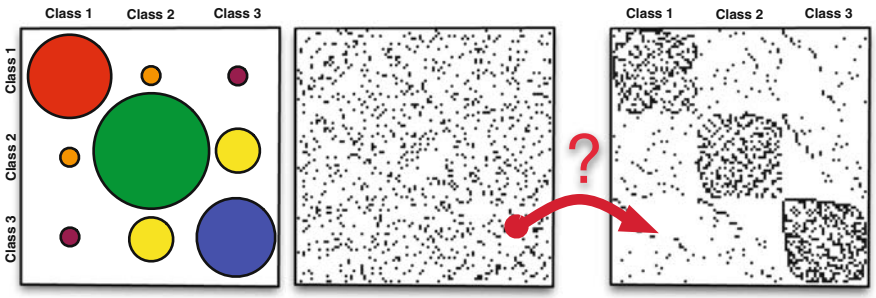


Fig. 7.1 The problem of community detection. Assume that the nodes in the network can be grouped into different classes. The probability of intra-class links is higher than for inter-class links. The *left* part of the figure represents the intra and inter-class link probabilities in the three classes. Larger *circles* correspond to higher probabilities of interaction. The *middle* part represents the adjacency matrix of a corresponding network in a random ordering of the rows and columns. The *right* part shows the same adjacency matrix, but this time with the rows and columns ordered, such that rows and columns corresponding to the nodes in the same class are next to each other. Community detection is about how to find such an ordering

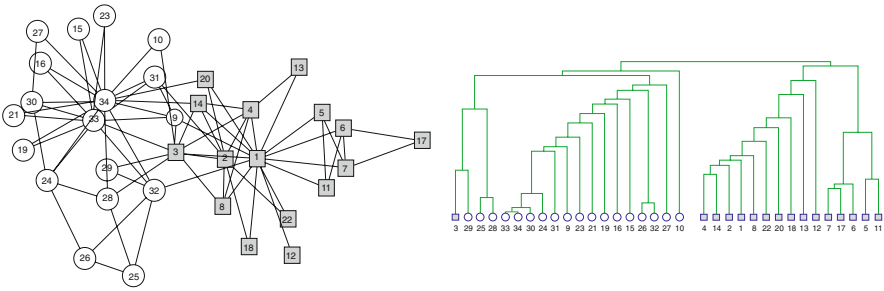


Fig. 7.2 The example network used by Girvan and Newman (GN) (2002). *Left*: The friendship network of a karate club is due to Zachary (1977). Over the course of the observation by Zachary, a dispute between the manager (33) and the instructor (1) of the club led to the break up of the club. The members represented by circles sided with the manager, while those represented by squares sided with the instructor. *Right*: The dendogram resulting from the application of GNs community detection algorithm reproduces the split up of the club almost perfectly. Only node 3 is misclassified. This apparent "predictive power" of the algorithm sparked the interest in community detection among physicists. (From Girvan and Newman, 2002)

remove the edge of largest betweenness from the network. In the now modified network, they repeat the process until the network is split into two disconnected components. It should be noted that the resulting dendogram can also be interpreted as an ordering of the rows and columns in the adjacency matrix corresponding to the nodes at the ends of the dendogram. Figure 7.2 shows the example GN used to demonstrate the power of their algorithm. The sociogram represents the friendship network of the members of a karate club obtained by Zachary who observed the club over a period of two years (Zachary, 1977). During the time, a dispute between the manager and the instructor of the club lead to the break up of the club. A number of members left with the instructor to form a new club. The merit of the GN algorithm is that it reproduces the actual split up of the club almost perfectly given only the sociogram and no further information. It seems as if the partition of the club was pre-determined by the structure of the friendship network.

The anecdote around Zachary's karate club is still a preferred example in introductory talks about community detection. The sociogram was also among the first benchmarks for other algorithms. However, it soon became clear that a more objective criterion is needed for the performance of an algorithm than whether or not it can reproduce the observed split of such a small network. An objective measure by which different assignments of nodes into communities could be compared was needed. The next section will introduce such a measure.

7.5 Assortative Mixing and Modularity

The algorithm of GN transforms the connections of the entire network into a more readable dendogram. Nevertheless, the question remains where to cut this dendogram and what exactly are the cohesive subgroups and communities. What is needed

is an objective criterion at which level the dendogram should be cut or the recursive partitioning stopped, such that all remaining connected components can be regarded as communities. Before we introduce this measure called "modularity" as proposed by GN, we shall first follow them on a small detour into assortative mixing that will clarify the meaning of modularity (Newman and Girvan, 2004; Newman, 2002).

So far, we have interpreted the problem of community detection as that of assigning nodes into previously unknown groups which represent previously unknown classes. Let us step back and assume we already knew the classes each node belongs to. We can then ask, whether our assumption does hold, that links between nodes in the same class are indeed more frequent than links between nodes in different classes. GN define the following quantities: e_{rs} as the fraction of edges that fall between nodes in class r and s. Further, they define $\sum_r e_{rs} = a_s$, the fraction of edges that are connected to at least one node in class s. The fraction a_s can also be interpreted as the empirical probability, that a randomly chosen edges has at least one end in class s. Hence, a_s^2 is the expected fraction of internal edges lying between nodes in class s. Comparing this expectation value to the true value e_{ss} for all groups s leads to the definition of the "assortativity coefficient" r_A:

$$r_A = \frac{\sum_s \left(e_{ss} - a_s^2\right)}{1 - \sum_s a_s^2}. \tag{7.1}$$

This assortativity coefficient r_A is one, if all links fall exclusively between nodes of the same type. The network is then perfectly assortative, but the different classes of nodes remain disconnected. It is zero if $e_{ss} = a_s^2$ for all classes s, i.e., no preference in linkage for either the same or a different class is present. It takes negative values, if edges lie preferably between nodes of different classes, in which case the network is called disassortative. The denominator corresponds to a perfectly assortative network. Hence, r_A can be interpreted as the percentage to which the network is perfectly assortative.

For the classes of the nodes, any measurable quantity may be used. Especially interesting are investigations into assortative mixing by degree, i.e., do nodes predominantly connect to other nodes of similar degree (assortative, $r_a > 0$) or is the opposite the case (disassortative, $r_a < 0$). Surprisingly, it was found that many social networks are assortative, while technological or biological networks are generally disassortative (Newman, 2003). Note that r_A may also be generalized to the case, where the class index s takes continuous values (Newman, 2003).

From the above considerations, it is now easy to define an objective quality measure for an assignment of nodes into communities. GN propose to simply calculate the numerator of the assortativity coefficient for a given community assignment. The higher its value, the better the assignment of nodes into communities. Formally, they define the "modularity" (Newman and Girvan, 2004):

$$\mathcal{Q} = \sum_s \left(e_{ss} - a_s^2\right), \tag{7.2}$$

where e_{ss} is the fraction of edges that fall within community s and a_s is the fraction of edges connected to at least one node in community s, just as before. This modularity measure can be used to compare different community assignments for the same network in an objective way. It also allows to compare the results of different algorithms.

In the following sections, we will see that the ad hoc introduction of $\mathcal{Q}$ can be based on and derived from a very general principle. We will show that the problem of finding a maximally assortative assignment of nodes into communities is equivalent to problems found in solid state physics and computer science, and we will see how insights from these two fields can help to interpret the results of social network analysis.

7.6 Spin Systems, Spin Glasses, and Combinatorial Optimization

In order to understand the approach for community detection developed in the following sections, it is helpful to take a little detour into Statistical Physics. Statistical Physics deals with the behavior of many particle systems and tries to explain the macroscopic behavior of a large ensemble of particles by their microscopic interactions. According to classical Statistical Physics, in thermal equilibrium, i.e., at temperature T, the probability that a system is observed to be in a particular (microscopic) configuration is proportional to the negative exponential of the energy in this configuration

$$p(\text{configuration}) \propto \exp\left(-\beta E(\text{configuration})\right). \tag{7.3}$$

This distribution of probabilities for configurations is called the Boltzmann distribution and deserves some explanation. First, by "configuration" we mean the set of variables that describe the system microscopically, e.g., for a gas this means the position and momentum of every single gas particle. The parameter β only simplifies the writing as it describes the inverse temperature $\beta \propto 1/T$. Large β correspond to small temperatures and small β to large temperatures. With E, we denote the energy of the system in a given configuration. It is important to note how β affects the sensitivity of $p(\text{configuration})$ to changes in E. First, it is clear that configurations with lower energies always have higher probabilities. The ratio of probabilities for two configurations c_1 and c_2 with energies E_1 and E_2 is $p(c_1)/p(c_2) = \exp(-\beta(E_1 - E_2))$. We see that for small enough β or high temperatures, the ratio of probabilities tends to one, i.e., all configurations are approximately equally probable. For large β or low temperatures, however, only those configurations with the smallest energies will have an appreciable probability to be observed in the system. For β tending to infinity or T tending to zero, the system is observed in the configuration with lowest energy with probability one. This state of lowest energy of a physical system is called the ground state.

In principle, we can understand nature as a minimizer from this theory. Physical systems are able to find configurations of minimal energy if they are cooled to

almost zero temperature. Understanding how this is achieved in nature and being able to check theoretical calculations against experiments has enabled physicists to obtain insights into this optimization process and has allowed for the development of methods to calculate, at least approximately, the configuration that minimizes energy. The methods developed in the study of systems that have an actual physical realization can of course also be used in the study of other combinatorial optimization problems, that do not necessarily have an exact analogue in the world of many particle physics. All that is necessary for this, is that we can reformulate the combinatorial optimization problem in a language amenable to methods from Statistical Physics.

To conclude our brief detour into statistical physics, we shall give two examples that illustrate the above. Especially, we will show how the energy of the systems is calculated for a particular configuration. In honor of the irish physicist Sir Willian Rowan Hamilton (1805–1865), the function that calculates the energy for a given configuration is called "Hamiltonian" $\mathcal{H}$ and we shall use this word, whenever we speak of the function and the word energy, whenever we speak of its value. The first is the model of a ferromagnet due to Ernst Ising. Ferromagnets, such as iron, have the property that they are magnetic below a certain critical temperature (the Curie temperature) and non-magnetic above this temperature. Ising assumed that ferromagnetic materials consist of lattices of "elementary magnets." Below the Curie temperature, these can align with their next neighbors in the same direction to add their magnetic fields, while above it, they are in a random orientation due to thermal fluctuations, such that the total magnetic field they produce is zero. Quantum mechanics shows that these "elementary magnets" are actually the manifestations of a quantum mechanical property of elementary particles called "spin" and that they can only have a finite number of orientations. In the Ising model, these spins have only two orientations called spin-up and spin-down. As a configuration, we understand the assignment of all spins (or elementary magnets) into one of the categories spin-up or -down. The Hamiltonian of the Ising ferromagnet can be written as:

$$\mathcal{H}(\{\sigma\}) = -J \sum_{<i,j>} \delta(\sigma_i, \sigma_j). \tag{7.4}$$

Here, σ_i denotes the spin of lattice site i. We simply adopt $\sigma_i = 1$ for spin-up and $\sigma_i = -1$ as spin-down. The function $\delta(\sigma_i, \sigma_j)$ is one, whenever $\sigma_i = \sigma_j$ and zero, whenever $\sigma_i \neq \sigma_j$ for the spins of lattice sites i and j. This function is called Kronecker-Delta. The sum runs over all pairs of neighboring sites in the lattice, denoted by the symbol $< i, j >$. We see that if two neighboring sites have the same spin, they contribute a constant value of $-J$ to the energy. The constant J, with $J > 0$, is called (ferromagnetic) coupling strength, because it determines how two neighboring spins are coupled to affect the total energy. The fact that the energy depends on the values of each of the lattice sites is denoted by the symbol $\{\sigma\}$, which we understand as configuration or one particular assignments of spin values to all lattice sites. From this simple model and the assumption of the Boltzmann

distribution (7.3), one can already derive the qualitative behavior of an actual ferromagnet. From (7.4), it is clear that the minimum energy is reached when all spins have the same value. The ground state is a homogeneous configuration. However, there are two possibilities for the ground state – spins can either be all up or all down, which gives the same energy. We say, the ground state is twofold degenerate. We can immediately see this minimal configuration from the Hamiltonian as the configuration that "satisfies" all couplings, i.e., all Kronecker-Deltas are one and the energy is simply $-J$ times the number of neighboring pairs of lattice sites. Figure 7.3a shows an example of a particular configuration of Ising spins on a lattice.

The second model we would like to consider is called Ising spin glass. The Hamiltonian is formally very similar to (7.4):

$$\mathcal{H}(\{\sigma\}) = -\sum_{<i,j>} J_{ij}\delta(\sigma_i, \sigma_j). \tag{7.5}$$

However, now, the coupling between lattice sites i and j is not a constant anymore. Rather, it is different for every pair i, j. As a simple example, we can assume that $J_{ij} = \pm J$ with equal probability independent of sites i and j. Again, we assume $J > 0$ to be some constant. Now, we can read off neither the ground state configuration nor the ground state energy directly from the Hamiltonian. The reason for this is the appearance of what is called "frustration." We cannot generally satisfy all Kronecker-Deltas due to conflicting couplings. Figure 7.3b shows an example of a part of a lattice with conflicting couplings. The calculation of ground state energies and the finding of ground state configurations now becomes a very difficult problem. However, established methods exist to find good approximative solutions to these problems. In the language of combinatorial optimization problems, the couplings represent local constraints on the variables, which are represented by the spins.

One obvious generalizations of the above models is their extension to more than two spin states, i.e., allowing σ_i to take values from $\{0, 1, 2, 3, \ldots., q\}$. Then, we speak of q-state Potts models instead of Ising models. Also, instead of a lattice topology and interactions only among the next neighbors, we can have arbitrary networks where the nodes carry a spin variable and we assign different couplings

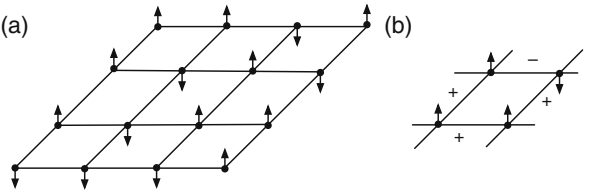


Fig. 7.3 (**a**) Example of an Ising model on a lattice with positive (ferromagnetic) couplings between all neighbors. The spin configuration shown does not lead to a minimal energy through (7.4), but it is clear that aligning all spins in the same direction leads to minimal energy. (**b**) The situation is different if we allow negative (anti-ferromagnetic) couplings. Shown is a "frustrated" plaquette, in which for all possible assignments of spins to lattice sites, at least one edge cannot be satisfied

to the edges. The problem of community detection, as we will show in Section 7.7, is of this spin glass type. Fortunately, we will be able to reuse some known results from general spin glass theory and apply them to the problem.

Now, with the very basics of combinatorial optimization by physical analogy being introduced, we shall return to our main problem of community detection.

7.7 Community Detection using Spin Glasses

7.7.1 Writing a Quality Function as a Hamiltonian

So far, we have not given a definition of what a community actually is. We only have an intuitive idea of a community as being a cohesive subgroup in a network, the members of which are more densely interconnected among themselves than with the rest of the network. We will start from a first principles *ansatz* that builds a quality function of an assignment of nodes into communities. This quality function should have the following properties: It should:

(a) reward internal links between nodes in the same group,
(b) penalize missing internal links,
(c) penalize external links between nodes in different groups, and
(d) reward missing external links.

These principles can be cast into a Hamiltonian similar to (7.5) which then acts as quality function. It is convention to write the Hamiltonian in a way, such that the "better" a particular configuration fulfills these four principles, the lower the energy. This way, the ground state of the Hamiltonian corresponds to the optimal configuration. This approach mapping combinatorial optimization problems to spin systems has been suggested for the first time by Fu and Anderson (1986) in the context of the problem of partitioning a graph into two equally sized parts. We write the following Hamiltonian:

$$\begin{aligned}\mathcal{H}(\{\sigma\}) = &-\sum_{i\neq j} a_{ij}\underbrace{A_{ij}\delta(\sigma_i,\sigma_j)}_{\text{internal links}} + \sum_{i\neq j} b_{ij}\underbrace{(1-A_{ij})\delta(\sigma_i,\sigma_j)}_{\text{internal non-links}} \\ &+\sum_{i\neq j} c_{ij}\underbrace{A_{ij}(1-\delta(\sigma_i,\sigma_j))}_{\text{external links}} - \sum_{i\neq j} d_{ij}\underbrace{(1-A_{ij})(1-\delta(\sigma_i,\sigma_j))}_{\text{external non-links}}.\end{aligned} \tag{7.6}$$

We assign a spin or community index $\sigma_i \in \{1, 2, 3, \ldots, q\}$ to every node i in the network. The matrix A_{ij} denotes the adjacency matrix of the network and a_{ij}, b_{ij}, c_{ij}, and d_{ij} denote the weights of the individual contributions for internal

and external links, respectively. The number of spin states q determines the maximum number of groups allowed and can, in principle, be as large as N, the number of nodes in the network. Note that not all group indices have to be used necessarily in the optimal assignment of nodes into communities, as some spin states may remain unpopulated in the ground state. If links or non-links are weighted equally, regardless whether they are external or internal, i.e., $a_{ij} = c_{ij}$ and $b_{ij} = d_{ij}$, then it is enough to consider the internal links and non-links. It remains to find a sensible choice of weights a_{ij} and b_{ij}, preferably such that the contribution of links and non-links can be adjusted through a parameter. As we will see, a convenient choice is $a_{ij} = 1 - \gamma p_{ij}$ and $b_{ij} = \gamma p_{ij}$, where p_{ij} denotes the probability that a link exists between node i and j. Note that p_{ij} is normalized, such that $\sum_{i<j} p_{ij} = M$. For $\gamma = 1$ this leads to the natural situation that the total amount of energy that can possibly be contributed by links and non-links is equal: $\sum_{i<j} A_{ij} a_{ij} = \sum_{i<j} (1 - A_{ij}) b_{ij}$. The average ratio α of the coupling strengths of present and missing links can then be expressed as:

$$\alpha = \frac{\langle a \rangle}{\langle b \rangle} = \frac{1 - \gamma p}{\gamma p}, \tag{7.7}$$

where p is the average connection probability in the network. For $\gamma = 1$, α equals the inverse ratio of the total number of links and non-links in the network. For any desired α, the corresponding γ is given by $\gamma = 1/p(1 + \alpha)$. Our choice of the weights allows us to further simplify the Hamiltonian (7.7):

$$\mathcal{H}(\{\sigma\}) = -\sum_{i \neq j} \left(A_{ij} - \gamma p_{ij} \right) \delta(\sigma_i, \sigma_j). \tag{7.8}$$

This represents a Potts spin glass with couplings $J_{ij} = A_{ij} - p_{ij}$ between all pairs of nodes: ferromagnetic ($J_{ij} > 0$) where links between nodes exist and antiferromagnetic ($J_{ij} < 0$) where links are missing.

Depending on the graph under study, one can assume different expressions for p_{ij}. Effectively, the Hamiltonian (7.8) is comparing the true distribution of links in the graph with the expected distribution given by a particular null model defined by p_{ij}. With this in mind, we can rewrite (7.8) in the following two ways:

$$\mathcal{H}(\{\sigma\}) = -\sum_{s} \left(m_{ss} - \gamma [m_{ss}]_{p_{ij}} \right) \tag{7.9}$$

and

$$\mathcal{H}(\{\sigma\}) = \sum_{s<r} \left(m_{rs} - \gamma [m_{rs}]_{p_{ij}} \right). \tag{7.10}$$

Here, the sum runs over the q spin states and m_{rs} denotes the number of edges between spins in group r and s. Consequently, the number of internal edges of group s is denoted by m_{ss}. The symbol $[\cdot]_{p_{ij}}$ denotes an expectation value under

the assumption of a link distribution p_{ij}, given the current assignment of spins. That is, $[m_{ss}]_{p_{ij}}$ is the expected number of internal edges in group s given the current assignment of spins, i.e., taking into account the size of group s or the degrees of the nodes in group s. Likewise, $[m_{rs}]_{p_{ij}}$ is the expected number of external edges between groups r and s. Note also that

$$\begin{aligned} 2\sum_s m_{ss} + \sum_{r\neq s} m_{rs} &= 2M, \\ 2\sum_s [m_{ss}]_{p_{ij}} + \sum_{r\neq s} [m_{rs}]_{p_{ij}} &= 2M \end{aligned} \tag{7.11}$$

and that we can always use the number of internal and external edges equivalently.

The number of edges within and between non-overlapping groups is an extensive quantity, i.e., $m_{13} + m_{23} = m_{1+2,3}$ for all choices of non-overlapping groups 1, 2, and 3 and $m_{33} = m_{11} + m_{22} + m_{12}$ for all groups 3 with proper subgroups 1 and 2 of empty intersection and union 3. We therefore require that the model of connection probability ensures the same for the expectation values: $[m_{13}]_{p_{ij}} + [m_{23}]_{p_{ij}} = [m_{1+2,3}]_{p_{ij}}$ and $[m_{33}]_{p_{ij}} = [m_{11}]_{p_{ij}} + [m_{22}]_{p_{ij}} + [m_{12}]_{p_{ij}}$.

Two exemplary choices of link distribution models p_{ij} shall illustrate the above. The simplest choice is to assume every link equally probable with probability $p_{ij} = p$ which leads naturally to

$$[m_{ss}]_p = p\frac{n_s(n_s-1)}{2} \quad \text{and} \quad [m_{rs}]_p = pn_r n_s, \tag{7.12}$$

with n_r and n_s denoting the number of spins in state r and s, respectively. This choice of model leads to the Hamiltonian (Reichardt and Bornholdt, 2004):

$$\mathcal{H}(\{\sigma\}) = -\sum_{i,j\in E} \delta(\sigma_i, \sigma_j) + \gamma p \sum_s^q \frac{n_s(n_s-1)}{2}. \tag{7.13}$$

Here, the first sum runs over all edges and only internal edges contribute. Equivalently, we can write (7.13) in terms of external edges:

$$\mathcal{H}(\{\sigma\}) = \sum_{i,j\in E} (1 - \delta(\sigma_i, \sigma_j)) - \gamma p \sum_{r<s}^q n_r n_s, \tag{7.14}$$

where only edges between different groups contribute to the first sum. We see that both, (7.13) and (7.14), compare the actual value of internal or external edges with its respective expectation value under the assumption of equally probable links and given community sizes.

A second choice for p_{ij} may take into account that the network exhibits a particular degree distribution. Since links are in principle more probable between nodes of high degree, links between these nodes should get a lower weight. Using

$$p_{ij} = \frac{k_i k_j}{2M} \tag{7.15}$$

takes this fact and the degree distribution into account. Note that it is possible to also include degree–degree correlations or any other form of prior knowledge about p_{ij} at this point. With these expressions we can write:

$$[m_{ss}]_{p_{ij}} = \frac{1}{2M}\frac{K_s^2}{2} \text{ and } [m_{rs}]_{p_{ij}} = \frac{1}{2M} K_r K_s. \tag{7.16}$$

Here, K_s is the sum of degrees of nodes in spin state s and plays the role of the occupation numbers in Eq. (7.13). Using these expressions, we can also write the Hamiltonian (7.8) in a form similar to (7.13):

$$\mathcal{H}(\{\sigma\}) = -\sum_{i,j\in E} \delta(\sigma_i, \sigma_j) + \frac{\gamma}{2M}\sum_{s}^{q} \frac{K_s^2}{2}. \tag{7.17}$$

Again, we give an equivalent formulation in terms of external, rather than internal edges, similar to (7.14):

$$\mathcal{H}(\{\sigma\}) = \sum_{i,j\in E} (1 - \delta(\sigma_i, \sigma_j)) - \frac{\gamma}{2M}\sum_{r<s}^{q} K_r K_s. \tag{7.18}$$

We see that even though we are dealing with a so-called "infinite range" spin glass, i.e., couplings between all pairs of nodes exist, it is only necessary to consider the interactions J_{ij} along the links A_{ij} of the network and the occupation numbers or the sum of node degrees of the individual spin states. This makes it easy to implement an efficient minimization routine for these Hamiltonians. It should be noted that both the formulations (7.13), (7.14) and (7.17), (7.18) are equivalent in case of a network with fixed connectivity, where all nodes have the same degree.

7.7.2 Equivalence with Newman Modularity

We have already introduced the modularity $\mathcal{Q}$ as an objective measure of how assortative or disassortative the links are distributed within and in between groups of nodes. It so happens that for our choice of weights a_{ij} and b_{ij} and $\gamma = 1$, the modularity $\mathcal{Q}$ is equivalent to our Hamiltonian. We can show this by writing the modularity in a slightly different way in the following manner (Clauset et al., 2004):

$$
\begin{aligned}
e_{ss} &= \frac{1}{2M} \sum_{i \neq j} A_{ij} \delta(\sigma_i, \sigma_j) \delta(\sigma_i, s) \\
a_s^2 &= \frac{1}{4M^2} \left(\sum_i k_i \delta(\sigma_i, s) \right)^2 \\
Q &= \frac{1}{2M} \sum_{i \neq j} \left(A_{ij} - \frac{k_i k_j}{2M} \right) \delta(\sigma_i, \sigma_j).
\end{aligned} \tag{7.19}
$$

This already resembles (7.8) when p_{ij} takes the form $k_i k_j / 2M$ and $\gamma = 1$. It is now clear that we can write:

$$
Q = -\frac{1}{M} \mathcal{H}(\{\sigma\}) \tag{7.20}
$$

with the Hamiltonian (7.8) and $\gamma = 1$. Therefore, maximum modularity is reached, when the Hamiltonian (7.8) with $p_{ij} = k_i k_j / 2M$ or equivalently (7.17) or (7.18) with $\gamma = 1$ are minimal. To maximize the modularity of a community structure is hence equivalent to finding the spin configuration that minimizes these Hamiltonians. This reformulation makes the problem amenable to efficient optimization routines.

7.7.3 Cohesion and Adhesion

From the above considerations and to simplify further developments, we will adopt the concept of "cohesion" and "adhesion" to our network terminology. We define the coefficient of cohesion:

$$
c_{ss} = m_{ss} - \gamma [m_{ss}]_{p_{ij}} \tag{7.21}
$$

as the difference between the number of internal links between nodes in group s and the respective expectation value. We see that c_{ss} is positive for groups of nodes that have more inner links than expected and negative for groups of nodes that have less inner links than expected. Equivalently, we define the coefficient of adhesion:

$$
a_{rs} = m_{rs} - \gamma [m_{rs}]_{p_{ij}} \tag{7.22}
$$

between nodes in group r and s with $r \neq s$. Note that a_{rs} is negative, when there are less links between group r and s than expected. Remember that both c_{ss} and a_{rs} depend on the global parameter γ and the assumed connection probability p_{ij}. For $\gamma = 1$ and the model $p_{ij} = k_i k_j / 2M$, we find from combining the Eq. (7.11)

$$2c_{ss} + \sum_{r,r\neq s} a_{sr} = 0. \tag{7.23}$$

For $p_{ij} = k_i k_j / 2M$ the cohesion is negative ($c_{ss} < 0$) if n_s consists of only one single node. We see that there must always exist a group of nodes n_r, to which this node has positive adhesion. We will see later that this means every node belongs to some group. Groups of only one node do not exist. We stress that relation (7.23) and the conclusions just drawn do not hold for $\gamma \neq 1$ or $p_{ij} = p$. Furthermore, we give a relation for the coefficient of cohesion of a group of nodes n_s and two proper subsets n_{s1} and n_{s2} with empty intersection and union n_s. It is easy to prove that

$$c_{ss} = c_{11} + c_{22} + a_{12}, \tag{7.24}$$

where c_{11} and c_{22} are the coefficients of cohesion of the respective subsets n_{s1} and n_{s2}, and a_{12} is the coefficient of adhesion between n_{s1} and n_{s2}. Equivalently, we can write for the adhesion coefficients with n_2 of two groups n_{r1} and n_{r2} with union n_r and empty intersection

$$a_{rs} = a_{1s} + a_{2s}. \tag{7.25}$$

For the last time, we shall give a reformulation of the Hamiltonian (7.8) in terms of the coefficients of cohesion and adhesion:

$$\mathcal{H}(\{\sigma\}) = -\sum_{s} c_{ss} = \sum_{r<s} a_{rs}. \tag{7.26}$$

So far, we have done nothing more than formulating the problem of community detection in mathematical terms and have gained some insight by reformulating it. All the different expressions we have given so far are nothing more and nothing less than a mathematical formulation of the four principles we have stated at the beginning of this section.

7.7.4 Extension to Weighted and Directed Networks

The above Hamiltonians and expressions for modularity can easily be extended by using a weighted instead of the 0, 1 adjacency matrix A_{ij}. The degrees of the nodes k_i are then calculated as $k_i = \sum_j A_{ij}$ and the average connection probability $p = \langle k \rangle / (N-1)$. The number of edges is then to be replaced by $M = 1/2 \sum_i k_i$, the total sum of link weights. Due to these simple relations, we will not differentiate between weighted and unweighted. In the same way, we can deal with directed networks by dropping the assumption of the symmetry of A_{ij}, p_{ij} and a_{rs}. The situation becomes more difficult when interpreting the results, since group r may have a strong adhesion with s, such that $a_{rs} > 0$, but not vice versa, such that $a_{sr} < 0$. The decision, whether to group r and s together then depends on whether

it is globally better to group them together or apart, i.e., based on the comparison of the absolute values of a_{rs} and a_{sr}. From this, we see that we can form a symmetric adjacency matrix $\tilde{A}_{ij}$ from the adjacency matrix of the directed network A_{ij} by writing: $\tilde{A}_{ij} = 1/2(A_{ij} + A_{ij})$ and a symmetric link probability $\tilde{p}_{ij}$ from the asymmetric of the directed network by writing $\tilde{p}_{ij} = 1/2(p_{ij} + p_{ji})$. After these transformations, we can work with the network in the same way as with an undirected network.

7.7.5 *Properties of the Hamiltonian and its Ground State*

Thus far, we have seen that the ground state spin configuration of our Hamiltonians can be interpreted as *the* community structure of a network. In other words, it is the "best" assignment of nodes into communities according to the quality function defined by the Hamiltonian. The lower the energy, the "better" the community structure. From the fact that the ground state is a configuration that is a minimum in the configuration space, we can derive a number of properties of the communities that apply to any local minimum of the Hamiltonian in the configuration space. If we take these properties as *defining* properties of what a community is, we then find valid alternative community structures also in the local minima of the Hamiltonian. The energy of these local minima will then allow us to compare these community structures. It may be that alternative but almost equally "good" community structures exist. Before we go on investigating the properties of spin configurations that represent local minima of the Hamiltonian, we will discuss a few properties of the Hamiltonians as such:

First, we note that for $\gamma = 1$ the Hamiltonians evaluate to zero in case of assigning all nodes into the same spin state due to the normalization constraint on p_{ij}, i.e., $\sum_{i<j} p_{ij} = M$, independent of the graph. Second, for a complete graph, any spin configuration gives the same zero energy at $\gamma = 1$. Third, for a graph without edges, e.g., only a set of nodes, any spin configuration gives zero energy independent of γ. Fourth, the expectation value of the Hamiltonians for a random assignment of spins at $\gamma = 1$ is zero. These considerations give an intuitive feeling for the fact that the lower the energy, the better the spin configuration is suited for a particular graph and that the choice of $\gamma = 1$ will result in what could be called a "natural partitioning" of the graph.

Since the Hamiltonians are all additive with respect to the different communities, i.e., the numbers of edges and the corresponding expectation values are extensive, they can be seen as independent entities and we can treat a single community independently from the rest of the network. The configuration space over which the Hamiltonian is minimized is a discrete space. Once we have defined a move set that is ergodic in this discrete space, a (local) minimum of the Hamiltonian (with respect to this move set) is defined as a configuration for which none of the steps from the move set leads to a lower energy. It is sufficient to consider only one move: change a group of nodes n_1 from spin state s to spin state r. The change in energy for this move in configuration space is:

$$\Delta\mathcal{H} = a_{1,s\setminus 1} - a_{1r}. \tag{7.27}$$

Here $a_{1,s\setminus 1}$ is the adhesion of n_1 with its complement in n_s and a_{1r} is the adhesion of n_1 with n_r. It is clear that if we move n_1 to a previously unpopulated spin state, then $\Delta\mathcal{H} = a_{1,s\setminus 1}$. This move corresponds to dividing group n_s. Furthermore, if $n_1 = n_s$, we have $\Delta\mathcal{H} = -a_{sr}$, which corresponds to joining groups n_s and n_r. For a spin configuration to be a local minimum of the Hamiltonian, there must not exist a move of this type that leads to a lower energy. It is clear that some moves may not change the energy and are hence called neutral moves. In case of equality $a_{1,s\setminus 1} = a_{1,r}$ and n_r being a community itself, we say that communities n_s and n_r have an overlap of the nodes in n_1.

For a community defined as a group of nodes with the same spin state in a spin configuration that makes the Hamiltonian minimal, we then have the following properties:

1. Every proper subset of a community has a maximum coefficient of adhesion with its complement in the community compared to the coefficient of adhesion with any other community ($a_{1,s\setminus 1} = \max$).
2. The coefficient of cohesion is non-negative for all communities ($c_{ss} \geq 0$).
3. The coefficient of adhesion between any two communities is non-positive ($a_{rs} \leq 0$).

The first property is proven by contradiction from the fact that we are dealing with a spin configuration that makes the Hamiltonian minimal. We also see immediately that every proper subset n_1 of a community n_s must have a non-negative adhesion with its complement $n_{s\setminus 1}$ in the community. This is especially true for every single node l in n_s ($a_{l,s\setminus l} \geq 0$). Then we can write $\sum_{l\in n_s} a_{l,s\setminus l} \geq 0$. Since $\sum_{l\in n_s} m_{l,s\setminus l} = 2m_{ss}$ and $\sum_{l\in n_s}[m_{l,s\setminus l}]_{p_{ij}} = 2[m_{ss}]_{p_{ij}}$, this implies $c_{ss} \geq 0$ for all communities s and proves the second property. The third property is proven by contradiction again. Again, we stress that for $\gamma = 1$ and $p_{ij} = k_i k_j/2M$, no community is formed of a single node due to condition (7.23). The last two properties can be summarized in the following inequality which provides an intuition about the significance of the parameter γ:

$$c_{ss} \geq 0 \geq a_{rs} \qquad \forall r, s. \tag{7.28}$$

Assuming a constant link probability, we can rewrite this inequality in order to relate the inner link density of a community and the outer link density between communities with an average link density:

$$\frac{2m_{ss}}{n_s(n_s-1)} \geq \gamma p \geq \frac{m_{rs}}{n_r n_s} \qquad \forall r, s. \tag{7.29}$$

We see that γp can be interpreted as a threshold between inner and outer link density under the assumption of a constant link probability.

Apart from giving an interpretation of the (local) minima of the Hamiltonian, the above properties also give a *definition* of what a community is, alternative to that of a set of nodes of equal spin value in a configuration that represents a minimum

of the Hamiltonian. When speaking of the community structure of a network, we generally refer to that obtained at lowest energy, i.e., in the ground state. We will use the term "community" denoting a subset of nodes that has all of the above properties. Note that this definition of community adapts itself naturally to different classes of networks, since a model p_{ij} is included in the definition of adhesion and cohesion. Since the assignment of nodes into communities changes with the value of γ, we shall further adopt the notion "community at level γ," in order to characterize possible hierarchies in the community structure.

7.7.6 Relation to Other Definitions of Communities

Let us put the above definition in perspective with other community definitions from the sociology literature. Wasserman and Faust (1994) give a comprehensive overview of possible definitions. Starting from the graph theoretical concepts of cliques as maximal complete subgraph, they go on to definitions based in reachability and diameter such as n-cliques, n-clans, or n-clubs. Other definitions given are based on nodal degree such as that of k-plexes and k-cores. All of these definitions have in common that they define a community or cohesive subgroup from the point of view of a single node. In contrast, one may also define cohesive subgroups on the level of the group as is done for lambda sets and LS-sets. Our definition of community falls into the latter class. It is closest to that of an LS set, which shall be repeated here: A set of nodes S in a network is an LS set, if each of its proper subsets has more links to its complement within S than to the outside of S. In our case, we have shown that for the communities we find every proper subset has a maximum coefficient of adhesion with its complement in the community. In this respect, our community definition is closest to that of an LS set.

7.7.7 Benchmarking Performance

With the above considerations, we have justified our approach theoretically to the point that the ground state has a number of properties, which one would intuitively associate with the notion of cohesive subgroup or community. Skipping the technical details of how to find the spin configuration that minimizes the energy, we shall directly present the benchmark results for our method on computer generated test networks with known community structure as suggested in (Newman, 2004). Nodes are assigned to communities and are randomly connected to members of the same community by an average of $\langle k_{in} \rangle$ and to members of different communities by an average of $\langle k_{out} \rangle$ links. Fixing the average degree of all nodes to $\langle k \rangle = \langle k_{in} \rangle + \langle k_{out} \rangle = 16$, it becomes more and more difficult for any algorithm to detect the communities as $\langle k_{out} \rangle$ increases on the expense of $\langle k_{in} \rangle$. When quantifying the performance of the algorithm, we define two measures. Sensitivity measures the percentage of all pairs of nodes that are correctly grouped together in the ground

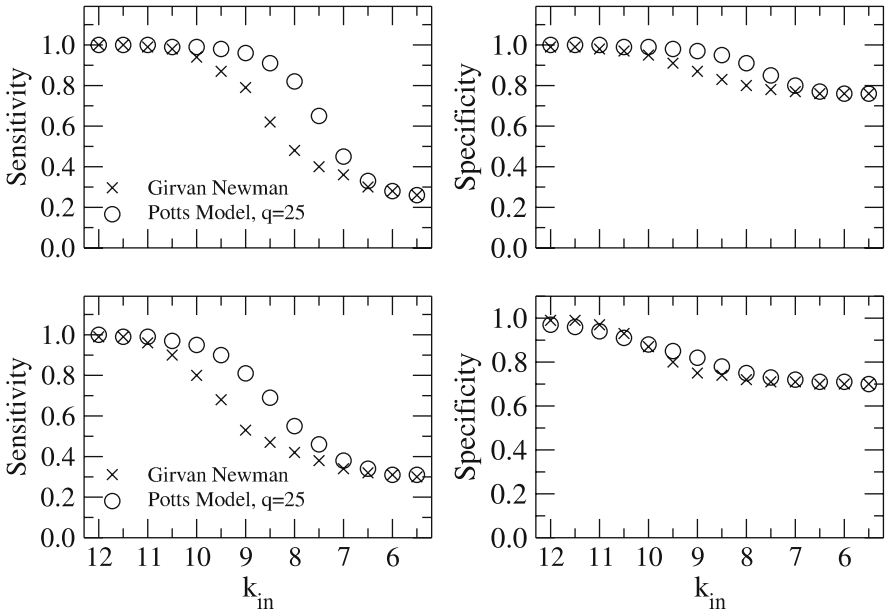


Fig. 7.4 Benchmark of the algorithm for networks with known community structure and comparison with Girvan and Newman. *Top row*: four communities of 32 nodes each, *bottom row*: four communities of 128, 96, 64, and 32 nodes, respectively. Symbol size corresponds to error bars

state. Specificity measures the percentage of all pairs of nodes that are correctly grouped apart in the ground state. Hence, sensitivity (specificity) is a measure of how good an algorithm performs in grouping nodes together (apart) that (do not) belong together. We tested two sets of networks. The first is composed of four equally sized communities of 32 nodes each and the second is composed of four communities of 128, 96, 64, and 32 nodes, respectively. Performance of our algorithm and, for comparison, the one by GN (Girvan and Newman, 2002) is shown in Fig. 7.4. Note the high sensitivity and specificity of our algorithm for both types of networks. We stress that the benchmarking performance does not depend on q, the number of possible spin states as long as it is sufficiently large (Reichardt and Bornholdt, 2007). The analysis was run with $\gamma = 1$.

7.7.8 Finding the Community Around a Given Node

So far in our analysis, we have partitioned the network into communities, such that every node belongs to exactly one community. Often, it is desirable not to find all communities in a network, but to find only the community to which a particular node belongs. This may be especially useful, if the network is very large and detecting all communities may be time consuming. In the framework presented in this article, we can do this using a fast, greedy algorithm. Starting from the node j we are

interested in, we successively add nodes as long as the adhesion of the community we are forming and the rest of the network decreases. Adding a node i from the rest of the network r to the community s around the start node, the adhesion between s and r changes by:

$$\Delta a_{sr}(i \to s) = a_{ir} - a_{is}. \tag{7.30}$$

For $p_{ij} = p$, this can be written as:

$$\Delta a_{sr}(i \to s) = k_{ir} - k_{is} - \gamma p(n_r - 1 - n_s), \tag{7.31}$$

where $n_r = N - n_s$ is the number of nodes in the rest of the network, and n_s the number of nodes in the community. For $p_{ij} = k_i k_j / 2M$, the change in adhesion reads:

$$\Delta a_{sr}(i \to s) = k_{ir} - k_{is} - \frac{\gamma}{2M} k_i \left(K_r - k_i - K_s\right). \tag{7.32}$$

Here, K_r and K_s are the sums of degrees of the rest of the network and the community under study, respectively, and k_i is the degree of node i to be moved from r to s, which has k_{is} links connecting it with s and k_{ir} links connecting it with the rest of the network. It is understood that only when the adhesion of i with s is larger than with r, the total adhesion of s with r decreases. Equivalent expressions can be found for removing a node i from the community s and rejoining it with r.

For $\gamma = 1$ and $p_{ij} = k_i k_j / 2M$, we have $a_{is} + a_{ir} + 2c_{ii} = 0$, and $c_{ii} < 0$ by definition and close to zero for all practical cases. Then, a_{is} and a_{ir} are either both positive and very small or have opposite sign. Choosing the node that gives the smallest Δa_{rs} will then result in adding a node with positive coefficient of adhesion to s. It is easy to see that this ensures a positive coefficient of cohesion in the set of nodes around j.

In order to benchmark the performance of this approach, we applied it again to computer generated test networks as done in the last section. As before, we used computer generated test networks of 128 nodes, which are grouped into four equal sized communities of size 32 with average degree $\langle k \rangle = \langle k_{in} \rangle + \langle k_{out} \rangle = 16$. Starting from a particular node, we are interested in the performance of the algorithm in discovering the community around it. We measure the percentage of nodes that are correctly identified as belonging to the community of the start node as sensitivity and the percentage of nodes that are correctly identified as *not* belonging to the community as specificity.

Figure 7.5 shows the results obtained for different values of $\langle k_{in} \rangle$ at $\gamma = 1$ and using $p_{ij} = k_i k_j / 2M$. We note that this approach performs rather well for a large range of $\langle k_{in} \rangle$ with good sensitivity and specificity. In contrast to the benchmarks for partitioning the entire network, we obtain a sensitivity that is generally larger than the specificity. This shows that partitioning the entire network tends to mistakenly group things apart, that do not belong apart by design, while constructing the community around a given node, tends to group things together, that do not belong

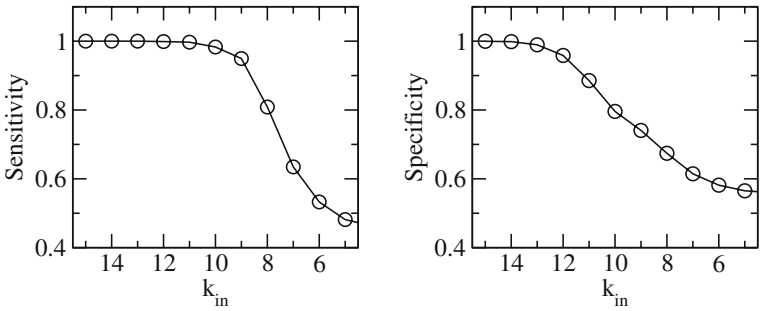


Fig. 7.5 Benchmark of the algorithm for discovering the community around a given node in networks with known community structure. We used networks of 128 nodes and four equal sized communities. The average degree of the nodes was fixed to 16, while the average number of intra-community links $\langle k_{in} \rangle$ was varied. Sensitivity measures the fraction of nodes correctly assigned to the community around the start node, while specificity measures the fraction of nodes correctly kept out of the community around the start node

together by design. This behavior is understandable, since working on the entire network amounts to effectively implementing a divisive method, while starting from a single node means implementing an agglomerative method.

7.7.9 Overlap, Stability, and Hierarchy of Community Assignments

Lastly, we turn to the problem of overlap of communities and alternative community structures, which has briefly been touched upon in the last section. One cannot generally assume that a community structure of a network is uniquely defined. There may exist several but very different partitions that all have a very high value of modularity. Palla et al. (2005) have introduced an algorithm to detect overlapping communities by clique percolation and Gfeller et al. have introduced the notion of nodes lying "between clusters" (2005). This uncertainty in the definition of the borders of a community is expected from the spin glass nature of the Hamiltonian, where generally many energy minima may exist that are comparably deep corresponding to comparably good assignments of nodes into communities. Additionally, the (local) minima of the Hamiltonian may be degenerate. The overlap of communities is linked to degeneracy of the minima of the Hamiltonian. Since the degeneracy can arise in several ways, we have to differentiate between two different types of overlap: overlap of community structure and overlap of communities.

We have already seen that it is undecidable if a group of nodes n_t should be member of community n_s or n_r, if the coefficients of adhesion are equal for both of these communities. Formally, we find $a_{t,s\setminus t} = a_{tr}$. In this situation, we speak of overlapping communities n_s and n_r with overlap n_t, since the number of communities in the network is not affected by this type of degeneracy. Nodes that do not form part of overlaps will always be grouped together and can be seen as the non-overlapping cores of communities. An example of this can be found in Fig. 7.6(a),

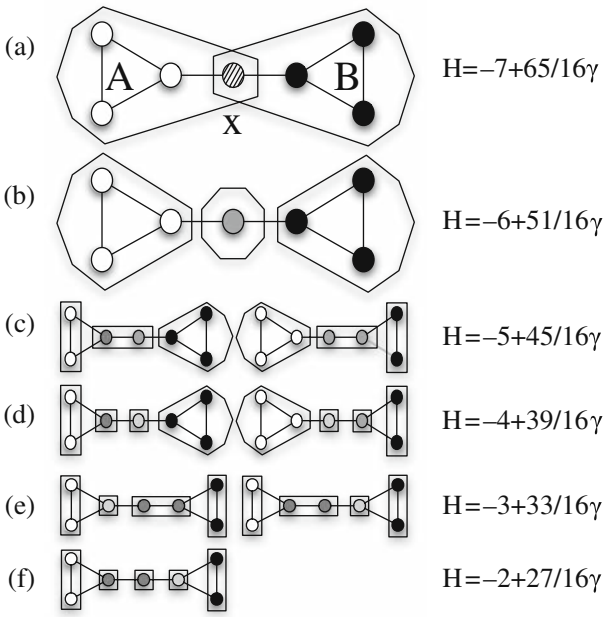


Fig. 7.6 For different values of γ, different spin configurations minimize the energy to form ground states. For $\gamma < 16/63$, the ground state is ferromagnetic. For $16/63 < \gamma < 8/7$, the twofold degenerate configuration (**a**) is the ground state, with node x belonging either to community A or B. For $8/7 < \gamma < 8/3$, configuration (**b**) shows the non-degenerate ground state. For $\gamma >= 8/3$, configurations (**b**), (**c**), (**d**), (**e**) and (**f**) all form ground states, but only (**f**) is ground state for $8/3 < \gamma < 4$

where communities A and B overlap in node x. The ground state at $\gamma = 1$ is twofold degenerate with node x belonging either to A or B.

On the other hand, it may be undecidable, if two groups of nodes should be grouped together or apart, if the coefficient of adhesion between them is zero, i.e., there exist as many edges between them as expected from the model p_{ij}. Similarly, it may be undecidable, if a group of nodes should form its own community or be divided and the parts joined with different communities, if this can be done without increasing the energy. In these situations, the number of communities in the ground state is not well defined and we cannot speak of overlapping communities, since communities do not share nodes in the degenerate realizations. We will hence refer to such a situation as overlapping community structures. An example of this can be found in Fig. 7.6(d), where the three nodes in groups A and B form either one community as in (a) or two distinct communities of 2 and 1 node each. In general, however, both types of overlap may be present in a network.

Since the coefficients of adhesion and cohesion depend on the value of γ chosen, one can assess the stability of community structures under the change of this parameter. The network shown in Fig. 7.6 illustrates this. It is made of seven nodes and eight links. The sum of degrees of the nodes in A and B is 7, each. For $\gamma = 1$,

we find the maximum modularity $Q = 47/128$. For values of $\gamma < 16/63$, we find a ferromagnetic ground state with all edges satisfied. At $\gamma = 16/63$, both the ferromagnetic and the two community state depicted in (a) are ground states. The twofold degenerate state of (a) will remain the ground state while $\gamma < 8/7$. For values of $8/7 < \gamma < 8/3$, we find the non-degenerate ground state shown in (b). Node x now forms its own community, while the nodes in A and B are still grouped together. This shows that node x is not as stably connected to the three remaining nodes in A and B as those are among themselves. At $\gamma = 8/3$, the ground state is eightfold degenerate and the configurations of (b), (c), (d), (e) and (f) all have the same minimal energy of $\mathcal{H}_{GS} = -5/2$. For $8/3 < \gamma < 4$, the configurations shown in (f) show the non-degenerate ground state.

The Hamiltonian is always constructed as a sum of a reward and a penalty term. The parameter γ controls the influence of this penalty term. We note that in order to go from a homogeneous configuration of high reward (many internal edges) to a more diverse configuration of lower reward term (less internal edges) by increasing γ, the penalty term also has to decrease. Otherwise, we cannot find a positive γ that favors the more diverse configuration. As illustrated in Fig. 7.6, the penalty term has to decrease as the number of satisfied edges decreases. A configuration that has less satisfied edges and larger penalty at the same time can never be a minimal configuration for any value of γ.

In order to study these questions more generally, we need to find a convenient way to characterize and compare different community structures. Different structures are best represented in a symmetric $N \times N$ co-appearance matrix, in which the entry i, j denotes in what fraction of all structures analyzed nodes i and j were grouped together. This allows the combined representation of the overlap of many different community structures necessary for the investigation of degenerate ground states. In addition, this type of representation allows the comparison of community structures obtained from a parameter variation of γ also for large networks and hence to study possible hierarchies and the stability of community structures for different values of γ when possible degenerate ground states can only be sampled in a stochastic manner.

We have already stressed that properties 1 through 3 are also valid for any local minimum of the energy landscape defined by the Hamiltonian and the graph. It may therefore be interesting to study also the local minima and compare them to the ground state. Local minima may be sampled by running greedy optimization algorithms using random initial conditions. For correlated energy landscapes, it is known that deeper local minima have larger basins of attraction in the configuration space. The Hamiltonian (7.8) induces such a correlated energy landscape on the graph, since the total energy is not affected drastically by single spin changes. We therefore expect that the deep local minima will be sampled with higher frequency and that pairs of nodes that are grouped together in deep minima will have larger entries in the co-appearance matrix. A number of examples of co-appearance matrices sampling local energy minima at different values of γ have been given in (Reichardt and Bornholdt, 2004).

Here, we shall instead investigate the possible hierarchies of the community structures directly from the adjacency matrix. The assignment of nodes into spin states can be interpreted as an ordering of the rows and columns corresponding to nodes of the network. The ordering is such rows and columns corresponding to nodes in the same spin state (community) are next to each other. The internal order among the nodes of the same spin state is random. The choice of the ordering of the communities is arbitrary, but some orderings may be more intuitive than others. The link density in the adjacency matrix is directly transformed into grey levels. We can distinguish communities as square blocks of darker grey. Different orderings may be combined into a consensus ordering. That is, starting from a super ordering given, we re-order the nodes within each community according to a second given sub-ordering, i.e., we only change the internal order of the nodes within communities of the super ordering. This leads to the formation of new communities of those nodes that are assigned together in one community in both orderings. We can then repeat the procedure to obtain further consensus orderings.

First, we give an example of a completely hierarchical network. By hierarchical, we mean that all communities found at a value of $\gamma_2 > \gamma_1$ are proper sub-communities of the communities found at γ_1. In our example, we have constructed a network made of four large communities of 128 nodes each. Each of these nodes have an average of 7.5 links to the 127 other members of their community and 5 links to the remaining 384 nodes in the network. Each of these four communities is composed of four sub-communities of 32 nodes each. Each node has an additional 10 links to the 31 other nodes in its sub-community. Figure 7.7 shows the adjacency matrix of this network in different orderings. At $\gamma = 1$, the ground state is composed of four large communities as shown in the left part of Fig. 7.7. Increasing γ above a certain threshold makes assigning different spin states to the 16 sub-communities of the ground state configuration. The middle part of Fig. 7.7 shows an ordering obtained with a value of $\gamma = 2.2$. We can see that some of these sub-communities

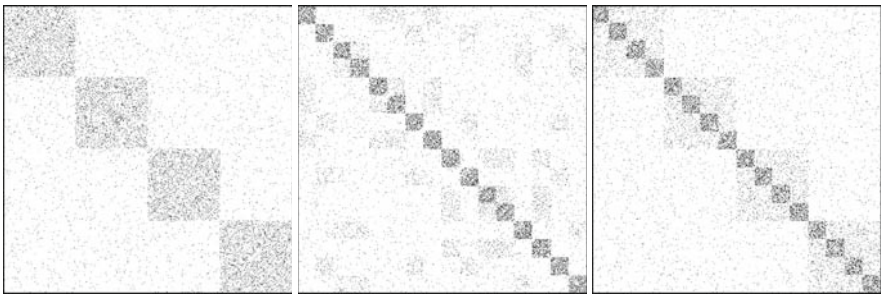


Fig. 7.7 Example of an adjacency matrix for a perfectly hierarchical network. The network consists of four communities, each of which is composed of four sub-communities. Using $\gamma = 1$, we find the four main communities (*left*). With $\gamma = 2.2$, we find the 16 sub-communities (*middle*). Link density variations in the off diagonal parts of the adjacency matrix already hint at a hierarchy. The consensus ordering (*right*) shows that each of the larger communities is indeed composed of four sub-communities each

Fig. 7.8 Example of an adjacency matrix for an only partially hierarchical network with overlapping community structure. The network consists of two large communities A and B, each of which contains a sub-community a and b, which are densely linked with each other. Using $\gamma = 0.5$, we find the two large communities (*left*). With a larger $\gamma = 1$, we find the two small sub-communities a and b grouped together. The consensus ordering (*right*) shows that most of the links that join A and B lie in fact between a and b

are more densely connected among each other. Imposing the latter ordering on top of the ordering obtained at $\gamma = 1$ then allows to display the full community structure and hierarchy of the network as shown in the right part of Fig. 7.7. Note that we have explicitly *not* used a recursive approach that applies the community detection algorithm to separate subgroups. Rather, we have obtained two *independent* orderings which are only compatible with each other, because the network has a hierarchical structure of dense communities composed of denser sub-communities.

In contrast to this situation, Fig. 7.8 shows an example of a network that is only partially hierarchical. The network consists of two large communities A and B containing 512 nodes, which have on average 12 internal links per node. Within A and B, a subgroup of 128 nodes exists, which we denote by a and b, respectively. Every node within this subgroup has 6 of its 12 intra-community links with the 127 other members of this subgroup. The two subgroups a and b have on average three links per node with each other. Additionally, every node has two links with randomly chosen nodes from the network. From Fig. 7.8, we see that we find the two large communities using $\gamma = 0.5$. Maximum modularity, however, is reached at $\gamma = 1$ when a and b are joined into a separate community. Only when using the consensus of the ordering obtained at $\gamma = 0.5$ and $\gamma = 1$, we can understand the full community structure with a and b being subgroups that are responsible for the majority of links between A and B. It is understood that this situation cannot be interpreted as a hierarchy, even though a and b are cohesive subgroups in A and B, respectively. We shall now turn to a real world example to see whether these situations can be found outside of artificially constructed examples.

7.7.10 A Real World Example

As a real world example, we study the co-authorship network (Warner, 2003) of the Los Alamos condensed matter preprint archive, considering articles published

between April 1998 and February 2004. This network has also been analyzed by Palla et al. in (2005). Every article induces a complete subgraph between the authors in this network. Since articles with a large number of authors induce very large cliques, every link induced by a single paper of n authors is only given a weight of $1/(n-1)$. After summing the weights for all papers, only links with a weight of 0.1 and greater were kept, transforming the network into a non-weighted one. The network consists of 30, 561 nodes connected by 125, 959 links. There are 668 connected components, the largest of which has 28, 502 nodes and 123, 604 links. We only work with the largest connected component. The average degree is $\langle k \rangle = 8.7$. We then minimize the Hamiltonian (7.8) using $p_{ij} = k_i k_j / 2M$ and $q = 500$. Three different values of γ were used. For each of the values of γ, some of the 500 spin states remained unpopulated, which makes us confident that we provided enough spin states. Figure 7.9 shows the adjacency matrix of the co-authorship network with rows and columns ordered according to the ground state at $\gamma = 0.5$. We can distinguish three major communities along the diagonal of the matrix and a large number of smaller communities. Off-diagonal entries in the matrix show where communities are connected with each other. Figure 7.10 shows the same adjacency matrix, but ordered according to the ground state obtained at $\gamma = 1$, while Fig. 7.11 was obtained ordering the adjacency matrix according to the ground state obtained at $\gamma = 2$. We see how the increase of γ leads to a higher number of smaller communities and a reduction in size of the major communities as expected. In Fig. 7.12, we show the adjacency matrix in a consensus ordering of the three single orderings. If the network was hierarchical with respect to γ, i.e., the communities found for larger values of γ are all complete sub-communities of those found at smaller γ, we should be able to distinguish this from the adjacency matrix in the same manner as in Fig. 7.7.

From the consensus ordering, we can see that community A from the $\gamma = 0.5$ ordering is composed of a number of smaller communities in a somewhat hierarchical manner, while community B seems to consist of a dense core and many adjacent nodes that are gradually removed as γ increases. Community C again is

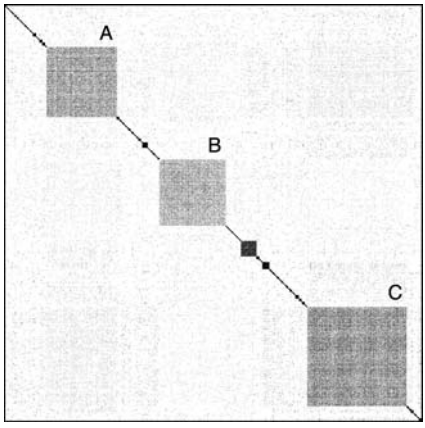


Fig. 7.9 Adjacency matrix of the co-author network ordered according to the ground state with $\gamma = 0.5$

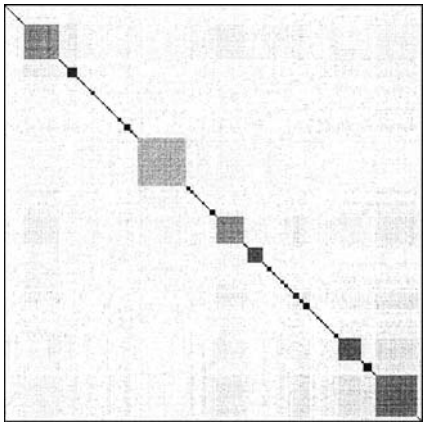


Fig. 7.10 Adjacency matrix of the co-author network ordered according to the ground state with $\gamma = 1$

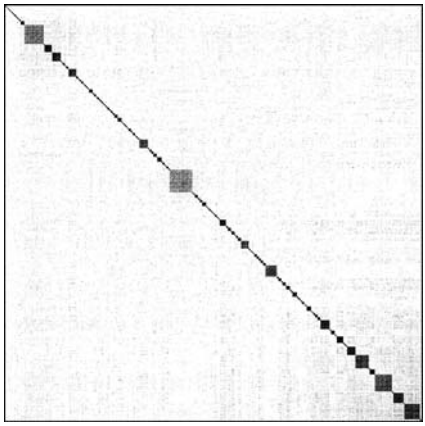


Fig. 7.11 Adjacency matrix of the co-author network ordered according to the ground state with $\gamma = 2$

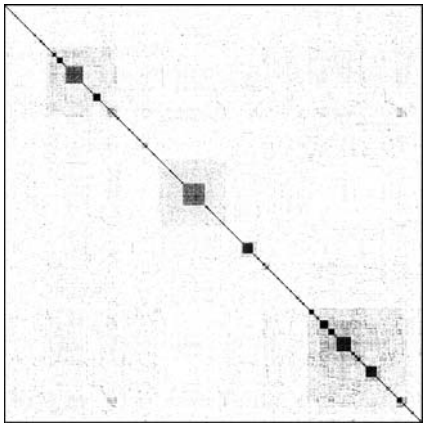


Fig. 7.12 Adjacency matrix of the co-author network ordered first according to the ground state with $\gamma = 0.5$. Within the clusters, the nodes were then ordered again according to the ground state with $\gamma = 1$ and within these clusters, the nodes were ordered according to the ground state with $\gamma = 2$

decomposed into several smaller subgroups by the consensus ordering that seems to show two levels of hierarchy.

The interpretation of the community structure and its hierarchy is beyond the scope of this article and shall not be attempted here. Rather, we intend to show that such structure exists in the link patterns of real world networks and how it can be uncovered. Furthermore, we'd like to stress that the cohesive subgroups or communities studied here are only a special case of general block models in networks (Doreian et al., 2005). The very same methods described here can be extended easily and used to discover general block structures in networks, too (Reichardt and White, 2007).

7.8 Relation of Other Algorithms to Spin Glasses

From our earlier considerations about adhesion and cohesion and the reformulations of the Hamiltonian as (7.26), it may be tempting to construct a heuristic algorithm that initially assigns different spin state to every node and then proceeds by grouping those nodes together, which have the highest coefficient of adhesion using as a model of connection probability $p_{ij} = k_i k_j / 2M$. This corresponds to a conventional agglomerative clustering procedure. As Fig. 7.13 shows, this approach fails, if the links between two communities connect nodes of low degree. The network consists of 14 nodes and 37 links. It is clearly seen that in the ground state, the network consists of two communities and edge x lies between them. However, when initially assigning different spin states to all nodes, the adhesion between the nodes connected by x is largest: $a = 1 - 16/2M$, since the product of degrees at this edge is lowest. Therefore, the agglomerative procedure described is misleading into grouping together the nodes connected by x already in the very first step. Furthermore, it is clear that in a network, where all nodes have the same degree initially, all edges connect nodes of the same coefficient of adhesion. In this case, it cannot be decided, which nodes to group together in the first step of the algorithm at all. This approach corresponds exactly to the fast community detection algorithm presented by Newman (2004). It was shown by Newman that the approach does deliver

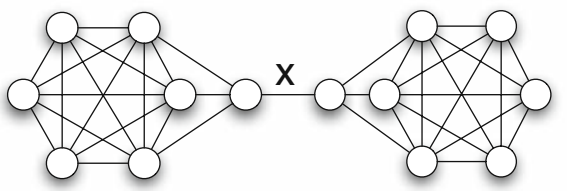


Fig. 7.13 Example network for which an agglomerative approach of grouping together nodes of maximal adhesion will fail. Starting from an assignment of different spin states to every node, the largest adhesion is found for the nodes connected by edge x and the nodes connected by x are grouped together first by the agglomerative procedure. However, it is clearly seen that x should lie between different groups

good results in benchmarks using computer generated test networks. The success of this approach depends of course on whether or not the misleading situations have a strong effect on the final outcome of the clustering. In the example shown, after grouping together the nodes at the end points of x, the algorithm will then proceed to further add nodes from only one of the two communities linked by x. Hence, the initial mistake persists, but does not completely destroy the result of the clustering.

Alternatively, it might be tempting to construct the ground state of the q-state Potts model by a recursive partitioning of the network into only two parts according to the ground state of a two-state Potts or Ising system. This procedure would be computationally simpler and result directly in a hierarchy of clusters due to repetition of the procedure on the parts until the total energy cannot be lowered anymore. Such a procedure would be justified, if the ground state of the q-state Potts Hamiltonian and the repeated application of the Ising system cut the network along the same edges. We will derive a condition under which this can be ensured.

In order for this recursive approach to work, we must ensure that the ground state of the two-state Hamiltonian never cuts through a community as defined by the q-state Hamiltonian. Assume a network made of three communities n_1, n_2, and n_3 as defined by the ground state of the q-state Hamiltonian. For the bi-partitioning, we now have two possible scenarios. Without loss of generality, the cut is made either between n_2 and $n_1 + n_3$ or between n_1, n_2, and $n_3 = n_a + n_b$, parting the network into $n_1 + n_a$ and $n_2 + n_b$. Since the former situation should be energetically lower for the algorithm to work, we arrive at the condition that

$$m_{ab} - [m_{ab}]_{p_{ij}} + m_{1b} - [m_{1b}]_{p_{ij}} > m_{2b} - [m_{2b}]_{p_{ij}}, \tag{7.33}$$

which must be valid for all subgroups n_a and n_b of community n_3. Since n_3 is a community, we further know that $m_{ab} - [m_{ab}]_{p_{ij}} > m_{1b} - [m_{1b}]_{p_{ij}}$ and $m_{ab} - [m_{ab}]_{p_{ij}} > m_{2b} - [m_{2b}]_{p_{ij}}$. Though $m_{ab} - [m_{ab}]_{p_{ij}} > 0$, since n_3 is a community, $m_{1b} - [m_{1b}]_{p_{ij}} < 0$ and $m_{2b} - [m_{2b}]_{p_{ij}} < 0$ for the same reason and hence condition (7.33) is not generally satisfied. Figure 7.14 illustrates a counter example. Assuming $p_{ij} = p$, the average link probability in the network, the upper part (a) of the figure shows the ground state of the system when using only two spin states. Part (b) of the figure shows the ground state of the system without constraints on the

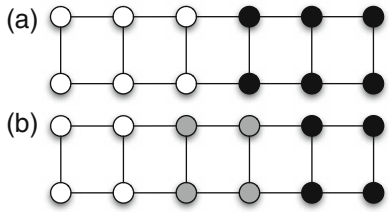


Fig. 7.14 Illustration of the problem of recursive bi-partitioning. The ground state of the Hamiltonian with only two possible spin states, as shown in (**a**), would cut through one of the communities that are found when allowing three spin states as shown in (**b**)

number of spin states, resulting in a configuration of three communities. We see that the bi-partitioning approach would have cut through one of the communities in the network. Recursive bi-partitioning cannot generally lead to an optimal assignment of spins that maximizes the modularity.

7.9 Statistical Significance and Theoretical Limits of Community Detection

In the previous sections, we have presented an approach to find possible cohesive subgroups in networks. The analysis, however, would not be complete, if we did not include some remarks about the statistical significance of the community structure we find. We will try to do so by comparison with community structures found in random networks. We will show that random networks may show surprisingly high modularities. If the modularity we find in a real world network is significantly larger than the modularity we find for an equivalent random network, then we may call the network truly modular.

The comparison with a random network is of course always possible by randomizing the original network by rewiring, keeping the degree distribution invariant (Maslov and Sneppen, 2002). One can then run a community detection algorithm in the randomized network, comparing the result to that of the original network. This method, however, can only give an answer to what a particular community detection algorithm may find in a random network and hence depends on the very method of community detection used. However, this approach always uses the correct null model of a random network for comparison, since it works on a randomized version of the original data. General results and insights, however, cannot be obtained by this method. Therefore, we intend to compare the results of a community detection algorithm with a theoretical result, obtained independently of any algorithm.

The results we present in this section allow us to compare the values of modularity obtained for a particular graph with an ER random graph that has the same number of nodes and edges. We admit that this is a very crude null model and may not be the best for comparison in many cases where the networks under study have a degree distribution very different from Poissonian. Nevertheless, it gives some general insights into the problem and one is able to give a number of very simple expressions for the expected modularity in a random graph. A more elaborate treatment can be found in (Reichardt and Bornholdt, 2007; Reichardt and Leone, 2008). The general conclusions, however, remain unaltered.

7.9.1 Statistical Significance

We have already seen that the problem of community detection can be mapped onto finding the ground state of an infinite range spin glass. For these models, in fact, a number of well-known results exist in the literature. Fu and Anderson (1986)

have calculated an expression for the number of edges to cut, when bi-partitioning a random graph into two equally sized parts using an Ising model. Kanter and Sompolinsky have done the same for partitioning a random graph into q equally sized parts (Kanter and Sompolinsky, 1987) using a Potts model. From these expressions, we can calculate expectation values for the modularity and the expected number of communities in random graphs. The approximations they use are valid for large, dense graphs. In this limit, the results depend only on the variance of the distribution of couplings. The couplings used in the study of modularity are $J_{ij} = A_{ij} - \gamma p_{ij}$ which have a mean independent of the particular form of p_{ij}:

$$J_0 = (1 - \gamma)p \tag{7.34}$$

which is zero in the case of the "natural partition" at $\gamma = 1$. If the mean of the couplings is zero, then it can be shown that the ground state is composed of equally sized communities. It is this fact that justifies the use of the results from (Fu and Anderson, 1986) and (Kanter and Sompolinsky, 1987), which are valid for equipartitions. The variance amounts to:

$$J^2 = p - (2\gamma - \gamma^2)\langle p^2 \rangle. \tag{7.35}$$

Now we can write immediately for the modularity at $\gamma = 1$ using the results of the q-partitioning problem:

$$\mathcal{Q} = -\frac{1}{M}\mathcal{H}_{GS} = \frac{N^{3/2}}{M} J \frac{U(q)}{q}, \tag{7.36}$$

where $U(q)$ is the ground state energy of a q-state Potts model with Gaussian distributed couplings of zero mean and variance J. For large q, we can approximate $U(q) = \sqrt{q \ln q}$. The exact formula for calculating $U(q)$ is given in (Kanter and Sompolinsky, 1987) and can be evaluated numerically. Table 7.1 gives the results for a few values of q. We see that maximum modularity is obtained at $q = 5$, though the value of $U(q)/q$ for $q = 4$ is not much different from it. This qualitative behavior that dense random graphs tend to cluster into only a few large communities is confirmed by our numerical experiments. By rewriting $M = pN^2/2$ and under the assumption of $p_{ij} = p$ as in the case of ER random graphs, we can further simplify Eq. (7.36) and write for the maximum value of the modularity of an ER random graph with connection probability p and N nodes:

Table 7.1 Values of $U(q)/q$ for various values of q which can be used to approximate the expected modularity with Eq. (7.36)

q	2	3	4	5	6	7	8	9	10
$\frac{U(q)}{q}$	0.384	0.464	0.484	0.485	0.479	0.471	0.461	0.452	0.442

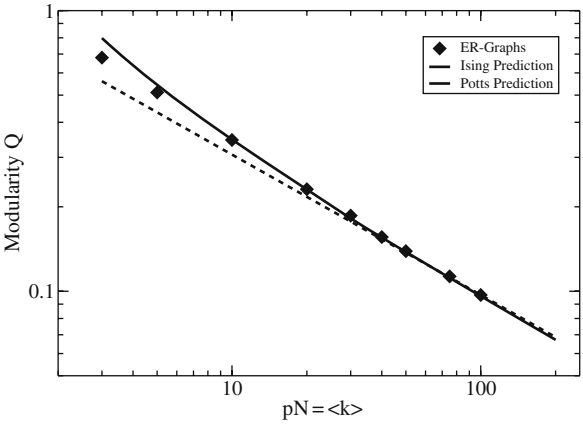


Fig. 7.15 Modularity of Erdős Rényi random graphs with average connectivity $pN = \langle k \rangle$ compared with the estimation from Eqs. (7.37) (dashed line) and (7.42) (solid line). For the experiment, random graphs with $N = 10,000$ were used

$$Q = 0.95\sqrt{\frac{1-p}{pN}} \tag{7.37}$$

where we have already made use of the fact that $q = 5$ makes the modularity maximal. Figure 7.15 shows the comparison of Eq. (7.37) and experiments where we have numerically maximized the modularity. We see that the prediction fits the data well for dense graphs and that modularity decays as a function of $(pN)^{-1/2}$.

Even though the estimations of the value of modularity for random graphs from the Potts spin glass are rather close to the actual situation for sparse random graphs, the number of communities, at which maximum modularity is achieved is not. In (Guimera et al., 2004), it is shown that the number of communities for which the modularity reaches a maximum is $\sqrt{N}$ for treelike networks with $\langle k \rangle = 2$. Therefore, one would expect the number of communities for which $\mathcal{Q}$ is maximal to rise for sparse graphs. Our numerical experiments on large Erdős Rényi random graphs also show that the number of communities found in sparse networks tends to increase as $\langle k \rangle$ decreases.

Even though we have seen that in general, recursive bi-partitioning will not lead to an optimal community assignment, we shall still use this approach for random graphs. We have seen that maximum modularity for random graphs is achieved for equipartitions. Hence, we should be able to partition the network recursively and at least find the number of communities in a random graph, for which further partitioning does not result in an improvement of the modularity. The number of cut edges $\mathcal{C} = \mathcal{C}(N, M)$ in any partition, will be a function of the number of nodes in the remaining part and the number of connections within this remaining part and their distribution. We note that the M connections will be distributed into internal and external links per node $k_{in} + k_{out} = k$. This allows us to write $\mathcal{C} = N\langle k_{out} \rangle/2$ for a bi-partition. After each partition, the number of internal connections a node has decreases due to the previous cut. We use these results in order to approximate the number of cut edges after b recursive bi-partitions which lead to 2^b parts:

$$\mathcal{C} = \sum_{t=1}^{b} 2^{t-1} \frac{N}{2^t} \langle k_{out,t} \rangle = \sum_{t=1}^{b} \frac{N}{2} \langle k_{out,t} \rangle \tag{7.38}$$

where $\langle k_{out,t} \rangle$ is the average number of external edges a node gains after cut t. In order to find $\langle k_{out,t} \rangle$, we go back to the results from Fu and Anderson (1986) again who find for a bi-partition:

$$\mathcal{C}_{\text{bi-part}} = \frac{M}{2} \left[1 - c \sqrt{\frac{1-p}{pN}} \right]. \tag{7.39}$$

with a constant of $c = 1.5266 \pm 0.0002$. We can write

$$\langle k_{out} \rangle = \frac{pN - c\sqrt{pN(1-p)}}{2} \tag{7.40}$$

$$\langle k_{in} \rangle = \frac{pN + c\sqrt{pN(1-p)}}{2} \tag{7.41}$$

from which we can calculate (7.38) substituting pN with the appropriate $\langle k_{in} \rangle$ in every step of the recursion. The modularity can then be written as:

$$Q = \frac{2^b - 1}{2^b} - \frac{1}{k} \sum_{t=1}^{b} \langle k_{out,t} \rangle. \tag{7.42}$$

Now we need to find only the number of recursions b that maximizes Q. Since the optimal number of recursions will depend on $pN = \langle k \rangle$, we also find an estimation of the number of communities in the network. Figure 7.15 shows a comparison between the theoretical prediction of the maximum modularity that can be obtained from Eq. (7.42). The improvement of (7.42) over (7.37) for sparse graphs must be due to the possibility of having larger numbers of communities, since otherwise the assumptions made are the same. Again, we find that the modularity behaves asymptotically like $(pN)^{-1/2}$ as already predicted from the Potts spin glass. Note that different expressions for the scaling of the modularity are necessary when the degree distribution is not Poissonian (Reichardt and Bornholdt, 2007; Reichardt and Leone, 2008).

Figure 7.16 shows the comparison of the number of communities estimated from (7.42) and the numerical experiments on random graphs. The good agreement between experiment and prediction is interesting, given the fact that (7.42) allows only powers of two as the number of communities. For dense graphs, the Potts limit of only a few communities is recovered. We see that sparse random graphs cluster into a large number of communities, while dense random graphs cluster in only a hand full of large communities. Most importantly, sparse random graphs exhibit very large values of modularity. These large values are only due to their sparseness and *not* due to small size. We also stress that statistically significant modularity must

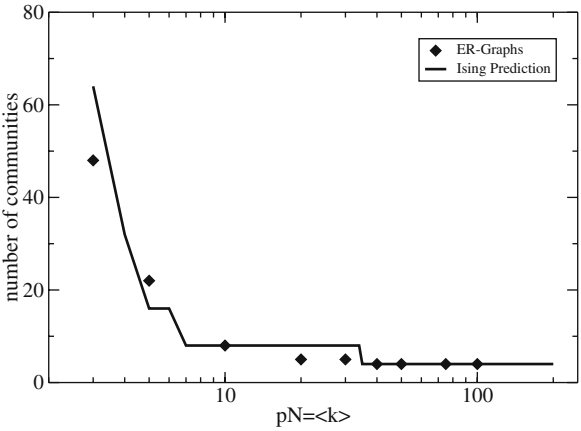


Fig. 7.16 Number of communities found in Erdős Rényi random graphs with average connectivity $pN = \langle k \rangle$ compared with the estimation from Eq. 7.42. For the experiment, random graphs with $N = 10,000$ were used

exceed the expectation values of modularity obtained from a suitable null model of the graph. If this null model is an Erdős Rényi random graph, then there is very little improvement possible over the values of modularity obtained for the null model for sparse graphs.

7.9.2 Theoretical Limits

In Fig. 7.4, we have seen that we are unable to recover completely the known community structure of the test networks with four equal sized communities for values of $\langle k_{in} \rangle < 8$. This is somewhat surprising, since the probability for an intra-community link is $p_{in} = k_{in}/31$ and for an inter-community link, it is $p_{out} = k_{out}/96$. This means, the two probabilities are equal $p_{in} = p_{out}$ only at $k_{in} = 4$ and we could hope that better algorithms exist that are able to detect the built-in community structure down to this limit. From the last section, we have seen how to calculate the modularity in a random network. Assuming that the test network with $N = 128$ and $p = 16/127$ was indeed random with $p_{in} = p_{out}$, we find an expectation value for the modularity of $\mathcal{Q} = 0.24$. The modularity we have built into the network by design can be calculated as $\mathcal{Q} = 3/4 - \langle k_{out} \rangle / \langle k \rangle$. From this, we see that for $\langle k_{out} \rangle > \langle k \rangle / 2$, the modularity we built into the system is smaller than what we find in a random graph. It is important to understand that we can understand "random graph" here as "any graph," i.e., the modularity we find for a random graph is that, which we can find almost surely in any graph with the same number of nodes and links. We can only recover a particular community structure, if its modularity is larger than that of the equivalent random graph. Otherwise, alternative configurations to the one built-in will be found that give higher values of modularity and hence are and should be preferred by any algorithm.

Looking back now at Fig. 7.15, we see that for sparse graphs, there is little room for improvement over random graphs, since any sparse graph already shows very

high values of modularity. For dense graphs, it is much easier to detect built-in community structures. For networks of four communities and $\langle k \rangle = 100$, $\langle k_{out} \rangle$ may increase up to 65 until the built-in modularity reaches that of an equivalent random graph. It can also be shown that it is easier to detect modules in networks which have a broad degree distribution (Reichardt and Leone 2008).

From these results, we can decide whether one can recover a built-in community structure from a network using any algorithm that finds a partition of the network that is maximally modular. We need to calculate the modularity of the built-in community structure and compare it to the expectation value of a random graph. The larger $\mathcal{Q}$ for the built-in community structure compared to the expectation value, the easier it will be to find this community structure for any algorithm. If it is smaller than the value for an equivalent random graph, no algorithm can exist that may recover the built-in community structure.

These considerations tell us that we are not able to find communities that arise from only a small difference in the connection probability of intra- and inter-community links. Due to the enormously huge number of possible groupings, highly modular or assortative assignments of nodes into classes exist also for completely uncorrelated random graphs. There exists a hard theoretical limit to the kind of community structures we may detect (Reichardt and Leone 2008).

7.10 Conclusion

We hope that this article, more than providing a brief introduction to the field of complex networks from a physics perspective, may serve as an appetizer for further reading to the social scientist. We could only cite a few papers and mention a few results, that we believe to be important. There are many more and we apologize for everything we had to leave out. Nevertheless, we hope to have given a flavor of the type of methods and results that can be provided from physics. The approach of mapping problems from social science onto problems from solid state physics is certainly unusual. We do not claim that it is superior to what has been done in decades of social science research. We do believe, however, that it offers an alternative perspective and maybe sheds some new light on certain problems. If the methods and results we provide may make a humble contribution to the cross-fertilization and interdisciplinary research of social and natural sciences, we would be most happy.

References

R. Albert and A.-L. Barabàsi. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002.

P. Arabie and L. J. Hubert. Combinatorial data analysis. *Annu. Rev. Psychol.*, 43:169–203, 1992.

A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree distribution. *J. Comb. Theory (A)*, 24:296, 1978.
B. Bollobas. *Random Graphs*. Cambridge University Press, New York, 2nd. edition, 2001.
S. Bornholdt and H. G. Schuster, editors. *Handbook of Graphs and Networks*. Wiley-Vch, Weinheim, 2003.
A. Clauset, M. E. J. Newman, and C. Moore. Finding community structure in very large networks. *Phys. Rev. E.*, 70:066111, 2004.
R. Cohen, K. Erez, and S. Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85:4626–4628, 2000.
P. Doreian, V. Batagelj, and A. Ferligoj. *Generalized Blockmodeling*. Cambridge University Press, New York, 2005.
S. N. Dorogovtsev and J. F. F. Mendes. The shortest path to complex networks. In N. Johnson, J. Efstathiou, and F. Reed-Tsochas, editors, *Complex Systems and Inter-Disciplinary Science*. World Scientific, 2005.
P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5:17–61, 1960.
Y. Fu and P. W. Anderson. Application of statistical mechanics to NP-complete problems in combinatorial optimisation. *J. Phys. A: Math. Gen.*, 19:1605–1620, 1986.
D. Gfeller, J.-C. Chappelier, and P. de los Rios. Finding instabilities in the community structure of complex networks. *Phys. Rev. E.*, 72:056135, 2005.
M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA*, 99(12):7821–7826, 2002.
R. Guimera, M. Sales-Pardo, and L. N. Amaral. Modularity from fluctuations in random graphs and complex networks. *Phys. Rev. E.*, 70:025101(R), 2004.
I. Kanter and H. Sompolinsky. Graph optimisation problems and the Potts glass. *J. Phys. A: Math. Gen.*, 20:L636–679, 1987.
F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and X. Aberg. The web of human sexual contacts. *Nature*, 411:907–908, 2001.
S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296:910–913, 2002.
S. Milgram. The small world problem. *Psychol. Today*, 2:60–67, 1967.
M. Molloy and B. Reed. A critical point for random graphs with given degree sequence. *Random Struct. Algor.*, 6:161–179, 1995.
M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89:208701, 2002.
M. E. J. Newman. Mixing patterns in networks. *Phys. Rev. E.*, 67:026126, 2003.
M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2): 167–256, 2003.
M. E. J. Newman. Fast algorithm for detecting community structure in networks. *Phys. Rev. E.*, 69:066133, 2004.
M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69:026113, 2004.
G. Palla, I. Derenyi, I. Farkas, and T. Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 435:814, 2005.
R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, 86:3200–3203, 2001.
R. Pastor-Satorras and A. Vespignani. Immunization of complex networks. *Phys. Rev. E.*, page 036104, 2002.
J. Reichardt and S. Bornholdt. Detecting fuzzy community structures in complex networks with a Potts model. *Phys. Rev. Lett.*, 93:218701, 2004.
J. Reichardt and S. Bornholdt. Graph partitioning and modularity of graphs with arbitrary degree distribution. *Phys. Rev. E*, 76:015102(R), 2007.
J. Reichardt and M. Leone. (Un) detectable cluster structure in sparse networks. *Phys. Rev. Lett.*, 101:078701, 2008.

J. Reichardt and D. R. White. Role models for complex networks. *Eur. Phys. J. B*, 60:217–224, 2007.
S. Warner. E-prints and the open archives initiative. *Library Hi Tech*, 21:151–158, 2003.
S. Wasserman and K. Faust. *Social Network Analysis*. Cambridge University Press, New York, 1994.
D. J. Watts and S. H. Strogatz. Collective dynamics in small world networks. *Nature*, 393:440–442, 1998.
W. Zachary. An information flow model for conflict and fission in small groups. *J. Anthr. Res.*, 33:452–473, 1977.

Chapter 8
Modeling Evolving Innovation Networks

Michael D. König, Stefano Battiston and Frank Schweitzer

8.1 Introduction

8.1.1 The Importance of Innovation Networks

Economists widely agree on technological change and innovation being the main components of economic growth (Aghion and Howitt, 1998; Tirole, 1988). In the absence of ongoing technological improvements, economic growth can hardly be maintained (Barro and Sala-i Martin, 2004). The close link between innovation and economic performance has become generally accepted. Following this insight, in recent years of economic growth, OECD countries have fostered investments in science, technology, and innovation (OECD, 2006).

Moreover, technologies are becoming increasingly complex. This increasing complexity of technologies can make an agent's "in-house" innovative effort insufficient to compete in an R&D intensive economy. Thus, agents have to become more specialized on specific domains of a technology and they tend to rely on knowledge transfers from other agents, which are specialized in different domains, in order to combine complementary domains of knowledge for production ("recombinant growth" (Weitzman, 1998)).

When one agent benefits from knowledge created elsewhere we speak of knowledge spillovers. Knowledge spillovers define "any original, valuable knowledge generated somewhere that becomes accessible to external agents . . . other than the originator"[1] (Foray, 2004, p. 95).

The knowledge-based economy is developing towards a state in which the costs for acquiring, reproducing, and transmitting knowledge are constantly decreasing,

M.D. König (✉)
Chair of Systems Design, ETH Zurich, Kreuzplatz 5, CH-8092 Zurich, Switzerland
e-mail: mkoenig@ethz.ch

[1] "Involuntary spillovers are a feature of market competition. Competition not only creates incentives to produce new knowledge but it also forces the other agents to increase their own performance through imitation, adoption and absorption of the new knowledge created elsewhere" (Foray, 2004, p. 91).

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_8,

spatial and geographical limitations on knowledge exchange are becoming less important, and attitudes change towards more open behavior of sharing knowledge instead of hiding it from others. In this state, knowledge externalities will play an increasingly important role.

When agents are using knowledge that is created elsewhere, they must have access to other agents across a network whose links represent the exchange or transfer of knowledge between agents. The importance of networks in innovative economies has been widely recognized, e.g., it has been observed that "the development of knowledge within industries is strongly influenced by the network structure of relations among agents" (Antonelli, 1996, p. 1). Subsequently, an ample body of empirical research has documented the steady growth of R&D partnerships among firms (Hagedoorn et al., 2006).

8.1.2 Markets for Knowledge Exchange

The exchange of knowledge is not unproblematic. Markets for knowledge exchange can exhibit serious market failures (Arora et al., 2004; Gerosky, 1995), which make it difficult for innovators to realize a reasonable return from trading the results of their R&D activities (the problem of appropriability (Gerosky, 1995)). This is due to the public good character of knowledge, which makes it different from products or services. Knowledge is non-rival, meaning its use by one agent does not diminish its usability by another agent, and sometimes (when knowledge spillovers cannot be avoided) non-excludable, meaning that the creator of new knowledge cannot prevent non-payers from using it. The problems associated with trading of knowledge can prevent agents from exchanging knowledge at all.

There are three generic reasons for failures of markets for technology (Arora et al., 2004; Arrow, 1962; Gerosky, 1995): (1) economies of scale/scope, (2) uncertainty and (3) externalities.

(1) R&D projects often require huge initial investments and they can exhibit economies of scale since the cost for useful technological information per unit of output declines as the level of output increases (Wilson, 1975). Besides, Nelson (1959) has shown that economies of scope can apply to innovative agents. The broader an agents' "technological base," the more likely it is that any outcome of its R&D activities will be useful for her. The result is that markets for knowledge exchange are often dominated by monopolies.
(2) Almost all economic investments bear a risk of how the market will respond to the new product (commercial success). Innovators face additional risks. First, their investment into R&D does not necessarily lead to a new technology. Second, if such a new technology is discovered, it has to be put into practice in a new and better product than the already existing ones. This inherent uncertainty of R&D projects often causes agents to invest "too little."
(3) Externalities are important when the action of one agent influences the profits of another agent without compensation through the market. Public goods are a

typical example of creating externalities. Knowledge is a public good and the returns innovators can realize often are far below their investments into R&D. This can seriously diminish an agent's incentive to do R&D.

In order to overcome the above mentioned problems associated with the returns on investment into R&D, appropriate incentive mechanisms have to be created that encourage agents to invest into R&D. In general, Von Hippel and Von Krogh (2003) suggest three basic models of encouraging agents to invest into R&D:

(1) The private investment model assumes that innovation is undertaken by private agents investing their resources to create an innovation. Society then provides agents with limited rights to exclusively use the results of their innovation through patents or other intellectual property rights (by creating a temporarily monopoly).[2]
(2) The so-called collective action model (Allen, 1983) assumes that agents are creating knowledge as a public good. Knowledge is made public and unconditionally supplied to a public pool accessible to everybody. The problem is that potential beneficiaries could wait until others provide the public good and thereby could free-ride. One solution to this problem is to provide contributors (in this case innovators) with some form of subsidy. Scientific research is such an example where reputation-based rewards are granted to scientists for their good performance.
(3) In the private-collective innovation model participants use their private resources in creating new knowledge and then make it publicly available. This is typically observed in open-source projects. There are several incentives (Lerner and Tirole, 2002; Von Hippel and Von Krogh, 2003) for agents to participants in open-source projects. These range from elevated reputations, the desire of building a community to the expectation of reciprocity from the community members for their efforts.

The collective action approach (2) gives a possible explanation for the willingness of agents to share knowledge if there are no costs associated with it. One can think of a pool of technologies that is accessible to everybody ("broadcasting" of technologies) (Allen, 1983). This can be the case where agents are non-rivals and shared information may have no competitive cost. Additionally, knowledge must be easily understandable and transferable. This assumes that knowledge is highly codified[3] such that the transfer of knowledge from one agent to the other is costless. But, if these assumptions do not hold, the costs for transferring knowledge can

[2] For a more detailed treatment of this issue we recommend Scotchmer (2004).

[3] The opposite case of codified knowledge is tacit knowledge. "Tacit knowledge is difficult to make explicit for transfer and reproduction. The exchange, diffusion, and learning of tacit knowledge require those who have it to take deliberate action to share it. This is difficult and costly to implement ... Knowledge can, however, be codified. It can be expressed in a particular language and recorded on a particular medium. As such, it is detached from the individual. When knowledge is codified, it becomes easily transferable" (Foray, 2004, p. 73).

often be considerable, and agents become more selective about whom to share their knowledge with. We study this situation in the next sections.

8.1.3 Economies as Evolving Networks

As we have already outlined above, modern economies are becoming increasingly networked, and this also affects the innovation process where information and knowledge are exchanged by interactions between agents (Gallegati and Kirman, 1999; Kirman, 1997). In the agent-based view, the aggregate behavior of the economy (macro-economics) cannot be investigated in terms of the behavior of isolated individuals. Not only there are different ways in which firms interact, learning over time, based on their previous experience; also interactions between them take place within a network and not in a all-to-all fashion.

The standard neoclassical model[4] of the economy assumes that anonymous and autonomous individuals take decisions independently and interact only through the price system, which they cannot influence at all. However, competition easily becomes imperfect because, if agents have only a minimal market power, they will anticipate the consequences of their actions and anticipate the actions of others.

Game theorists have tried to integrate the idea of strategically interacting agents into a neoclassical[5] framework. But still they leave two questions unanswered. First, it is assumed that the behavior is fully optimizing. This leads to agents with extremely sophisticated information processing capabilities. Such ability of passing these enormous amounts of information in short times cannot be found in any realistic setting of human interaction. Advances in weakening that assumption are referred to as "boundedly rationality" (Gigerenzer and Selten, 2002). Second, the problem of coordination of activities is not addressed in the standard equilibrium model of the economy. Instead it is assumed that every agent can interact and trade with every other agent, which becomes quite unrealistic for large systems.

One has to specify the framework within the individual agents take price decisions, and thus limit the environment within which they operate and reason. An obvious way is to view the economy as a network in which agents interact only with their neighbors. In the case of technological innovation, neighbors might be similar firms within the same industry, but these firms will then be linked either through customers or suppliers with firms in other industries. Through these connections innovations

---

[4] A standard neoclassical model includes the following assumptions (Gabszewicz, 2000): (1) perfect competition, (2) perfect information, (3) rational behavior, (4) all prices are flexible (all markets are in equilibrium). The resulting market equilibrium (allocation of goods) is then efficient. See Hausman (2003) for a discussion of these assumptions.

[5] The individual decision making process is represented as maximizing a utility function. A utility function is a way of assigning a number to every possible choice such that more-preferred choices have a higher number than less-preferred ones (Varian, 1996). The gradients of the utility function are imagined to be like forces driving people to trade, and from which economic equilibria emerge as a kind of force balance (Farmer et al., 2005).

will diffuse through the network. The rate and extent of this diffusion then depends on the structure and connectivity of the network. The evolution of the network itself should be made endogenous where the evolution of the link structure is dependent on the agents' experience from using the links available to them. In this framework the individuals learn and adapt their behavior and this in turns leads to an evolution of the network structure. The economy then becomes a *complex evolving network*.

8.1.4 Complex Networks

Although no precise mathematical definition exists for a *complex network*, it is worth to elaborate the notion associated with it. In general, a network is a set of items some of which are linked together by pairwise relationships. The structure of the relationships can be represented mathematically as a graph in which nodes are connected by links (possibly with varying strength). However, a network is usually also associated with some dynamic process on the nodes which in turn affects the structure of the relationships to other nodes. A wide variety of systems can be described as a network, ranging from cells (a set of chemicals connected by chemical reactions), to the Internet (a set of routers linked by physical information channels). It is clear that the structure of the relationships co-evolves with the function of the items involved.

As a first step, a network can be described simply in terms of its associated graph.[6] There are two extreme cases of relatively simple graphs: regular lattices on one side and random graphs[7] on the other side. During the last century, graph theory and statistical physics have developed a body of theories and tools to describe the behavior of systems represented by lattices and random graphs. However, it turns out that, at least for physical scales larger than biomolecules, most systems are not structured as lattices or as random graphs. Moreover, such a structure is not the result of a design, but it emerges from self-organization. In some cases self-organization results from the attempt to optimize a global function. In other cases, as it is typical in economics, it results from nodes locally trying to optimize their goals, e.g., an individual utility function.

Large networks are collectively designated as complex networks if their structure (1) is coupled to the functionality, (2) emerges from self-organization, and (3) deviates from trivial graphs. This definition includes many large systems of enormous technological, intellectual, social, and economical impact (Frenken, 2006).

---

[6] In general, a graph represents pairwise relations between objects from a certain collection. A graph then consists of a collection of nodes and a collection of links that connect pairs of nodes.

[7] The classical Erdös-Reny random graph is defined by the following rules (Bollobas, 1985):

(1) The total number of nodes is fixed.

(2) Randomly chosen pairs of nodes are connected by links with probability p.

The construction procedure of such a graph may be thought of as the subsequent addition of new links between nodes chosen at random, while the total number of nodes is fixed.

8.1.5 The Statistical Physics Approach

As we will discuss later, many of the theoretical tools developed in economics and specifically in game theory to characterize the stability of small networks of firms cannot be used for large networks. Asking which is the optimal set of connections that a firm should establish with other firms has little meaning in a large network if strategic interaction is taken into account (with more than say, 100 nodes it is simply not feasible to compute). On the other hand, it makes sense to ask what are connectivity properties of the nodes a firm should try to target in order to improve its utility with a certain probability. It is then necessary to turn towards a statistical description of these systems, where one is no longer interested in individual quantities but only in averaged quantities.

There exists an arsenal of such tools developed within *statistical physics* in the last century that allow to predict the macroscopic behavior of a system from the local properties of its constituents (Durlauf, 1999; Amaral et al., 1999). Such tools work very well for systems of identical particles embedded in regular or random network structures in which interactions depend on physical distance. Both a regular and a random structure have a lot of symmetries, which one can exploit to simplify the description of the system. However, in complex networks many of those symmetries are broken: individuals and interactions are heterogeneous. Moreover the physical distance is often irrelevant (think for instance of knowledge exchange via the Internet). Therefore, a satisfactory description of such systems represents a major challenge for statistical physics (Amaral et al., 2001).

In the last few years, we have thus witnessed an increasing interest and effort within the field of statistical physics in studying complex networks that traditionally were object of investigation by other disciplines, ranging from biology to computer science, linguistics, politics, anthropology, and many others. One of the major contributions of statistical physics to the field of complex networks has been to demonstrate that several dynamic processes taking place on networks that deviate from random graphs, exhibit a behavior dramatically different from the ones observed on random graphs.

An example for all is the case of virus spreading: it has been shown that while for random networks a local infection spreads to the whole network only if the spreading rate is larger than a critical value, for scale-free networks[8] any spreading rate leads to the infection of the whole network. Now, technological as well as social networks are much better described as scale-free graphs than as random graphs. Therefore all vaccination strategies for both computer and human viruses, which have been so far designed based on the assumption that such networks were random graphs, need to be revised. This highly unexpected result goes against volumes previously written on this topic and is due to the presence of a few nodes with very

[8] A scale-free network is characterized by a degree distribution which follows a power-law, $f(d) = \alpha d^{\gamma}$. The degree distribution gives the number of agents with a certain number, d, of in- or outgoing links (in- or out-degree), see the next section for a definition of degree of a node.

large connectivity. In this case, the rare events (infection of highly connected nodes) and not the most frequent ones matter.

Explaining the macroscopic behavior of a system in terms of the properties of the constituents has been a major success of the physicist's reductionist approach. But, while in physical systems the forces acting on single constituents can be measured precisely, this is not the case in a socio-economic system where, moreover, each agent is endowed with high internal complexity. Today, the physicist's approach to socio-economic systems differs from the nineteenth century positivist approach in so far as it does not aim at predicting, for instance, the behavior of individual agents. Instead, taking into account the major driving forces in the interactions among agents at the local level we try to infer, at a system level, some general trends or behavior that can be confirmed looking at the data. This is also very different from taking aggregate quantities and infer a macroscopic behavior from a "representative" agent[9] as it is done in several approaches in mainstream economics.

8.1.6 Dynamics Versus Evolution in a Network

After discussing the notion of a complex network which has been strongly influenced by physics, we now try to classify different complex networks.

The nodes in an economic network are associated with a state variable, representing the agents' wealth, a firm's output or, in the case of innovation networks, knowledge. There is an important difference between the evolution of the network and the dynamics taking place on the state variables. In the first, nodes or links are added to/removed from the network by a specific mechanism and in the latter, the state variables are changed as a result of the interactions among connected nodes (see also Gross and Blasius, 2007 for a review). Consequently, there are four aspects that can be investigated in complex economic networks (Battiston, 2003).

1. Statistical characterization of the static network topology without dynamics of state variables,
2. network evolution without dynamics of the state variables,
3. dynamics of state variables in a static network and
4. dynamics of state variables and evolution of the network at the same time.

This can be incorporated in the following Table 8.1.

[9] The concept of the representative agent assumes an economy which consists of a sufficiently homogeneous population of agents. Because all the agents are equivalent, the aggregate quantities of the system can be calculated by multiplying the average agent, or the representative agent, by the number of agents (the system size). For example, the total production of an economy is obtained by summing up the production levels of the individual firms that constitute the economy. To determine the behavior of the system, it is therefore sufficient to know the characteristics of the representative agent.

Table 8.1 Overview of the different ways in which a network and the state variables of the nodes can be related

Case	State variables	Network
1.	Static	Static
2.	Dynamic	Static
3.	Static	Dynamic
4.	Dynamic	Dynamic

In socio-economic systems as well as in biological systems, dynamics and evolution are often coupled, but do not necessarily have the same time scale. In Sect. 8.3.9, we will show how the coupling of fast knowledge growth (dynamics) and slow network evolution can lead to the emergence of self-sustaining cycles in a network of knowledge sharing (cooperating) agents.

8.1.7 Outline of This Chapter

In this chapter, we focus on (i) the emergence and (ii) the performance of different structures in an evolving network. The different scenarios we develop shall be applied to firms exchanging knowledge in a competitive, R&D intensive economy. In the existing literature reviewed in the following section, there are two different lines of research addressing these problems: (i) Models of network formation were developed based on individual utility functions, e.g., by Jackson (2003), in which simple architectures emerge in the equilibrium. (ii) In another group of models, e.g., Padgett et al. (2003), firms have specific skills and take actions based on goals or learning, and innovation is associated with the emergence of self-sustaining cycles of knowledge production. Although both lines of work address the problem of network emergence and performance, they differ significantly in terms of methods and results. We try to bridge them by introducing a novel model of evolving innovation networks that combines the topological evolution of the network with dynamics associated with the network nodes.[10]

We start our approach by giving a short introduction to graph theory in Sect. 8.2.1. Here we restrict ourselves only to the most important terms and definitions that are necessary in the following sections.[11]

We then proceed by giving an overview of the existing literature on economic networks. In the first part of our literature review, we explore some basic models of innovation networks. The selection of these models is by no means unique nor exhaustive, but points to important contributions to the growing literature on economic and innovation networks.[12] Similar to our own approach, these models make considerably simple assumptions and thus allow for analytical insights. This holds

[10] The approach of combining a dynamics of the network with a dynamics in the nodes is discussed in Gross and Blasius (2007).

[11] The reader interested in a good introduction in graph theory can consult West (2001).

[12] For an excellent introduction, see Jackson (2008), Vega-Redondo (2007), and Goyal (2007).

in particular for the connections model in Sect. 8.2.2. The model in Sect. 8.2.3 can be considered as an extension of the basic connections model where "small-world" networks emerge. In the subsequent Sect. 8.2.4, we discuss a model that takes heterogeneous knowledge into account, as a further extension. In the second part of the literature review, we briefly sketch models in which we observe cyclic network topologies. We show that in certain cases the stability of a network and its performance depends critically on its cyclic structure. The critical role of cycles in a networked economy has already been identified by Rosenblatt (1957) and many succeeding authors, e.g., Bala and Goyal (2000), Kim and Wong (2007), Maxfield (1994). In this chapter, we review some recent models in which cyclic networks emerge: in Sect. 8.2.5, we first introduce a model of production networks with closed loops and next we discuss a model of cycles of differently skilled agents.

Finally, in Sect. 8.3 we develop a novel framework, which we call *Evolving Innovation Networks*, to study the evolution of innovation networks. We show how different modalities of interactions between firms and cost functions related to these interactions can give rise to completely different equilibrium networks. We have studied the case of linear cost and bilateral interactions in König et al. (2008a,b). There we find that, depending on the cost, the range of possible equilibrium networks contains complete, intermediate graphs with heterogeneous degree distributions as well as empty graphs. Here, we focus on a type of non-linear cost and both, on unilateral and bilateral interactions. In the unilateral case, we find that, in a broad range of parameter values, networks can break down completely or the equilibrium network is very sparse and consists of few pairwise interactions and many isolated agents. Equilibrium networks with a higher density can be reached if (i) the utility function of the agents accounts for a positive externality resulting from being part of a technological feedback loop or if (ii) all interactions are bidirectional (direct reciprocal). Otherwise, the network collapses and only few, if any, agents can beneficially exchange knowledge.

The results found in our novel approach to evolving innovation networks are summarized in Sect. 8.4.1. The appendices shall be useful for the reader interested in more numerical results and the parameters and explanation of the algorithms used.

8.2 Basic Models of Innovation Networks

8.2.1 Graph Theoretic Network Characterization

Before we start to describe specific models of economic networks, we give a brief introduction to the most important graph theoretic terms used throughout this chapter to characterize networks. For a broader introduction to graph theory see West (2001). In this chapter, we will use the terms graph and network interchangeably, i.e., both refer to the same object.

A *graph* G is a pair, $G = (V, E)$, consisting of a set of node $V(G)$ and a set of links $E(G)$. K_n is the *complete graph* on n nodes. C_n the *cycle* on n nodes. Nodes i and j are the endpoints of the link $e_{ij} \in E(G)$.

The *degree*, d_i, of a node i is the number of links incident to it. A graph can either be *undirected* or *directed*, where in the latter case one has to distinguish between *in-degree*, d_i^-, and *out-degree*, d_i^+, of node i. In the case of an undirected graph, the neighborhood of a node i in G is $N_i = \{w \in V(G) : e_{wi} \in E(G)\}$. The degree of a node i is then $d_i = |N_i|$. The first-order neighborhood is just the neighborhood, N_i, of node i. The second-order neighborhood is, $N_i \cup \{N_v : v \in N_i\}$. Similarly, higher order neighborhoods are defined. In the case of a directed graph, we denote the out-neighborhood of node i by N_i^+ and the in-neighborhood by N_i^-. A graph G is *regular* if all nodes have the same degree. A graph G is *k-regular* if every node has degree k.

A *walk* is an alternating list, $\{v_0, e_{01}, v_1, ..., v_{k-1}, e_{k-1k}, v_k\}$, of nodes and links. A *trail* is a walk with no repeated link. A *path* is a walk with no repeated node. The shortest path between two nodes is also known as the *geodesic distance*. If the endpoints of a trail are the same (a closed trail) then we refer to it as a *circuit*. A circuit with no repeated node is called a *cycle*.

A *subgraph*, G', of G is the graph of subsets of the nodes, $V(G') \subseteq V(G)$, and links, $E(G') \subseteq E(G)$. A graph G is *connected*, if there is a path connecting every pair of nodes. Otherwise, G is disconnected. The *components* of a graph G are the maximal connected subgraphs.

The *adjacency matrix*, $\mathbf{A}(G)$, of G, is the n-by-n matrix in which the entry a_{ij} is 1 if the link $e_{ij} \in E(G)$, otherwise a_{ij} is 0. For an undirected graph, $\mathbf{A}$ is symmetric, i.e., $a_{ij} = a_{ji}\ \forall i, j \in V(G)$. An example of a graph and its associated adjacency matrix is given in Fig. 8.1. For example, in the first row with elements, $a_{11} = 0, a_{12} = 1, a_{13} = 0, a_{14} = 0$, the element $a_{12} = 1$ indicates that there exists a link from node 1 to node 2.

In a *bipartite* graph G, $V(G)$ is the union of two disjoint independent sets V_1 and V_2. In a bipartite graph, if $e_{12} \in E(G)$ then $v_1 \in V_1$ and $v_2 \in V_2$. In other words, the two endpoints of any link must be in different sets. The *complete bipartite graph* with partitions of size $|V_1| = n$ and $|V_2| = m$ is denoted $K_{n,m}$. A special case is the *star* which is a complete bipartite graph with one partition having size $n = 1$, $K_{1,m}$.

There exists an important class of graphs, *random graphs*, which are determined by their number of nodes, n, and the (independent) probability p of each link being present in the graph (Bollobas, 1985).

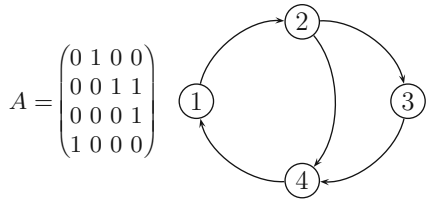


Fig. 8.1 (*Right*) A directed graph consisting of four nodes and five links. (*Left*) The corresponding adjacency matrix A

We now introduce two topological measures of a graph, the clustering coefficient and the average path length. For further details see, e.g., Newman (2003) and Costa et al. (2007). The following definitions assume undirected graphs.

For each node, the *local clustering coefficient*, C_i, is simply defined as the fraction of pairs of neighbors of i that are themselves neighbors. The number of possible neighbors of node i is simply $d_i(d_i - 1)/2$, where d_i is the degree of node i. Thus we get

$$C_i = \frac{|\{e_{jk} \in E(G) : e_{ij} \in E(G) \wedge e_{ik} \in E(G)\}|}{d_i(d_i - 1)/2}. \tag{8.1}$$

The *global clustering coefficient* C is then given by

$$C = \frac{1}{n} \sum_{i=1}^{n} C_i. \tag{8.2}$$

A high clustering coefficient C means (in the language of social networks), that the friend of your friend is also likely to be your friend. It also indicates a high redundancy of the network.

The *average path length*, l, is the mean geodesic (i.e., shortest) distance between node pairs in a graph:

$$l = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i \geq j}^{n} d_{ij}, \tag{8.3}$$

where d_{ij} is the geodesic distance from node i to node j.

In Sect. 8.2.3, we will show a model of innovation networks that produces "small-worlds" which combine the two properties of a high clustering coefficient and a small average path length.

In the following sections, we will describe some basic models of economic network theory, where we shall use the definitions and notations introduced above.

8.2.2 The Connections Model

The *connections model* introduced by Jackson and Wolinsky (1996) is of specific interest, since it allows us to compute equilibrium networks analytically. The succeeding models can then be considered as extension of the connections model. Since these models are more complicated than the basic connections model they can, to a large extent, only be studied via computer simulations.[13] Nevertheless, they are of interest because they show a wider range of possible network configurations and associated performance of the agents in the economy.

[13] For the use of computer simulations in economics, see Axelrod and Tesfatsion (2006).

In the following, we discuss the (symmetric) connections model proposed by Jackson and Wolinsky (1996).[14] In this model, agents pass information to other agents to whom they are connected to. Through these links they also receive information from those agents that they are indirectly connected to, that is, through the neighbors of their neighbors, their neighbors, and so on.

The individual incentives to form or severe links determine the addition or deletion of links. Incentives are defined in terms of the utility of the agents which depends on the interactions among agents, i.e., the network. The utility functions assigns a payoff to each agent and this payoff depends on the network the agents are embedded in.

The utility, $u_i(G)$, agent i receives from network G[15] with n agents is a function $u_i : \{G \in \mathcal{G}_n\} \to \mathbb{R}$ with

$$u_i(G) = \sum_{j \neq i} \delta^{d_{ij}} - \sum_{j \in N_i} c, \tag{8.4}$$

where d_{ij} is the number of links in the shortest path between agent i and agent j. $d_{ij} = \infty$ if there is no path between i and j. $0 \leq \delta < 1$ is a parameter that takes into account the decrease of the utility as the path between agent i and agent j increases. $N(i)$ is the set of nodes in the neighborhood of agent i. c is a positive parameter capturing the fact that direct links are costly. This implies that agents want to have short paths to other agents while maintaining as few links as possible.

A measure of the global performance of the network is introduced by its efficiency. The total utility of a network is defined by

$$U(G) = \sum_{i=1}^{n} u_i(G). \tag{8.5}$$

A network is considered efficient if it maximizes the total utility of the network $U(G)$ among all possible networks, $\mathcal{G}(n)$ with n nodes.

Definition 1 A network G is strongly efficient if $U(G) = \sum_{i=1}^{n} u_i(G) \geq U(G') = \sum_{i=1}^{n} u_i(G')$ for all $G' \in \mathcal{G}(n)$

Under certain conditions no new links are accepted or old ones deleted. This leads to the term pairwise stability.

Definition 2 A network G is pairwise stable if and only if

1. for all $e_{ij} \in E(G)$, $u_i(G) \geq u_i(E \backslash e_{ij})$ and $u_j(G) \geq u_j(E \backslash e_{ij})$
2. for all $e_{ij} \notin E(G)$, if $u_i(G) < u_i(E \cup e_{ij})$ then $u_j(G) > u_j(E \cup e_{ij})$

[14] For a good introduction and discussion of related works, we recommend the lecture notes of Zenou (2006). There one can find the proofs given here and related material in more detail. For a general introduction to economic networks, see also Jackson (2006).

[15] In this model the network is undirected.

In words, a network is pairwise stable if and only if (i) removing any link does not increase the utility of any agent, and (ii) adding a link between any two agents, either does not increase the utility of any of the two agents, or if it does increase one of the two agents' utility then it decreases the other agent's utility.

The point here is that establishing a new link with an agent requires the consensus (i.e., a simultaneous increase of utility) of both of them. The notion of pairwise stability can be distinguished from the one of Nash equilibrium,[16] which is appropriate when each agent can establish or remove unilaterally a connection with another agent.

There exists a tension between stability and efficiency in the connections model. This will become clear, after we derive the following two propositions.

Proposition 3 *The unique strongly efficient network in the symmetric connections model is*

1. *the complete graph K_n if $c < \delta - \delta^2$,*
2. *a star encompassing everyone if $\delta - \delta^2 < c < \delta + \frac{n-2}{2}\delta^2$,*
3. *the empty graph (no links) if $\delta + \frac{n-2}{2}\delta^2 < c$.*

Proof 1. We assume that $\delta^2 < \delta - c$. Any pair of agents that is not directly connected can increase its utility (the net benefit for creating a link is at least $\delta - c - \delta^2 > 0$) and thus the total utility, by forming a link. Since every pair of agents has an incentive to form a link, we will end up in the complete graph, where all possible links have been created and no additional links can be created any more.

2. Consider a component of the graph G containing m agents, say G'. The number of links in the component G' is denoted by k, where $k \geq m - 1$, otherwise the component would not be connected. For example, a path containing all agents would have $m - 1$ edges. The total utility of the direct links in the component is given by $k(2\delta - 2c)$. There are at most $\frac{m(m-1)}{2} - k$ left over links in the component that are not created yet. The utility of each of these left over links is at most $2\delta^2$ (it has the highest utility if it is in the second-order neighborhood). Therefore, the total utility of the component is at most

$$k2(\delta - c) + \left(\frac{m(m-1)}{2} - k\right) 2\delta^2. \tag{8.6}$$

Consider a star with m agents. See as an example a star containing four agents in Fig. 8.2. The star has $m - 1$ agents which are not in the center of the star. The utility of any direct link is $2\delta - 2c$ and of any indirect link $(m - 2)\delta^2$, since any agent is two links away from any other agent (except the center of the star). Thus, the total utility of the star is

[16] Considering two agents playing a game (e.g., trading of knowledge) and each adopting a certain strategy. A Nash equilibrium is characterized by a set of strategies where each strategy is the optimal response to all the others.

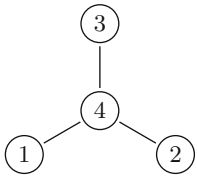


Fig. 8.2 A star encompassing 4 agents

$$\underbrace{(m-1)(2\delta-2c)}_{\text{direct connections}}+\underbrace{(m-1)(m-2)\delta^2}_{\text{indirect connections}}. \tag{8.7}$$

The difference in total utility of the (general) component and the star is just $2(k-(m-1))(\delta-c-\delta^2)$. This is at most 0, since $k \geq m-1$ and $c > \delta-\delta^2$, and less than 0 if $k > m-1$. Thus, the value of the component can equal the value of the star only if $k = m-1$. Any graph with $k = m-1$ edges, which is not a star, must have an indirect connection with a distance longer than 2, and getting a total utility from indirect connections less than $2\delta^2$. Therefore, the total utility of the indirect links will be below $(m-1)(m-2)\delta^2$ (which is the total utility from indirect connections of the star).

If $c < \delta-\delta^2$, then any component of a strongly efficient network must be a star. In a similar fashion, it can be shown (Jackson and Wolinsky, 1996) that a single star of $m+n$ agents has a higher total utility than two separate stars with m and n agents. Accordingly, the star is a strongly efficient network.

3. A star encompassing every agent has a positive value only if $\delta + \frac{n-2}{2}\delta^2 > c$. This is an upper bound for the total achievable utility of any component of the network. Thus, if $\delta + \frac{n-2}{2}\delta^2 < c$ the empty graph is the unique strongly efficient network.

Proposition 4 *In the connections model in which the utility of each agent is given by (8.4), we have*

1. *A pairwise stable network has at most one (non-empty) component.*
2. *For $c < \delta-\delta^2$, the unique pairwise stable network is the complete graph K_n.*
3. *For $\delta-\delta^2 < c < \delta$, a star encompassing every agent is pairwise stable, but not necessarily the unique pairwise stable graph.*
4. *For $\delta < c$, any pairwise stable network that is non-empty is such that each agent has at least two links (and thus is efficient).*

Proof 1. Let us assume, for the sake of contradiction, that G is pairwise stable and has more than one non-empty component. Let u^{ij} denotes the utility of agent i having a link with agent j. Then, $u^{ij} = u_i(G+e_{ij}) - u_i(G)$ if $e_{ij} \notin E(G)$ and $u^{ij} = u_i(G) - u_i(G-e_{ij})$ if $e_{ij} \in E(G)$. We consider now $e_{ij} \in E(G)$. Then $u^{ij} \geq 0$. Let e_{kl} belong to a different component. Since i is already in a component with j, but k is not, it follows that $u^{jk} > u^{ij} \geq 0$, because agent k will receive an additional utility of δ^2 from being indirectly connected to agent i. For similar reasons $u^{jk} > u^{lk} \geq 0$. This means that both agents in the separate component would have an incentive to form a link. This is a contradiction to the assumption of pairwise stability.

2. The net change in utility from creating a link is $\delta - \delta^2 - c$. Before creating the link, the geodesic distance between agent i and agent j is at least 2. When they create a link, they gain δ but they lose the previous utility from being indirectly connected by some path whose length is at least 2. So if $c < \delta - \delta^2$, the net gain from creating a link is always positive. Since any link creation is beneficial (increases the agents' utility), the only pairwise stable network is the complete graph, K_n.
3. We assume that $\delta - \delta^2 < c - \delta$ and show that the star is pairwise stable. The agent in the center of the star has a distance of 1 to all other agents and all other agents are separated by two links from each other. The center agent of the star cannot create a link, since she has already maximum degree. She has no incentive to delete a link either. If she deletes a link, the net gain is $c - \delta$, since there is no path leading to the then disconnected agent. By assumption, $\delta - \delta^2 < c < \delta$, $c - \delta < 0$ and the gain is negative, and the link will not be removed. We consider now an agent that is not the center of the star. She cannot create a link with the center, since they are both already connected. The net gain of creating a link to another agent is $\delta - \delta^2 - c$, which is strictly negative by assumption. So she will not create a link either. The star is pairwise stable.

   The star encompasses all agents. Suppose an agent would not be connected to the star. If the center of the star would create a link to this agent, the net gain would be $\delta - c > 0$ and the benefit of the non-star agent is again $\delta - c > 0$. So both will create the link.

   The star is not the unique pairwise stable network. We will show that for four agents, the cycle, C_4 is also a pairwise stable network. Consider Fig. 8.3.

   If agent 3 removes a link to agent 4, then her net gain is $c - \delta - \delta^3$. For the range of costs of $\delta - \delta^2 < c < \delta - \delta^3 < \delta$, she will never do it. If agent 3 adds a link to agent 1, Fig. 8.4, the net gain is $\delta - \delta^2 < 0$. Thus, for $n = 4$ and $\delta - \delta^2 < c < \delta - \delta^3$, then there are at least two pairwise stable networks: the star and the cycle.
4. For $\delta < c$, the star is not a pairwise stable network because the agent in the center of the star would gain $c - \delta$ from deleting a link. Moreover, it can be

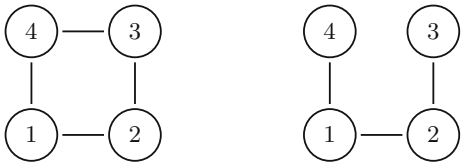


Fig. 8.3 A cycle of four agents (*left*) and the resulting graph (*right*) after the deletion of a link from agent 3 to agent 4

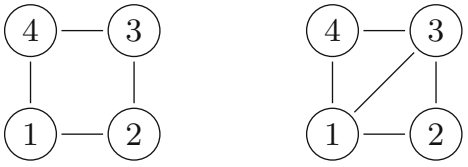


Fig. 8.4 A cycle of four agents (*left*) and the resulting graph (*right*) after the creation of a link from agent 3 to agent 1

shown (Jackson and Wolinsky, 1996) that any connected agent has at least two links.

One can see, from the two propositions described above, that a pairwise stable network is not necessarily efficient. For high cost, i.e., $c > \delta$ there are non-empty pairwise stable networks but they are not efficient.

We now come to the evolution of the network as described in Jackson and Watts (2002). The network changes when agents create or delete a link. At every time step an agent is chosen at random and tries to establish a new link or delete an already existing one. If a link is added, then the two agents involved must both agree to its addition, with at least one of them strictly benefiting (in terms of a higher utility) of the new link. Similarly, a deletion of a link can only be in a mutual agreement. This adding and deleting of links creates a sequence of networks. A sequence of networks created by agents myopically adding and deleting links is called an *improving path*[17] (Jackson and Watts, 2002).

There is a small probability, ϵ, that a mistake occurs (trembling hand) and the link is deleted if present or added if absent. ϵ goes to zero in the long run, $\lim_{t\to\infty} \epsilon(t) = 0$. By introducing this decreasing error ϵ in the agent's decisions, the evolution of the network becomes a Markov process[18] with a unique limiting stationary distribution of networks visited (Jackson and Watts, 2002).

The following definition is important to describe the stochastic evolution of the network.

Definition 5 A network is evolutionary stable if it is in the limiting stationary distribution of networks of the above mentioned Markov process.

We have already investigated the structure and stability of the star, Fig. 8.2, and the cycle, Fig. 8.3. In Jackson and Watts (2002) it is shown that for the case of four agents, the evolutionary stable networks indeed are the stars and cycles. So the network of agents evolves into a quite simple equilibrium configuration.

8.2.3 The Connections Model and Small-World Networks

Carayol and Roux (2005, 2003) propose a model of innovation networks in which networks emerge that show the properties of a "small world."[19] This model is an

---

[17] Each network in the sequence of network updates differs in one link from the previous one. An improving path is a finite set of networks $G_1, ..., G_k$ in which one agents is better off by deleting a link (G_{k+1} has one link less than G_k) or two agents are better off by adding a link (G_{k+1} has one link more than G_k).

[18] A Markov process is a random process whose future states are determined by its present state and not on the past states, i.e., it is conditionally independent on the past states given the present state.

[19] A small-world network combines high clustering (high probability that your acquaintances are also acquaintances to each other) with a short characteristic path length (small average distance between two nodes) (Watts and Strogatz, 1998).

extension of the above described connections model, Sect. 8.2.2, and it uses the same notion of pairwise stability and efficiency.

We now give a sketch of the model. Agents are localized on a cycle and benefit from knowledge flows from their direct and indirect neighbors. Knowledge transfer decays along paths longer than one link. This means that less knowledge is received, the longer the path between the not directly connected agents is. The transfer rate is controlled by an exogenous parameter, δ. Each agent has a probability to innovate that is dependent on her amount of knowledge. The knowledge level of an agent is dependent on two factors. (i) the in-house innovative capabilities of the agent and (ii) the knowledge flows coming directly from the neighbors or indirectly (with a certain attenuation factor) from those agents that are connected to the neighbors.

Agent i supports costs, for direct connections which are linearly increasing with geographic distance, that is the distance on the cycle on which they recede. Agent i's utility u_i at a time t is given by the following expression:

$$u_i(G(t)) = \sum_{j \neq i} \delta^{d_{ij}} - c \sum_{j \in N_i} d'_{ij}, \tag{8.8}$$

where d_{ij} is the geodesic distance between agent i and agent j. $\delta \in (0, 1)$ is a knowledge decay parameter and $\delta^{d_{ij}}$ gives the payoffs resulting from the direct or indirect connection between agent i and agent j. c is a positive constant. d'_{ij} describes the geographic distance between agent i and agent j, that is the distance on the cycle. This is the main difference in the assumptions compared to the connections model discussed in Sect. 8.2.2.

Agents are able to modify their connections. This is where the network becomes dynamic. Pairs of agents are randomly selected. If the two selected agents are directly connected they can jointly decide to maintain a link or unilaterally decide to sever the link. If they are not connected, they can jointly decide to form a link. The decision is guided by the selfishness of the agents, which means that they only accept links from which they get a higher utility.

The stochastic process of adding links to the network can be seen as a Markov process where each state is the graph structure at a certain time step. The evolution of the system is a discrete time stochastic process with the state space of all possible graphs. A small random perturbation where the agents make mistakes in taking the optimal decision to form a link or not is introduced. Agents are making errors with a probability $\epsilon(t)$. This error term decreases with time, $\lim_{t \to \infty} \epsilon(t) = 0$.

The introduction of ϵ enables us to find long-run stationary distributions that are independent of initial conditions (the ergodicity of the system) (Jackson and Watts, 2002). Simulations are used in order to find these stationary distributions. Agents are forming and severing links until the network reaches a pairwise stable configuration, where the agents have no incentive to create or delete links any more. The set of stochastically stable networks selected in the long run is affected by the rate of knowledge transfer, δ. The authors find critical values of this parameter for which stable "small-world" networks are dynamically selected (Carayol and Roux,

2005). This is the main difference in the resulting equilibrium network structure to the connections model, in which simpler network configurations are obtained.

8.2.4 Introducing Heterogeneous Knowledge

Ricottilli (2005, 2006) studies the evolution of a network of agents that improve their technological capabilities through interaction while knowledge is heterogeneously distributed among agents. In addition to the sharing of knowledge, each agent is assumed to have an "in-house" innovative capability. Considerable effort is necessary for this "in-house" research and as research is not always successful, it is assumed to change stochastically.

An agent i's innovation capability, V_i, is given by

$$V_i(t) = \sum_{j=1}^{n} a_{ij} b_{ij}(t) V_j(t) + C_i(t) \tag{8.9}$$

with an economy consisting of n agents. a_{ij} =const. is the broadcasting capacity of agent j to agent i and $a_{ii} = 0$ since no agent can broadcast information to herself. The matrix A with elements a_{ij} indicates the total technological information broadcasting capability of this economy. The proximity matrix elements $b_{ij}(t)$ are either 0 or 1 according to whether agent j is identified from agent i as an information supplier. This is the neighborhood of agent i. $C_i(t) \in (0, 1)$ is the in-house capability of agent i. This is a stochastic variable.

Each agent i assesses the value of knowledge of its neighbors (where $b_{ij} \neq 0$), which are the addends of the first term in (8.9). From this function the least contributing one, denoted by $\gamma_i(t)$, is selected.

$$\gamma_i(t) = \min_{1 \leq j \leq N} \{a_{ij} b_{ij}(t-1) V_j(t-1)\}. \tag{8.10}$$

In a random replacement procedure (search routine) an agent selects either its neighbors and second neighbors (local, weak bounded rationality) or the entire economy excluding its first and second neighbors (global, strong bounded rationality). By doing so, agent i assigns a new member j to the set of information suppliers, setting b_{ij} from 0 to 1. This selection is only accepted if

$$V_i(t) > V_i(t-1). \tag{8.11}$$

The population of agents is classified according to the size of the set of other agents by which they are observed. Global paradigm setters are agents that are observed by almost all agents in the economy. Local paradigm setters are observed by almost all agents belonging to the same component.

Simulations of the evolution of the network show that stable patterns emerge. When the knowledge-heterogeneity of the economy is not very high, global paradigm

setters emerge. For high levels of heterogeneity the economy becomes partitioned into two separate halves. In each homogeneous one, local paradigm setters emerge. Ricottilli (2006, 2005) shows that the highest technological capabilities are achieved neither with a local search routine in which only the second neighbors are included nor in global search routines that span the whole economy. Rather, a combination of both improves the system's innovative efficiency the most.

8.2.5 Emerging Cyclic Network Topologies

When studying multi-sector trading economies and input–output systems, Rosenblatt (1957) already identified the importance of circular flows and "feedback" input dependencies between industries (realized by subgraphs called "cyclic nets"). A sufficient condition for a strongly connected network (in which there exits a path from every agent to every other agent and that has an irreducible adjacency matrix) is the existence of a cycle. Subsequent works (Baldry and Ghosal, 2005; Maxfield, 1994) have further incorporated cyclic network topologies (or strong connectivity which implies the existence of a cycle in the network) for the existence of a competitive economy. More recently, Kim and Wong (2007) studied a generalized model of Bala and Goyal (2000) and found that the equilibrium networks consist of cycles (so-called "sub-wheel partitions").

In the following sections, we will focus on some recent network models of knowledge transaction and innovation (the creation of new knowledge) in which cyclic interactions of agents emerge. In Sect. 8.3, we will study a new model of evolving innovation networks. Similarly to the above mentioned authors, we find that the existence of equilibrium networks with a positive knowledge production depends critically on the existence of cycles in the network.

8.2.5.1 Production Recipes and Artifacts

We start by reviewing a model by Lane (2005) in which agents try to produce and sell artifacts. These artifacts can be manufactured according to a production recipe. Such a recipe can either be found independently or through the sharing of knowledge with other agents, which in turn can lead to an innovation, that is the discovery of a new recipe.

Let us denote with r_{ik} the kth recipe of agent i. There is an external environment which consists of external agents (customers) and artifacts which are not produced in the model. At each time period t one agent i is randomly chosen. Then the following steps are taken:

1. The agent tries to get the input required for each recipe r_{ik}. If it is in the agent's own stock then she can produce immediately. If it is not, she buys it from another agent and if it cannot be bought she moves to another recipe.

2. The agent chooses a goal, i.e., the product she wants to produce (one that gives high sales). Therefore she has to find the right recipe for the goal. She produces the product if a successful recipe is found. This can be achieved in two ways. The agent either can try to innovate by herself or she can try to innovate together with another agent.
3. The wealth of agent i at time $t+1$, $w_i(t)$, is calculated according to

$$w_i(t+1) = w_i(t) + \sum_{k=1}^{N_k} n_{ik}(t) - w_i(t) \sum_{l=1}^{N_l} p_{il} c_{il} - \lambda w_i(t), \tag{8.12}$$

   where N_k is the number of products sold, $n_{ik}(t)$ is the number of units sold of product k belonging to agent i, p_{il} is the number of products produced with recipe r_{il}, and c_{il} is the production cost.
   The last term $-\lambda w_i(t)$ guarantees that the wealth of an agent that has not sold any products and does not have any active recipes vanishes.
4. The recipes that could not be successfully used to produce products are canceled.
5. The set of acquaintances of an agent is enlarged. This is possible when two or more agents that have goals which are close in artifact space (i.e., they require similar inputs) cooperate to produce that artifact.
6. With a certain probability dead agents are substituted.

The basic dynamics, absent innovation, is one of production and sales, where the supply of raw materials is external as well as final product demand. There are two main differences to most agent-based innovation models. First, here the agents try to develop new recipes in order to produce products with high sales, as opposed to many agent-based models where the generation of novelty is driven by some stochastic process. Second, in simulations Lane (2005) shows that the network of customers and suppliers often forms closed, self-sustaining cycles.

8.2.5.2 An Autocatalytic Model with Hypercycles

Padgett (1996) and Padgett et al. (2003) introduces an autocatalytic model, based on a hypercycle[20] model. Here agents are represented as skills and these skills are combined in order to produce. Skills, like chemical reactions, are rules that transform products into other products.

In the following, we will give a short overview of the model.[21] There are two main aspects in the dynamic interactions between the agents: The process of production and the process of learning. The process of production includes three entities: skills, products, and agents. Skills transform products into other products. The skills are features of the agents. On a spatial grid the agents are arrayed with periodic

[20] A hypercycle is a system which connects self-replicative units through a cycle linkage (Eigen and Schuster, 1979).

[21] This agent-based model is publicly available on the website http://repast.sourceforge.net/examples/index.html under the application module hypercycle.

boundaries. Each agent has eight possible neighbors. At each asynchronous iteration a random skill is chosen. An agent with that skill randomly chooses an input product. If this product fits to the skill then the product is transformed. The transformed product is passed randomly to the neighbors of the agent. If the trading partner has the necessary skill it transforms the product further and passes it on. If the agent doesn't have the compatible skill, the product is ejected into the output environment and a new input product is selected.

One can look at the production process from a wider perspective. An input product comes from the environment, then passes through production chains of skills until it is passed back as output to the environment. These chains self-organize because of a feedback mechanism of the agents. This mechanism is learning through the trade of products.

The process of learning is modeled as learning by doing. If a skill transforms a product and then passes it on to another transforming skill, then the skill is reproduced (learned). Whenever one skill is reproduced anywhere in the system then another one is deleted at random to keep the overall number of skills constant. The agents are able to learn new skills by practicing them and they can forget skills they did not use for a certain period of time. This procedure of learning introduces a feedback mechanism. When an agent loses all its skills, then it is assumed to never recover.

In Padgett (1996) and Padgett et al. (2003) the emergence of self-reinforcing hypercycle production chains is shown. In these hypercycles agents reproduce each other through continuous learning. Such cycles generate a positive growth effect on the reproduction of skills. Thus, even in a competitive environment the sharing of knowledge is crucial to the long-run performance of the system.

8.3 A New Model of Evolving Innovation Networks

8.3.1 Outline of the Modeling Framework

In this section, we study the evolution of networks of agents exchanging knowledge[22] in a novel framework. The network can evolve over time either, by an external selection mechanism that replaces the worst performing agent with a new one or, by a local mechanism, in which agents take decisions on forming or removing a link. In the latter case, we investigate different modalities of interaction between agents, namely bilateral interactions, representing R&D collaborations (Hagedoorn et al., 2006, 2000) or informal knowledge trading (Von Hippel, 1987), versus unilateral interactions (similar to Bala and Goyal (2000) agents decide unilaterally whom to connect to), representing a generalization of informal knowledge trading. We further study the impact of varying costs for maintaining links and the

[22] See also the chapter of Robin Cowan and Nicolas Jonard in this book as well as Cowan and Jonard (2004), Cowan et al. (2004).

impact of augmenting or diminishing effects on the value of knowledge with the number of users associated with different types of knowledge. Our model exhibits equilibrium networks and we compare their structure and performance. Similar to the models discussed in the last section we will show that cyclic patterns in the interactions between agents play an important role for the stability (permanence) and performance of the system.

We study different assumptions on the behavior of agents. In the most simple case, denoted by *Extremal Dynamics*, agents form links at random and, through an external market selection mechanism, the worst performing agent (this is where the denotation extremal stems from) is replaced with a new one. In this setting, agents are completely passive and they are exposed to a least-fit selection mechanism.

In a more realistic setting, called *Utility Driven Dynamics*, agents choose with whom to interact, but their behavior is still boundedly rational and does not consider strategic interaction. The way in which agents create or delete links to other agents is a trial and error process for finding the right partner. Here we study two different modes of interaction. In the first interaction mode, agents are creating bilateral links. Bilateral links represent formal R&D collaborations among agents (Hagedoorn et al., 2000), or informal knowledge trading (Von Hippel, 1987). In the second interaction mode, agents are transferring knowledge unilaterally, which means that one agent may transfer her knowledge to another but the reverse is not mandatory. In this setting, the transfer of knowledge may be reciprocated, but knowledge can also be returned from a third party. In the latter case, we speak of indirect reciprocity. If knowledge is transferred unilaterally, the innovation network can be represented as a directed graph comprising unilateral links, while if all interactions are bilateral, the innovation network can be represented as an undirected graph.

In the setting of unilateral links, we also investigate the impact of additional benefits from network externalities. These benefits consider specific structural properties of networks which have an augmenting effect on the value of knowledge. We study two different types of network properties which increase the value of knowledge. We call these types *Positive Network Externalities*. The first Positive Network Externality considers the factor that the more the centrality of an agent rises with the creation of a link, the higher is the benefit from that link. A high centrality indicates that an agent is connected to other agents through short paths. This means that, when knowledge travels along short distances between agents, it has a higher value than knowledge that has to be passed on between many agents. This effect can be captured by introducing an attenuation of knowledge with the distance it has to travel (by getting passed on from one agent to the next) until reaching an agent. The second Positive Network Externality captures an opposite effect when knowledge is passed on from one agent to the next. Here the value of knowledge increases with the number of transmitters (who are also user) of that knowledge. More precisely, we assume that feedback loops create an increase in the value of knowledge of the agents that are part of the loop. The more the agents absorb and pass on knowledge the higher is the value of that knowledge. This means that a link that is part of a long

feedback loop increases the value of the knowledge passed on from one agent to the next.[23]

We can summarize the different settings that are studied in this section as follows. We investigate the performance and evolution under the two aforementioned assumptions on the behavior of agents, namely Extremal Dynamics and Utility Driven Dynamics. In the latter setting, we further study the effect of different modes of interaction, i.e., bilateral and unilateral knowledge transactions among agents. When studying unilateral interaction among agents, we introduce different augmenting processes on the value of knowledge depending on the structure of the network, called Positive Network Externality. We study the impact of an attenuation of the value of knowledge by the distance from the giver to the receiver as well as the contrary effect of an increase of the value of knowledge with the number of users of that knowledge depending on the type of knowledge under investigation. Finally, we discuss the networks obtained under these different settings with respect to their topologies and performance.

8.3.2 Bilateral Versus Unilateral Knowledge Exchange

We interpret bilateral interactions as R&D collaborations on a formal or informal basis (Hagedoorn et al., 2000). Both parties involved share their knowledge in a reciprocal way, that means one agent is giving knowledge to another if and only if the other agent is doing this as well and both agents benefit from this transaction.

We then compare bilateral interaction with the case of agents sharing knowledge in a unidirectional way with other agents. They then maintain only those interactions that are in some form reciprocated (and this way lead to an increase in their knowledge levels after a certain time) but not necessarily from the agent they initially gave their knowledge to (indirect reciprocity). The latter is referred to unilateral knowledge exchange which can be seen as a generalization of informal knowledge trading.

In the case of informal knowledge trading, agents exchange knowledge if both strictly benefit. Instead, in the case of generalized informal knowledge trading, one agent transfers knowledge to another one without immediately getting something back. After a certain time (time horizon T) an agent evaluates its investment by assessing its total net increase in knowledge. By introducing unilateral knowledge exchange we relax two requirements: (i) we do not require that the investment in sharing ones knowledge has to be reciprocated instantaneously and in a mutually concerted way. And (ii) the reciprocation does not necessarily have to come from the same agent. With this generalization we introduce that (i) agents have only limited information on the value of knowledge of others and on the network of interactions. (ii) Agents proceed in a trial and error fashion to find the right partners

[23] We study closed loops, because we assume that knowledge issued from one agent has to return to that agent in order for her to take advantage of this added value of knowledge (created by the multiplicity of other users).

for exchanging their knowledge. In this setting reciprocity emerges either directly or indirectly.

If the total knowledge level of an agent at the time horizon T is higher than it was when the agent started to share her knowledge with another agent, this interaction is evaluated beneficial, otherwise it is not. Only if the interaction is evaluated beneficial, the agent continues sharing its knowledge with the other agent, otherwise it stops the interaction. This procedure requires only limited information on the other agents, since the agent cares for its own total increase in knowledge and does not need to evaluate the individual knowledge levels of others. We describe this link formation mechanism in more detail in Sect. 8.3.9.3.

8.3.3 Unilateral Knowledge Exchange and Reciprocity

If the interaction of agents are unilateral then agents invest into innovation by sharing knowledge with other agents. An investment is an advance payment with the expectation to earn future profits. When one agent transfers knowledge to another one without immediately getting something back, this can be regarded as an investment. There are usually two ways in which an investment can be expected to bring in reasonable returns.

One way is the creation of contracts. As a precondition for contracts technologies must be protectable by intellectual property rights (IPR). Otherwise agents cannot trade them (once the technology is offered, i.e., made public, everybody can simply copy it and there is no more need to pay for it). Contracts must be binding and complete (Dickhaut and Rustichini, 2001). The contract has to be binding or agents may not meet their agreement after the payment has been made. It has to be complete, or uncertain agreements may lead agents to interpret it in a way most favorable to their position and this can cause agents to retreat from the contract.

The requirements for contracts can be difficult to realize. Another way is to expect reciprocative behavior to the investment. The beneficiary can either directly or indirectly reciprocate the benefit. Direct reciprocity means to respond in kind to the investor, and indirect reciprocity to reward someone else than the original investor.

One of the possible explanations for reciprocal behavior (Bolton and Ockenfels, 2000; Fehr and Fischbacher, 2003; Fehr and Schmidt, 1999; Nowak and Sigmund, 1998, 2005) (see, e.g., Dieckmann, 2004, for a survey) is to assume the existence of reputation. Agents believe that if they invest into another agent they will increase their reputation and then realize a reasonable return coming back to them directly or indirectly ("strategic reputation building").

In reality, only partial information about reputation is available and experimental works show that, even in the absence of reputation, there is a non-negligible amount of reciprocal cooperative behavior among humans (Bolton et al., 2005). As Dickhaut and Rustichini (2001) put it, "...investment occurs even though agents cannot create binding contracts nor create reputation." Thus, agents invest into each other by transferring their knowledge even if they cannot immediately evaluate the benefit from this investment.

We assume that agents are not a priori reciprocating if they receive knowledge from others. But they perceive that interactions that are reciprocated in some way are beneficial (increasing their own knowledge) and these are the interactions that they maintain in the long run.

The problems associated with bilateral exchange of knowledge (direct reciprocity) and experimental evidence suggest that unilateral knowledge exchange, in which indirect reciprocity can emerge, is a relevant mode of interaction between agents. Moreover, the fact that interactions between anonymous partners become increasingly frequent in global markets and tend to replace the traditional long-lasting mutual business relationships poses a challenge to economic theory and is one of the reason for the growing interest about indirect reciprocity in the economic literature.

8.3.4 Indirect Reciprocity, Directed Graphs, and Cycles

An R&D network can be described as a graph in which agents are represented by nodes, and their interactions by directed links. Indeed, as mentioned above, if agent i transfers knowledge to agent j (e.g., by providing a new technology), the reverse process, i.e., that agent j in turn transfers knowledge to i, is in principle not mandatory. This means that the links representing the transfer of knowledge are directed. The underlying graph can be represented by an adjacency matrix, $\mathbf{A}$ with elements $a_{ij} \in [0, 1]$, which is not symmetric, $a_{ij} \neq a_{ji}$. In other words, directed means that we distinguish the pairs (i, j) and (j, i) representing the links from i to j and from j to i, respectively. On the other hand, if the adjacency matrix is symmetric, it means that any two agents are connected both by a link from i to j and by a link from j to i. We say, in this case, that they are connected by a *bidirectional* link. Notice that the symmetry also implies that the two links have identical weights.

Reciprocity requires the presence of cycles. In particular, direct reciprocity corresponds to a cycle of order $k = 2$, while indirect reciprocity corresponds to a cycle of order $k \geq 3$ (see Figs. 8.5 and 8.6). Therefore, the emergence and permanence of

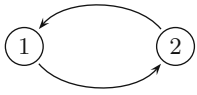


Fig. 8.5 A cycle of length 2 represents an interaction between agents that is direct reciprocal

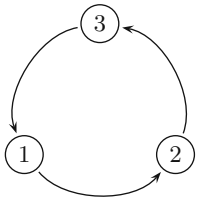


Fig. 8.6 A cycle of length 3 (or longer) represents an interaction between agents that is indirect reciprocal

direct/indirect reciprocity is deeply connected to the existence of cycles and in the graph of interactions.

8.3.5 Formal Modeling Framework

In this section, we formalize the general framework for the investigation of evolving networks of selfish agents engaged in knowledge production via the sharing of knowledge. In such a framework, it is possible to investigate how the emergence and permanence of different structures in the network is affected by (1) the form of the **growth function** of the value of knowledge, (2) the length of **time horizon** after which interactions are evaluated and (3) the **link formation/deletion rules**. At a first glance, this problem includes a multitude of dimensions, as the space of utility functions and link formation/deletion rules is infinite. However, some natural constraints limit considerably the number of possibilities and make a systematic study possible. In the following, we present the general framework. We then focus on a subset of the space of utility functions and link formation rules. For these, we present briefly some analytical results, but since the value of knowledge of an agent is assumed to be a non-linear function of the neighboring agents, we illustrate them in terms of computer simulations. We finally summarize the results and discuss them in relation to the context of innovation.

We consider a set of agents, $N = \{1, ..., n\}$, represented as nodes of a network G, with an associated variable x_i representing the value of knowledge of agent i. The value of knowledge is measured in the units of profits an agent can make in a knowledge-intensive market. It has been shown that the growth of such knowledge-intensive industries is highly dependent on the number and intensity of strategic alliances in R&D networks (Powell and Grodal, 2006). In our model we bring the value of knowledge of an agent, denoted by $x_i(t)$, at time t in relation with the values of knowledge of the other agents $x_j(t)$ at time t in the economy, that are connected to the current agent i. A link from i to j, e_{ij}, takes into account that agent i transfers knowledge to agent j. The idea is, that through interaction, agents transfer knowledge to each other which in turn increases their values of knowledge.

We focus here only on the network effects on the value of knowledge of an agent. We therefore neglect the efforts of agents made to innovate on their own, without the interaction with others.[24] In particular, we assume that the growth of the value of knowledge of agent i depends only on the value of knowledge of the agents, j, with outgoing links pointing to him (those who transfer knowledge to her), $j \in V(G)$ such that $e_{ji} \in E(G)$.

[24] The "in-house" R&D capabilities of an agent could be introduced by an additional (stochastic) term $S_i(x_i)$. Similar to Ricottilli (2006) in Sect. 8.2.5, $S_i(x_i)$ captures the innovation activities of agent i without the interaction with other agents. We assume that the "in-house" capabilities of agents are negligible compared to network effects. Thus, we concentrate only on network effects on the increase or decrease in the value of knowledge.

In a recent study on the dynamics of R&D collaboration networks in the US IT industry, Hanaki et al. (2007) have shown that firms form R&D collaborations in order to maximize their net knowledge (information) flow. Cassiman and Veugelers (2002) suggested that this knowledge flow can be decomposed into incoming and outgoing spillovers capturing the positive and negative effects of R&D collaborations.

We try to incorporate these positive and negative effects into a differential equation that describes the change (increase or decrease) in the value of knowledge of an agent through R&D collaborations with other agents. We assume that the knowledge growth function can be decomposed into a decay term, a benefit term, and a cost term depending on the interactions of an agent. The equation for knowledge growth reads

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -D_i(x_i) + B_i(\mathbf{A}, \mathbf{x}) - C_i(\mathbf{A}, \mathbf{x}), \tag{8.13}$$

where

$\dot{x}_i \ldots$ growth of the value of knowledge of agent i
$\mathbf{A} \ldots$ adjacency matrix (representing the network)
$\mathbf{x} \ldots$ vector of agents' values of knowledge
$D_i(\dot{x}_i) \ldots$ knowledge decay (obsolescence of knowledge)
$B_i(\mathbf{A}, \mathbf{x}) \ldots$ interaction benefits of agent i
$C_i(\mathbf{A}, \mathbf{x}) \ldots$ interaction costs of agent i

$\mathbf{B} \geq 0$ and $\mathbf{C} \geq 0$ are benefit and cost terms, respectively, while $\mathbf{D} \geq 0$ is a decay term which includes the fact that a technology loses its value over time (obsolescence). In our setting, only through R&D collaborations with other agents, an agent can overcome the obsolescence of knowledge. This ensures that agents who do not interact with others have necessarily vanishing value of knowledge in our model (since $B_i = C_i = 0 \Leftrightarrow a_{ij} = 0 \, \forall j$ and thus $\dot{x}_i < 0$). In other words, we investigate an R&D intensive economy in which an agent's performance is critically depending on its R&D collaborations.

Interaction is described by the adjacency matrix $\mathbf{A}$ that contains the elements a_{ij} in terms of 0 and 1. This dynamics can be interpreted as a *catalytic network* of R&D interactions (passing a technology to another agent, R&D collaborations), where the different agents are represented by nodes, and their interaction by links between these nodes, cf. Fig. 8.1. More precisely,

$$a_{ij} = \begin{cases} 1 & \text{if agents } i \text{ transfers knowledge to agent } j \\ 0 & \text{otherwise} \end{cases} \tag{8.14}$$

We noted already that the network of interactions is modeled on a directed graph, which means that the adjacency matrix is not generally symmetric: $a_{ij} \neq a_{ji}$.

The benefit term, $B_i(\mathbf{x}, \mathbf{A})$, accounts for the fact that an agent's value of knowledge increases by receiving knowledge from other agents. The cost term, $C_i(\mathbf{x}, \mathbf{A})$, accounts for the fact that transferring knowledge to other agents is costly. Such a cost can vary in magnitude depending on the technological domain, but, in general, to make someone else proficient in whatever new technology requires a non-null effort.

In the following, we will further specify the growth of the value of knowledge in (8.13). We will make simple assumptions on benefits, $B_i(\mathbf{x}, \mathbf{A})$, and costs, $C_i(\mathbf{x}, \mathbf{A})$, which allow us to derive some analytical results and thus gain some insight on the behavior of the system.

8.3.6 Pairwise Decomposition

Networks are sets of pairwise relationships. In systems of interacting units in physics, a superposition principle holds, such that the force perceived by a unit is due to the sum of pairwise interactions with other units. Similarly, one could think of decomposing both benefits and costs of each agent i in a sum of terms related to the agents j interacting with i. However, this would imply to ignore network externalities[25] (it is very important to note this fact). We will see in the following that externality does play an important role. So far, in the literature on complex networks one has considered only the pairwise interaction term, while the literature on economic networks has focused on some simple externalities such as the network size, or the distance from other agents, see Sect. 8.2.5.

Our approach is to assume that benefit and cost are each decomposable in two terms: one term related to the direct interaction, further decomposable in pairwise terms, and another term related to externality (corresponding to positive and negative externality):

$$B_i(\mathbf{A}, \mathbf{x}) = \sum_j b_{ji}(x_j, a_{ji}) + b^e_{ji}(x_j, \mathbf{A}), \tag{8.15}$$

$$C_i(\mathbf{A}, \mathbf{x}) = \sum_j c_{ij}(x_i, a_{ij}) + c^e_{ij}(x_i, \mathbf{A}, \mathbf{x}), \tag{8.16}$$

where b stands for benefit, c for cost, e for externality. The effect of network externalities will be explained in Sect. 8.3.9.9. Benefit, $b_{ji}(x_j, a_{ji})$, and cost, $c_{ij}(x_i, a_{ij})$, terms are monotonically increasing with the value of knowledge, x_i . They have the following properties:

[25] In our model we define a network externality as a function of the network that affects the utility of an agent.

$$b_{ji}(x_j, a_{ji}) = \begin{cases} 0 & \text{if } a_{ji} = 0 \vee x_j = 0 \\ > 0 & \text{if } a_{ji} = 1 \wedge x_j > 0 \end{cases}, \tag{8.17}$$

$$c_{ij}(x_i, a_{ij}) = \begin{cases} 0 & \text{if } a_{ij} = 0 \vee x_i = 0 \\ > 0 & \text{if } a_{ij} = 1 \wedge x_i > 0. \end{cases} \tag{8.18}$$

We assume that benefits are linear functions of the value of knowledge of agent i which shares its knowledge with agent j. We introduce the linear assumption $b_{ji}(x_j, a_{ji}) = a_{ji}x_j$.

In the most simple case costs for transferring knowledge can be neglected, $c_{ij}(x_i, a_{ij}) = 0$. This means that knowledge is fully codified (Foray, 2004) and it can be transferred to another agent without any losses. Further, null costs imply that knowledge is non-rivalrous, meaning that the value of knowledge is not reduced by the use of that knowledge by another agent. When costs are neglected, the growth in the value of knowledge of agent i is given by the following equation (the case of **Null Interaction Costs**, further analyzed in Sect. 8.3.8).

$$\frac{dx_i}{dt} = -dx_i + b\sum_{j=1}^{n} a_{ji}x_j. \tag{8.19}$$

In more realistic setting, costs cannot be neglected. In order to come up with a reasonable expression for these costs, we make some further assumptions. We assume that the higher the value of knowledge of an agent is, the more complex it is. Moreover, the more the complex knowledge is, the more difficult it is to transfer it (Rivkin, 2000; Sorenson et al., 2006). The coordination and processing capabilities of agents are constrained ("managerial breakdown"). Thus, the more complex knowledge gets the higher are the costs for transferring it. The cost, $c_{ij}(x_i, a_{ij})$, for transferring knowledge from agent i to agent j is an increasing function of the value of the knowledge that is to be transferred, x_i. We assume that costs increase by more than a proportional change in the value of knowledge that is being transferred.

$$c_{ij}(\alpha x_i) > \alpha c_{ij}(x_i). \tag{8.20}$$

This characteristic is closely related to decreasing returns to scale and convex cost functions.[26] The most simple setting for such a function is a quadratic term of the form $c_{ij}(x_i, a_{ij}) = ca_{ij}x_i^2$. The growth in the value of knowledge of agent i is then governed by the following equation (the case of **Increasing Interaction Costs**, further analyzed in Sect. 8.3.8):

[26] In the standard economic theory of the agent the extent to which a given input can increase output is usually assumed to be a decreasing function of the input. The output increases at a decreasing rate when the input in production increases (Hausman, 2003).

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\mathrm{d}x_i + b\sum_{j=1}^{n} a_{ji}x_j - c\sum_{j=1}^{n} a_{ij}x_i^2. \tag{8.21}$$

This is an ordinary differential equation with a linear decay, a linear benefit, and quadratic costs.

Equation (8.21) can be interpreted as an extension of a logistic equation. In a complete graph every agent shares her knowledge with every other agent. Starting with the same initial values, this symmetry implies that all knowledge values are identical, i.e., $x_i = x$, equation (8.21) then becomes

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}t} &= -dx + b(n-1)x - c(n-1)x^2 \\ &\xrightarrow{\frac{d}{b}\ll n} b(n-1)x\left(1 - \tfrac{c}{b}x\right)\end{aligned}. \tag{8.22}$$

Equation (8.22) is similar to the logistic function $\dot{x} = \alpha x(1 - \frac{x}{\beta})$ with parameters $\alpha = b(n-1)$ and $\beta = b/c$.

In the following section, we relate the topology (cyclic topologies in particular) of the network with the long-run values of knowledge of the agents.

8.3.7 Non-permanence of Directed Acyclic Graphs

The study of the relation between the performance of an economy and the underlying network of interactions has already a long tradition, see, e.g., Rosenblatt (1957) ("cyclic nets"). More recently Maxfield (1994) has shown that the existence of a competitive equilibrium is related to the strong connectedness of the network of relations between users and producers in a market economy. Strong connectedness means that there exists a closed walk or a cycle in the network. On the other hand, if there does not exist such a cycle, then the network is not strongly connected. In a similar way in our model strong connectedness is critically influencing the performance of the agents. In the main result of this Sect. 8.11, we show that in our model all values of knowledge vanish if the underlying network of interactions does not contain a cycle.

For the general equation (8.13) we can identify the topology of the network in which agents cannot be permanent. Hofbauer and Sigmund (1998) give the following definition of permanence:

Definition 6 A dynamical system is said to be permanent if there exists a $\delta > 0$ such that $x_i(0) > 0$ for $i = 1, ..., n$ implies $\lim_{t\to\infty} \inf x_i(t) > \delta$.

We generalize the above notion of permanence to networks in which the nodes have a state variable attached (that depends on the state variable of their neighbors). If the state variables are non-zero the network is said to be permanent, otherwise it is not. This is justified, since nodes with vanishing state variables have no interactions at all.

First, we have to introduce the definition of graphs which do not contain any closed walks or cycles.

Definition 7 A directed acyclic graph is a directed graph with no directed cycles.

More general, if a graph is a directed acyclic graph then it does not contain a closed walk.

For several proofs in this section we need the following lemma (denoted by the *comparison principle* (Khalil, 1995)).

Lemma 8 *If we consider two time-dependent variables, $x(t)$ and $y(t)$ with different growth functions $g(x)$ and $f(x)$ (continuous, differentiable)*

$$\dot{x} = f(x) \tag{8.23}$$

$$\dot{y} = g(x) \tag{8.24}$$

$$x(0) = y(0) \tag{8.25}$$

and $g(x) \geq f(x)$ then it follows that $y(t) \geq x(t)$. Similarly, if $g(x) \leq f(x)$ then $y(t) \leq x(t)$.

Proof Using Cauchy's mean value theorem for the two continuous, differentiable functions, $x(t)$ and $y(t)$, we have

$$\frac{x'(\tau)}{y'(\tau)} = \frac{x(t) - x_0}{y(t) - y_0} \geq 1 \tag{8.26}$$

with $\tau \in (0, t)$. The inequality holds since $x'(\tau) = f(x(\tau)) \geq y'(\tau) = g(y(\tau))$ $\forall \tau \in (0, t)$. It follows that

$$x(t) - x_0 \geq y(t) - y_0 \tag{8.27}$$

$$x_0 = y_0 \tag{8.28}$$

and thus $x(t) \geq y(t)$. □

If a network is a directed acyclic graph then it does not contain a closed walk. For a directed acyclic graph we can make the following observation.

Proposition 9 *In every directed acyclic graph, there is at least one node v with no incoming links, i.e., a source.*

Proof (Godsil and Royle, 2001) We give a proof by contradiction. We assume that every node has an incoming link. We start with some node u and find an incoming link (x, u) – by assumption every node has at least one incoming link. We go to the destination of the link, x. Again, we can find an incoming link (y, x). We then proceed to node y. There is an incoming link (z, y). We consider node z. After at most $n + 1$ steps, we will visit some node in the graph twice. This is a contradiction to the assumption that the graph is acyclic. □

We can partition the nodes in the network into specific sets which take into account from which other nodes there exists an incoming path to these nodes. We will show that this is important to obtain a result on the permanence of the values of knowledge of the agents.

Definition 10 We denote the set of sources of a directed acyclic graph G by S_0. We say that S_0 is the 0-th order sources of G. The nodes that have only incoming links from S_0 are denoted by S_1, the 1-st order sources of G. We consider the graph $G \backslash S_0$. The nodes that have only incoming links from S_1 in $G \backslash S_0$ (obtained by removing the nodes in S_0 and their incident links from G) are denoted by S_2. Accordingly, the nodes having only incoming links from S_{k-1} in the graph $G \backslash (S_{k-2} \cup \ldots \cup S_0)$ are denoted by S_k, the k-th order sources of G, where $k \leq n$.

We can have at most n such sets in the graph G with n nodes. In this case G is a directed path P_k. Moreover, we have that

Proposition 11 *The nodes in a directed acyclic graph G can be partitioned in the sets S_0, S_1, ..., S_k, $k \leq n$ defined in (8.10).*

Proof From Proposition (9) we know that the directed acyclic graph G has at least one source node. All the sources form the set S_0. If we remove the nodes in S_0 (as well as their incident links) from G then we obtain again a directed acyclic graph $G_1 := G \backslash S_0$ (since the removal of links cannot create cycles). Therefore, Proposition (9) also holds for G_1. We consider the source nodes in G_1. These nodes have not been sources in G and they have become sources by removing the incident links of the sources in G. Thus, the source nodes in G_1 have only incoming links from nodes in S_0. Further on, the sources in G_1 form the set S_1. We can now remove the nodes S_1 from G_1 and obtain the graph G_2 with new sources S_2. We can consider the k-th removal of source nodes. We make the induction hypothesis that the sources of G_{k-1} form the set S_{k-1}. Removing the sources from G_{k-1} gives a directed acyclic graph G_k which contains the sources S_k. One can continue this procedure until all nodes have been put into sets $S_0, S_1, \ldots, S_k$ with at most $k = n$ sets. □

There exists a relationship between the set (defined in (8.8)) a node belongs to and the nodes from which there exists an incoming path to that node.

Corollary 12 *Consider a node $i \in S_j$. Then there does not exist a path from nodes $k \in S_m$, $m \geq j$, to node i. Conversely, node i has only incoming path from nodes in the sets S_0, ..., S_{j-1}.*

Proof Assume for contradiction that there exists such a path from a node $k \in S_m$, $m \geq j$ to a node $i \in S_j$. By the construction of the sets S_j (8.8), node i must be a source with no incoming links after the removal of the sets $S_0, \ldots, S_{j-1}$ from G. But this is a contradiction to the assumption that node j has an incoming link from a node $k \in S_m$, $m \geq j$. □

From the above definition and observations we can derive an upper bound on the values of knowledge of the nodes in a directed acyclic graph.

Proposition 13 *Consider* (8.13) *with a linear decay* $D_i(x_i) = dx_i$*, a linear benefit* $B_i(\mathbf{A}(G), \mathbf{x}) = b\sum_{j\in N_i^-} x_j$ *and a non-negative cost* $C_i(\mathbf{A}(G), \mathbf{x}) \geq 0$ *where* $d \geq 0$*,* $b \geq 0$*. Then for every node* i *in* G *there exists a* $k \leq n$ *such that*

$$x_i(t) \leq (a_k t^k + a_{k-1}t^{k-1} + ... + a_0)e^{-dt}. \tag{8.29}$$

Proof From Proposition (11) we know that the directed acyclic graph G has a partition of nodes into sources $S_0, \ldots, S_k$, $k \leq n$. Consider a node $x_0 \in S_0$. With (8.13) the time evolution of her value of knowledge is given by

$$\dot{x}_0 = -dx_0 - C_0 \leq -dx_0. \tag{8.30}$$

Here we use the fact that $C_0 \geq 0$. The function solving the equation $\dot{x} = -dx$ is an upper bound for $x_0(t)$ (with identical initial conditions), see (8.1).

From Proposition (9) we know that there are first-order sources S_1 in G that have only incoming links from nodes in S_0. The evolution of the value of knowledge for a node $x_1 \in S_1$ is given by

$$\dot{x}_1 = -dx_1 + \sum_{j\in S_0} x_j - C_1. \tag{8.31}$$

The second term on the right-hand side of the above equation contains the sum of all values of knowledge of all nodes in S_0. We know that they are bounded from above by $x(t) \leq x(0)e^{-dt}$. Thus, (8.31) has an upper bound

$$\dot{x}_1 \leq -dx_1 + a_1 e^{-dt} \tag{8.32}$$

with an appropriate constant a_1. The solution of the equation $\dot{x} = -dx + a_1 e^{-dt}$ is given by $x(t) = (a_1 + a_0 t)e^{-dt}$. It follows that

$$x_1 \leq (a_1 + a_0 t)e^{-dt}. \tag{8.33}$$

In the following, we make a strong induction. We have the induction hypothesis that for the $(k-1)$-th order sources there exists an upper bound

$$\dot{x}_{k-1} \leq (a_{k-1}t^{k-1} + a_{k-2}t^{k-2} + \ldots + a_0)e^{-dt} \tag{8.34}$$

and this holds also for all nodes in the sets of sources with order less than $k-1$. We consider the nodes in S_k with $l \in S_k$. We have that

$$\dot{x}_l(t) = -dx_l + b\sum_{j\in N_l^-} x_j - C_l, \tag{8.35}$$

where the in-neighborhood N_l^- contains only nodes in the sets $S_0, \ldots, S_{k-1}$. For these nodes an upper bound is given by (8.34), and thus we get an upper bound for (8.35)

$$\dot{x}_l(t) \leq -dx_l + (a_{k-1}t^{k-1} + a_{k-2}t^{k-2} + \ldots + a_0)e^{-dt}. \tag{8.36}$$

We can now use the following lemma:

Lemma 14 *For an ordinary differential equation of the form*

$$\dot{y} + dy = (a_k t^k + a_{k-1}t^{k-1} + ... + a_2 t + a_1)e^{-dt} \tag{8.37}$$

there exists a solution of the form

$$y(t) = \left(\frac{a_k}{k+1}t^{k+1} + ... + a_0\right)e^{-dt} \tag{8.38}$$

with the limit $\lim_{t\to\infty} y(t) = 0$

Solving for the upper bound from above gives the desired result.

$$x_l(t) \leq (a_k t^k + a_{k-1}t^{k-1} + ... + a_0)e^{-dt}. \tag{8.39}$$

□

With the last Proposition (13) it is straightforward to obtain the following proposition, which is the main result of this section.

Proposition 15 *Consider* (8.13) *with a linear decay* $D_i(x_i) = dx_i$*, a linear benefit* $B_i(\mathbf{A}(G), \mathbf{x}) = b\sum_{j\in N_i^-} x_j$*, and a non-negative cost* $C_i(\mathbf{A}(G), \mathbf{x}) \geq 0$ *where* $d \geq 0$*,* $b \geq 0$*. If the network G is a directed acyclic graph then the values of knowledge vanish. This means that G is not permanent.*[27]

Proof From Proposition (13) we know that each node k in the graph G has a value of knowledge which is bounded by $x_k(t) \leq (a_k t^k + a_{k-1}t^{k-1} + \cdots + a_0)e^{-dt}$ for some finite $k \leq n$. Since any finite polynomial grows less than an exponential function we have that $\lim_{t\to\infty} x_k(t) = 0$. This holds for all nodes in G. This completes the proof that for all $i = 1, \ldots, n$ in a directed acyclic graph G we have that $\lim_{t\to\infty} x_i(t) = 0$ and therefore G is not permanent. □

Thus, if agents are permanent, the graph contains a closed walk (or a cycle). If agents get their links attached at random, only those survive, who are part of a cycle. If agents can chose, whom to transfer their knowledge to, then they have to form

[27] Remember that the definition of permanence in (8.6) requires that all nodes have non-vanishing state variables. On the other hand, vanishing state variables would imply that nodes do not interact with each other and the network would not be permanent.

cycles in order to survive. Others (Bala and Goyal, 2000; Kim and Wong, 2007) have found similar results in which the equilibrium network consists of cycles.

There exist a convenient way to identify if a network contains a cycle without actually looking at the permanence of the network which would require to compute the long-run values of knowledge (usually by numerical integration). Instead, from the eigenvalues of the adjacency matrix, $\mathbf{A}(G)$, of a graph, G, one can determine if G contains a cycle. The Perron-Frobenius eigenvalue of a graph G, denoted by $\lambda_{\mathrm{PF}}(G)$, is the largest real eigenvalue of $\mathbf{A}(G)$. The following properties hold (Godsil and Royle, 2001)

Proposition 16 *If a graph G*

1. *has no closed walk, then* $\lambda_{PF}(G) = 0$,
2. *has a closed walk, then* $\lambda_{PF}(G) > 1$.

Thus, if the graph contains permanent agents, then $\lambda_{\mathrm{PF}}(G) > 1$. Hofbauer and Sigmund (1998), Stadler and Schuster (1996) have found similar conditions under which populations are permanent in a network of replicators.[28]

Finally, we can compute the probability of a network to contain a cycle if links were attached at random.

Proposition 17 *The probability of a random graph* $G(n, p)$ *with n nodes containing a cycle is given by (Jain and Krishna, 2002)*

$$P = \left(1 - (1-p)^{n-1}\right)^n \tag{8.40}$$

which is 0 *if* $p = 0$ *and* 1 *if* $p = 1$.

Proof We can compute the probability of having a closed walk in a random graph $G(n, p)$. Each link is created with probability p. Thus we have a Bernoulli process for the adjacency matrix elements a_{ij} (which indicate if a link exists or not).

$$a_{ij} = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}. \tag{8.41}$$

For every node we have $n-1$ events to create a link and we are asking for the probability of having at least one of them being created (every node should have at least one incoming link). This is a binomial cumulative function of the form (Durrett, 2004; Casella and Berger, 2001):

$$P = \sum_{k=1}^{n-1} \binom{n}{k} p^k (1-p)^{n-k} \tag{8.42}$$

[28] The replicator equation (in continuous form) is given by: $\dot{x}_i = x_i\,(f_i(\mathbf{x}) - \phi(\mathbf{x}))$, where $\phi(\mathbf{x}) = \sum_i x_i\, f_i(\mathbf{x})$ and $f_i(\mathbf{x})$ is the fitness of species i.

which is equivalent to (8.40), if we use the Binomial theorem

$$(x+y)^n = \sum_{i=1}^{n} \binom{n}{i} x^i y^{n-i}. \tag{8.43}$$

□

A similar result to (8.11) has been found by Kim and Wong (2007). The authors study a generalized version of the network formation model introduced by Bala and Goyal (2000).[29] The equilibrium networks in their model are so-called "minimal" graphs, which are graphs that maximize the number of agents that are connected while maintaining only as few links as possible. It is intuitively clear that the most sparse connected graph is a cycle. Thus, the authors find stable equilibrium networks that consist of cycles. However, in Sect. 8.3.9.8 we will show that the network evolution can reduce the set of possible cycles in the equilibrium network such that only the smallest cycles survive.

Thus, cycles play an important role in the evolution of the network and the ability of agents to have non-vanishing knowledge levels. Before we define the evolution of the network in Sect. 8.3.9, we study the dynamics of the values of knowledge for a static network in the next Sect. 8.3.8. There we will further specify the cost functions under investigation: null costs and non-linear costs for maintaining links.

8.3.8 Static Network Analysis

In the following, we analyze the growth functions for the value of knowledge and study two cases separately. In the first, costs are set to zero while in the second costs are a quadratic function of the values of knowledge of the agents.

8.3.8.1 Null Interaction Costs

The most simple case of our general framework is the one of linear benefit and null costs.[30]

---

[29] For a further study of Bala and Goyal (2000) applied to information networks see Haller et al. (2007), Haller and Sarangi (2005).

[30] This model has been studied by Jain and Krishna (1998b), Krishna (2003) to explain the origin of life from the perspective of interacting agents. The model of Jain and Krishna intends to describe the catalytic processes in a network of molecular species (which we will denote in the following by agents). However, it was very soon suggested to be applicable to an economic innovation context of interacting agents. In the next sections we will present a more general framework encompassing some of the limitation of the present one. In their model the **x** were interpreted as concentrations of chemical species. The a_{ij} are the kinetic coefficients that describe the replication of agents i resulting from *binary* interactions with other agents j.

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -dx_i + \sum_{i=1}^{n} a_{ji}x_j. \tag{8.44}$$

In vector notation (8.44) reads:

$$\dot{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}} - d\mathbf{I})\mathbf{x}, \tag{8.45}$$

where $\mathbf{A}^{\mathrm{T}}$ is the transposed of the adjacency matrix and $\mathbf{I}$ is the identity matrix. The solution of the set of equations (8.45) depends on the properties of the matrix $\mathbf{A}$ and has the general form (matrix exponential):

$$\mathbf{x}(t) = e^{-dt}e^{\mathbf{A}^{\mathrm{T}}t}\mathbf{x}(0) \tag{8.46}$$

representing an exponential increase in time of the vector of knowledge values. The relative values of knowledge (shares) are given by

$$y_i = \frac{x_i}{\sum_j x_j}; \quad \sum_j y_j = 1. \tag{8.47}$$

Rewriting (8.44) by means of (8.47) gives us the dynamics of the shares:

$$\dot{y}_i = \sum_{j}^{n} a_{ji}y_j - y_i \sum_{k,j}^{n} a_{jk}y_j. \tag{8.48}$$

Equation (8.48) has the property of preserving the normalization of $\mathbf{y}$. Note that the decay term does not appear in this equation for the relative values. It can be shown (Horn and Johnson, 1990; Boyd, 2006; Krishna, 2003) that the eigenvector to the largest real eigenvalue of $\mathbf{A}^{\mathrm{T}}$ ($\mathbf{A}$ respectively) is the stable fixed point of (8.48).[31] If we consider an eigenvector $\mathbf{y}^{(\lambda)}$ associated with the largest real eigenvalue λ of matrix $\mathbf{A}^{\mathbf{T}}$ (identical to the largest real eigenvalue of $\mathbf{A}$) we have

$$\sum_{j=1}^{n} a_{ji}y_j^{(\lambda)} = \lambda y_i^{(\lambda)}. \tag{8.49}$$

Inserting $\mathbf{y}^{(\lambda)}$ into (8.48) yields

$$\dot{y}_i^{(\lambda)} = \sum_{j}^{n} a_{ji}y_j^{(\lambda)} - y_i^{(\lambda)} \sum_{k,j=1}^{n} a_{jk}y_j^{(\lambda)} \tag{8.50}$$

[31] If the largest real eigenvalue has multiplicity more than one then the stable fixed point can be written as a linear combination of the associated eigenvector and generalized eigenvectors (Braun, 1993).

$$= \lambda y_i^{(\lambda)} - y_i^{(\lambda)} \underbrace{\sum_{k,j=1}^{n} a_{jk} y_j^{(\lambda)}}_{\lambda \sum_k^n y_k^{(\lambda)} = \lambda} \tag{8.51}$$

$$= \lambda y_i^{(\lambda)} - \lambda y_i^{(\lambda)} = 0. \tag{8.52}$$

Thus, $y_i^{(\lambda)}$ is a stationary solution of (8.48). For the proof of stability see, e.g., Krishna (2003).

8.3.8.2 Increasing Interaction Costs

In the following, we study the evolution of the values of knowledge under a given network structure and we try to compute the fixed points wherever possible. We first show that the values of knowledge are non-negative and bounded. For graphs with two nodes, for regular graphs (including the complete graph), cycles, and stars with an arbitrary number of nodes, we can compute the equilibrium points analytically. For generic graphs with $n \geq 3$ nodes, we have to rely on numerical integrations.

The non-linear (quadratic) dynamical system is given by

$$\dot{x_i} = -dx_i + b \sum_{j=1}^{n} a_{ji} x_j - c \sum_{j=1}^{n} a_{ij} x_i^2 \tag{8.53}$$

with initial conditions, $x_i(0) > 0$. a_{ij} are the elements of the adjacency matrix, $\mathbf{A}$, of a graph G. This can be written as

$$\dot{x_i} = -dx_i + b \sum_{j=1}^{n} a_{ji} x_j - c d_i^+ x_i^2 \tag{8.54}$$

where $d_i^+ = \sum_{j=1}^{n} a_{ij}$ is the out-degree of node i. In the case of increasing costs we know that the values of knowledge are bounded. We have that

Proposition 18 *For the dynamical system* (8.53) *the values of knowledge are non-negative and finite, i.e.,* $0 \leq x_i < \infty$, $i = 1, \ldots, n$.

Proof For the lower bound $x_i \geq 0$, we observe that

$$\dot{x_i} \geq -dx_i - c(n-1)x_i^2. \tag{8.55}$$

The lower bound is the solution of the equation $\dot{x} = -dx - c(n-1)x^2$. The solution of this equation can be found by solving the corresponding equation for the transformed variable $z = \frac{1}{x}$. We get $x(t) = \frac{de^{da}}{e^{dt} - c(n-1)e^{da}}$ with an appropriate constant $a = \frac{1}{d} \ln \frac{x(0)}{d+(n-1)c}$. Starting from non-negative initial values $x(0) \geq 0$ this lower

bound is non-negative as well and approaches null for large t, i.e., $\lim_{t\to\infty} x(t) = 0$. We conclude that $x_i(t) \geq 0$.

In order to compute an upper bound, $x_i \leq \text{const.} < \infty$ we first make the following observation. The nodes of a graph, $G = (V, E)$, can be partitioned into nodes without outgoing links, $V_f \subseteq V$ ("free-riders"), and nodes with at least one outgoing link, $V_s \subseteq V$ (sources).

Since the "free-riders" in V_f have no outgoing links, the benefit terms of the sources in V_s are independent of the values of knowledge of the free-riders. Accordingly, a source node $i \in V_s$ has the following knowledge dynamics.

$$\dot{x}_i = -dx_i + b \sum_{j \in V_s \setminus i} a_{ji} x_j - c x_i^2 d_i^+, \tag{8.56}$$

where d_i^+ is the out-degree of node i. We can give an upper bound of

$$\dot{x}_i \leq -dx_i + b \sum_{j \in V_s} x_j - c x_i^2. \tag{8.57}$$

This upper bound has a (finite) fixed point and so does $x_i(t)$. The fixed point is given by

$$dx_i + cx_i^2 = b \sum_{j \in V_s} x_j. \tag{8.58}$$

This is a symmetric equation and therefore all x_i are identical, $x_i = x$. For contradiction assume that there would be $x_i \neq x_j$. Then we have that

$$\underbrace{dx_i + cx_i^2}_{b \sum_{k=1}^n x_k} \neq \underbrace{dx_j + cx_j^2}_{b \sum_{k=1}^n x_k} \tag{8.59}$$

But the left and right side of the equation are identical and so two different x_i, x_j cannot exist.

When all solutions are identical we get $x_i = x = \frac{bn-d}{c}\ \forall i$. Thus, we have shown that there exists an upper bound with a finite fixed point for the source nodes, that is $x_i(t) \leq \infty$, $i \in V_s$.

We now consider the nodes with no outgoing links ("free-riders"). A node $i \in V_f$ follows the dynamics

$$\dot{x}_i = -dx_i + b \sum_{j \in V_s} a_{ji} x_j. \tag{8.60}$$

We have shown already that the source nodes are bounded by some constant, $\sum_{j \in V_s} x_j \leq \text{const}$. Thus, we have that

$$\dot{x}_i \leq -dx_i + \text{const.} \tag{8.61}$$

We have an upper bound of the x_i, $i \in V_f$, given by

$$x_i(t) \leq x_0 e^{-dt} + \frac{\text{const.}}{d} \tag{8.62}$$

with $\lim_{t\to\infty} x_i(t) = \frac{\text{const.}}{d}$. We have shown that for all nodes (sources V_s as well as "free-riders" V_f) $0 \leq x_i < \infty$, $i \in V(G)$. □

For special types of graphs we can deduce further results on the values of knowledge of the agents. First, we can compute the fixed points (given by $\dot{x}_i = 0$) for regular graphs.

Proposition 19 *For any k-regular graph G the fixed point of the values of knowledge is given by* $x^* = \frac{kb-d}{kc}$. *In particular, the complete graph* K_n *has the highest total value of knowledge among all regular graphs with* $x^* = \frac{(n-1)b-d}{(n-1)c}$.

Proof The dynamics of the values of knowledge of the nodes in a regular graph with degree $d_i^+ = d_i^- = k$ is given by

$$\dot{x}_i = -dx_i + b\sum_{j\in N_i} x_j - ckx_i^2. \tag{8.63}$$

Starting with homogeneous initial conditions we make the Ansatz $x_i = x$ $i = 1, ..., n$. We get the positive stable fixed points $x^* = \frac{kb-d}{kc}$. □

Second, we can compute the fixed points for cycles.

Proposition 20 *For any cycle* C_n *the fixed point of the values of knowledge is given by* $x^* = \frac{b-d}{c}$.

Proof The dynamics of the values of knowledge of a cycle C_k of length k is given by

$$\dot{x}_i = -dx_i + bx_{i-1} - ckx_i^2. \tag{8.64}$$

Starting with homogeneous conditions we make the Ansatz $x_i = x$, $i = 1, \ldots, n$. We get the positive stable fixed points $x^* = \frac{b-d}{c}$. □

Third, the fixed points for a star can be computed (the proof can be found in the Appendix (A)).

Proposition 21 *For a star* $K_{n,n-1}$ *there exists a fixed point which increases with the number of nodes. For* $d = 0$ *the star has a fixed point of* $x^* = \frac{b}{c}$.

Proof The dynamics of the values of knowledge of a star $K_{1,n-1}$ is given by

$$\dot{x}_1 = +b\sum_{i=2}^{n} x_i - c(n-1)x_i^2, \tag{8.65}$$

$$(\dot{x}_i)_{i>1} = -dx_i + bx_1 - cx_i^2, \tag{8.66}$$

where we assume that all links are bidirectional. Starting with homogeneous initial conditions we make the Ansatz $x_i = x_2$, $i = 2, ..., n$. Then x_2 is determined by the root of the polynomial

$$x_2^3 + \frac{2d}{c}x_2^2 + \frac{d(cb+(n-1)cd)}{(n-1)c^3}x_2 + \frac{b(d^2-(n-1)b^2)}{(n-1)c^3} = 0. \tag{8.67}$$

And $x_1 = \frac{d}{b}x_2 + \frac{c}{b}x_2^2$. For $d = 0$, we obtain $x_1^* = x_2^* = \frac{b}{c}$. □

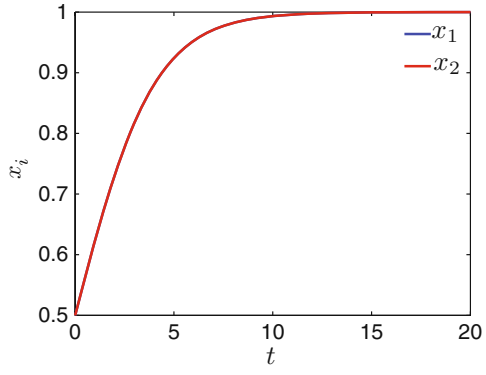
The fixed point increases with the benefit b and decreases with the decay d and the cost c.

We observe that, for vanishing decay, $d = 0$, the fixed point of the system is identical for the regular graph, the cycle, and the star and given by $\frac{b}{c}$. As expected, this fixed point is increasing with the benefit and decreasing with the cost. In a regular graph the fixed point is increasing with the degree k and the asymptotic value (for large k) is $\frac{b}{c}$. Thus, in a regular graph the fixed point ranges for increasing k from $\frac{b-d}{c}$ to $\frac{b}{c}$. Similarly, for the star the fixed point also increases with the number of nodes (i.e., the degree of the central node) but we cannot provide an analytical expression here. On the other hand, the fixed point of the cycle is independent of the length of the cycle. This means that there is no incentive for nodes to be part of larger cycles. And, as we will see in the next section, this limits the growth of the network.

Example 22 We numerically integrate (8.53) for $n = 2$ nodes. We set $d = 0.5$, $c = 0.5$, and $b = 1$. Fixed points are denoted x_i^* for $i = 1, 2$. $x_i^* = 0$ is a fixed point for all graphs.

(1)

$$A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

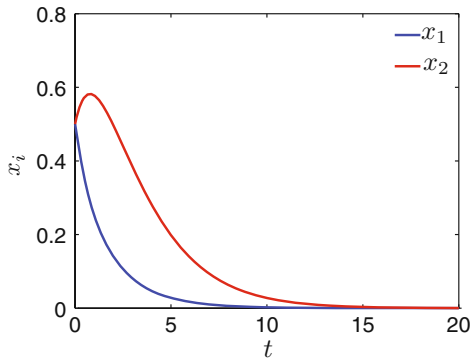


The fixed points are given by

$$x_i^* = \frac{b-d}{c},\ i = 1, 2, 3. \tag{8.68}$$

(2)

$$A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

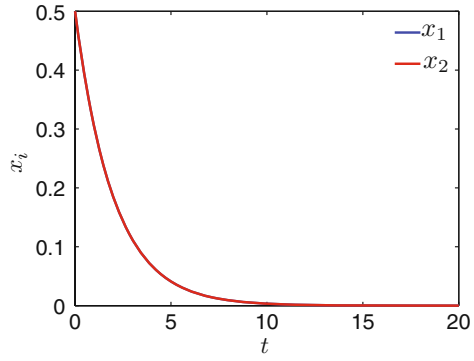


The fixed points are given by

$$x_i^* = 0,\ i = 1, 2, 3. \tag{8.69}$$

(3)

$$A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$



The fixed points are given by

$$x_i^* = 0,\ i = 1, 2, 3. \tag{8.70}$$

In general, the fixed points of (8.53) can only be computed numerically. As an example, we compute the fixed points for all graphs with $n = 3$ nodes for a specific choice of parameters. The results can be found in Appendix (A). In our model we numerically integrate (8.53) for a large time T (and we find that in our simulations the system always reaches a stable fixed point).

8.3.9 Dynamics of Network Evolution

8.3.9.1 Network Evolution as an Iterative Process

After providing the static equilibrium analysis, in this section we turn now to the *dynamics* of the network evolution by investigating different assumptions for the *creation* and *deletion* of links in the network. In particular, we compare two different scenarios, namely the so-called extremal dynamics, where agents do not decide themselves about the link creation and deletion, and the utility driven dynamics, where agents make this decision themselves based on different rules discussed below.

We first define the utility of the agents in our model for a given network G.

Definition 23 Consider a (static) network G. The utility of agent i is given by

$$u_i = \begin{cases} y_i(T), & \text{for Null Interaction Costs} \\ x_i(T), & \text{for Increasing Interaction Costs} \end{cases}, \tag{8.71}$$

where the value of knowledge $x_i(t)$ is given by (8.53) and $\mathbf{A}(G)$, the relative value of knowledge $y_i(t)$ by (8.48) and $\mathbf{A}(G)$. T is called the time horizon.

We assume that the accumulation of knowledge is faster than the frequency of the agents creating or deleting links.[32] With this assumption, we can introduce a *time-scale separation* between the accumulation of knowledge and the evolution of the network.

The evolution of the system is then defined by an alternating sequence of knowledge accumulation, where we keep the network fixed for a given time T, $\mathbf{A}(G)$ =const., and changes in the links (asynchronous updating of the nodes) (see Fig. 8.7). When the knowledge accumulation has reached time T, the network structure is changed. A change in the network takes place by either link addition between two agents i and j, $a_{ij} = 0 \rightarrow a_{ij} = 1$, or by link removal, $a_{ij} = 1 \rightarrow a_{ij} = 0$. When the network has changed, the new utility, determined by (8.18), can be computed for time $2T$. This iterative procedure of knowledge accumulation and link

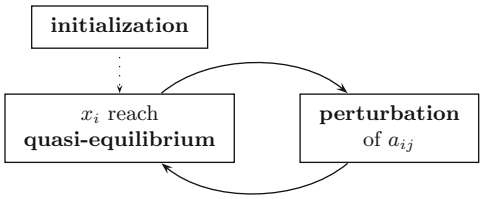


Fig. 8.7 Schematic representation of the network evolution as an iterative process

[32] This means that the value of knowledge on the market (which is not explicitly modeled here) reaches a stationary state determined by the R&D collaborations of each agent (and her neighbors). Only after this adaptation of the evaluation of the stocks of knowledge is finished, i.e., it has reached a stationary state, agents asynchronously change their links.

changes continues for $3T$, $4T$, ... and so on until the network reaches an equilibrium. One can schematically represent this iteration by the following algorithm:

1. initialization: Random graph $G(n, p)$.
2. **quasi-equilibrium**: fast knowledge growth/decline
   With **A** fixed, agents evolve according to (8.13) for a given (large) time T.
3. **perturbation**: slow network evolution
   After time T, the network evolves according to two alternative selection processes:
   1. *Extremal Dynamics.*[33]
      The agent with the minimum utility is chosen (if there are more than one agent with the same minimum value, then one of them is chosen at random). The utility of that agent is set to its initial value and all its outgoing and ingoing links are replaced with new random links drawn with probability p from and to all other agents in the system.
   2. *Utility Driven Dynamics*
      An agent is randomly chosen to create or delete one link (unidirectional or bidirectional link formation mechanisms, see Sect. 8.3.9.3). More specifically:
      (i) Either a pair or a single agent is randomly chosen to create or remove a link.
      (ii) The effect of this link decision (creation or deletion) is evaluated at time T. The evaluation can have the following consequences on the link decision.
      - If the utility has increased, then sustain the link decision.
      - If the utility has decreased, then undo the link decision.
4. Stop the evolution, if the network is stable (stability is defined in Sect. 8.3.9.3, otherwise go to **2**

8.3.9.2 Extremal Dynamics Versus Utility Driven Dynamics

Extremal dynamics intends to mimic natural selection (the extinction of the weakest) and the introduction of novelty, which is a global selection mechanism. In contrast, utility driven dynamics is a local selection mechanism that mimics the process by which selfish agents improve their utility through a trial and error process.

The decision upon to add or to remove a link implies a certain level of information processing capabilities (IPC) of the agents. IPC is usually bounded in a complex environment consisting of many other agents and a complex structure of interactions between these agents. In our approach we assume that the agents have no information on the knowledge values of the other agents and only limited information on their links (alliances). They only know with whom they interact directly (their neighborhood). In Table 8.2 we give a short overview of levels of increasing IPC.

[33] See Bak and Sneppen (1993).

Table 8.2 Increasing levels of agents' information processing capabilities (IPC)

0	Least fit addition/removal of links, e.g Jain and Krishna (1998b)
1	Reactive (passive) acceptance/refusal of link changes.
2	Deliberate decision upon to add/remove a link based on an individual utility function depending on the network, without considering the possible decision of others. An example would be the Connections model discussed in Sect. 8.2.2 with a utility function given by (8.4).
3	Strategic interaction, e.g., Bala and Goyal (2000), considering the possible actions of others

Extremal dynamics refers to a situation in which agents are exposed to link changes that they cannot influence and thus to level 0 in Table 8.2. Utility driven dynamics instead requires a higher level of IPC than a mere acceptance or refusal of link changes. But it requires less IPC than an approach assuming strategic interactions of agents. This follows from the fact that in our model, agents do not estimate how other agents could react on their decisions to change their links. This situation refers to level 2 in Table 8.2. In this chapter, we compare two different settings, level 0 and level 2. In the following paragraphs, we describe them in more detail.

0 Extremal Dynamics: At time T the agent with the smallest utility is removed from the system and replaced with a new one (market entry). The new agent is randomly connected to the already existing agents and a small initial value of knowledge is assigned to it. This process is a least fit replacement (extinction of the weakest) and the new agent introduces a kind of novelty in the system (innovation).

2 Utility Driven Dynamics: The main difference between local link formation (utility driven dynamics) compared to global link formation (extremal dynamics) is that agents are now individually taking decisions upon their interactions and they do that on the basis of a utility function (their values of knowledge at time T). Agents are bounded rational since they explore their possible interaction partners in a trial and error process. At every period, that is after time T, an agent is selected at random to create and delete links (asynchronous update). We distinguish two possible link formation mechanisms which we study separately, namely unilateral and bilateral link formation. In the former, unilateral link formation (i), the agent optimally deletes an old link and randomly creates a new link. Optimal means that either for creation or deletion of links the action is taken only if it increases the value of knowledge of the agent at time T in the range of all possible actions. In the latter, bilateral formation (ii), the selected agent optimally deletes a bilateral connection that she currently has or she randomly creates a new bilateral connection. Here optimal (in the range of all possible actions) means, that links are deleted if the initiator of the deletion, i.e., the selected agent, can increase its value of knowledge at time T with the deletion of the link, while for the bilateral creation both agents involved have to strictly benefit from the

creation of the mutual connection.[34] In the following two sections, we give a description of mechanisms (i) and (ii).

To compare the two levels, for utility driven dynamics the evolution of the network follows from local, utility driven, actions, as opposed to extremal dynamics, where the evolution follows from a global stochastic process (least fit selection plus random link formation). To be more specific about the latter, the rules for the network evolution, i.e., the creation and deletion of links under extremal dynamics, are the following:

Step 1 After a given time T the *least fit agent*, i.e., the one with the smallest $u_i = y_i(T)$, is determined. This agent is removed from the network along with all its incoming and outgoing links.

Step 2 A new agent is added to the network with some small initial value of knowledge y_0. The new agent will take the place of the old one (it gets the same label), and randomly links itself to the other nodes in the network with the same probability p. Each of the other nodes can in turn link itself to the newcomer node with a probability p.

These rules for the network evolution are intended to capture two key features: *natural selection*, in this case, the extinction of the weakest; and the *introduction of novelty*. Both of these can be seen as lying at the heart of natural evolution. The particular form of selection used in this model has been inspired by what *Bak and Sneppen* have called "extremal dynamics" (Bak and Sneppen, 1993).

8.3.9.3 Rules for Link Creation and Deletion Using Utility Driven Dynamics

In this section, we introduce the process of the formation and deletion of links by agents that maximize a local utility function (depending on the agent and its neighbors). After time T, long enough such that the system reaches a quasi-equilibrium in the values of knowledge, an agent is randomly chosen to create or delete a link, either unidirectional or bidirectional.

Unilateral Link Formation

If agents unilaterally delete or create links, it is possible that the interactions they form create a feedback loop, i.e., a closed cycle of knowledge sharing agents, that involves more than two agents. This introduces the concept of indirect reciprocity (see Sect. 8.3.3). Unilateral formation of links (we then have a directed network) is necessary for indirect reciprocity to emerge, since if all interactions were bilateral they would be direct reciprocal by definition. We now describe the procedure of unilateral link creation and deletion.

---

[34] This behavior is individually optimal and thus may also be called rational.

1. *Random Unilateral Creation*
   An agent creates a link to another one to which it is not already connected at random and evaluates the creation of the link by comparing the change in their values of knowledge before and after the creation. Only if the change is positive, the link is maintained, otherwise the agent does not create the link. In this way agents explore possible partners for sharing their knowledge in a trial and error procedure (Fig. 8.8).

   Step 1 An agent i is selected at random.
   Step 2 Another agent j is selected at random which is not already an out-neighbor of i.
   Step 3 Agent i creates an outgoing link to agent j.
   Step 4 The new utility (for the old network plus the new link e_{ij}) of agent i is computed and compared with the utility before the creation.
   Step 5 Only if agent i's utility strictly increases compared to her old utility, then the link is created.

2. *Optimal Unilateral Deletion*
   An agent deletes one outgoing link if this increases her utility (Fig. 8.9).

   Step 1 Agent i is selected at random s.t. it has at least one outgoing link.
   Step 2 Agent i deletes separately each of its outgoing links to its neighbors $v_j \in N_i^+$ and records the change in her utility, Δu_i. Before the next link is deleted, the previous one is recreated.
   Step 3 Agent i computes the maximum change Δu_i and if it is positive, deletes the referring link. This means that only one link is finally deleted. The deletion only takes place if the current agent strictly increases her utility.

To characterize the equilibrium networks under this link formation and deletion mechanism, we introduce the following characterization of stability.[35]

Definition 24 A network is **unilaterally stable** if and only if (i) no agent can create a link to (strictly) increase her utility and (ii) no agent can remove a link to (strictly) increase her utility.

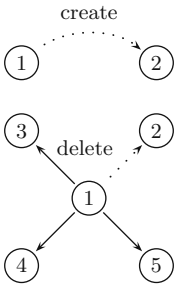


Fig. 8.8 Random unilateral creation

Fig. 8.9 Optimal unilateral deletion

[35] Compare this to the definition of bilateral stability (8.20)

Bilateral Link Formation

If agents form links bilaterally then all interactions are direct reciprocal by definition. We describe the process of bilateral link creation and deletion in the following paragraphs:

1. *Random Bilateral Creation*
   In this link creation process, a pair of agents is selected at random and given the possibility to form a bilateral connection (Fig. 8.10).

   Step 1 Two agents are uniformly selected at random such that they are not connected already.
   Step 2 Both agents create an outgoing link to each other and therewith create a 2-cycle.
   Step 3 The new utilities (for the old network plus the new 2-cycle) of both agents are computed and compared with the utilities before the creation.
   Step 4 Only if both agents strictly benefit in terms of their utilities compared to their old utilities, then the bilateral connection is created.

2. *Optimal Bilateral Deletion*
   An agent deletes one of its outgoing links to another agent from which the agent also has an incoming link if this deletion increases her utility (Fig. 8.11).

   Step 1 Agent i is selected at random such that it has at least one mutual link to another agent.
   Step 2 From all bilaterally connected neighbors agent i deletes separately each of its outgoing links to its neighbors (and so does each neighbor j to agent i). For each, the change in the utility, Δu_i is recorded. Before new links are deleted, the old ones are recreated.
   Step 3 Agent i computes the maximum change Δu_i and, if it is positive, the referring bilateral connection is deleted. The deletion only takes place if agent i strictly increases her utility.

In order to characterize the equilibrium outcomes of our simulations, we will introduce a characterization of network stability. This definition has been introduced already in Sect. 8.2.2 and we repeat it here for expository reasons.

Definition 25 A network G is *pairwise stable* if (i) removing any link does not increase the utility of any agent and (ii) adding a link between any two agents, either doesn't increase the utility of any of the two agents, or if it does increase one of the two agents' utility then it decreases the other agent's utility.

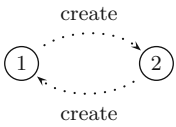


Fig. 8.10 Random bilateral creation

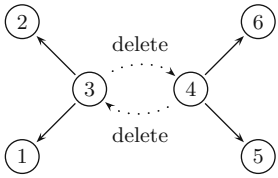


Fig. 8.11 Optimal bilateral deletion

8.3.9.4 The Role of the Time Horizon for Unilateral Link Formation

So far, we have assumed that the time horizon T (after which agents evaluate their decisions to create or delete links) is long enough such that the values of knowledge reach a stationary state and the utilities of the agents are given by the fixed points of the values of knowledge. In this section, we discuss the effect of a time horizon that is smaller than the time to convergence to the stationary state of the values of knowledge. For related works that incorporate a finite time horizon in the evaluation of the actions of agents see, e.g., Huberman and Glance (1994) or Lane and Maxfield (1997).

If we consider utility driven dynamics, we will show that permanent networks with positive values of knowledge emerge if agents wait long enough (with respect to the time the values of knowledge need in order to reach a stationary state) in evaluating their decisions. This is a necessary condition. Otherwise networks are not able to emerge or, if a network with positive knowledge values is existing already, it gets destroyed over time (network breakdown). This effect is important in the case of null as well as increasing costs.

To illustrate this point, we consider a 5-cycle of agents and the deletion of one link in this cycle which creates a linear chain of five nodes, Fig. 8.12. The evolution of value of knowledge for null costs and for costs $c = 0.5$ can be seen in Fig. 8.13.

More formally we can give the following proposition:

Proposition 26 *Consider the dynamical system* (8.21)*. For a directed path P_k of length k the value of knowledge of node k is larger than ϵ for $t \leq \tau(\epsilon)$, i.e., $x_k(t \leq \tau(\epsilon)) \geq \epsilon$ while $\lim_{t\to\infty} x_k(t) = 0$.*

Proof Consider a directed path P_k of length k (Fig. 8.14).

For node 1 (the source has no incoming links) in (8.21) we get

$$\dot{x}_1(t) = -dx_1 - cx_1^2. \tag{8.72}$$

By introducing the variable $z = \frac{1}{x_1}$ and solving for z, one can find the solution for x_1

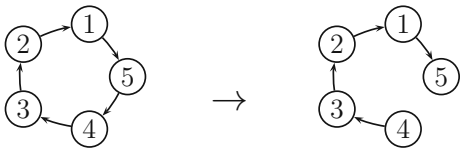


Fig. 8.12 A 5-cycle and a linear chain of five nodes (obtained from the cycle by removing one link)

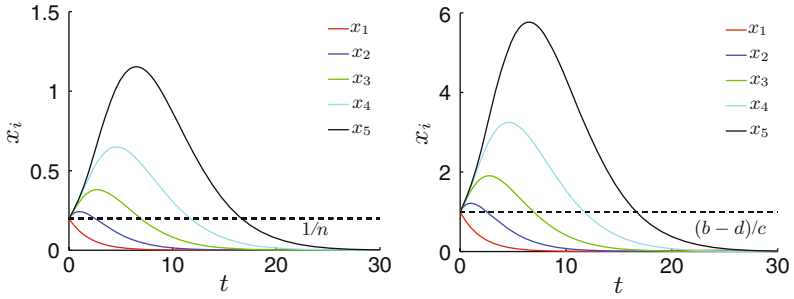


Fig. 8.13 Numerical integration of the value of knowledge for $d = 0.5$, $b = 1.0$, null cost $c = 0.0$ (*left*) and cost $c = 0.5$ (*right*): evolution of knowledge values for a linear chain of five nodes (obtained from the C_5 by removing a link). The agent that removes the link (black upper curve) initially experiences an increase in the value of knowledge. After an initial increase she experiences a decline and at a certain time her value of knowledge reaches her initial value ($1/n$ in the case of null cost and $(b-d)/c$ in the case of increasing cost) and then it further decreases. After a time long enough her value of knowledge vanishes completely

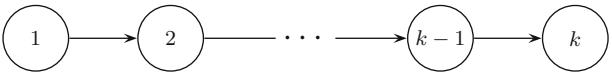


Fig. 8.14 A directed path P_k of length k

$$x_1(t) = \frac{de^{da}}{e^{dt} - ce^{da}} \tag{8.73}$$

with a constant $a = \frac{1}{d} \ln \frac{x_1(0)}{d+c}$ and the limit $\lim_{t\to\infty} x_1(t) = 0$. Accordingly, for the k-th node we have that

$$\dot{x}_k = -dx_k + bx_{k-1} - cx_k^2. \tag{8.74}$$

Since $x_k \geq 0$, from Proposition (18), the following inequality holds

$$\dot{x}_k \geq -dx_k - cx_k^2 \tag{8.75}$$

and

$$x_k(t) \geq \frac{de^{da'}}{e^{dt} - ce^{da'}} \tag{8.76}$$

with a proper constant $a' = \frac{1}{d} \ln \frac{x_k(0)}{d+c}$. Equating with $\epsilon(t)$ at $t = \tau$ we get

$$\epsilon = \frac{d}{e^{dt-da'} - c} \tag{8.77}$$

which yields

$$\tau(\epsilon) = \frac{\ln\left(\frac{d}{\epsilon} + c\right) + a'd}{d}. \tag{8.78}$$

Thus, we have found an $\epsilon(\tau)$ such that for $t \leq \tau(\epsilon)$ $x_k(t) \geq \epsilon$. The limit $\lim_{t\to\infty} x_k(t) = 0$ follows directly from the fact that the directed path P_k is a directed acyclic graph and we can apply Proposition (15). □

With Proposition (26) one can readily infer the following. If an agent in a cycle C_k of length k removes a link unilaterally then a path P_k is created. If the time horizon after which the agent evaluates this link removal is smaller then $\tau(\epsilon)$) the agent's value of knowledge satisfies $x_k(t \leq \tau(\epsilon)) \geq \epsilon$ (this gives the utility of the agent, see (23)). From Proposition (20) we know that the value of knowledge of the agent in the cycle is given by $x_k(0) = \frac{b-d}{c}$. Choosing $\tau(\epsilon)$ such that $\epsilon > \frac{b-d}{c}$ gives $x_k(t \leq \tau(\epsilon)) \geq x_k(0)$ and the agent experiences an increase in her utility by removing the link. The agent removes the link in order to increase her utility. This destroys the cycle. The time horizon of the agent in this case is too short in order to anticipate the vanishing long-run values of knowledge of all the agents in the resulting path, $\lim_{t\to\infty} x_(t) = 0$.

From this observation we conclude that if the time horizon is too short, then all cycles would get destroyed and no network would ever be able to emerge nor sustain, since only cyclic networks can be permanent. Agents who remove their links because, in the short run, their utility increases therewith, can be considered as free-riders. The value of knowledge is maintained by their predecessors in the cycle while they refuse themselves to contribute to knowledge sharing and production in the network since they do not have any outgoing links. In the short run they benefit from the knowledge shared and produced in the network without contributing to it. However, as we have seen from the discussion above, in the long run this causes the total value of knowledge of the network to vanish. Therefore, we can say that the free-riding behavior of agents leads to the breakdown of the economy.

8.3.9.5 Simple Equilibrium Networks for Unilateral Link Formation

In this section, we identify the most simple equilibrium networks for unilateral link formation. There exists a multitude of other equilibrium networks which usually cannot be computed analytically and which depend on the parameter values for decay, benefit, and cost.

The most simple equilibrium network is the empty network.

Proposition 27 *The empty graph is unilaterally stable.*

Proof In an empty graph all nodes have vanishing values of knowledge. Creating a link does not create a cycle (which would be the case however if links were formed bilaterally) and thus the empty graph plus a link is a directed acyclic graph with vanishing values of knowledge, see Proposition (15). The creation of a link does not increase the utility of an agent. Thus, the agents do not form any links. □

Moreover, if all agents form disconnected cycles then we have an equilibrium network.

Proposition 28 *The set of disconnected cycles* $\{C^1, ..., C^k\}$*, and possibly isolated nodes is unilaterally stable.*

Proof We give a proof for 2-cycles. The proof can easily be extended to cycles of any length. Consider the two cycles C_2^1 and C_2^2 in Fig. 8.15.

From Proposition (20) we know that the fixed points are given by $x_i = \frac{b-d}{c}$, $i = 1, ..., 4$. In order to show that we have a unilaterally stable equilibrium, we (i) first show that no link is created and in the following (ii) that no link is deleted.

(i) If a link is created (w.l.o.g.) from node 2 to 4 we get from the dynamics on the value of knowledge in the case of increasing interaction costs given by (8.21)

$$\begin{aligned} \dot{x}_1 &= -dx_1 + bx_2 - cx_1^2 \stackrel{!}{=} 0 \\ \dot{x}_2 &= -dx_1 + bx_1 - cx_2^2 \stackrel{!}{=} 0 \end{aligned}. \tag{8.79}$$

From the first order conditions for the fixed points we get for node 1

$$x_1 = \frac{bx_2 - c}{d}. \tag{8.80}$$

And inserting this into the fixed point of node 2 gives

$$x_2 = \frac{b^2 - d^2 + \sqrt{b^4 - 8bc^3d - 2b^2d^2 + d^4}}{4cd} \tag{8.81}$$

If the last inequality is fulfilled, then the creation of the link would decrease the utility of agent 2. The inequality holds if $c^3 \geq \frac{(d-b)^3}{b}$ which is certainly true for $b > d$ and $c > 0$. Thus, no link is created between the cycles.

(ii) If a link is deleted in a C_2 then we get vanishing steady-state values of knowledge. Since $\frac{b-d}{c} \geq 0$ this would reduce the utility of the agent. Therefore, the link is not removed.

If there are $k \leq \lfloor \frac{n}{2} \rfloor$ 2-cycles in G then the above argument holds for any pair of cycles. Similarly, no isolated node can create a link in order to increase her utility nor can a node in a cycle create a link to an isolated node. Neither link creation

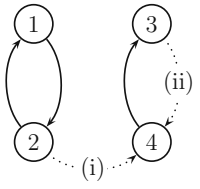


Fig. 8.15 Two cycles C_2^1 and C_2^2 and the cases of link creation (i) and deletion (ii)

nor removal increases the utility of the initiating agent and so the set of 2-cycles is unilaterally stable. □

We further conjecture that a set of disconnected autocatalytic sets, where an autocatalytic set is defined as a set of nodes each having an incoming link from a node of that set (Jain and Krishna, 2001), stays disconnected under unilateral link formation. Thus, the size (in terms of nodes) is stable.

With (28) we know that a cycle is unilaterally stable. In Sect. 8.3.9.4, however, we have shown that this result is critically depending on the time horizon T after which the action of an agent is evaluated (and it is true for cycles of any length only if $T \to \infty$).

For parameter values $d = 0.5$, $b = 0.5$, and $c = 0.1$ also, the complete graph with three nodes K_3 and the path P_3 is unilaterally stable. We observe this in simulations in Fig. 8.22. However, by computing the fixed points numerically for $d = 0.5$, $b = 0.5$, and $c = 0.1$ in Appendix (A) one can see that K_3 is no longer unilaterally stable (because removing a link increases the utility of an agent).

In the next section, we investigate if the dynamic processes of link formation and deletion lead to the simple equilibrium structures suggested above (and indeed we show that they are not obtained).

8.3.9.6 Simulation Studies Using Different Growth Functions

In the remainder of this chapter, we study simulations with different growth functions (for the value of knowledge) and different link formation mechanisms. We assume that the time horizon T is long enough such that the values of knowledge reach their stationary state. The dynamics of the value of knowledge is given by (8.19) with null costs or by (8.21) with increasing costs. The different link formation mechanisms are described in Sect. 8.3.9.2. We compare the equilibrium networks obtained from different costs and link formation rules in terms of their structure and performance. Finally, we study the effect of different positive network externalities on the equilibrium networks.

Table 8.3 gives an overview of the simulations that we study in the following. We set $d = 0.5$, $b = 0.5$, and $c = 0.1$. The complete set of parameter values used throughout this section can be found in Table 8.5 in the Appendix (C).

8.3.9.7 Null Interaction Costs

In the following, we briefly discuss the evolution of the network with least fit link formation and null link costs. This model has been studied in detail by Seufert and Schweitzer (2007); Jain and Krishna (2001). Later Saurabh and Cowan (2004) have applied it to an innovation model where new ideas are created and destroyed in a network of ideas.

In this model agents do not have to pay costs for maintaining interactions. Accordingly, the dynamics on the values of knowledge is given by (8.19)

Table 8.3 Overview of the simulation studies in the following sections with different knowledge growth functions and different link formation mechanisms

Knowledge Dynamics	Network Dynamics	Section
Null costs, $c_{ij} = 0$: $\frac{dx_i}{dt} = -dx_i + b\sum_{j=1}^{n} a_{ji}x_j$	Least fit replacement	(8.3.9.7)
Quadratic cost, $c_{ij} \propto x_i^2$: $\frac{dx_i}{dt} = -dx_i + b\sum_{j=1}^{n} a_{ji}x_j - c\sum_{j=1}^{n} a_{ij}x_i^2$	Least fit replacement unilateral link formation bilateral link formation	(8.3.9.8)
Quadratic cost, $c_{ij} \propto x_i^2$, and externality, w_{ji}: $\frac{dx_i}{dt} = -dx_i + \sum_{j=1}^{n}(ba_{ji} + b_e w_{ji})x_j - c\sum_{j=1}^{n} a_{ij}x_i^2$	Unilateral link formation	(8.3.9.9)

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -dx_i + b\sum_{j=1}^{n} a_{ji}x_j$$

and the dynamics in the shares of the values of knowledge $y_i = x_i / \sum_{j=1}^{n} x_j$ is given by (8.48).

$$\dot{y}_i = \sum_{j}^{n} a_{ji}y_j - y_i \sum_{k,j}^{n} a_{jk}y_j.$$

The utility is given by $u_i = y_i(T)$. We have described in Sect. 8.3.8 that the fixed point (stationary solution) of the relative values of knowledge in (8.48) exists and is given by the eigenvector to the largest real eigenvalue of the adjacency matrix. We assume that the time horizon T (after which links get created or deleted) is large enough such that the system has reached this stationary state before links are changed.

Extremal Dynamics: Least Fit Replacement

After time T the worst performing agent (in terms of her share of value of knowledge $y_i(T)$) is replaced with a new one. We have described this global link formation mechanism in Sect. 8.3.9.2. Jain and Krishna (1998a, 2001), Seufert and Schweitzer (2007) have extensively studied the behavior of the dynamics on $\mathbf{y}$ and the network G represented by $\mathbf{A}(G)$. They showed that strongly connected sets of nodes with free-riders (that are receiving knowledge from the strong component but are not contributing knowledge back to the strong component) attached[36] appear and get destroyed in the process of repeatedly removing the worst performing node (with minimum y_i) and replacing it with a new one.

[36] Jain and Krishna (2001) denote this set of nodes the *autocatalytic set (ACS)*: it is a subgraph of nodes in which every node has at least one incoming link from that subgraph.

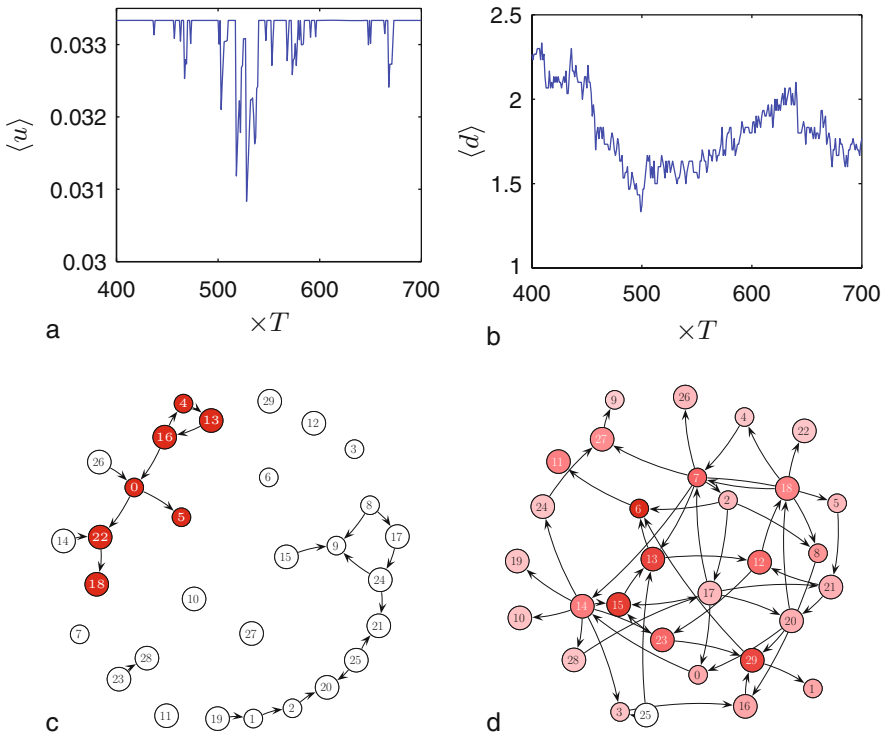


Fig. 8.16 Least fit replacement: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 5000 iterations

In computer simulations we can reproduce the results of Jain and Krishna (2001), Seufert and Schweitzer (2007). We observe crashes and recoveries in the average utility and degrees of the agents over time as can be seen in Fig. 8.16. Thus, no stable equilibrium network can be realized with this type of network dynamics.

In the model of Jain and Krishna (2001) links are costless. In the next Sect. 8.3.9.8 we assume that links have a cost attached, that is an increasing function of the value of knowledge that is being transferred (Sect. 8.3.9.8).

Moreover, the least fit network dynamics treats agents as completely passive units that are exposed to an external selection mechanism. In a more realistic approach one should take into account that agents are deliberately deciding upon with whom to engage in an R&D collaboration or to share their knowledge with. These decisions are taken on the basis of increasing a utility function, that is their value of knowledge.[37] We introduce local link formation rules in Sect. 8.3.9.3. Moreover, as a further extension we study the effect of positive network externalities in Sect. 8.3.9.9.

[37] A model in which the eigenvector associated with the largest real eigenvalue is used as a utility function is studied in Ballester et al. (2006).

8.3.9.8 Increasing Interaction Costs

In this section we study the effect of increasing costs for maintaining interactions with other agents on the resulting equilibrium networks. The evolution of the value of knowledge is given by (8.53) and the utility of the agents by (8.18). The cost of a link depends quadratically on the value of knowledge of the agent that initiates the interaction. We study three different link formation mechanisms. The first is a least fit replacement. We will compare the results of the simulation with the preceding section where links were costless. In the following two sections link formation mechanisms are studied in which agents decide locally upon to create or delete links either unilaterally or bilaterally based on their utility (8.18). We assume that the time horizon is long enough such that the utility of the agents is given by the fixed points of the value of knowledge. We will show that least fit replacement of agents leads to a total network breakdown eventually from which the system cannot recover. Moreover, we show that bilateral link formation leads to a complete graph while with unilateral link formation this is not the case. For unilateral link formation only a small number of agents have non-vanishing knowledge values in the resulting equilibrium network and these cluster together in bilateral connections. Depending on the link formation mechanism and the parameter values (for decay, benefit, and cost) the equilibrium networks can vary considerably.

The evolution of the value of knowledge of agent i (8.53)

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -dx_i + b\sum_{i=1}^{n} a_{ji}x_j - c\sum_{i=1}^{n} a_{ij}x_i^2$$

and her utility is given by $u_i = x_i(T)$.

Extremal Dynamics: Least Fit Replacement

Similarly to the preceding section, links are formed and removed by a least-fit selection mechanism (introduced in Sect. 8.3.9.2). The agent with the smallest utility (8.18) is replaced with a new agent. But in this section costs for maintaining links are an increasing function of the knowledge value of the transmitting agent.

In this setting, it is possible that the system breaks down completely. A simulation run exhibiting such a crash can be seen in Fig. 8.17. If the network is sparse enough the link removal mechanism can destroy the cycles in the network and thus creates a directed acyclic graph. As soon as the network evolution hits a directed acyclic graph, all value of knowledge vanish (and accordingly the utilities of the agents) and the network entirely breaks down.

We do not experience a breakdown of the network in the case of null costs in the last section since there we were considering relative values of knowledge only. The normalization of the relative values, $\sum_{i=1}^{n} y_i = 1$ prevents all the shares to become 0 at the same time, $y_i = 0\ \forall i$. Thus, we do not get a total breakdown of the network in which all values of knowledge vanish. Instead, there the system can always recover from a crash of the network.

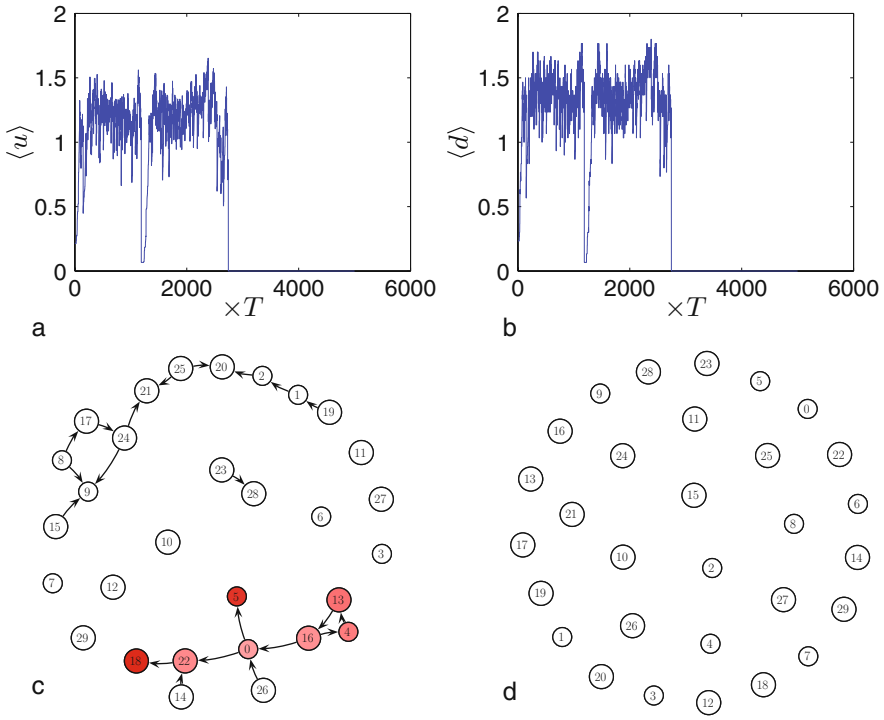


Fig. 8.17 Extremal dynamics: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 5000 iterations (in the equilibrium). The network experiences a total breakdown eventually

Utility Driven Dynamics: Bilateral Link Formation

In this section, agents are creating or deleting links bilaterally. All interactions are therefore direct reciprocal. In simulations we observe the following effect. Bilateral creation and deletion results in a complete subgraph (the average degree is $1/n \sum d_i = 1/20 \times 8 \times 7 = 2.8$, see Fig. 8.20) of the agents that were part of a permanent set in the initial graph[38] (the stability criterion which defines an equilibrium network is given in (25)).

Utility Driven Dynamics: Unilateral Link Formation

The mechanisms of unilateral creation and deletion of links has been introduced in Sect. 8.3.9.3. In our simulations we observe the following effect. When we allow for unilateral link formation, large cycles get reduced to a small set of 2-cycles. In the equilibrium network (the stability criterion which defines an equilibrium network is

[38] The creation of the initial random graph with a given link creation probability has been chosen rather small such that only a few nodes are permanent.

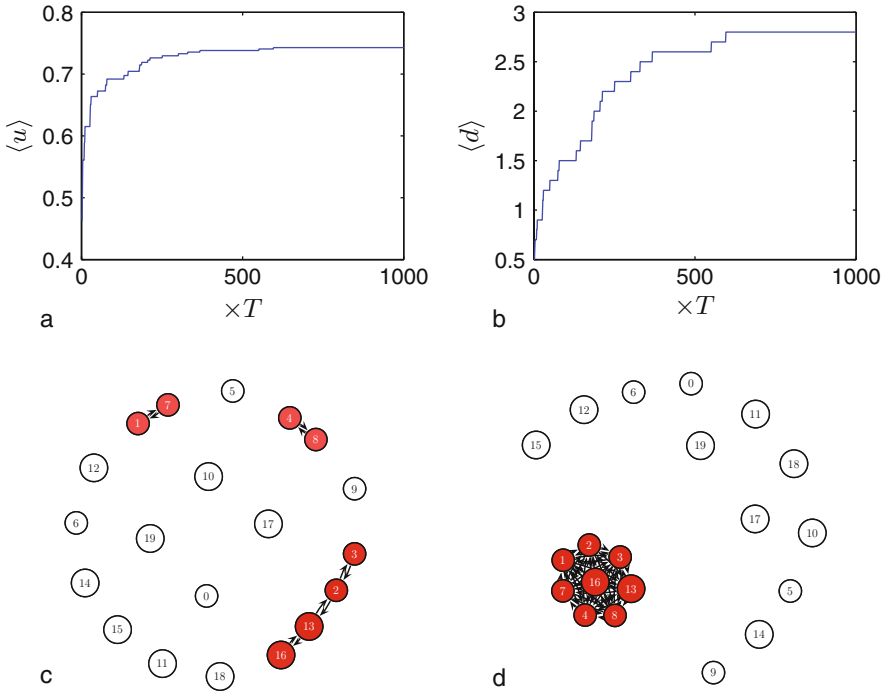


Fig. 8.18 Bilateral link formation: (**a**) Average value of knowledge. (**b**) Average degree. (**c**) Initial random graph (for reasons of visualization we have chosen a rather sparse random graph). (**d**) Graph after 1000 iterations (in the equilibrium)

given in (8.19)) most of the agents are isolated nodes and thus have vanishing values of knowledge. Only a few of them are organized in 2-cycles and small subgraphs consisting of multiple 2-cycles. As we will show, the reason for this is that as soon as there exists a shortcut (a smaller cycle) in a larger cycle agents try to free-ride and, after the other agents have realized that and sopped sharing their knowledge with them, they get isolated and experience vanishing values of knowledge. One can interpret this result as follows: Even though agents could in principal form indirect reciprocal interactions the resulting equilibrium network consists only of direct reciprocal interactions (2-cycles and clusters of 2-cycles).

We can give an example of the process of the reduction of cycles in a graph G with three nodes for parameter values $d = 0.5$, $b = 1$, and $c \in (0, 1)$. By numerically comparing utilities (the fixed points of the value of knowledge) before and after a link is created or deleted, we show that there exists a sequence of link deletions and creations which transform a 3-cycle into a 2-cycle while every link change is associated with an increase in the utility (the fixed point in the value of knowledge) of the initiating agent (Jackson (2003) calls this sequence of graphs an "improving path").

In Fig. 8.19 (left) agent 3 creates a link to agent 1 because in this range of parameters this increases her value of knowledge. This can be seen in Fig. 8.19

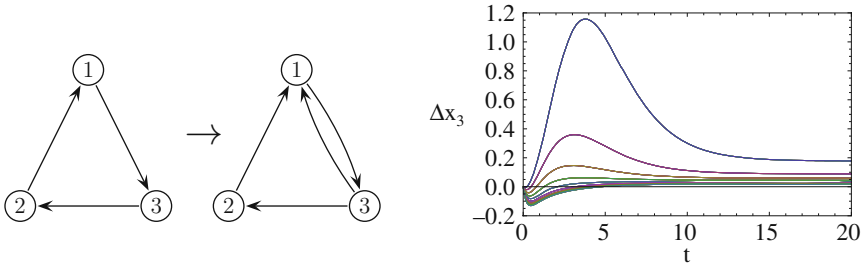


Fig. 8.19 Agent 3 forms a link to agent 1, e_{31}, and thus a 2-cycle is created inside a 3-cycle. The situation is illustrated on the *left hand side*. On the *right*, the evolution of the values of knowledge for different values of cost is shown

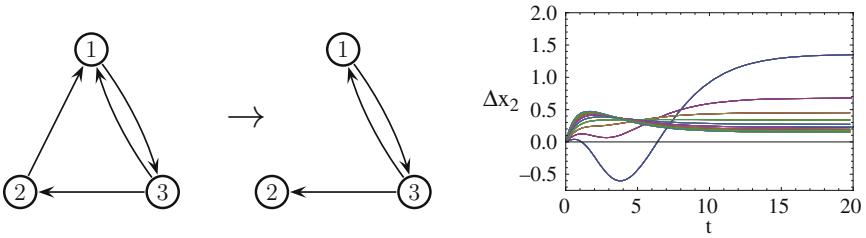


Fig. 8.20 Deletion of the link e_{21}. Agent 2 is not sharing any knowledge with others but only receiving knowledge from agent 3. Thus, agent 2 is free-riding. The situation is illustrated on the *left hand side*. On the *right*, the evolution of the values of knowledge for different values of cost are shown

(right), where the increase Δx_3 different costs $c \in (0, 1)$ are plotted and $\Delta u_3 = \lim_{t\to\infty} \Delta x_3 > 0$.

In Fig. 8.20 (left) agent 2 removes her link to agent 1 and thus she stops contributing knowledge but instead is only receiving knowledge from agent 3. We say that agent 2 is free-riding. This increases her utility, since $\Delta u_2 = \lim_{t\to\infty} \Delta x_2 > 0$, as can be seen in Fig. 8.20 (right) for different costs c.

Finally, in Fig. 8.21 (left) agent 3 removes her link to agent 2 because she is better off, as illustrated in Fig. 8.21 (right), when she stops contributing knowledge

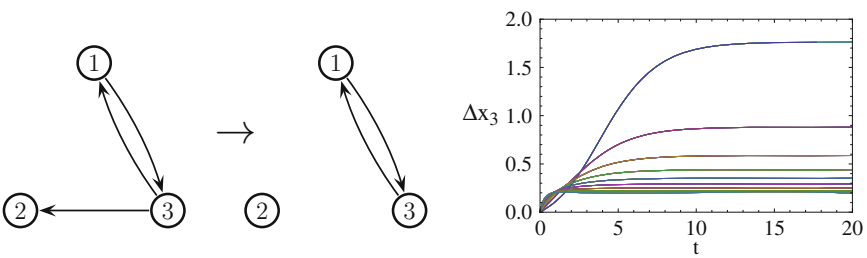


Fig. 8.21 Deletion of the link e_{32} by agent 3. Agent 3 realizes that she is better off by not sharing her knowledge with agent 2. Agent 2, who was free-riding before now gets isolated and experiences a vanishing value of knowledge in the long run. The situation is illustrated in the figure to the left. On the right, the evolution of the values of knowledge for different values of cost are shown

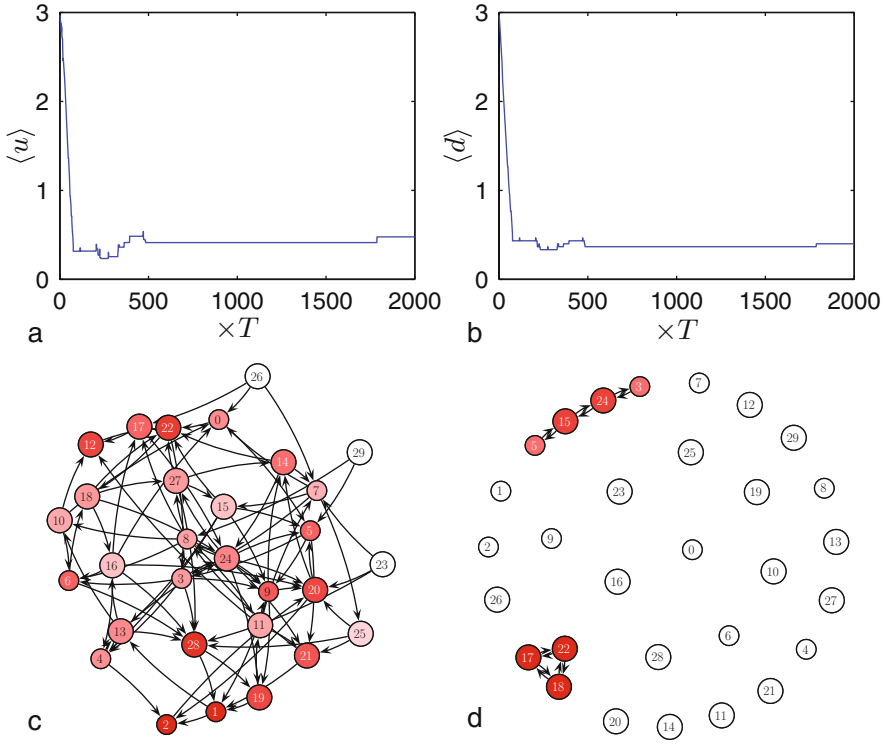


Fig. 8.22 Unilateral link formation: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 2000 iterations (in the equilibrium). For the parameter values $d = 0.5$, $b = 0.5$, $c = 0.1$, used in this simulation, the complete graph K_3 is an equilibrium. Note from Appendix (A) one can see that for parameter values $d = 0.5$, $b = 1$, $c = 0.5$ this is no longer the case and K_3 would be reduced to a 2-cycle C_2

to an agent that is nothing contributing in return. This is actually true for agent 2. Agent 2 therefore gets isolated and experiences a vanishing value of knowledge in the long run, $\lim_{t\to\infty} x_2 = 0$. Her utility is null.

We end up in a setting where out of a cooperation of many (the sharing of knowledge) only a small set of cooperators remains and all the remaining agents vanish, i.e., have vanishing values of knowledge and utility. We can see this in a simulation starting from an initial random graph with 30 agents and the resulting equilibrium network in Fig. 8.22 (bottom right).

Since the performance of the system in terms of the total value of knowledge is very low, we investigate in the next section the conditions under which the performance can be increased (with more agents being permanent in the equilibrium). We find that the existence of a positive network externality (explained in the next section) can enhance the performance of the system.

8.3.9.9 Introducing Positive Network Externalities

In this section, we study the growth of the value of knowledge which includes an additional benefit term contributing to an increase in the value of knowledge. This additional benefit depends on the network structure itself. In the economic literature (Mas-Colell et al., 1995; Tirole, 1988), "positive network externalities arise when a good is more valuable to a user the more users adopt the same good or compatible goods." In our model we define a network externality simply as a function of the network structure that affects the utility of an agent. Including the externality in the benefit can yield more complex structures with non-vanishing knowledge values as equilibrium networks. More precisely, we introduce weights for the connections between the agents that depend on a measure of network externalities that we will introduce in the following sections. This means that, if we have strong network externalities between agent i and agent j, then the weight w_{ij} will represent this effect and attain a high value. Taking into account the existence of such network externalities, the growth of the value of knowledge of agent i is given by the following equation:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = -dx_i + b\sum_{i=1}^{n} a_{ji}x_j + \underbrace{b_e\sum_{i=1}^{n} w_{ji}x_j}_{\text{positive network externality}} - c\sum_{i=1}^{n} a_{ij}x_i^2 \tag{8.82}$$

and the utility is again given by $u_i = \lim_{t\to\infty} x_i(t)$. Link changes are based on the increase in utility. The network benefit incorporates the fact that the value of knowledge can change with the number of users of that knowledge, (8.82). But the number of users can either enhance or diminish the value of knowledge that is being transferred between agents, depending on the type of knowledge under investigation. On one hand, the value can decrease with the number of agents that pass on that knowledge. Knowledge is attenuated with the distance from the creator to the receiver. We study this type of knowledge with a link weight defined in (8.82) and denoted by w_{ji}^{c}. In the next section, we study the opposite effect: the value of knowledge increases with the number of users. This holds for example for general purpose technologies that get more valuable the more they are applied and used in different contexts (and users). The link weights used for this type of knowledge in (8.82) are denoted by w_{ji}^{ccn}, w_{ji}^{cce}, where the first measures the number of agents using that knowledge and the second the number of interactions.

We introduce different link weights, denoted by w_{ji}^{c}, w_{ji}^{ccn}, w_{ji}^{cce}. Moreover, agents are creating and deleting links unilaterally (utility driven dynamics). We then study the effect of different weights on the equilibrium networks obtained.

Shortest-Path Centrality

The growth function of the value of agent i is given by

$$\frac{dx_i}{dt} = -dx_i + \sum_{j=1}^{n}(ba_{ji} + b_e w_{ji}^{c})x_j - c\sum_{j=1}^{n} a_{ij}x_i^2. \tag{8.83}$$

The utility is given by $u_i = \lim_{t\to\infty} x_i(t)$ (large T). Link changes are accepted on the basis of an increase in utility. The shortest-path centrality measure computes the sum of the inverse lengths of all the shortest paths containing the link for which the centrality is computed. If two agents are not connected then the length of the path is assumed to be infinity and thus its weight is zero. Instead, if two agents are directly connected via a link, then the weight is one. The weight values links more that bring agents closer to each other. This is a similar approach to the Connections Model introduced in Sect. 8.2.2 with a utility given by (8.4). The centrality link weight, w_{ij}^{c}, is then computed as follows:

$$w_{ij}^{c} = \sum_{v\in V} \frac{1}{(d_{jv}+1)},\ w_{ij}^{c} \in [0,1], \tag{8.84}$$

d_{jv} is the shortest path between node j and node v. If there exists no path between two nodes, then the distance between them is infinity.[39]

In simulations, Fig. 8.23, we observe that in the equilibrium network only a few agents have non-vanishing utilities (the asymptotic knowledge values) and most of them are isolated nodes with zero utilities. This result does not differ too much from the studies in Sect. 8.3.9.8 where no externality is considered. Apparently, if more agents should have non-vanishing utilities induced by an additional benefit depending on the network structure, this cannot be realized with the centrality link weight.

Circuit-Centrality

The *circuit-centrality* measure puts a weight on the links that depends on the number of distinct nodes that are contained in all the circuits going through the link under consideration. The motivation is that, if many agents are involved in the transfer of knowledge and this knowledge then comes back to the agent (thus creating a feedback on the technology issued by the agent), it gets an added value (e.g., for general purpose technologies (GPT) (Bresnahan and Trajtenberg, 1995; Karshenas and Stoneman, 1995; Cohen, 1995)). The more agents use a technology the more it is improved and so the more agents are involved in such a feedback loop the higher is the value of the technology. We can either count the number of different agents involved in this feedback loop or the number of interactions (links). Either possibility is explored in the next sections. This is an alternative way to study the

[39] We use a standard *depth-first-search* algorithm to compute the shortest paths. More details on this algorithm and further discussion is given in Ahuja et al. (1993), Cormen et al. (2001), Steger (2001), Steger and Schickinger (2001).

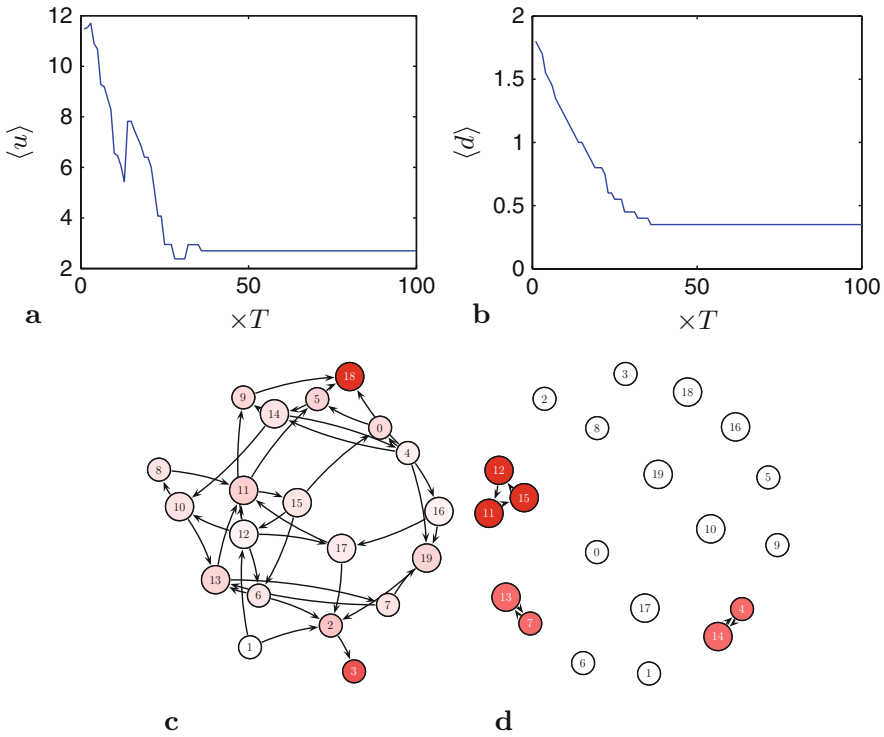


Fig. 8.23 Shortest-path Centrality: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 500 iterations (in the equilibrium)

emergence of indirect reciprocity where others Nowak and Sigmund (2005) have studied it by introducing a (global) reputation mechanism.

We then define the weight of a link w_{ij} as (i) the number m_n of distinct *nodes* that are in the circuits from node i to j,

$$w_{ij}^{\text{ccn}} = \frac{m_n}{n}, \; w_{ij}^{\text{ccn}} \in [0, 1] \tag{8.85}$$

and (ii) the number m_e of distinct *links* that are in the circuits from node i to j,

$$w_{ij}^{\text{cce}} = \frac{m_e}{n(n-1)}, \; w_{ij}^{\text{cce}} \in [0, 1]. \tag{8.86}$$

An example of the different link weights can be seen in Fig. 8.23.

In order to compute all circuits in a directed graph G, one needs to compute the trails in G. The closed trails then are the circuits in G. We use an algorithm to compute all trails in G from a given source node s. An explanation of the algorithm is given in Appendix (B).

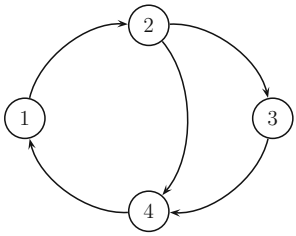


Fig. 8.24 The link e_{12} is contained in two circuits with nodes 1, 2, 4 and 1, 2, 3, 4. The number of distinct nodes in these circuits is 4 and the number of distinct links is 5. Accordingly, $w_{ij}^{\text{ccn}} = \frac{4}{4} = 1$ and $w_{ij}^{\text{cce}} = \frac{5}{12} = 0.42$

By introducing the circuit-centrality externality, we will show that more agents are permanent in the equilibrium network. The performance of the system is increased compared to the equilibrium networks that emerge with unilateral link formation without this externality.

Using circuit-centrality measure with the number of nodes (8.85), the growth of the value of knowledge is given by

$$\frac{dx_i}{dt} = -dx_i + \sum_{j=1}^{n}(ba_{ji} + b_e w_{ji}^{\text{ccn}})x_j - c\sum_{j=1}^{n} a_{ij}x_i^2. \tag{8.87}$$

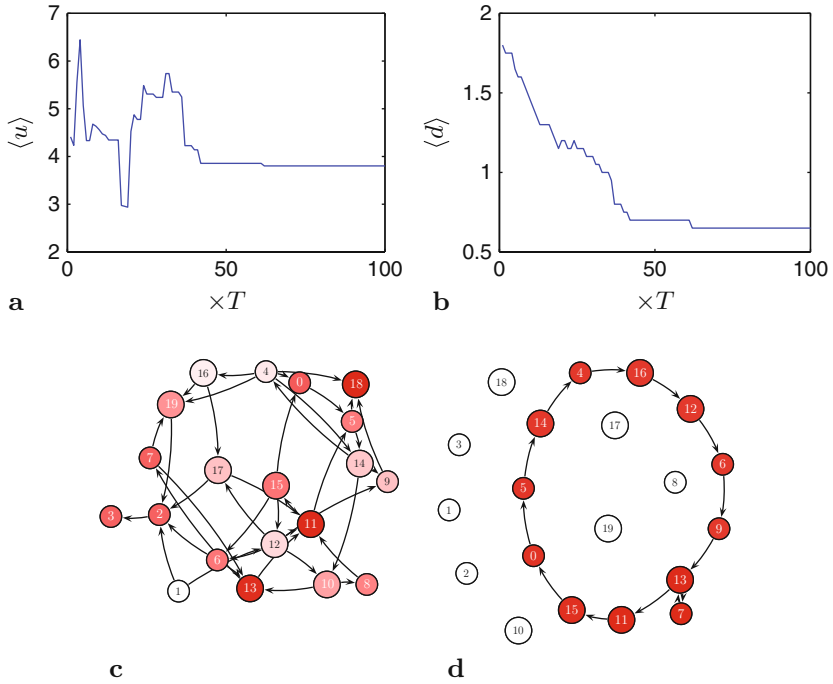


Fig. 8.25 Circuit centrality as a function of the number of nodes: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 500 iterations (in the equilibrium)

The utility is given by $u_i = \lim_{t\to\infty} x_i(t)$ (large T). Link changes are accepted on the basis of an increase in utility. Different to the centrality externality (Sect. 8.3.9.9) we observe larger cycles as the equilibrium networks. This can be seen in the simulation run in Fig. 8.25. This positive externality allows for more agents to be permanent in the equilibrium network than without an externality or with the centrality-externality. We thus obtain a higher performance of the system.

Using the circuit-centrality measure with the number of links (8.86), the growth of the value of knowledge is given by

$$\frac{dx_i}{dt} = -dx_i + \sum_{j=1}^{n}(ba_{ji} + b_e w_{ji}^{\mathrm{cce}})x_j - c\sum_{j=1}^{n} a_{ij}x_i^2. \tag{8.88}$$

The utility is given by $u_i = \lim_{t\to\infty} x_i(t)$ (large T). Link changes are accepted on the basis of an increase in utility. The circuit-centrality with the number of links values the number of interactions instead the number of agents, that take part in the transfer of knowledge. We still observe (Fig. 8.26) the emergence of circuits as equilibrium networks and a similar level of performance (in terms of the total value

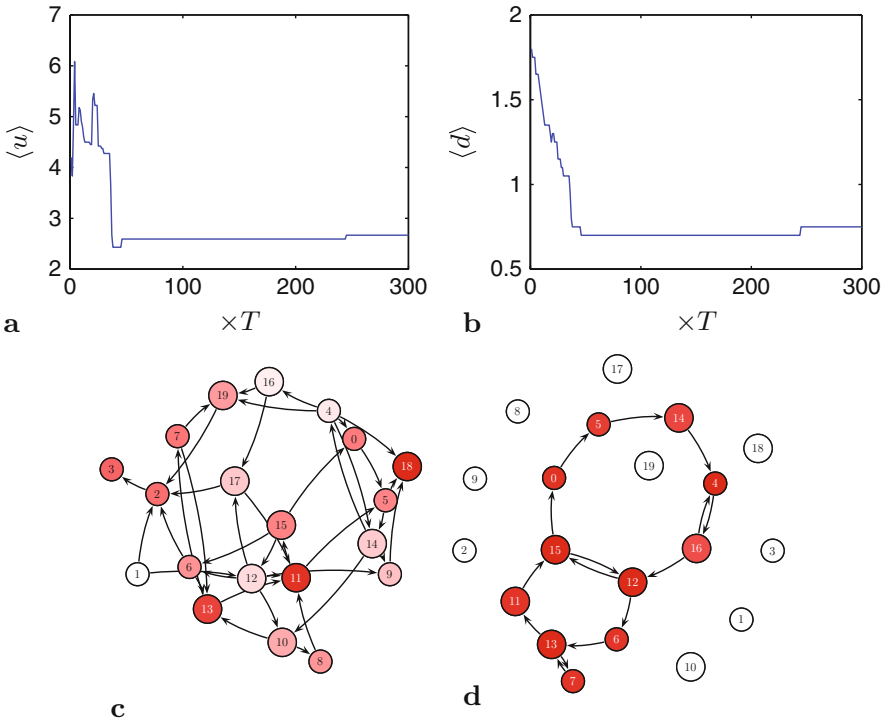


Fig. 8.26 Circuit centrality as a function of the number of links: (**a**) Average utility. (**b**) Average degree. (**c**) Initial random graph. (**d**) Graph after 500 iterations (in the equilibrium)

of knowledge of the system). But now there are more circuits (more links) in the subgraph containing the set of permanent agents. The equilibrium network has a higher level of redundancy (since its has more circuits and links) and is therefore more robust against the destruction of a single circuit (induced by a node or link failure).

8.4 Discussion and Conclusion

8.4.1 Results from the Novel Modeling Approach

In the following, we summarize the results found by studying our model of innovation dynamics, as described in Sects. 8.3. Let us start by looking at the dynamics of the value of knowledge in a static network, in Sect. (8.3.8). If we assume that growth occurs only through interaction among agents (thus neglecting "in-house" R&D capabilities), then the network sustains itself only through cycles (more precisely through closed walks or strongly connected components). Agents survive and grow only if they are part of a cycle (strongly connected component) or if they are connected to such a cycle through an incoming path.[40] We have shown that an innovation network which is acyclic will have vanishing knowledge values for all agents in the network. However, if agents form cycles they have permanent knowledge values.

Considering the evolution of the network we have studied two different settings, Extremal Dynamics and Utility Driven Dynamics. If the network evolves according to a least fit selection mechanism (extremal dynamics) then we observe crashes and recoveries of the knowledge values of the agents and the network itself. Thus, an extremal market selection mechanism which replaces the worst performing agent with a new agent cannot generate equilibrium networks nor does it sustain a high performance in the value of knowledge of the individual agents or the economy as a whole. Notice also that extremal dynamics means that agents are completely passive and have no control on whom they interact with.

In a more realistic setting (utility driven dynamics), agents decide with whom they interact and they do so in order to increase their utility. In the context of innovation, this corresponds to their value of knowledge. The information processing capabilities of agents may be limited, especially if there is a large number of agents in the economy. Thus, we allow agents to decide themselves to create or delete links on a trial and error basis. Those interactions that prove to be beneficial are maintained while detrimental ones are severed. We find that, under these conditions, the evolution of the network depends on the cost, c_{ij}, of an interaction between the agents, the type of link formation (unilateral versus bilateral), and the time horizon T after which interactions are evaluated.

[40] Jain and Krishna (2001) have denoted this set of nodes the autocatalytic set (ACS).

In the case of null cost, agents always form new links and thus the complete graph is eventually realized.

We have shown in Sect. 8.3.7 that the knowledge values of agents vanish if the underlying network does not contain a cycle (similar to the results obtained by Maxfield (1994), Rosenblatt (1957). Equilibrium networks contain cycles, a result which is similar to the findings of Bala and Goyal (2000), Kim and Wong (2007)). However, the evolution of the network driven by the selfish linking process of agents can lead to the destruction of these vital cycles. For a short time horizon T, and unilateral link formation, cycles get destroyed because agents free-ride and delete their outgoing links as it is beneficial in the short run to save the costs of supporting other agents. As a result, the whole innovation network is destroyed. On the other hand, if the time horizon is long enough, agents do not delete the cycles they are part of.

However, even when the time horizon is long, large cycles get destroyed in favor of smaller ones when agents unilaterally form or delete links. The network, starting from an initial state of high density, evolves into an absorbing state in which most of the nodes are isolated and few pairs of nodes are connected by bilateral links. These pairs are trivial cycles of length $k = 2$.[41]

Recall that pairwise connections are direct reciprocal interactions. This means that even though agents are unilaterally forming links and therefore indirect reciprocal interactions would be possible in principle (this is equivalent to interaction taking place on a cycle of length $k \geq 3$) no relation of indirect reciprocity is able to emerge nor to survive. From the point of view of the global performance of the innovation network, this is a very unsatisfactory situation.

In Sect. 8.3.9, we have studied situations in which even unilateral knowledge exchange can have a higher performance in terms of the number of permanent agents and their total value of knowledge. We introduce an externality in the knowledge growth function which increases the value of knowledge of the agent depending on their position in the network. We study a type of technology where the value of knowledge decreases with the number of agents transferring the knowledge. Here unilateral knowledge exchange still leads to equilibrium networks with a low performance and only a few permanent agents.

However, for a type of knowledge where its value increases with the number of agents that transfer and use it, more agents can be permanent in the equilibrium network and the system performance is increased. Moreover, if the number of interactions instead of the number of users determines the added value of the knowledge that is being transferred, then the equilibrium has not only a higher performance than in the setting, where knowledge is attenuated with the number of users or where no externalities are considered, but it is also more robust against node or connection failures. We observe that, in our framework, indirect reciprocity emerges

[41] This is different to the results obtained by Kim and Wong (2007) since there the benefit term in the utility of the agents depends on the size of the connected component but not on its structure.

Table 8.4 Overview of the equilibrium networks that are realized under different assumptions on the network evolution and a quadratic cost function

Network evolution / Cost function	Extremal dynamics	Utility driven dynamics		
			Unilateral	
		Bilateral	Without externality	With externality
$c_{ij} = cx_i^2$	Network breakdown	K_n	Set of C_2	Set of $C_{k\geq 2}$

if it is associated with a positive externality, taking into account the structure of the network.

If agents form or delete links bidirectionally, that means, every exchange of knowledge is direct reciprocal, the network evolves into a complete graph. This equilibrium network has a high performance and all agents are permanent. In our study, we find that unilateral knowledge exchange is always inferior to bilateral knowledge exchange. But the above discussion has shown that, when bilateral knowledge exchange is not possible and agents are sharing their knowledge unilaterally, innovation networks are still able to emerge.

The different cases studied in this section have shown that the equilibrium innovation network that is realized in the evolution of the system depends critically on the assumptions made on the behavior of agents (extremal dynamics versus utility driven dynamics), on their time horizon for evaluating their decisions and on the cost associated with the sharing of knowledge. These results are summarized in Table 8.4

8.4.2 General Conclusions

In this chapter, we studied a variety of different models for innovation networks. We started by discussing the importance of networks in economics and emphasized that these networks are intrinsically dynamic and composed of heterogeneous units. The notion of a *complex network* was used in the beginning to briefly explain how statistical physics can be involved to study them. We tried to classify different approaches to modeling economic networks, in particular we considered the connection between the state variables associated with the nodes of a network, e.g., the productivity level of a firm, and the dynamics of the network itself, i.e., the interactions between firms.

Before developing our own modeling framework, we discussed some basic models of economic networks with agents engaged in knowledge production. These models show that the economy can evolve into equilibrium networks which are not necessarily efficient. Moreover, the equilibrium networks that emerge in these models are rather simple. We briefly introduced some models in which more complicated network structures emerge, which may be closer to real-world innovation networks. We then discussed models in which cycles, i.e., closed feedback loops,

play an important role in the network formation and the performance of the system (similar to Rosenblatt, 1957).

The major part of the chapter was devoted to the development of our own modeling framework, which is based on catalytic knowledge interactions. In this setting, there are permanent agents (with non-vanishing knowledge values) only if the underlying network contains a cycle. We investigated the evolution and performance of the system under different selection mechanisms, i.e., a least fit selection mechanism, denoted by Extremal Dynamics, versus Utility Driven Dynamics in which agents decide upon their interaction partners in a trial and error procedure. We observe that a least fit mechanism cannot generate stable networks nor sustain high performance in knowledge production. Moreover, such a mechanism assumes that agents are completely passive entities. In the case of Utility Driven Dynamics, agents choose their actions in order to increase their utility but their information processing capabilities are limited. If agents are evaluating their interactions after a time long enough, we obtain equilibrium networks with non-vanishing (permanent) knowledge production.

In our framework, we investigated different assumptions about the behavior of agents, that is, we either assume that agents share knowledge bilaterally or unilaterally. If all interactions are bilateral, the equilibrium network is a complete graph and it has the highest performance. However, if direct reciprocal interactions cannot be enforced (which means that links are not necessarily bilateral), we still observe the emergence of networks of knowledge sharing agents. But in the equilibrium network only bilateral interactions remain. Moreover, only a few agents are permanent and the system has a low performance compared to the case of purely bilateral interactions. However, for unilateral interactions, the number of permanent agents can be significantly increased, for a type of technology where the number of users increases its value.

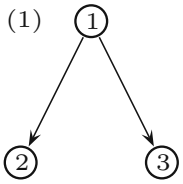
Our studies show that the range of innovation networks that can emerge in this general framework is affected by various parameters. Amongst these are information processing capabilities of agents, their time horizon, their behavior in interaction with others, the cost associated with the sharing of knowledge, and the type of technology which agents produce and transfer.

The variety of possible networks is quite large and the network model appropriate for a given application should be determined based on the specificities of the problem under investigation.

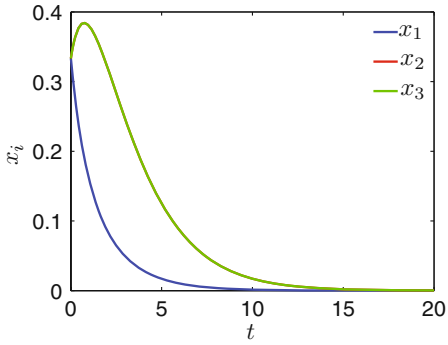
Appendix A: Stationary Solutions for Three Agents

Example 29 We compute the fixed points for all graphs (auto-morphisms) with $n = 3$ nodes and initial values $\mathbf{x}(0) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$. For the numerical integration we set $d = 0.5$, $c = 0.5$, and $b = 1$. The fixed points (stationary solutions) are denoted x_i^* for $i = 1, 2, 3$. Where possible, we give the analytical solutions for the positive fixed points. $x_i^* = 0$ is a fixed point for all graphs.

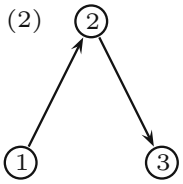
(1)



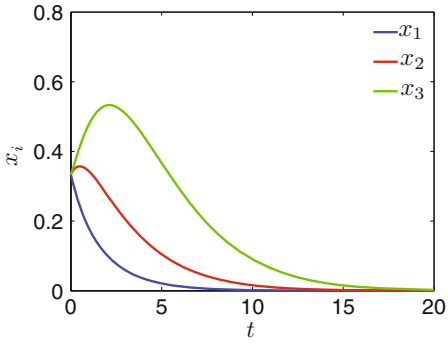
$$A_1 = \begin{pmatrix} 0\,1\,1 \\ 0\,0\,0 \\ 0\,0\,0 \end{pmatrix}$$



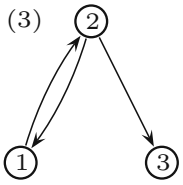
(2)



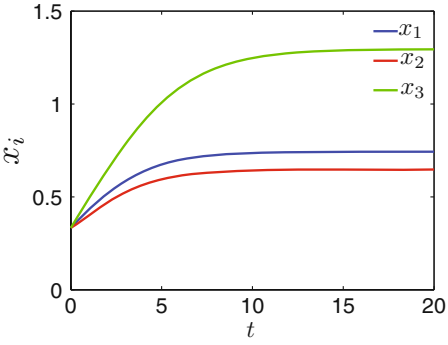
$$A_2 = \begin{pmatrix} 0\,1\,0 \\ 0\,0\,1 \\ 0\,0\,0 \end{pmatrix}$$



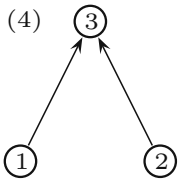
(3)



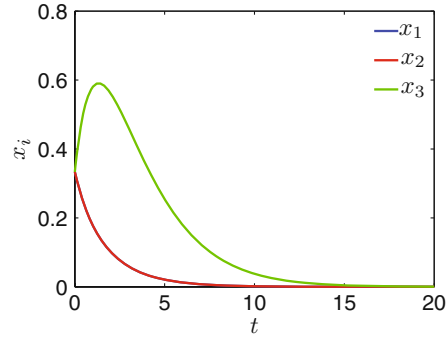
$$A_3 = \begin{pmatrix} 0\,1\,0 \\ 1\,0\,1 \\ 0\,0\,0 \end{pmatrix}$$



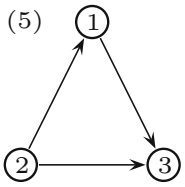
(4)



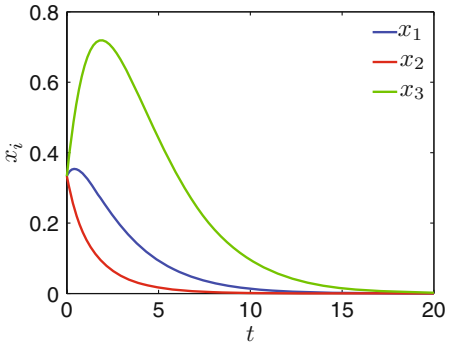
$$A_4 = \begin{pmatrix} 0\,0\,1 \\ 0\,0\,1 \\ 0\,0\,0 \end{pmatrix}$$



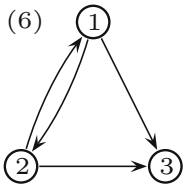
(5)



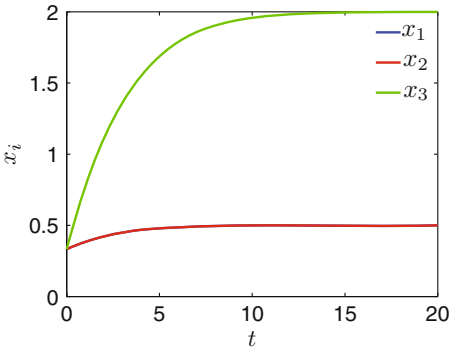
$$A_5 = \begin{pmatrix} 0\,0\,1 \\ 1\,0\,1 \\ 0\,0\,0 \end{pmatrix}$$



(6)



$$A_6 = \begin{pmatrix} 0\,1\,1 \\ 1\,0\,1 \\ 0\,0\,0 \end{pmatrix}$$

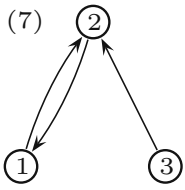


Non-trivial fixed points are given by

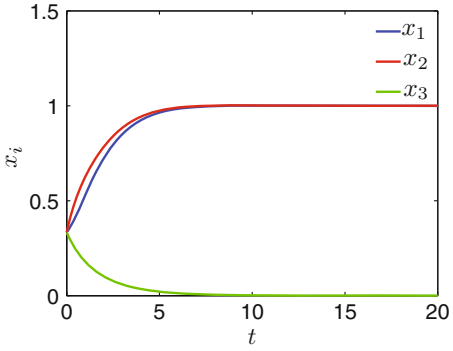
$$x_1^* = x_2^* = \frac{b-d}{2c}, \tag{8.89}$$

$$x_3^* = \frac{b(b-d)}{cd}. \tag{8.90}$$

(7)



$$A_7 = \begin{pmatrix} 0\,1\,0 \\ 1\,0\,0 \\ 0\,1\,0 \end{pmatrix}$$

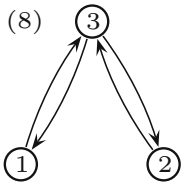


Non-trivial fixed points are given by

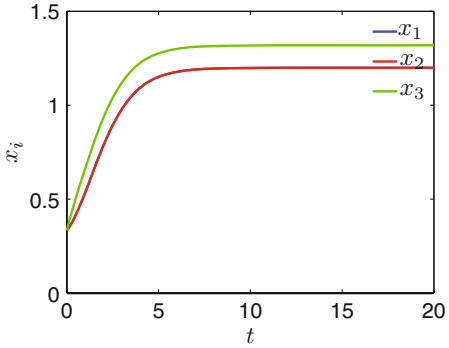
$$x_1^* = x_2^* = \frac{b-d}{c}, \tag{8.91}$$

$$x_3^* = 0. \tag{8.92}$$

(8)



$$A_8 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

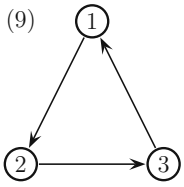


$x_1^* = x_2^*$ are the roots of the polynomial

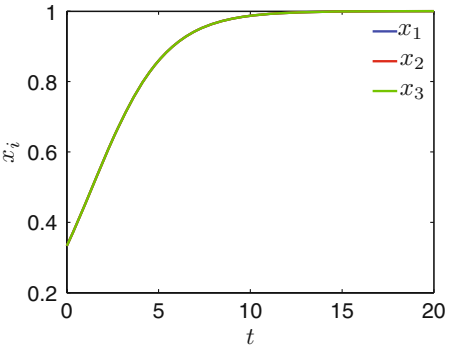
$$x^3 + \frac{2d}{c}x^2 + \frac{d(b+2d)}{2c^2}x + \frac{b(d^2-2b^2)}{2c^3} = 0$$

and $x_3^* = \frac{d}{b}x + \frac{c}{b}x^2$.

(9)



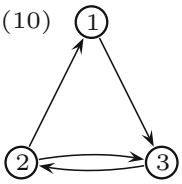
$$A_9 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$



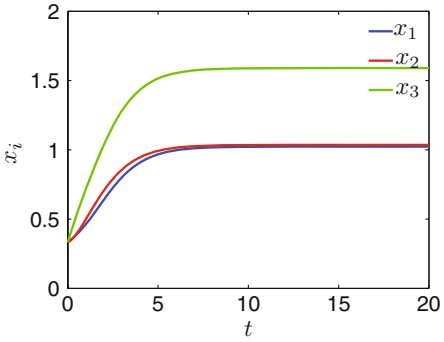
Non-trivial fixed points are given by

$$x_1^* = x_2^* = x_3^* = \frac{b-d}{c}. \tag{8.93}$$

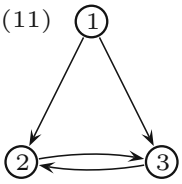
(10)



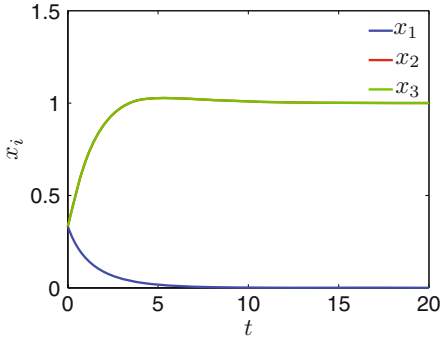
$$A_{10} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$



(11)



$$A_{11} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

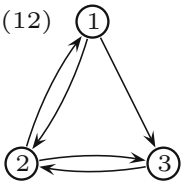


Non-trivial fixed points are given by

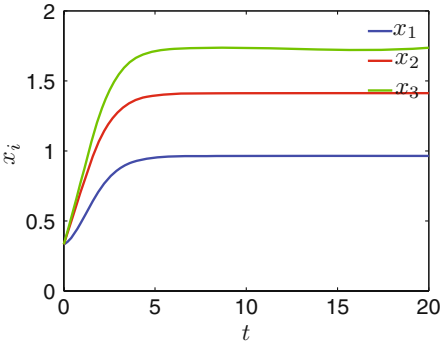
$$x_2^* = x_3^* = \frac{b-d}{c}, \tag{8.94}$$

$$x_1^* = 0. \tag{8.95}$$

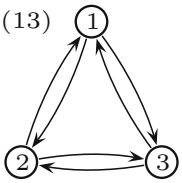
(12)



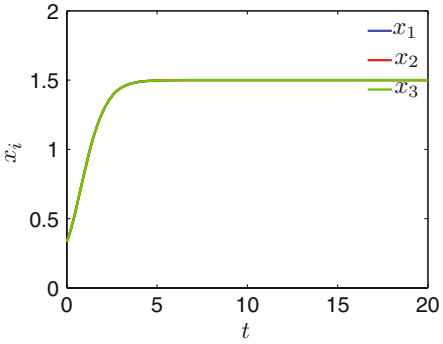
$$A_{12} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$



(13)



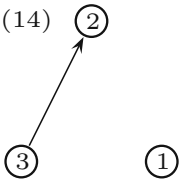
$$A_{13} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$



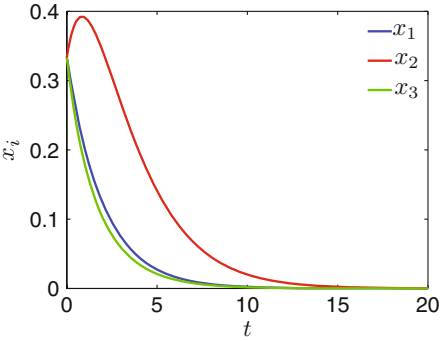
Non-trivial fixed points are given by

$$x_1^* = x_2^* = x_3^* = \frac{2b - d}{2c}. \tag{8.96}$$

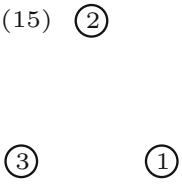
(14)



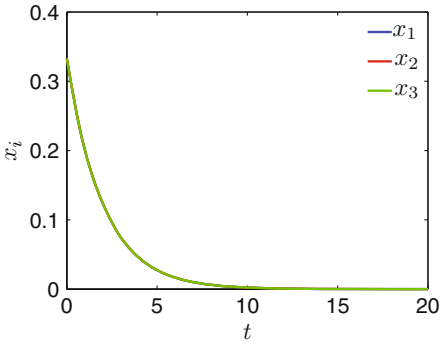
$$A_{14} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

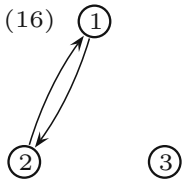


(15)

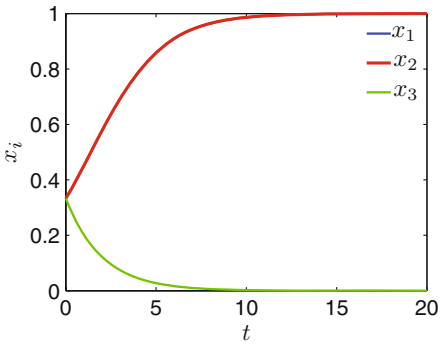


$$A_{15} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$





$$A_{16} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$



Non-trivial fixed points are given by

$$x_3^* = 0, \tag{8.97}$$

$$x_1^* = x_2^* = \frac{b-d}{c}. \tag{8.98}$$

Appendix B: All-Trails-Single-Source Algorithm

With algorithm All-Trails-Single-Source we want to compute all trails from a given node $s \in V$ to all other nodes in a directed graph $G = (V, E)$. From these trails we can extract the trails which end in node s and thus form circuits. The pseudo-code for the All-Trails-Single-Source algorithm is given in Algorithm 1. $N^+(v)$ denotes the out-neighborhood of node v. In the following, we will give a short description of the algorithm.

Algorithm 1 All-Trails-Single-Source

```
S ← newStack();
v ← s;
W ← {}; {initialization of empty list of trails}
loop
  if ∃u ∈ N+(v)\{s} s.t. the link e_vu cannot be appended to W to create a new trail then
    S.push(v);
    W[u].addEdge((v, u));
    v ← u;
  else if S.isEmpty() == false then
    v ← s.pop();
  else
    break;
  end if
end loop
return W;
```

Similar to a depth-first-search algorithm, links are explored out of the most recent discovered node v that still has unexplored links leaving it. This procedure of exploring links can be represented by a search tree T. The tree T explored by Algorithm 1 contains all trails starting at the source s to every node in G.

At every node i, a list $W[i]$ of trails leading from the source s to i is assigned. When the next link e_{ij} from i to j is processed, to all trails in $W[i]$ the link e_{ij} is appended (if this is possible, meaning that no link repetition is allowed), denoted by $W[i]+e_{ij}$. At node j these trails are added, that is $W'[j] = W[j] \cup \{W[i]+e_{ij}\}$. This procedure is continued until the algorithm terminates. The algorithm terminates, if there are no further links available for exploration. The progress of Algorithm 1 on a directed graph with four nodes and two circuits containing the nodes (2, 3, 4, 1) and (2, 4, 1), respectively, is shown in Fig. 8.27.

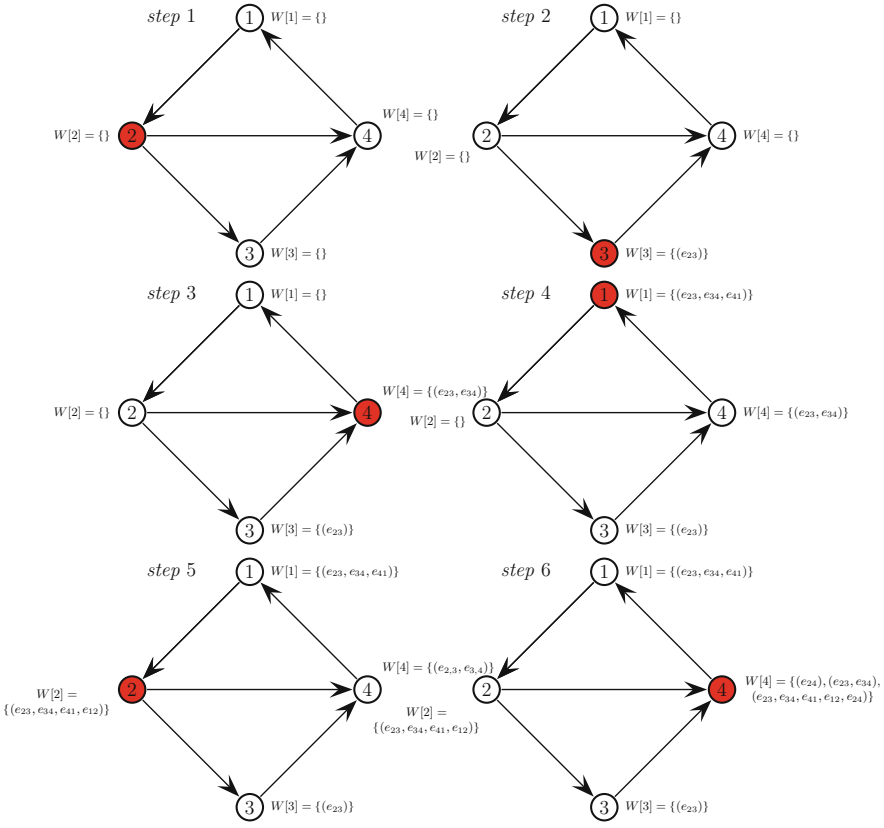


Fig. 8.27 The progress of Algorithm 1 on a directed graph with node 2 as the source node. The node indicated in red is visited in the succeeding steps. After step 6 no further trails are added to the list of trails

Appendix C: Simulation Parameters

The parameters shown in Table 8.5 have been used for the simulation runs presented in Sect. 8.3.9.6.

Table 8.5 Simulation parameters

Description	Variable	Value
Initial link creation probability	p	0.1
Initial value of knowledge	$\mathbf{x}(0)$	1.0
Number of agents (without externality)	n	30
Number of agents (with externality)	n	20
Max. numerical integration time (time horizon)	T	100
Numerical integration time step	Δt	0.05
Max. number of network updates	N	[100, 5000]
Benefit	b	0.5
Decay	d	0.5
Cost	c	0.1

Acknowledgments We are much obliged to Koen Frenken who helped a lot to improve the model presented in this chapter. Moreover, we would like to thank Giorgio Fagiolo for the clarifying discussions. In writing the chapter we have always been conscious of our debt to colleagues who have helped us in bringing this work to the present state. Among these are in particular Kerstin Press and Mauro Napoletano who have pointed us to some inconsistencies and sections which needed a more extended explanation of the material treated there in the early versions of this chapter. The discussion with Dimo Brockhoff on the algorithms were of great help. We would like to thank the discussant and the critical audience at the First International Conference on Economic Sciences with Heterogeneous Interacting Agents in Bologna, 2006. Finally, we would like to thank Peter Howitt, Robert Axtell, Herbert Gintis, and Leigh Tesfatsion for taking so much time in discussing the model at the Seventh Trento Summer School on Agent-Based Computational Economics in Trento, 2006 and Axel Leijonhufvud for making these discussions possible.

References

Aghion, P., Howitt, P. (1998). *Endogenous Growth Theory*. Cambridge, MA: MIT Press.

Ahuja, Magnanti; Orlin (1993). *Network Flows: Theory, Algorithms, and Applications*. Upper Saddle River, NJ: Prentice-Hall.

Allen, R. C. (1983). Collective invention. *Journal of Economic Behavior & Organization* **4(1)**, 1–24.

Amaral, L. A. N., Cizeau, P., Gopikrishnan, P., Liu, Y., Meyer, M., Peng, C.-K., Stanley, H. E. (1999). Econophysics: Can statistical physics contribute to the science of economics? *Computer Physics Communications,* **121(122),** 145–152.

Amaral, L., Gopikrishnan, P., Matia, K., Plerou, V., Stanley, H. (2001). Application of statistical physics methods and concepts to the study of science & technology systems. *Scientometrics* **51**, 9–36.

Antonelli, C. (1996). Localized knowledge percolation processes and information networks. *Journal of Evolutionary Economics* **6(3)**, 281–295.

Arora, A., Fosfuri, A., Gambardella, A. (2004). *Markets for Technology: The Economics of Innovation and Corporate Strategy*. Cambridge, MA: MIT Press.

Arrow, K. J. (1962). Economic welfare and the allocation of resources for innovation. *The Rate and Direction of Inventive Activity*, National Bureau of Economic Research, Princeton: Princeton U. Press. pp. 609–625.

Axelrod, R., Tesfatsion, L. S. (2006). *A Guide for Newcomers to Agent-Based Modeling in the Social Sciences*. Staff General Research Papers 12515, Iowa State University, Department of Economics.

Bak, P., Sneppen, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Physical Review Letters* **71(24)**, 4083–4086.

Bala, V., Goyal, S. (2000). A non-cooperative model of network formation. *Econometrica* **68(5)**, 1181–1230.

Baldry, R., Ghosal, S. (2005). Irreducible economies and strongly connected graphs. *Journal of Mathematical Economics* **41(8)**, 937–956.

Ballester, C., Calvo-Armengol, A., Zenou, Y. (2006). Who's who in networks. Wanted: The key player. *Econometrica* **74(5)**, 1403–1417.

Barro, R. J., Sala-i Martin, X. (2004). *Economic Growth*. Cambridge, MA: MIT Press.

Battiston, S. (2003). *Dynamics and Evolution of Complex Networks in Socio-Economical Systems*. Ph.D. thesis, Ecole Normale Superieure, Ecole Doctorale de Paris 7.

Bollobas, B. (1985). *Random Graphs*. 2nd edn. Cambridge, NY: Cambridge University Press.

Bolton, G. E., Katok, E., Ockenfels, A. (2005). Cooperation among strangers with limited information about reputation. *Journal of Public Economics* **89**, 1457–1468.

Bolton, G. E., Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *The American Economic Review* **90(1)**, 166–193.

Boyd, S. (2006). Linear Dynamical Systems. Lecture Notes, Stanford University.

Braun, M. (1993). *Differential Equations and Their Applications*. Texts in Applied Mathematics. 4th edn. Ney York Springer.

Bresnahan, T. F., Trajtenberg, M. (1995). General purpose technologies "Engines of growth"? *Journal of Econometrics* **65(1)**, 83–108.

Carayol, N., Roux, P. (2003). *Self-Organizing Innovation Networks: When do Small Worlds Emerge? Working Papers of GRES – Cahiers du GRES 2003–8*, Groupement de Recherches Economiques et Sociales.

Carayol, N., Roux, P. (2005). *A Strategic Model of Complex Networks Formation*. *Technical Report*, Universite Louis Pasteur, Strasbourg and Universite Paris Sud.

Casella, G., Berger, R. L. (2001). *Statistical Inference*. Pacific Grove, CA: Duxbury Press.

Cassiman, B., Veugelers, R. (2002). R&D cooperation and spillovers: Some empirical evidence from belgium. *The American Economic Review* **92(4)**, 1169–1184.

Cohen, W. (1995). Empirical Studies of Technological Activity. *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell Publishers Ltd.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. (2001). *Introduction to Algorithms*. Cambridge, MA: MIT Press.

Costa, L. F., Rodrigues, F. A., Travieso, G., Boas, P. V. (2007). Characterization of complex networks: A Survey of measurements, Arxiv preprint cond-mat/0505185.

Cowan, R., Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control* **28(8)**, 1557–1575.

Cowan, R., Jonard, N., Ozman, M. (2004). Knowledge dynamics in a network industry. *Technological Forecasting and Social Change* **71(5)**, 469–484.

Dickhaut, J., Rustichini, A. (2001). *Investment Game*. *Technical Report* Forthcoming in the Encyclopedia of Cognitive Sciences. Edited by Nature, available at the website: http://www.econ.umn.edu/arust/InvestmentGame_9_ (1).Pdf

Dieckmann, A. (2004). The power of reciprocity: Fairness, reciprocity, and stakes in variants of the dictator game. *Journal of Conflict Resolution* **48(4)**, 487–505.

Durlauf, S. N. (1999). *How Can Statistical Mechanics Contribute to Social Science? Technical Report*, Wisconsin Madison – Social Systems.

Durrett, R. A. (2004). *Probability. Theory and Examples*. Thomson Learning.

Eigen, M., Schuster, P. (1979). *The Hypercycle: A Principle of Natural Self-Organization*. Berlin. Springer.
Farmer, J. D., Smith, E., Shubik, M. (2005). Economics: The next physical science? Arxiv preprint physics/0506086.
Fehr, E., Fischbacher, U. (2003). The nature of human altruism. *Nature* **425**, 785–791.
Fehr, E., Schmidt, K. M. (1999). A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics*, **114**, 817–868.
Foray, D. (2004). *The Economics of Knowledge*. Cambridge, MA: MIT Press.
Frenken, K. (2006). Technological innovation and complexity theory. *Economics of Innovation and New Technology* **15(2)**, 137–155.
Gabszewicz, J. (2000). *Strategic Interaction and Markets*. Center for Operations Research and Econometrics (CORE), Catholic University of Louvain: Oxford University Press.
Gallegati, M., Kirman, A. P. (eds.) (1999). *Beyond the Representative Agent*. Cheltenham, GL: Edward Elgar Publishing.
Gerosky, P. (1995). *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell Publishers Ltd., chap. Markets for technology: knowledge, innovation and appropriability.
Gigerenzer, G., Selten, R. (eds.) (2002). *Bounded Rationality: The Adaptive Toolbox (Dahlem Workshop Reports)*. Cambridge, MA: MIT Press.
Godsil, C. D., Royle, G. F. (2001). *Algebraic Graph Theory*. Springer.
Goyal, S. (2007). *Connections: An Introduction to the Economics of Networks*. Princeton: Princeton University Press.
Gross, T., Blasius, G. (2007). Adaptive Coevolutionary Networks – A Review. Eprint arXiv: 0709.1858.
Hagedoorn, J., Cloodt, M., Roijakkers, N. (2006). Patterns In Inter-Firm R&D Networks in the Global Computer Industry: A Historical Analysis Of Major Developments During the Period 1970–1999. Paper presented at SPRU 40th Anniversary Conference – The Future of Science, Technology and Innovation Policy, SPRU, Brighton, East Sussex, UK.
Hagedoorn, J., Link, A. N., Vonortas, N. S. (2000). Research partnerships. *Research Policy* **29(4–5)**, 567–586.
Haller, H., Kamphorst, J., Sarangi, S. (2007). (Non-)existence and scope of nash networks. *Economic Theory* **31**, 597–604.
Haller, H., Sarangi, S. (2005). Nash networks with heterogeneous links. *Mathematical Social Sciences* **50**, 181–201.
Hanaki, N., Nakajima, R., Ogura, Y. (2007). The Dynamics of R&D Collaboration in the IT Industry. Working Paper.
Hausman, D. M. (2003). *Inexact and Separate Science of Economics*. Cambridge: Cambridge University Press.
Hofbauer, J., Sigmund, K. (1998). *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press.
Horn, R. A., Johnson, C. R. (1990). *Matrix Analysis*. Cambridge: Cambridge University Press.
Huberman, B., Glance, N. (1994). Beliefs and cooperation. Arxiv eprint arXiv: adap-org/9405001.
Jackson, M. O. (2003). *A Survey of Models of Network Formation: Stability and Efficiency*. *Working Papers 1161*, California Institute of Technology, Division of the Humanities and Social Sciences.
Jackson, M. O. (2006). Cambridge University Press. The Economics of Social Networks. *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*. New York: Vol. I
Jackson, M. (2008). *Social and Economic Networks*. Princeton: Princeton University Press.
Jackson, M. O., Watts, A. (2002). The evolution of social and economic networks. *Journal of Economic Theory* **106(2)**, 265–295.
Jackson, M. O., Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory* **71(1)**, 44–74.

Jain, S., Krishna, S. (1998a). Autocatalytic sets and the growth of complexity in an evolutionary model. *Physical Review Letters* **81(25)**, 5684–5687.

Jain, S., Krishna, S. (1998b). Emergence and growth of complex networks in adaptive systems. E-print adap-org/9810005.

Jain, S., Krishna, S. (2001). A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proceedings of the National Academy of Sciences* **98(2)**, 543–547.

Jain, S., Krishna, S. (2002). Graph theory and the evolution of autocatalytic networks. *Handbook of Graphs and Networks*. Berlin: John Wiley and VCH.

Karshenas, M., Stoneman, P. (1995). Technological Diffusion. *Handbook of the Economics of Innovation and Technological Change*. Oxford: Blackwell Publishers Ltd.

Khalil, K. H. (1995). *Nonlinear Systems*. Upper Saddle River, NJ: Prentice-Hall.

Kim, C., Wong, K.-C. (2007). Network formation and stable equilibrium. *Journal of Economic Theory* **133**, 536–549.

Kirman, A. (1997). The economy as an evolving network. *Journal of Evolutionary Economics* **7(4)**, 339–353.

König, M. D., Battiston, S., Napoletano, M., Schweitzer, F., (2008a). Efficiency and stability of dynamic innovation networks, CER-ETH Working Paper No. 08/95, ETH Zurich.

König, M. D., Battiston, S., Napoletano, M., Schweitzer, F. (2008b). On algebraic graph theory and the dynamics of innovation networks. *Networks and Heterogeneous Media*, **3(2)**, 201–219.

Krishna, S. (2003). *Formation and Destruction of Autocatalytic Sets in an Evolving Network Model*. Ph.D. thesis, Centre for Theoretical Studies, Indian Institue of Science. E-print arxiv:nlin.AO/0403050 v1.

Lane, D. (2005). *A Theory Based Dynamical Model of Innovation Processes*. *Technical Report*, Department of Political Economy, University of Modena.

Lane, D., Maxfield, R. (1997). Foresight, Complexity and Strategy. Chapter in book: Economy As an Evolving Complex System II, Westview Press.

Lerner, J., Tirole, J. (2002). Some simple economics of open source. *The Journal of Industrial Economics* **L(2)**, 197–234.

Mas-Colell, A., Whinston, M. D., (1995). *Microeconomic Theory*. New York: Oxford University Press.

Maxfield, R. R. (1994). General Equilibrium and the theory of directed graphs. *Journal of Mathematical Economics* **27**, 23–51.

Nelson, R. (1959). The simple economics of basic research. *Journal of Political Economy* **67**, 297–306.

Newman, M. (2003). The structure and function of complex networks. *SIAM Review* **45(2)**, 167–256.

Nowak, M., Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature* **393**, 573–577.

Nowak, M., Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* **437**, 1291–1298.

OECD (2006). *OECD Science, Technology and Industry Outlook 2006*. OECD Publishing.

Padgett, J. F. (1996). *The Emergence of Simple Ecologies of Skill: A Hypercycle Approach to Economic Organization*. Working Papers 96-08-053, Santa Fe Institute.

Padgett, J. F., Lee, D., Collier, N. (2003). Economic production as chemistry. *Industrial Corporate Change* **12(4)**, 843–877.

Powell, W., Grodal, S. (2006). *Oxford Handbook of Innovation*. USA: Oxford University Press. chap. Networks of Innovators.

Ricottilli, M. (2005). *Firms and Network Formation Through the Transmission of Heterogeneous Knowledge*. *Technical Report*, Department of Economics, University of Bologna.

Ricottilli, M. (2006). *Constraints and Freedom of Action: A Fitness Trade-Off*. *Technical Report*, Department of Economics, University of Bologna.

Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science* **46(6)**, 824–844.

Rosenblatt, D. (1957). On linear models and the graphs of minkowski-leontief matrices. *Econometrica* **25**, 325–338.

Saurabh, A., Cowan, R. (2004). *The Growth of Knowledge and Complexity in an Evolving Network Model of Technological Innovation*. *Technical Report*, UNU-INTECH and MERIT, University of Maastricht, The Netherlands. Paper presented at "Organisations, Innovation and Complexity: New Perspectives on the Knowledge Economy", 9th–10th September 2004, CRIC, University of Manchester, Manchester, England, UK.

Scotchmer, S. (2004). *Innovation and Incentives*. Cambridge, MA: MIT Press.

Seufert, A., Schweitzer, F. (2007). Aggregate behavior in an evolutionary network model. *International Journal of Modern Physics C* **18(10)**, 1659–1674.

Sorenson, O., Rivkin, J. W., Fleming, L. (2006). Complexity, networks and knowledge flow. *Research Policy* **35(7)**, 994–1017.

Stadler, P., Schuster, P. (1996). Permanence of sparse autocatalytic networks. *Mathematical Biosciences*, **131(2)**, 111–133.

Steger, A. (ed.) (2001). *Diskrete Strukturen 2.: Kombinatorik, Graphentheorie, Algebra*. Berlin: Springer Verlag.

Steger, A., Schickinger, T. (eds.) (2001). *Diskrete Strukturen 1.: Kombinatorik, Graphentheorie, Algebra*. Berlin: Springer Verlag.

Tirole, J. (1988). *The Theory of Industrial Orgnization*. Cambridge, MA: MIT Press.

Varian, H. R. (1996). *Intermediate Microeconomics: A Modern Approach*. New York: WW Norton.

Vega-Redondo, F. (2007). *Complex Social Networks*. Series: Econometric Society Monographs. Cambridge, MA: Cambridge University Press.

Von Hippel, E. (1987). Cooperation between Rivals: Informal know-how trading. *Research Policy* **16**, 291–302.

Von Hippel, E., Von Krogh, G. (2003). open source software and the private-collective innovation model: Issues for organization science. *Organization Science* **14(2)**, 209–223.

Watts, D. J., Strogatz, S. H. (1998). Collective dynamics of small-world networks. *Nature* **393**, 440–442.

Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics* **113(2)**, 331–360.

West, Douglas, B. (2001). *Introduction to Graph Theory*. Upper Saddle River, NJ: Prentice-Hall. 2nd edn.

Wilson, R. (1975). Informational economics of scale. *Bell Journal of Economics* **6**, 184–195.

Zenou, Y. (2006). Course on networks: Theory and applications. Lecture notes.

Chapter 9
Propagation of Innovations in Complex Patterns of Interaction

Albert Diaz-Guilera, Sergio Lozano and Alex Arenas

9.1 Introduction

In recent times the possibility of using the tools of statistical physics to analyze the rich dynamical behaviors observed in social, technological, and economical systems has attracted a lot of attention from the physics community (Arthur et al. 1997, Mantegna and Stanley 1999, Bouchaud and Potters 2000). So far, one of the main contributions to these fields has been the analysis of simple models that capture the basic features of the investigated phenomena. The goal is to identify the relevant parameters as well as the essential mechanisms governing their dynamics with the hope that this information will help us to understand the physical behavior of real complex systems. A real part of this effort has been devoted to the characterization of real networks, identifying their main features, and understanding how they arise (Watts and Strogatz 1998, Barabasi and Albert 1999, Strogatz 2001, Dorogovtsev and Mendes 2002, Albert and Barabasi 2002).

In particular, we tackle the problem of diffusion of innovations in a social network, and we try to understand how the stimulus for change spreads by gradual local interaction between the individual nodes (agents) forming the network. Most of the times these 'waves' of change come in terms of intermittent bursts separating relatively long periods of quiescence (Krugman 1996, Arenas et al. 2000).

There are two mechanisms involved in the diffusion of innovations that a mathematical model should take into account. On the one hand, there is a pressure for adopting a new technology coming from marketing campaigns and mass media (Guardiola 2001). These external processes are essentially independent of the social network structure and one can view their effects as a random independent process on the agents. On the other hand, one should take into account the effect of the surrounding agents who define the social network. Once an agent has decided to adopt a new technology, those who are in contact with him can evaluate the new payoff the agent has got by acquiring the new technology and compare it with the

A. Diaz-Guilera (✉)
Departament de Física Fonamental, Universitat de Barcelona, Marti i Franques 1, 08028 Barcelona, Spain

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_9,

current benefits. This propagating mechanism stands for inter-agent communication processes. By balancing the payoff increment with the associated upgrading cost, they may decide to adopt, or not, the new technology.

In the first part of this contribution, we review some of the results obtained by our group in the last years concerning models of propagation of innovations within different frameworks. Most of this work was initiated in regular lattices, but the last attempts have taken into account that social environments are not well modeled by regular patterns of interactions but by heterogeneous ones. This has led us to investigate a practical case reported in the social sciences literature (Saxenian 1994), where not only the resistance of the agents to innovate is important, but also the structure of the social network plays a crucial role in the development of a more or less innovative society. To the description of this analysis, we devote the final part of the contribution.

9.2 The Model

Our starting point is a model of diffusion of innovations by imitation proposed in Guardiola et al. (2002) and also studied in Llas et al. (2003a). Although it is simple it displays a very rich collective behavior that can be related to well-known theories of self-organized criticality and its avalanche dynamics (Jensen 1998).

The dynamics of the model are implemented such that each site in the network (agent) is characterized by a real and continuous variable a_i. In a general way, we can consider this quantity as a characteristic of an agent that other agents might want to imitate. When an agent has adopted a new characteristic, her neighbors become aware of the change and balance their interest (quantified as a_i–a_j) with their resistance to change C to decide if they would like to imitate this change. In this way, C controls the mechanism of imitation and it is constant in time and the same for all the agents in the current scenario.

The dynamics can be summarized as follows:

- The system is asynchronously updated. At each time step a randomly selected agent updates her state a_i

$$a_i \to a_i + \Delta_i, \tag{9.1}$$

  where Δ_i is a random variable with mean λ.
- All agents $j \in \Gamma(i)$, where $\Gamma(i)$ is the set of neighbors of agent i, decide whether they want to upgrade or not, according to the following rule:

$$a_i - a_j > C \Rightarrow a_j = a_i. \tag{9.2}$$

- If any $j \in \Gamma(i)$ has decided to imitate, we also let her neighbors decide whether they want to imitate this behavior or not. In this way, the information of an update may spread beyond the first neighbors of the originally perturbed site.

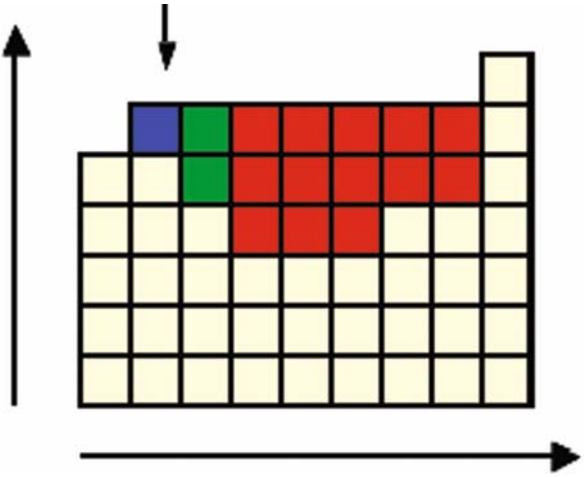


Fig. 9.1 Schematic representation of the dynamical rules in a one-dimensional system with a cost $C = 1$. Initially, the state of the system corresponds to the *white area*. The agent where the *vertical arrow points* receives an external update. The agent at its right now is two units below it and decides to imitate. This generates an avalanche that involves most of the neighbors at the right, since any agent has at its given time a level that is below its left neighbor, corresponding to the *darker area*

This procedure is repeated until no more agents want to change, concluding an *avalanche* of imitation events. In this way, we have assumed that the time scale of the imitation process is much shorter than that corresponding to the external updates. In Fig. 9.1 we present schematically these dynamical rules.

9.3 Dynamical Regimes

According to the cost value C it is possible to distinguish several regimes. In Fig. 9.2, we can see the technology profile (the interface defined by technology of all the agents) of two extreme cases. For $C < 1$, once there is an external update, a system size avalanche is immediately triggered and all agents end up in a very close technological level, and the profile is hence very planar. For $C >> 1$, upgrading is so expensive that agents never care about their neighbors technology, and most avalanches are of size just 1. In this regime the profile, as seen in Fig. 9.2 (right), is quite rough.

In between these two regimes (see Fig. 9.3, left), there is a region showing a rich dynamics where one finds technological avalanches of all possible sizes. Actually, for some values of C the probability density of having an avalanche of size s shows a power-law behavior

$$P(s) \approx s^{-\tau}. \tag{9.3}$$

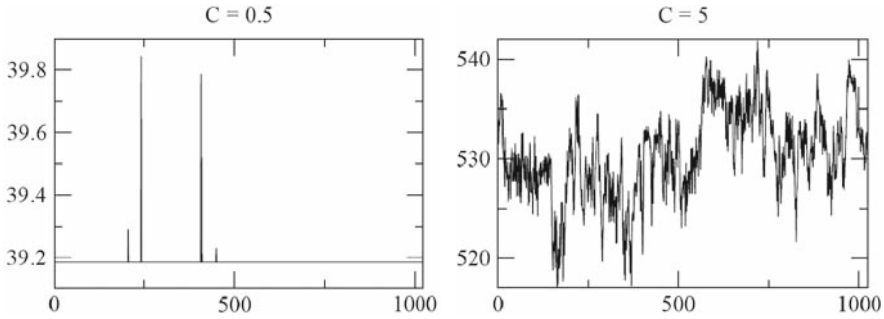


Fig. 9.2 Technological profile of a one-dimensional system in two extreme cases. *Left*: supercritical that corresponds to a dynamical state where avalanches are always very large. *Right*: subcritical, which corresponds to avalanches of unit size

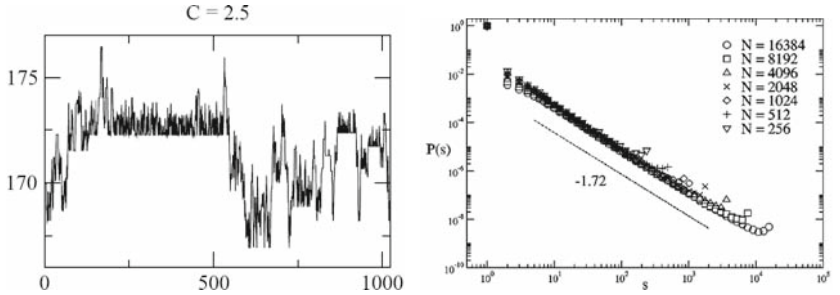


Fig. 9.3 *Left*: technological profile for the critical region. *Right*: probability density of having a technological avalanche of size s for $C = 3$ in a log–log scale

Figure 9.3 right shows $P(s)$ for several system sizes and $C = 3$ in a log–log scale. We can observe there a power-law distribution over four decades for the largest system size.

9.4 The Observables

9.4.1 Mean Rate of Progress

In the model introduced so far, we have considered that each individual adjustment elicits a certain cost. This cost is fixed and does not depend on the advanced technological level. In economic terms, an optimal situation would be one in which the system reaches a certain average global technological level with the minimum cost, that is, the minimum number of adjustments. According to such social perspective, the problem is to find the most advantageous regime of advance, leading to the largest global advance given a fixed cost.

For these reasons, it is convenient to define a macroscopic observable measuring the advance rate of the system. A useful magnitude for such purposes is the following:

$$\rho \equiv \lim_{T\to\infty} \rho(T) = \lim_{T\to\infty} \frac{\sum_{t=1}^{T} H(t)}{\sum_{t=1}^{T} s(t)}, \tag{9.4}$$

where $s(t)$ is the size of the avalanche occurring at time t, i.e., the number of agents whose technological level has been updated, and $H(t)$

$$H(t) = \sum_{i=1}^{N} [a_i(t) - a_i(t-1)] \tag{9.5}$$

is the total advance achieved during the same event (the interface area increment caused by an avalanche). According to the definition of the mean rate of progress, it is easy to see that for the two extreme cases its value is very small. For low cost C, any external update is followed by a large avalanche in which all agents advance the same amount, since the profile is extremely smooth, and hence ρ is equal to the mean value of the external update λ. On the other hand, for large C, all the avalanches are very small (most of them of size one) and hence the advance of the avalanche is equal to the mean value of the external update again. It is interesting to realize that for intermediate values of the cost, the system is optimal since it reaches a maximum for the mean rate of progress. In Fig. 9.4, we show several plots of ρ versus the cost, for several system sizes (Guardiola et al. 2002).

In previous works in similar models, it was shown that the system reaches its maximum mean rate of progress in a self-organized way (Arenas et al. 2000, 2000a). It is the result of the aggregated dynamics that the system reaches by itself its optimal outcome, for a constant and uniform value of the cost. In any case, we can say that there exists a certain value of the cost which, following the usual terminology in statistical physics (Stanley 1971), we will call the critical value, for which the technology profile grows more efficiently. This leads to the following paradoxical result: upgrading costs should be neither cheap nor expensive in order to have an optimal technological growth (Arenas et al. 2002). Obviously, our concept of efficiency is related to the number of times cost is paid, that is, from the point of view of the population, but not of the companies who sell the product. Sellers will always look for a scenario where agents acquire as many new products as often as possible (Guardiola 2001).

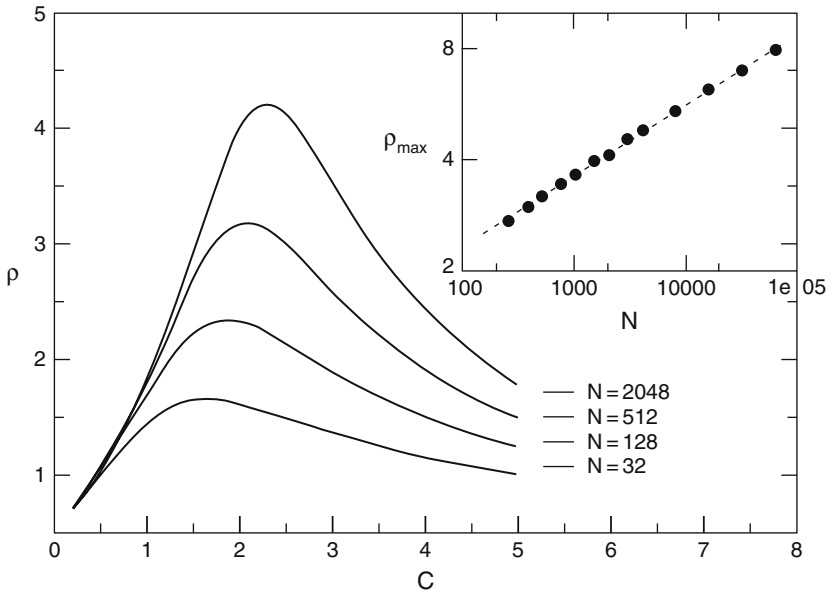


Fig. 9.4 Mean rate of progress as a function of the cost C for several system sizes. At the extremes, the value of ρ goes to the mean value of the update. There is a peak, $\rho_{\max}$, that diverges in the large system size limit. In the inset, we plot $\rho_{\max}$ against system size N. Dashed line shows $\rho_{\max} \approx N^{0.20}$

9.4.2 Profile Roughness

In the previous section, we showed several snapshots of the population technological profile. This profile corresponds to the continuous interface defined by the technological level a_i of all agents. We have also analyzed the technological profile from the point of view of the statistical mechanics of non-equilibrium growing interfaces (Barabasi and Stanley 1995).

The most straightforward way of giving a quantitative measure of the roughening of the interface is to look at the interface height fluctuations around its mean value. The interface width is defined by the variance of the technological profile

$$w(N,t) \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}[a_i(t) - \overline{a(t)}]^2}, \tag{9.6}$$

where $\overline{a(t)}$ is the spatially averaged technological profile. From this definition one can define many different properties, as for instance how this interface width scales with time and with the system size, showing different scaling properties (Guardiola 2001). In terms of this scaling some anomalous properties have been found (Llas et al. 2003b, Llas et al. 2007).

9.5 Interplay Between Dynamics and Structure

Let us now put some attention to the influence of the connectivity pattern on the collective properties of the system and, in particular, in macroscopic observables such as ρ. There are two simple cases that have been systematically analyzed: a fully connected network and a regular lattice. In the globally coupled case, the information referent to any change elicited in an arbitrary position of the network is immediately available to every other agent. As a consequence, the state of all agents a_i is bound in gaps of width C. This limits the advance of any agent to a maximum of C. On the other hand, when the system is defined on a 1D ring this limitation only applies to the nearest neighbors. If the information spreads beyond these neighbors, the advance achieved can exceed C. In this extreme cases different qualitative behaviors are observed, which can be quantified by the exponent by which the maximum value of the mean advance rate scales with the system size

$$\rho_{\max} \approx N^{\alpha}. \tag{9.7}$$

When one considers the dynamics of the model on a ring $\alpha = 0.20$, while mean-field calculations and numerical simulations of the dynamics of the model on a fully connected network show that $\alpha = 0.50$ (Guardiola et al. 2002).

The two particular cases considered so far are usually chosen either for their numerical simplicity, as in the ring, or because they allow for simple mean-field calculations as in the fully connected network. Clearly, the structure of both cases is far from the much more complex pattern of interactions observed in realistic systems (Strogatz 2001). However, they appear as paradigms of two opposite generic situations either a scenario where the propagation of information can be constrained to a local neighborhood, or one where it may reach the whole population in just one step. One wonders which will be the dominant features in a more general situation. In this section, we will consider this issue. In particular, we will study the relation between the salient features present in a more general structure and the dynamical properties of the system.

It is well known that, starting from a ring, the random addition of a few links produces changes in the properties of the network, such as a rapid drop in the average distance between nodes, maintaining the local structure (Watts and Strogatz 1998). It has been shown that this feature can be related to significant changes in the dynamical properties in some systems. In order to study what effects may be present in our model, we will consider the dynamics on a ring lattice with N vertices and $k = 2$ edges per vertex, adding a new link at random with probability p per edge. In general, when one modifies the underlying structure, a quantitative variation of the numerical values of the dynamical magnitudes that characterize the system is observed. However, we will focus our attention on whether the scaling properties are modified, since they are an indication of qualitative changes in these properties. In order to quantify the changes in the scaling behavior of the system, we have computed the exponent for different values of p, as shown in Fig. 9.5.

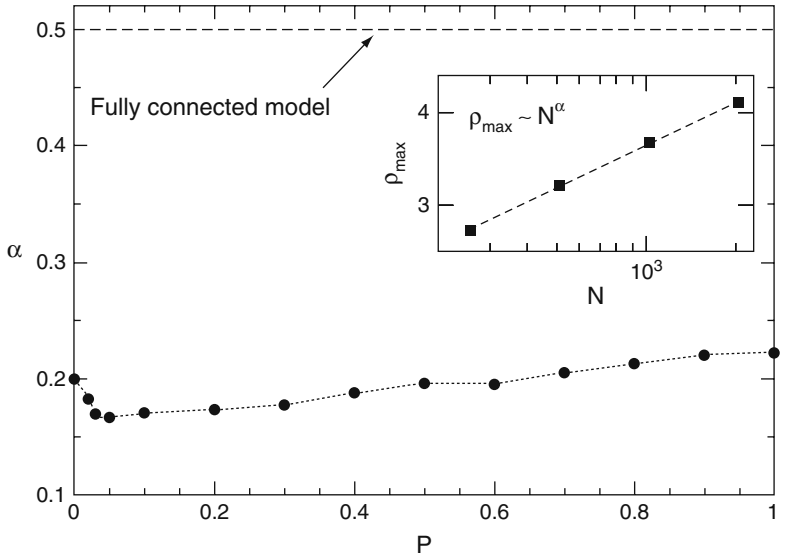


Fig. 9.5 The exponent α characterizes the power-law divergence. The figure displays α vs. p. The *inset* shows the fit used to obtain α when $p = 0.5$

The main conclusion that can be extracted from the figure is that the addition of new links does not modify substantially the scaling properties of the system. For small p we observe a small decrease in the value of α, an effect that seems to be correlated with the rapid drop in the average distance between agents. As p increases to $p = 1$ a slight growth of α is reported. It is important to understand why the qualitative behavior of the system is similar to the one characteristic of the ring. In this case, an avalanche propagates through steps in which the information reaches the nearest neighbors of a modified site. To generate a large event or simply to reach agents far away from the initial updated unit, a large number of steps are required. In this sense, the information propagates by a local process. The addition of new links allows the information to reach agents through shortcuts, but it does not change the mean mechanism of diffusion, i.e., in order to proceed further, an avalanche still requires a large number of steps, which is still dominated by a local process. One cannot under-stress the fact that for the imitation strategy described in this contribution, in contrast to what is observed in other models, the characteristic behavior of the ring is dominant even when a more general structure, such as a small world network, is considered (Llas et al. 2003a).

The local character of the diffusion process typical of low connectivity networks is lost when considering densely connected networks. In this situation, the information about the state of a given agent is available to any other agent in just a few steps. The small world construction is far away from this limit, even when $p = 1$ the mean number of links per site, i.e., the connectivity of the system, has increased

very slightly. In order to reach a highly connected state, we added links randomly to the ring up to a fraction f of the total possible links in the system. This means adding $f[N(N–1)/2]–N$ connections to the ring. This recipe allows us to interpolate between the ring ($f = 0$) and the fully connected model ($f = 1$). For finite values of f, the system corresponds to a ring with a superimposed random graph. For this construction the connectivity of the system will be fN, and thus, for a fixed value of f, the connectivity will increase as N grows. If this plays a significant role in the dynamical behavior of the system, we expect that it will be reflected in an important quantitative change in α. In Fig. 9.6, we present the behavior of the peak of the mean rate of advance $\rho_{\max}$ as a function of system size N when $f = 0.005$. For low values of N, the behavior resembles the one observed in the small world case. As N grows there is a crossover to the fully connected behavior and $\rho_{\max} \approxeq N^{0.5}$ (Llas et al. 2003a).

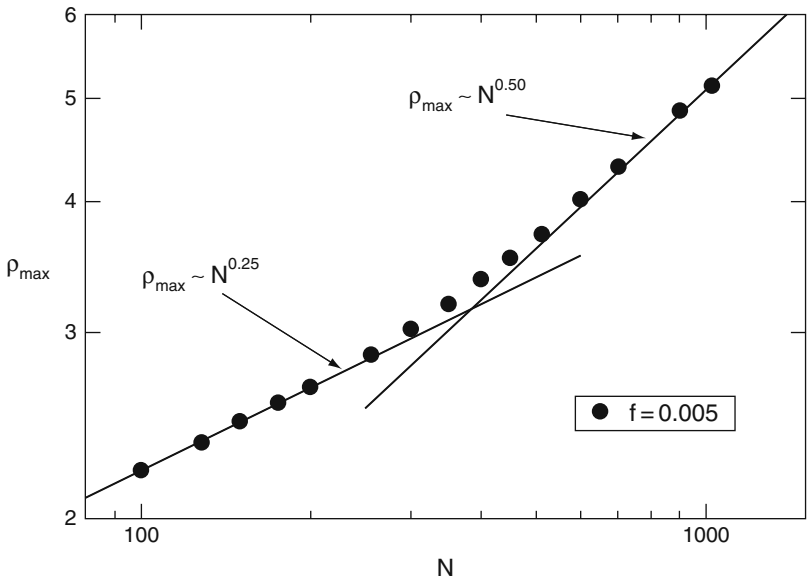


Fig. 9.6 The peak of the mean rate of advance in terms of the system size for a fixed value of the fraction of added links. A clear crossover is observed for a sufficiently large system size

9.5.1 Identification of an Optimal Behavior

Up to now, we have analyzed different mechanisms for the spreading of the information. On one hand, C appears as a parameter that can be tuned in order to reach an optimal regime for a fixed underlying network. On the other hand, we have also seen that diverse structures lead to different ways in which the information spreads. This line of reasoning naturally leads to the question of which is the structure that gives

the optimal regime for a fixed value of C. For small C there is almost no resistance to change and the information easily spreads. In fact, we expect that if $C \to 0$ then ρ will also decrease independently of the underlying network. On the other extreme, for $C \to \infty$ the behavior of the system will resemble a random deposition process, and again the behavior of ρ is expected to decrease when considering any general structure. To analyze what happens for intermediate values of C, we should proceed with care, as we will see immediately.

In order to do this, we consider a system with a certain distribution of couplings between agents and a fixed C. In general, the value of ρ will be smaller than $\rho_{\max}$. Now, supposing that we can modify the connectivity pattern, it is natural to ask which structure leads to the optimal behavior. The question cannot be answered properly without looking at the parameter that control the dynamics, i.e., the strategy to follow depends precisely on C. To optimize the behavior of the system, one cannot split the problem into two independent parts; dynamics and the underlying structure must be considered as a whole (Llas et al. 2003a).

Let us analyze some features associated to the mechanism of imitation and consider the difference between the highest ($a_{\max}$) and the lowest ($a_{\min}$) characteristic values in the system. For sufficiently high connectivity C will bound this gap and, as a consequence, for these systems, the value of ρ cannot exceed this value. On the other hand, when the propagation of the information is constrained to advance through local processes, a more heterogeneous profile of characteristic values can be formed, allowing for a larger gap between $a_{\max}$ and $a_{\min}$. In this case, avalanche events consisting of a large number of steps will produce a large advance, allowing the value of ρ to exceed C. In fact, by using the probability distribution of avalanches $P(s)$ and advances $P(H)$ obtained using numerical simulations in Guardiola et al. (2002) this can be easily verified analytically. In this situation, a sparsely connected network will necessarily have a greater ρ than a highly connected one. For increasing values of C, avalanche events will easily get blocked in a few steps. In this context, large advances in a sparsely connected network will become very rare. More frequent advances will be observed in a highly connected network since many agents are permitted to find out about an update in any step. This situation offers the possibility for a higher ρ to be observed in highly connected structures.

Following this analysis, we have considered a general system of fixed size N. The evolution of the mean rate of advance ρ versus C has been studied for two different situations: the ring and another structure where the number of links (measured in terms of f) is large enough to observe fully connected behavior. The results are illustrated in Fig. 9.7. Note that as f is varied the qualitative shape of the ρ curve is similar. However, the position of the peak corresponds to different values of C. For increasing values of f the peak corresponds to larger values of C, eventually reaching the curve corresponding to the globally coupled case. It is important to stress that, as C is varied, the optimal network may change from a highly connected to a sparsely connected one. This behavior is clearly reflected in the inset, where we present the behavior of ρ versus f for two different values of C. For $C = 2$ a decrease in ρ is observed as f grows. The addition of links is harmful to the system. On the other

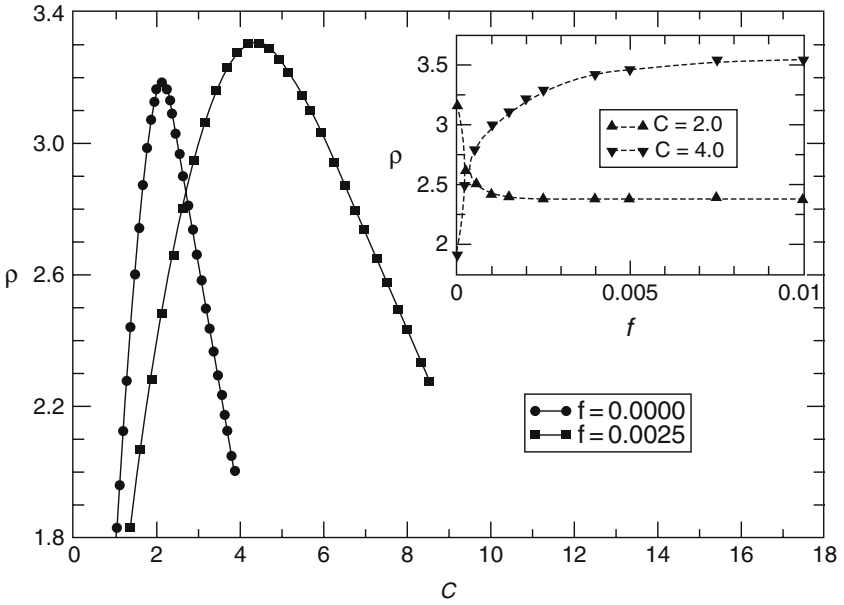


Fig. 9.7 Mean rate of advance as a function of the cost for a fixed system size $N = 512$ and for two different values of the fraction of added links. The *inset* represents the behavior of the mean rate of advance as a function of the fraction of added links for two values of the cost

hand, for $C = 4.0$ the opposite behavior is observed, and ρ increases its value as f grows. Clearly, in this case, the addition of links is beneficial.

These results show that in order to optimize the behavior of the system a non-trivial combination of dynamical rules and underlying structure should be considered. When the interplay between both allows for $\rho C > 1$, a sparsely connected structure performs better than a highly connected one. On the other hand, when $\rho C < 1$ the opposite is true. Note that when $\rho/C \sim 1$ the behavior should be independent of the underlying network. In fact, Fig. 9.7 shows that when $\rho/C \sim 1$ both curves intersect. Numerical simulations for different N and f also show these qualitative behaviors.

9.6 Practical Case

It is a common claim of scholars from social disciplines that computational models, like the one we have presented above, are too simple to represent the richness of real scenarios. In order to deepen on this matter, the aim of the second part of this contribution is to present an example of how this sort of models can be used as a quantitative tool to support a previous research made by social scientists.

9.6.1 Anna Lee Saxenian's Network Perspective

About a decade ago, Anna Lee Saxenian authored a book which pointed out the key role of relational aspects as success factors of regional economic development (Saxenian 1994). Her theories were supported by historical data from two North American hi-tech industrial poles (Silicon Valley and Boston's Route 128), which had presented significantly divergent behaviors during the late 1980s. In her opinion, certain cultural features could perfectly justify that divergent evolution.

On the one side, she observed that Silicon Valley presented a densely networked and flexible organization; within this community information was shared quite freely between companies (even competitors) and with research institutes and local government institutions. Moreover, the extremal high values of job mobility and business creation rates, was interpreted by Saxenian as an indicator of individual initiative and independence.

On the other hand, she found that Boston Route 128 was dominated by large and autarkic corporations with a very rigid hierarchy, where people involved in research and development were expected to be strictly loyal to the company and behave synchronized as a block.

After publication of Saxenian's book, many scholars have positioned either close to her position or clearly opposite to it, following a viewpoint presented in Florida and Kenney (1990), which minimizes those cultural traits compared with technological ones. This situation has grown up a debate that is still alive.

Here, we would like to contribute to this discussion from a completely different point of view. Using a variation of the model presented before, we simulate characteristics of both technological sites and obtain numerical results that agree with Saxenian's hypothesis.

9.6.2 The Experiment

More concretely, we have modified some characteristics of the model of diffusion of innovations presented in order to make it realistic enough to capture two concrete cultural features, that summarize the differences pointed out by Saxenian to be determinant. These characteristics are the topology of the social substrate of each pole and the diversity of individual behavior.

About the social topology, our model should reflect that people in Silicon Valley used their formal and informal links to share information with other people beyond their organizational boundaries, while in Boston, because of the autarky and secrecy imposed, information exchange was mainly restricted to the same group (company, division, department, or team).

In relation to individual behavior of actors in our model, individual initiative and independence observed in Silicon Valley should be represented as a wide diversity of individual behaviors (to highlight the idea that each actor had its own perspective). On the contrary, rigid hierarchy and employee obedience in Boston would correspond to a configuration with low diversity, where decisions are taken centrally and actors act as synchronized as possible.

After 'translating' observations made by Saxenian into concrete constraints, the next step consists on adapting the original model to them.

In order to obtain social topologies that agree with the first constraint, we have used an algorithm proposed in Boguñá et al. (2004) that is able to build up social-like network topologies from the homophile characteristic, which is everyone's preference to establish and maintain relations with people that have any common characteristic with (profession, hobbies, age, or political feelings, for example). In our study case, network of Boston 128 is considered to have a large homophile, because actors tend to exchange information only with people in closed homogeneous groups. On the contrary, Silicon Valley is supposed to have a small one, due to people's inter-organizational interactions.

Since, as it was explained above, in our model the resistance to change value of an actor defines its role as innovation adopter, the incorporation of the other feature to be treated (diversity on individual behavior) has been solved by introducing differences on the resistance value of actors. To explain this second adaptation more detailed, we need to introduce a new concept called 'resistance profile.'

The resistance profile of a system is the representation as a histogram of the frequency of occurrence in the system of each possible value of resistance to change. In the original model, where all agents had the same resistance to change value C, the resistance profile would be 0 for all values except C. This particular performance of the resistance profile is far from reality. If we calculated the resistance profile of a real scenario, we would find some kind of distribution of resistance values around a central or average one. Consequently, if we want to reproduce artificially a real resistance profile we can assign resistance values to actors following a statistical distribution. Moreover, if we want to reproduce a situation with a wide diversity of behaviors (a wide resistance profile) we need a distribution with a high standard deviation. On the contrary, a low standard deviation leads us to a narrow resistance profile simulating a homogeneous scenario.

In our particular case, we have used a Gaussian distribution to generate actor's resistance with a high deviation to simulate Silicon Valley heterogeneity and a low one for Boston Route 128.

Figure 9.8 shows, graphically, particularities introduced to the original model in order to simulate Silicon Valley and Boston Route 128 scenarios, respectively.

9.6.3 Results

Once both technological poles have been simulated as two configurations of the original model, we can obtain some results and analyze them. Before this, however, we have to define the variables we will play with. Our intention is to have a quantitative measure of the fitness of each technological pole to different values of economic dynamism.

First, we need an observable that indicates how well an innovation system (an industrial district) works in a certain situation. In this case, we have chosen the mean rate of progress (ρ), an observable defined and explained in the first part of this contribution.

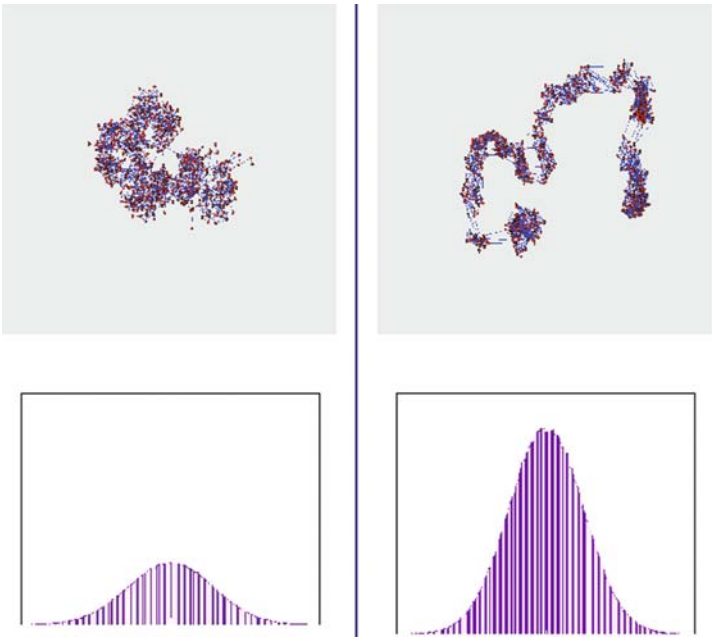


Fig. 9.8 Modifications made to the original model to introduce Silicon Valley (*left*) and Boston Route 128 (*right*) particularities. For each configuration, the topology (*top*) and the resistance to change profile (*bottom*) are shown

We also have to define an independent variable to quantify economical dynamism. The idea of using resistance to change to represent the behavior of actors is also useful here. In a dynamic situation, people tend to adopt novelties easily (their resistance to change is low) and, on the contrary, become more conservative in a more quiet situation (their resistance to change increases). Taking this into account, we can express the dynamism of the environment as the inverse of the average of the resistance to change values of all actors in the system.

The evolution of ρ as a function of the economical dynamism for both configurations is shown in Fig. 9.9. We can see that the maximum value of each model's ρ corresponds to different values of dynamism. This means that each configuration works better with a particular economical environment. Route 128 model gains its peak for a medium value and, if environment conditions change to a more dynamic scenario, its ρ value drops down. This agrees strictly with Saxenian (1994): '...Route 128 system flourishes in an environment of market stability and slowly-changing technologies.'

On the other hand, plot for Silicon Valley's model presents a stronger peak at a dynamism value a half upper. Once again, this behavior has its correspondence in Saxenian (1994): 'The region, if not all the firms in the region, is organized to innovate continuously....'

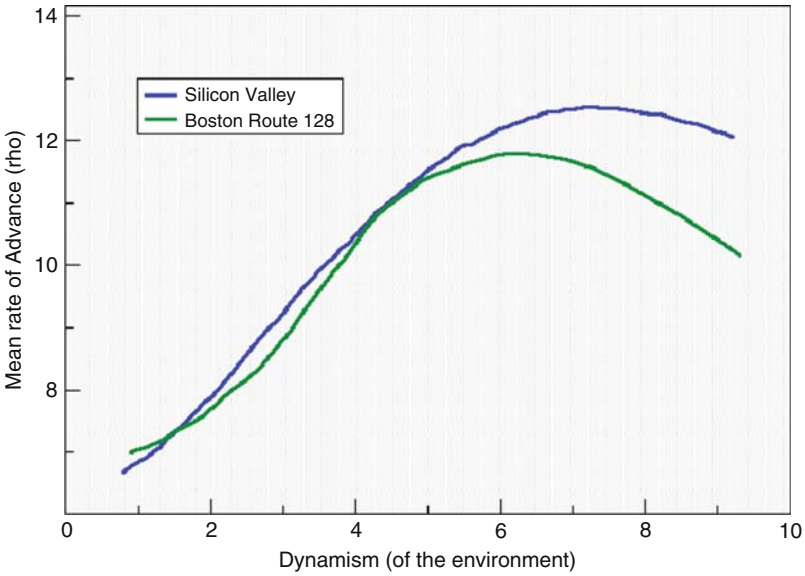


Fig. 9.9 Evolution of the fitness of each innovation system as a function of the dynamism of the situation

Finally, if environment moves to an extremely quiet economical situation, both values fall down. This should be read as a consequence of the loss of dynamism of regional economy, that affects in a similar way all industrial complexes independently from their particular characteristics.

9.7 Conclusions

We have presented results obtained in different scenarios about propagation of innovations, in regular as well as in heterogeneous patterns of connectivity. On the one hand, we have defined relevant observables and reviewed some aspects. On the other hand, we have shown that, in a particular case of study in the social science literature, structural features play a fundamental role in explaining the success of innovation systems. In order to tackle this second topic, we have adapted a computational model and obtained quantitative results that supports the conclusions of the previous qualitative analysis of the real raw data. We feel that this sort of two-step approach (from raw data to qualitative conclusions and from these conclusions to simple modeling) could be very useful in the study of different social and economical issues, not only innovation diffusion but also other kind of phenomena with a dependence on the underlying social structure.

References

Albert R, Barabasi A-L (2002) Statistical mechanics of complex networks. Rev Modern Phys 74:47.

Arenas A, Diaz-Guilera A, Guardiola X, Oron G, Llas M, Perez-Vicente CJ, Vega-Redondo F (2002a) New results in a self-organized model of techno logical evolution. In: Schweitzer F (ed.) Modeling Complexity in Economic and Social Systems, World-Scientific, Singapore.

Arenas A, Diaz-Guilera A, Pérez-Vicente CJ, Vega-Redondo F (2000) Self-organized evolution in a socio-economic environment. Phys Rev E 61:3466–3469.

Arenas A, Díaz-Guilera A, Pérez-Vicente CJ, Vega-Redondo F (2002b) Self-organized criticality in evolutionary systems with local interaction. J Econ Dyn Control 26:2115–2142.

Arthur WB, Durlauf SN, Lane DA (eds.) (1997) The economy as an evolving complex system II. SFI Studies in the Sciences of Complexity, Addison Wesley, Longman.

Barabasi A-L, Albert R (1999) Emergence of scaling in random networks. Science 286:509.

Barabasi A-L, Stanley HE (1995) Fractal Concepts in Surface Growth, Cambridge University Press, Cambridge, UK.

Boguñá M, Pastor-Satorras R, Díaz-Guilera A, Arenas A (2004) Models of social networks based on social distance attachment. Phys Rev E 70:056122.

Bouchaud JP, Potters M (2000) Theory of Financial Risk, Cambridge University Press, Cambridge, UK.

Dorogovtsev SN, Mendes JFF (2002) Evolution of networks. Adv Phys 51:1079–1187.

Florida R, Kenney M (1990) The Breakthrough Illusion, Basic Books, New York, NY.

Guardiola X (2001) Emergent phenomena in networks of interacting units: Synchronization, phase-locking and criticality. Ph.D. Thesis, Universitat de Barcelona.

Guardiola X, Díaz-Guilera A, Arenas A, Perez-Vicente CJ, Llas M (2002) Modelling diffusion of innovations in a social network. Phys Rev E 66:026121.

Jensen HJ (1998) Self-Organized Criticality, Cambridge University Press, Cambridge, UK.

Krugman PR (1996) The Self-Organizing Economy, Blackwell, Malden, MA.

Llas M, Diaz-Guilera A, Lopez JM, Gleiser PM (2007) The role of time scale separation in a nonequilibrium roughening transition. Physica A 374:289–292.

Llas M, Gleiser PM, Diaz-Guilera A, Perez-Vicente CJ (2003a) Optimization as a result of the interplay between dynamics and structure. Physica A 326:567–577.

Llas M, Gleiser PM, Lopez JM, Diaz-Guilera A (2003b) Nonequilibrium phase transition in a model for the propagation of innovations among economic agents. Phys Rev E 68:066101.

Mantegna RN, Stanley HE (1999) An Introduction to Econophysics, Cambridge University Press, Cambridge, UK.

Saxenian A (1994) Regional Advantage: Culture and Competition in Silicon Valley and Route 128, Harvard University Press, Cambridge, UK.

Stanley HE (1971) Introduction to Phase Transition and Critical Phenomena, Oxford University Press, Oxford, UK.

Strogatz SH (2001) Exploring complex networks. Nature 410:268.

Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. Nature 393:440.

Chapter 10
Sensitive Networks – Modelling Self-Organization and Innovation Processes in Networks

Ingrid Hartmann-Sonntag, Andrea Scharnhorst and Werner Ebeling

10.1 Introduction

This contribution is devoted to the interdisciplinary theory of self-organization processes, paying particular attention to stochastic effects connected with innovations in network systems. On our understanding "self-organization" is the spontaneous formation of structures (Ebeling and Feistel, 1982, 1994; Feistel and Ebeling, 1989). An "innovation", on a general system-theoretical understanding, is the appearance of, for example, a new species, a new mode of behaviour, a new technology, a new product or a new idea (Ebeling and Sonntag, 1986; Bruckner et al., 1989, 1990, 1996; Ebeling et al., 1999).

Technological innovation is considered to be the basic driving force for economic evolution and growth. In economics, innovation networks (Frenken, 2000; Kowol and Küppers, 2005) and networks economies (Nagurney, 2003) have been widely discussed in recent decades. At the beginning of the 21st century, it was noted that an unified understanding of socio-economic networks is still missing (Saviotti, 2001). Kirman remarked: "For many economists the study of networks is limited to the analysis of the functioning of physical networks such as the railway, the telephone system or the internet for example" (Kirman, 2003). In the meantime, as shown in the present volume, networks of innovative economic activities have gained more attention. Although networks are thought to be constituted by sets of actors and links (Saviotti, 2001), the very nature of these actors and the links between them vary according to different authors and approaches. Nations, institutions, firms, products or individuals, for instance, may represent the actors. The links can be defined quite differently as well, for example, as exchange of knowledge between firms or as exchange of goods between nations. Innovation in such contexts is usually seen as an outcome of complex networks with heterogeneous actors.

In this contribution, we define innovation neither as a product of network activity nor as a process of knowledge diffusion, but rather as a specific process in the

I. Hartmann-Sonntag (✉)
Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, 12489 Berlin, Federal Republic of Germany
e-mail: ebeling@physik.hu-berlin.de

A. Pyka, A. Scharnhorst (eds.), *Innovation Networks*, Understanding Complex Systems,
DOI 10.1007/978-3-540-92267-4_10,

formation of a complex network. Thus, innovation is conceptualized as a process in the structure formation of complex networks. We describe innovation as a process which dramatically and decisively changes the composition of a network. The network we have in mind is a network of interaction possibilities, which serves as the structural backbone for various activities of actors. Accordingly, a network will behave differently depending on which innovation has been introduced. More specifically, we propose a network description for dynamic processes in a system. On this system-theoretical perspective, the appearance of innovations is related to certain mechanisms in the growth and change of networks. So, we create a network picture for dynamic processes including the emergence, survival or extinction of innovations. This is the same approach which has also been taken in the mathematical theory of biological evolution, population biology and other fields of complexity theory.

In this contribution, we rely on a special approach to describe innovations in evolutionary systems. We start in Sect. 10.2 with an abstract definition of an innovation. We will show in Sect. 10.3 that our approach is related to recent developments in statistical physics, an area which best can be described as an emerging field of complex networks theory (Scharnhorst, 2003). Thus, we will first give a survey of network approaches ranging from chemical networks (Temkin et al., 1996; Fell, 1997) over biochemical webs (Fell, 1997; Fell and Wagner, 2000), protein webs (Jeong et al., 2000) and food webs (Drossel and McKane, 2003), to the structure of the internet (Faloutsos et al., 1999) and the world wide web (Albert et al., 1999) up to general approaches from statistical mechanics (Albert and Barabási, 2000, 2002; Bianconi and Barabási, 2001a).

In the second part of the contribution (Sect. 10.4) and in connection with our special interest in innovation processes, we develop a dynamic network theory – in particular the theory of "sensitive networks" – in more detail. The term "sensitive networks" denotes networks that are sensitive to the introduction or removal of one or a few nodes or edges, or in a more general context, to the occupation of a node. Specifically, sensitivity is linked to the question of whether a node (a species, mode of behaviour, an idea or a technology) is occupied by at least one individual or not. We will show that this problem is relevant for the modelling of innovation processes. The dynamics of innovation processes may be described by stochastic equations which can be formally treated within the framework of statistical physics. Some generalizations are possible.

10.2 Innovation and Sensitive Networks

In economics, innovation is mainly understood as technological innovation, describing the introduction of new technologies, products and production processes. The differentiation between invention and innovation relates innovation to the economic exploitation of new ideas. However, it is also possible to look at innovations from a more general, evolutionary point of view. Ziman gives one example for such an approach when he writes "Go to a technological museum and look at the bicycles.

Then go to a museum of archaeology and look at the prehistoric stone axes. Finally, go to a natural history museum and look at fossil horses. In each case, you will see a sequence, ordered in time of changing but somewhat similar objects." (Ziman, 2000, p. 3).

Figure 10.1 illustrates our approach. The system is composed of a large set of enumerable types which are represented as nodes in a network. At a certain point in time, only a small part of these nodes are populated, that is, they are active. A populated node is represented by a grey node and a non-occupied node by an empty circle. This way our network consists of grey/occupied and empty/non-occupied nodes with certain links. The whole picture changes over time. The occupation of nodes changes due to the interaction, and occupied nodes can vanish. The pattern of interaction between them (including processes of self-influence) determines the dynamic composition of the system. It is visualized in terms of (active) links between the nodes. This active network will produce a dynamics which has a certain set of stable stationary states. We assume that the activated part of the network is embedded in a much larger network of inactive nodes and links. The inactive nodes represent future possibilities in the evolution of the system. An innovation appears when an unoccupied node becomes occupied for the first time. With this

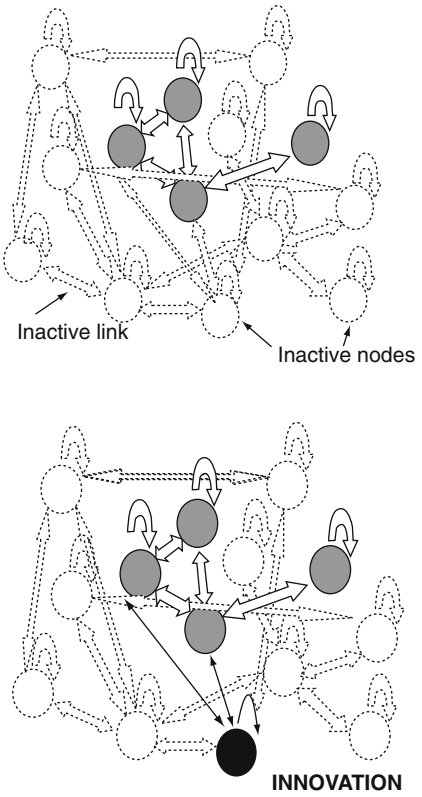


Fig. 10.1 Illustration of a sensitive network. Embedded in a larger network of inactive nodes and links; the activated part of the network (*upper part* of the figure) changes its composition when an innovation emerges (*lower part* of the figure)

initial occupation, the set of links connecting the "new" node with already occupied nodes also becomes activated. It is readily apparent that such an event changes the whole composition of the system. Accordingly, the stable state that the system might have already reached becomes instable, and the system searches for a new stable state. If we assume that the interaction between the nodes (types) is a competitive one, the stable state of a certain activated network can also include the deactivation of certain nodes. Types (nodes) which are selected out will transit to a non-occupied inactive status. However, not every change is an innovation.

In this contribution, we follow a system-theoretical approach to innovation. In this framework, innovation is something new to the system and most importantly the emergence of an innovation changes the state of the system dramatically. In other words, the actual state of the system becomes unstable and a transition to a new state occurs. To define an innovation, we first have to define the state of the system. Here, we again choose a very specific approach. We represent the state of the system as a point in the high-dimensional occupation number space (Ebeling and Feistel, 1982; Ebeling and Sonntag, 1986). In this space, a coordinate axis is attached to a certain type i of elements (with $i = 1, 2, \ldots, s$, natural numbers). The occupation numbers are represented on this axis.

To describe technological innovation we have to ask for a re-specification of this abstract concept. For socio-economic systems, the axes of the state space refer to different possible taxonomies. For instance, an axis i can represent a certain technology from a set s, of different technologies present in the system. With such a technological taxonomy, competition processes between technologies can be described (Saviotti and Mani, 1995; Bruckner et al., 1996). The carriers of this competition process are firms using different technologies and competing with their products on a market. This way, we link back to an economic understanding of an innovation process that "requires insight into system dynamics grounded in a variety of firm competencies and behaviour and a variety of demand" (Saviotti and Nooteboom, 2000, p. 5). Let us note that the state space concept can be applied to quite different processes. The type i might also stand for the size class that a certain firm belongs to. Then, growth processes of firms are in the focus of the description. Moreover, the type i may represent a certain group in society. Formation of political opinions (Weidlich, 2000) or emergence of norms and violence in groups (Nachtigall, 1998) are then considered. Innovation in these cases covers new forms of collective behaviour.

We can find each type i in N_i elements in the system. The elements may be individuals, but they may also be organizational and institutional units like firms and groups. N_i, the occupation numbers, are functions of time. They are positive or zero. A complete set of occupation numbers $N_1, N_2, \ldots, N_s$ at a fixed time, characterizes the occupation state of this system. The time dependent change of the occupation numbers is described by the movement of this point in the space. The whole motion takes place on the non-negative cone K of the space. In this picture, we can describe the case that a type i is not present in the system at time t. That means the type i is occupied with the number zero ($N_i = 0$). We call a system an under-occupied system if we can make the assumption that the sum of the

occupation numbers is essentially smaller than the total number of possible elements (Ebeling and Sonntag, 1986).

According to this picture, an innovation is an occupation of a non-occupied type. In deterministic systems, zero occupation can only be achieved in the limit of infinite time $t \to \infty$, if a sort died out (zero can be a stable stationary state). For finite times $t > 0$ types cannot arise, if they are not in the system at time $t = 0$, and present types cannot die out. The situation is different if we use the stochastic picture. In stochastic systems the zero state can be reached in finite times t. A stochastic description offers the advantage that at finite times new sorts (innovations) can arise or die out. Therefore, the stochastic description is especially suited for evolutionary processes and in particular for innovation processes.

Let us note here that any innovation will change the taxonomy of types in the system (Allen, 1994). Innovation has to do with uncertainty and its prediction is impossible. With the notion of an under-occupied system, we escape the problem of determining a priori the place or kind of an innovation. Instead, we equip the system with a reservoir of possible innovations. Which of these possibilities will eventually become reality remains uncertain. In some respects this is a trick to avoid the problem with a changing taxonomy. There are other possibilities to escape this problem, e.g. so-called continuous models operating on a characteristic space as we discussed elsewhere (Ebeling et al., 1999; Ebeling and Scharnhorst, 2000; Ebeling et al., 2001). However, in our view the discrete approach we are using has certain advantages; these we will discuss later.

In an under-occupied system most elements have, at a given time t, the occupation number zero. So we can pass from the high-dimensional cone K to a low-dimensional cone K^+. Accordingly, the time-dependent variation of the system can be described as a switching of the state point on the edges of the cone K. If K^+ is an element of the set of all possible cones, we observe a switch from one sub-cone to another. Because the process is discrete, it will appear as hopping on the edges of different positive cones.

In Fig. 10.2, we visualize such a process for three dimensions. At any point in time, the state of the system is represented by a certain vector $\boldsymbol{N}(t)$. In a stationary

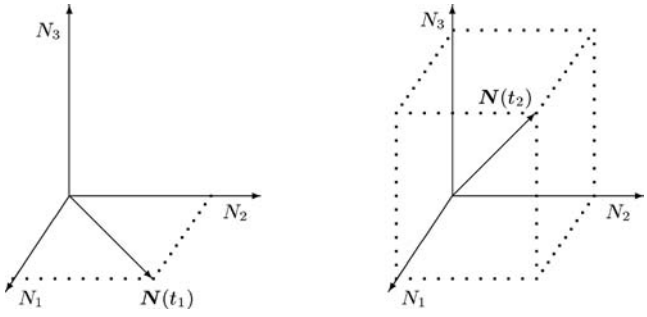


Fig. 10.2 The occupation number space: the distribution of individuals over different types is represented by a vector in the positive cone. The appearance or extinction of a type can be described as approaching or leaving one of the edges of the positive cone

state, the endpoint of this vector defines a positive cone. In our example, the vector moves in the plane spanned by N_1 and N_2. An innovation opens up a new dimension of the system. In our example, a new third type is introduced into the system. After the innovation, the vector moves in the space defined by N_1, N_2 and N_3. In general, we can assume that the system operates in a multidimensional space where the cone can have a very complicated shape, and the vector $\boldsymbol{N}(t)$ jumps between the edges of this cone.

The hopping process visualizes the transition between one stable stationary state and another stable stationary state. In this sense, innovation is the outcome of a process of destabilization. Within the framework we propose in this contribution, innovations are seen as stochastic instabilities. The changes occurring in the occupation number space result from interactions of the different types present in the system. These interactions can be visualized as graphs or networks where the nodes represent the types, and the links between them represent different forms of interactions. In the network picture, an innovation corresponds to the appearance of a new node and the activation of a link to this node. The models that we present in Sect. 10.3 allow us to differentiate between different processes which finally introduce such a new node.

The conceptualization of types as elements (nodes) of a network represents a graph-theoretical approach to the dynamics of the system. Therefore, other network approaches are of particular relevance in order for us to develop our theoretical approach further (Hartmann-Sonntag et al., 2004; Ebeling et al., 2006). Section 10.3 will introduce the newly emerging specialty of complex network theory and position our approach within this field.

10.3 Sensitive Networks and the Emergence of the Field of Complex Networks

In recent years, complex systems in nature and society have been carefully investigated. In the 1970s, already theories of self-organization were used to build a bridge between social and natural systems investigations (Ebeling and Feistel, 1982; Prigogine and Sanglier, 1987). As part of this development, complex networks were examined. Recently, as a new branch in complexity theory (Schweitzer, 1997) complex networks have been reconsidered and extensively studied (Scharnhorst, 2003). They seem to be particularly relevant for the study of innovation processes.

In the context of complexity theory, the concept of networks has been employed as an easy-to-use metaphor. As Bornholdt and Schuster note: “Recent advances in the theory of complex networks indicate that this notion may be more than just a philosophical term. Triggered by recently available data on large real world networks (e.g. on the structure of the internet or on the molecular networks in the living cell) combined with fast computer power on the scientist’s desktop, an avalanche of quantitative research on network structure and dynamics currently stimulates

diverse scientific fields" (Bornholdt and Schuster, 2003, p. V). Social networks form one important area of application of complex network theory (Harary et al., 1965; Scott, 2000). The structures found cover networks of collaboration (Newman, 2000; Barabási et al., 2002), networks of recognition (citation networks) (Vazquez, 2001) and networks of corporate directors (Strogatz, 2001). In economic theories, innovation is increasingly understood as the outcome of the interaction between scientific, economic and political systems (Pyka and Küppers, 2002). Instead of considering an innovation as a singular event, the network character of innovations is stressed. Innovation networks seem to be a new organizational form of knowledge production.

The structural analysis of systems is of great interest. Albert and Barabási (2002) give a very good presentation of this subject and its development. In the very beginning, investigations of large complex systems were done by using random graph theory. To a growing extent it became possible to analyse real complex systems and large systems as well. With the development of computer capacity, the amount of available digital data and the possibilities to analyse and visualize this data also increase (Scharnhorst, 2003). It becomes possible to compare the theoretical results generated by random graph theory with those of the real data analysis. Obviously more than pure randomness exists. Organizational principles and the rules of system evolution play a decisive role, leading to small-world behaviour and scale-free networks. Our world is not a random world. Other evolutionary principles are of great interest. In addition, it is evident that a theory of evolving networks may give a more realistic approach to real systems. This is why we give special consideration to evolving networks here.

From the analysis of empirical data we learn that many real networks have a small-world character. The small-world concept describes the fact that despite their often large size, in most networks there is a relatively short path between any two nodes. The small-world property characterizes most complex networks (Hartmann-Sonntag et al., 2004). For example, the chemicals in a cell are typically separated by only three reactions, or in a more exotic case, the actors in Hollywood are on average within three co-stars from each one another. The small-world concept corresponds to our observations; it is a structural, not an organizing, principle (Watts, 1999; Buchanan, 2002).

Not all nodes in a network have the same number of edges (the same node degree). The spread in node degrees is characterized by a distribution function $P(k)$, which gives the probability that a randomly selected node has exactly k edges. Since in a random graph the edges are placed randomly, the majority of nodes have approximately the same degree, close to the average degree of the network. The degree distribution of a random graph is a Poisson distribution with the peak over the average degree (Albert and Barabási, 2002). In real networks the distributions of the edges are more complicated. Important results were obtained by the analysis of large, real systems: the degree distribution deviates significantly from a Poisson distribution and follows general structural rules in many cases. Many large networks are scale-free, that is, their degree distribution follows a power law. In addition, even for those networks for which $P(k)$ has an exponential character, the degree

distribution significantly deviates from a Poisson distribution achieved by random graph theory for such systems (Albert and Barabási, 2002; Barabási, 2002).

Scale-free networks express a hierarchy between the nodes. Not every node is important at the same level. Accordingly, not every link between the types has the same importance. There are very sensitive relations or elements too. As mentioned previously, we consider networks as "sensitive" if their properties depend strongly on the introduction or removal of one or a few nodes or edges, or changes to the occupation of nodes. We will show that for the evolutionary character, the description of time behaviour by master equations on occupation number spaces is an appropriate tool. The discrete character of occupation number description allows for an appropriate description of the introduction or, respectively, the removal of relations, edges, etc. We will analyse not only the steady states of our stochastic systems, but also the time evolution. Albert and Barabási (2002) also refer to approaches with master and rate equations. In addition, they claim that these methods, without using a continuum assumption, appear more suitable for obtaining exact results in more challenging network models. Further, the authors mentioned that the functional form of the degree distribution, $P(k)$, cannot be guessed until the microscopic details of the network evolution are fully understood. According to our point of view, the method of master equations is an excellent tool to use in the investigation of many open questions; it is able to bring much more light to bear on this subject. For example, by using the discrete approach, we have a chance to obtain statements about the nature of fluctuations – one of the most important issues.

Let us come back now to the question of the distribution of the graph. Again, random graph theory leads to a Poisson degree distribution. Albert and Barabási (2002) give a near exhaustive survey of empirical data sets for real complex networks and show that the real degree distributions are not Poisson distributions, but scale-free distributions, or exponential distributions. These authors write: "The high interest in scale-free networks might give the impression that all complex networks in nature have power-law degree distributions. ... It is true that several complex networks of high interest for scientific community, such as the world wide web, cell networks, the internet, some social networks, and the citation network are scale-free. However, others, such as the power grid or the neural networks of *c. elegans*, appear to be exponential. ... Evolving networks can develop both power-law and exponential degree distributions. While the power-law regime appears to be robust, sub-linear preferential attachment, aging effects, and growth constraints lead to crossovers to exponential decay. ... If all processes shaping the topology of a certain network are properly incorporated, the resulting $P(k)$ often has a rather complex form, described by a combination of power laws and exponentials".

In this respect, our aim here is to calculate the role of fluctuations by using the master equation approach. In this way, we can make statements about how the systems differ from linear systems which obey a Poisson distribution. In principle, by investigating the fluctuations, correlations and spectral densities, we are able to study several microscopic events. One of the questions to answer is which fluctuation effects produce power-law distributions. We find a deep connection of these network systems to systems which produce $1/f$-noise. Again, $1/f$-noise is

a stochastic process with a specific power-law spectrum (Klimontovich, 1995). A characteristic property of processes which produce $1/f$-noise are long range-correlations. We suppose that the scale-free networks and small-world behaviour may have some relation to this. We reiterate that in small-world networks the degree function obeys a power-law; there exists a small pathway between each two of the elements.

In investigating self-organization and evolutionary processes in networks, our basic approach is to understand the corresponding networks as dynamic, or more precisely, as evolutionary systems. This dynamic and evolutionary approach allows us to make statements about innovation processes, special competition effects, the sensitivity of networks, the constraints of growth and the fitness of network systems.

Socio-economic systems are complex systems, which consist of many connections between the elements. Therefore, complexity is a further concept to be defined. Ebeling et al. (1998) write:

> As complex we describe holistic structures consisting of many components, which are connected by many (hierarchically ordered) relations respectively operations. The complexity of a structure can be seen in the number of equal respectively distinct elements, in the number of equal respectively distinct relations and operations, as well as in the number of hierarchical levels. In the stricter sense, complexity requires that the number of elements becomes very large (practically infinite).

We are especially interested here in the origin of complex structures, and in the development of order (information). In the end, we have to answer the question of which parameter relations (order parameters) determine the qualitative behaviour of the system (Hartmann-Sonntag et al., 2004). Prigogine (1955) in collaboration with his coworkers, did pioneering work in the investigation of self-organizing systems (Nicolis and Prigogine, 1977). Further important work has been done in this field by Eigen (1971) on the self-organization of macromolecules and by Eigen (1971) and Eigen and Schuster (1977, 1978) on the hypercycle model. The mechanisms of self-organization are clearly worked out by Nicolis and Prigogine (1977). These authors give a stringent physical and mathematical formulation of these processes, in particular with respect to the energetic and entropic aspects. A somewhat different view on this was developed in the formulation of synergetics by Haken (1978). The investigation of such systems shows that the formation of order in complex systems can be allocated to physical processes, which play a role far from equilibrium (Ebeling and Feistel, 1982). We underline the fact that biological, just as socio-economic processes, can be investigated with the help of the theory of self-organization because they obey valid physical and chemical laws. However, processes which include real-life (biological and socio-economic) systems also obey other rules and laws that are not solely determined by physics. This is already evident from the very general character of the structures we consider here. As stated above, we formulate the idea of structure mathematically and discuss the meaning of structure in the context of technological evolution.

10.4 Stochastic Models of Innovation Processes

10.4.1 Overview of Stochastic Effects in Networks

10.4.1.1 Birth and Death Processes

A special stochastic process called "birth and death process" is of particular importance to our study of stochastic effects in economic processes, and especially in innovation processes. A birth process is a random appearance of a new element in a system. A death process is the disappearance of an element. Processes of this kind play a big role in biology, ecology and sociology.

Processes of this type are also relevant in economy (Kaniovski, 2000). Economic growth is characterized by structural changes based on the introduction of new technologies in the economic world. To describe technological evolution, we must determine the system, its elements and their interactions. Here we consider production plants and technologies. We consider firms composed of different manufacturing plants. The plants are introduced as elementary units which function as decision carriers according to market conditions (choosing a new technology or not). Plants also play the role of users of a particular technology. The technologies are understood as the different types present in the system. The production plants are the elements or representatives of these technologies. On this perspective, technologies eventually compete for the plants using them. This perspective differs from the way one usually thinks about technological change, where the firm is central (Hartmann-Sonntag et al., 2004). The underlying process is still a decision made by plants or firms. However, the model approach constructs an inverse perspective on it. Let us note here that this perspective is quite normal for any population dynamic approach which deals with types (groups or species) and elements (individuals). However, in contrast to biological processes, human beings, organizations and firms are not bound to a certain type or group to which they belong initially. In contrast to individuals of biological species, they have the opportunity to change the group they belong to. It is this kind of transition behaviour that makes the model particularly relevant for socio-economic applications. Further, we deliberately use the notion of a plant or production unit as the simplest element in the system. By assuming that firms consist of several plants or production units, growth processes of firms are also covered by the model approach. Technological change is usually considered as a macro-economic change process. However, in order to describe it as an evolutionary process, one must consider this process at the microscopic level. This means we have to consider the micro-economic carriers of technological change. In the framework we present here, these are the plants (Bruckner et al., 1996).

The basic ideas for the modelling of these processes go back to so-called urn models. In 1907, the physicists, Paul and Tatyana Ehrenfest, developed a simple model for the diffusion of N molecules (Ehrenfest and Ehrenfest, 1907). The Ehrenfests studied two urns, A and B, which were isolated with respect to exchange with their surroundings. With respect to exchange between urns, the Ehrenfests assumed

permeable connections between A and B. Because of the isolation of the two urns, the total number of molecules in A and B remains constant. At regular time intervals, a molecule is randomly (that means with the probability $(1/N)$) chosen and changes from one urn to the other.

In 1926, Kohlrausch and Schrödinger (1926) gave a continuous diffusion approximation for such processes. Feller (1951) formulated a realistic variant of this model. He used a discrete Markovian process with continuous time. The time between the molecule crossings was exponentially distributed.

Originally developed for molecular processes, the model soon found many applications to biological processes. Surveys of biological applications of birth and death processes were given by Bartholomay (1958a,b, 1959) and Eigen (1971).

The Ehrenfest model represents the prototype for the investigation of decision processes in a group between the possibilities A and B, respectively, between yes and no (Ebeling et al., 2000). For example, we may consider the decision of whether to accept a new technology or not. Applications to social and economic processes were surveyed by Weidlich (2000).

10.4.1.2 Stochastic Effects in Small and Sensitive Networks

Many complex systems display a surprising degree of tolerance to errors. The results indicate a strong correlation between robustness and network topology. In particular, scale-free networks are more robust than random networks against random node failures (Hartmann-Sonntag et al., 2004), but are more vulnerable when the most connected nodes are targeted (Albert et al., 2000).

In small networks any nodes or edges play a specific role. Their addition or removal drastically changes the properties of the whole system. Another problem where stochastic effects play a big role is the question of how a single new mutant can win the selection process. If one considers, as an example, networks of web sites and competition processes between them about attracting visitors one would ask how can a new web site become a giant cluster among other already important web sites (clusters)? Once a dominant regime or design is in place one can further ask: Is it possible to overcome the “once-forever” selection behaviour by stochastic effects?

The addition or removal of sensitive nodes or edges is a subject of investigation also in big networks. With sensitive we mean here elements which play a special role in the network. Let us consider some examples.

Cellular networks can be subject to random errors as a result of mutations or protein misfolding, as well as harsh external conditions eliminating essential metabolites. Jeong et al. (2000) studied the responses of the metabolic networks of several organisms to random and preferential node removal. Removing up to 8% of the substrates, they found that the average path length did not increase when nodes were removed randomly. However, it increases rapidly after the removal of the most connected nodes and up to 500% when only 8% of the nodes are removed. Similar results have been obtained for the protein network of yeast as well (Jeong and Mason, 2001; Vogelstein et al., 2000).

Solé and Montoya (2001) studied the response of the food webs to the removal of species (nodes) (see also Montoya and Solé, 2002). The results indicate that random species removal causes the fraction of species contained in the largest cluster to decrease linearly. However, when the most connected (keystone) species are successively removed, the relative size of the largest cluster quickly decays.

The error and attack tolerance of the internet and the world wide web was investigated by Albert et al. (2000). The internet is occasionally subject to hacker attacks targeting some of the most connected nodes. They show that the average path length on the internet is unaffected by the random removal of as many as 60% of the nodes, while if the most connected nodes are eliminated (attack), the average path length peaks at a very small fraction of removed nodes. Albert, Jeong and Barabási investigated the world wide web (Albert et al., 1999) and showed that the network survives as a large cluster under high rates of failure, but under attack, the system abruptly crashes. These authors write: "The result is that scale-free networks display a high degree of robustness against random errors, coupled with a susceptibility to attacks".

Wagner and Fell (2000) studied the clustering coefficient, focusing on the energy and biosynthesis metabolism of the *Escherichia coli* bacterium. They found that in addition to the power-law degree distribution, the undirected version of this substrate graph has a small average path length and a large clustering coefficient.

Hartmann-Sonntag et al. (2004) investigated in detail the structure and the dynamic behaviour of networks. In particular, they considered fundamental properties of graphs as, for instance, the probability distribution for their components. They introduced a stochastic model by which the role of fluctuations for the survival of innovations in such non-linear systems can be explained.

Bianconi and Barabási (2001b) showed the existence of a closed link between evolving networks and an equilibrium Bose gas. According to them, the mapping onto a Bose gas predicts the existence of two distinct phases as a function of the energy distribution. In the fit-get-rich-phase, the fitter nodes acquire edges at a higher rate than older but less fit nodes. In the end, the fittest node will have the most edges, but the richest node is not an absolute winner, because its share of the edges decays to zero for large system size.

Maurer and Huberman (2000) present a dynamic model of web site growth in order to explore the effects of competition among web sites. They show that under general conditions, as the competition between sites increases, the model exhibits a sudden transition from a regime in which many sites thrive simultaneously, to a "winner takes all market" in which a few sites grab almost all of the users, and most of the other sites become nearly extinct. This prediction is in accord with empirical data measurements on the nature of electronic markets.

Dorogovtsev et al. (2003) developed a statistical mechanical approach for random networks. They summarize as follows: "Using the traditional formalism of statistical mechanics, we have constructed a set of equilibrium statistical ensembles of random networks without correlation and have found their partition function and main characteristics. We have shown that a 'scale-free' state in equilibrium

networks without condensate may exist only in a single marginal point, so that in such an event this state is an exhibition". They underline the important fact that this differs crucially from the situation for growing networks. The latter, in the process of growing, self-organize into scale-free structures in a wide range of parameters without condensation.

10.4.2 Stochastic Analysis of Innovation Processes

10.4.2.1 A General Formulation of the Model

The stochastic approach given here is based on a model, which was developed in the context of general models of evolutionary processes, and in particular biochemical processes (Ebeling et al., 1981; Schimansky-Geier, 1981; Ebeling and Feistel, 1982; Ebeling and Sonntag, 1986). Later, the model found numerous applications in modelling scientific evolution (Ebeling and Scharnhorst, 1986; Bruckner et al., 1990) and technological evolution (Bruckner et al., 1994, 1996). In this contribution, we first introduce the model framework in a general form and later concentrate on its application to economic innovation and competition processes.

In the stochastic picture we use the ideas developed in Sect. 10.2 and in Sect. 10.4.1.1 about the occupation number space. Contrary to deterministic models, the stochastic description offers the advantage that at finite times new types (sorts, fields, species, technologies) can arise or "die out".

Let us introduce a set of types numbered by $i = 1, 2, \ldots, s$. We denote by $N_i(t)$ the number of elements belonging to a certain type. For an economic application $N_i(t)$ represents the number of production plants using the technology i (Note that a type corresponds to an urn in the Ehrenfest problem formulation. For more details see Hartmann-Sonntag et al., 2004). These numbers are called occupation numbers. They are a function of time. The occupation numbers are positive or zero.

$$N_i(t) = \{0, 1, 2, \ldots\} \,. \tag{10.1}$$

Now, the state of the system at time t can be described by the probability distribution of the occupation numbers

$$P(N_1, N_2, \ldots, N_s; t) = P(N; t). \tag{10.2}$$

We consider as elementary processes, processes during which only one occupation number can change, and processes during which at most two occupation numbers can change (transition processes):

$$\begin{array}{ccc} (N_i) & \longrightarrow & (N_i + 1) \\ (N_i) & \longrightarrow & (N_i - 1) \\ \begin{pmatrix} N_i \\ N_j \end{pmatrix} & \longrightarrow & \begin{pmatrix} N_i - 1 \\ N_j + 1 \end{pmatrix}. \end{array} \tag{10.3}$$

For the time being, we also assume that growth and decline processes of the total number of elements in the system are possible. Noteworthy, however, here is that the original formulation of the Ehrenfest model only contains transition processes. This is due to the fact that in the Ehrenfest model the total number of elements remains constant and exchanges occur only between the urns. Decision processes in the model occur as transitions of elements between types.

If we assume that all decisions which lead to a change in the set of occupation numbers depend mostly on the present state, we can apply the concept of Markov processes. We can thus describe the dynamics of the system with the help of the master equation. This equation is a balance equation between building and reduction processes:

$$\frac{\partial P(N;t)}{\partial t} = W(N|N')\, P(N') - W(N'|N)\, P(N) \tag{10.4}$$

with

$$N = \{N_1, N_2, \ldots, N_s\}\ . \tag{10.5}$$

The transition probabilities per time unit that the system turns from the state N' to the state N or vice versa are expressed by $W(N|N')$ and $W(N'|N)$, respectively.

The transition probabilities are supposed as follows (Feistel and Ebeling, 1978; Ebeling and Feistel, 1982; Jiménez-Montaño and Ebeling, 1980; Ebeling and Sonntag, 1986; Heinrich and Sonntag, 1981; Ebeling et al., 1990a,b; Hartmann-Sonntag et al., 2004; Ebeling et al., 2006):

1. Spontaneous generation

$$\begin{aligned} &W(\ldots, N_i + 1, \ldots, N_j, \ldots, N_k, \ldots | \ldots, N_i, \ldots, N_j, \ldots, N_k, \ldots) \\ &\quad = A_i^{(0)}. \end{aligned} \tag{10.6}$$

2. Self-reproduction

$$\begin{aligned} &W(\ldots, N_i + 1, \ldots, N_j, \ldots, N_k, \ldots | \ldots, N_i, \ldots, N_j, \ldots, N_k, \ldots) \\ &\quad = A_{ij}^{(1)}\, N_j + E_i^{(1)}\, N_i \end{aligned} \tag{10.7}$$

$$E_i^{(1)} = A_i^{(1)} + B_{ij}^{(1)}\, N_j + C_{ijk}^{(1)}\, N_j N_k. \tag{10.8}$$

3. Decay

$$W(\ldots, N_i - 1, \ldots, N_j, \ldots, N_k, \ldots | \ldots, N_i, \ldots, N_j, \ldots, N_k, \ldots) = E_i^{(2)} N_i \tag{10.9}$$

$$E_i^{(2)} = A_i^{(2)} + B_{ij}^{(2)} N_j. \tag{10.10}$$

4. Conversion/Transition/Exchange/Mutation

$$W(\ldots, N_i + 1, \ldots, N_j - 1, \ldots, N_k, \ldots | \ldots, N_i, \ldots, N_j, \ldots, N_k, \ldots) = E_j^{(3)} N_j \tag{10.11}$$

$$E_j^{(3)} = A_{ij}^{(3)} + B_{ij}^{(3)} N_i + \bar{B}_{ik}^{(3)} N_k + C_{ijk}^{(3)} N_i N_k \tag{10.12}$$

with $j \neq i; k \neq i; j$.

The coefficients can be introduced differently for special cases. For instance, they can be considered as constant or as functions of the total number of elements and other the system parameters; this latter approach introduces an additional nonlinearity to the system.

The content of the four elementary processes introduced above will be quite different according to the nature of the system under consideration. For instance, for catalytic networks, self-reproduction may appear as a result of a process of spontaneous self-reproduction (term related to $A_i^{(1)}$), error reproduction ($B_{ij}^{(1)}$) or catalytic self-reproduction ($C_{ijk}^{(1)}$). The decay can appear in the form of spontaneous decay ($A_i^{(2)}$) and decay related to catalytic help ($B_{ij}^{(2)}$). Transition or conversion processes in catalytic networks correspond to mutation processes with reproduction and ternary reproduction processes; these in turn are important for processes with a constant overall particle number.

In the case of technological evolution, self-reproduction appears as a growth process of firms expanding their number of production units and plants using the same technology. Synergetic effects from the surrounding network of firms are supposed to occur when the transition rate also depends on the number of firms using another technology. Spontaneous generation stands for startups. Decay processes refer to both a decrease in firm size (closing of production units) and a decrease in the total number of firms (complete closing down of whole enterprises) (Hartmann-Sonntag et al., 2004). The most interesting process is related to conversion or transition. Here, the use of another technology by a production unit is described. The use of a technology may be new for an individual firm only (firm-specific innovation) or for the whole system of firms (system-specific innovation). Triggered by R&D, both invention of a technology and imitation behaviour are covered by this process. Also, in this case, other technologies can influence a firm's decision to adopt a certain

technology. The advantage of this type of model is that technological change is considered as the outcome of the development of a network of technologies and firms influencing one another.

The transition probabilities are formulated generally. As a special case, we can obtain from this *Ansatz* the stochastic equations, which correspond to the deterministic EIGEN-model (Eigen and Schuster, 1978) with the condition of constant overall particle number (Jiménez-Montaño and Ebeling, 1980; Ebeling et al., 1981; Heinrich and Sonntag, 1981).

With the help of the s-dimensional generation function

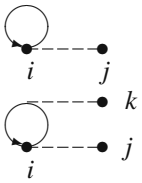
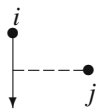
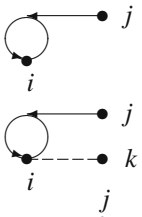
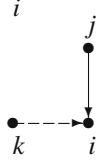
$$F(s_1, s_2, \ldots, s_s; t) = \sum_N s_1^{N_1}\, s_2^{N_2}\, \ldots\, s_s^{N_s}\, P(N_i, t) \quad \text{with} \quad |s_i| < 1 \tag{10.13}$$

we can write the master equation with the transition probabilities as follows:

$$\begin{aligned}
\dot{F}(s;t) = \sum_{i \neq j} \Bigg\{ & A_i^{(0)}\,(s_i - 1)\,F + A_i^{(1)}\, s_i (s_s - 1)\,\frac{\partial F}{\partial s_i} \\
& + A_{ij}^{(1)}\, s_j (s_i - 1)\,\frac{\partial F}{\partial s_j} + B_{ij}^{(1)}\, s_i s_j (s_i - 1)\,\frac{\partial^2 F}{\partial s_i \partial s_j} \\
& + C_{ijk}^{(1)}\, s_i s_j s_k (s_i - 1)\,\frac{\partial^3 F}{\partial s_i \partial s_j \partial s_k} + A_i^{(2)}\,(1 - s_i)\,\frac{\partial F}{\partial s_i} \\
& + B_{ij}^{(2)}\, s_j (1 - s_i)\,\frac{\partial^2 F}{\partial s_i \partial s_j} + A_{ij}^{(3)}\,(s_i - s_j)\,\frac{\partial F}{\partial s_j} \\
& + B_{ij}^{(3)}\, s_i (s_i - s_j)\,\frac{\partial^2 F}{\partial s_i \partial s_j} + \bar{B}_{ik}^{(3)}\, s_k (s_i - s_j)\,\frac{\partial^2 F}{\partial s_k \partial s_j} \\
& + C_{ijk}^{(3)}\, s_i s_k (s_i - s_j)\,\frac{\partial^3 F}{\partial s_i \partial s_j \partial s_k} \Bigg\}.
\end{aligned} \tag{10.14}$$

10.4.2.2 A Network Representation of the Model

In order to relate the model to the idea of sensitive networks we now introduce a network, representation of transition probabilities given in (10.6), (10.7), (10.8), (10.9), (10.10), (10.11), and (10.12). We consider a system with s interacting types (sorts, fields, plants). This system can be described by a graph in which each element i corresponds to a vertex of the number i (Hartmann-Sonntag et al., 2004). We mark the transition probabilities for different processes by edges of different types in order to distinguish these probabilities:

1. Spontaneous generation (simple innovation)	$A_i^{(0)}$	
2. Self-reproduction	$A_i^{(1)} N_i$	
Error reproduction	$A_{ij}^{(1)} N_j$	
Catalytic self-reproduction (sponsored self-reproduction)	$B_{ij}^{(1)} N_i N_j$ $C_{ijk}^{(1)} N_i N_j N_k$	
3. Spontaneous decay	$A_i^{(2)} N_i$	
Catalytic decay	$B_{ij}^{(2)} N_i N_j$	
4. Mutation (innovation)	$A_{ij}^{(3)} N_j$	
Mutation (innovation) with reproduction	$B_{ij}^{(3)} N_i N_j$ $C_{ijk}^{(3)} N_i N_j N_k$	
Mutation without reproduction	$\bar{B}_{ik}^{(3)} N_k N_j$	

For under-occupied systems, the transition from a non-occupied ($N_i = 0$) state to an occupied ($N_i > 0$) state is of special interest. We will call such a transition an innovation. To describe this in the network picture, we distinguish two states for one vertex. The states are marked as follows:

$\circ_{\,\mathrm{i}}$	non-occupied	$N_i = 0$
$\bullet_{\,\mathrm{i}}$	occupied	$N_i > 0$

We also distinguish two states for the edges of the graph. Edges emanating out from unoccupied vertices ($N_i = 0$) are omitted because they are inactive (they cannot work). Active (working) edges are characterized by the graphic representation shown above. For socio-economic networks, non-occupied vertices stand for new technologies which have not yet been discovered. These can be understood as hidden possibilities. The model does not allow for the prediction of any specific new technology. But, it can make statements about favourable or unfavourable conditions for the emergence of new technologies. One can also ask how the system handles the appearance of new technologies in general.

Eventually, we obtain a graph that describes the whole system at a fixed time t. If the colour of a vertex is changed, new connections (elementary processes) can flare up or former connections discontinue. Processes can spread the elements over the non-occupied vertices or they can select them from the occupied vertices. The basic structure of the network is the "maximal" graph (all vertices i are occupied, all interactions can occur).

In a network with a small overall number of elements (individuals, organizations, plants) ($N \ll s$), many vertices are not occupied. Let us consider a graph with ($N \gg s$) and reduce the overall number of elements to a state ($N \ll s$). In general, vertices such as i and j become non-occupied by the following processes: $A_i^{(2)} N_i$, $B_{ij}^{(2)} N_i N_j$, $A_{ij}^{(3)} N_j$, $B_{ij}^{(3)} N_i N_j$, $\bar{B}_{ik}^{(3)} N_k N_j$, $C_{ijk}^{(3)} N_i N_j N_k$. These are the processes of decay and the processes of conversion. If only a certain part of the processes work because of the small overall number, we obtain a graph with many components. With a decreasing number of elements (individuals, organizations, plants) the maximal graph with few components develops into a graph with a large number of components (Ebeling and Feistel, 1982). The "minimal" graph we obtain (maximally decomposed) consists of self-reproduction processes, spontaneous generation (simple innovations), sponsored innovation processes and sponsored innovation processes with self-reproduction (self-reproducing process). Conversely, vertices can be occupied by the following processes: spontaneous generation (simple innovation) $A_i^{(0)}$, error reproduction $A_{ij}^{(1)} N_j$, mutation without reproduction (innovation without self-reproduction) $A_{ij}^{(3)} N_j$ and mutation without reproduction but with catalytic help (sponsored innovation without self-reproduction) $\bar{B}_{ik}^{(3)} N_k N_j$.

We assume that the processes through which the vertices can be occupied are very rare. This means spontaneous generations (simple innovations) and mutations have a small probability. After a relatively short time, the components are in a local equilibrium (first process). Then, sponsored innovation processes begin to become significant starting from an occupied component which is in equilibrium (Hartmann-Sonntag et al., 2004). New components (non-occupied) can be occupied (second process). This is a hopping process between components. Certain components die out under selection pressure, others survive and grow.

10.4.3 Application to Technological Innovations

10.4.3.1 Technologies with Linear Growth Rates

In the following, we consider a system in which the total number of elements is constant. As mentioned above, a useful instrument to model such processes is the theory of stochastic transitions between urns, established by Ehrenfest. Compared with the general model introduced so far, in the case of a closed system, all elementary processes appear as transition processes. Further, let us restrict ourselves to the case of technological evolution. Then, the urns stand for different technologies. Symbolic spheres travelling between the urns stand for production plants looking for technologies.

N_i is the number of plants, which uses the technology i, this means they belong to the urn i. These numbers are called occupation numbers:

$$N_i(t) = \{0, 1, 2, \ldots\} \,. \tag{10.15}$$

First, we formulate a simple model for the binary decision process:

$$N = N_1 + N_2 \,. \tag{10.16}$$

We assume that during elementary processes the occupation number only changes by ± 1. This is the so-called one-step process. During transition processes, at most two occupation numbers can change in this way:

$$\begin{pmatrix} N_1 \\ N_2 \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} N_1 - 1 \\ N_2 + 1 \end{pmatrix}. \tag{10.17}$$

For instance, we assume that E_1 is the growth rate of plants using technology 1. For a new technology 2 the growth rate is E_2. We assume:

$$E_2 \; > \; E_1 \,. \tag{10.18}$$

In this case, technology 2 has a greater growth potential or will grow faster than technology 1. We will assume that this is an expression of technology 2 being "better" or more suitable for production plants. In the stochastic picture, the transition from urn 1 to urn 2 will simply have a higher probability than the reverse transition. So, plants will change their technology more often to technology 2. They will replace the old technology by a better (new) one.

We write the transition probability for this process as follows (Hartmann-Sonntag et al., 2004):

$$W^+(N_2 + 1|N_2) = E_2 \, N_1 \left(\frac{N_2}{N}\right) + E_{21} = W^+_{N_2}. \tag{10.19}$$

In particular, we assume that the transition probability is proportional to the number of plants that use the old technology:

$$W^+(N_2) \sim N_1. \tag{10.20}$$

Furthermore, the probability is also proportional to the relative number of plants that use the new technology

$$W^+(N_2) \sim \frac{N_2}{N}. \tag{10.21}$$

We introduce an additional process of spontaneous change from technology 1 to technology 2, which is described by the coefficient E_{21}. The opposite spontaneous transition appears with the coefficient E_{12}. In total, the opposite transition process has the probability:

$$W^-(N_2 - 1|N_2) = E_1\, N_2 \left(\frac{N_1}{N}\right) + E_{12} = W^-_{N_2}. \tag{10.22}$$

We can formulate the master equation for this discrete process. This equation describes the time–behaviour of the probability distribution of the occupation numbers.

$$P(N_1, N_2; t). \tag{10.23}$$

With the mentioned transition probabilities follows:

$$\begin{aligned}\frac{\partial}{\partial t} P(N_1, N_2; t) = {} & W^+_{N_2-1}(N_2|N_2 - 1)\; P(N_2 - 1; t) \\ & + W^-_{N_2+1}(N_2|N_2 + 1)\; P(N_2 + 1; t) \\ & - W^+_{N_2}(N_2 + 1|N_2)\; P(N_2; t) \\ & - W^-_{N_2}(N_2 - 1|N_2)\; P(N_2; t)\end{aligned} \tag{10.24}$$

and (Hartmann-Sonntag et al., 2004; Ebeling et al., 2006)

$$\begin{aligned}&\frac{\partial}{\partial t} P(N_1, N_2; t) \\ &\quad = \left[E_{21} + \frac{E_2}{N}\,(N_2 - 1)(N_1 + 1)\right]\; P(N_1 + 1, N_2 - 1; t) \\ &\quad + \left[E_{12} + \frac{E_1}{N}\,(N_2 + 1)(N_1 - 1)\right]\; P(N_1 - 1, N_2 + 1; t) \\ &\quad - \left[E_{21} + E_{12} + (E_1 + E_2)N_2(N - N_2)\right]\; P(N_1, N_2; t).\end{aligned} \tag{10.25}$$

If we use the relation $N_1 + N_2 = N$ (i.e. $N_1 = N - N_2$), we can write:

$$\begin{aligned}\frac{\partial}{\partial t} P(N_2;t) &= \left[E_{21} + \frac{E_2}{N}\,(N_2-1)(N-N_2+1)\right] P(N_2-1;t) \\ &+ \left[E_{12} + \frac{E_1}{N}\,(N_2+1)(N-N_2-1)\right] P(N_2+1;t) \\ &- \left[E_{21} + E_{12} + \frac{(E_1+E_2)}{N} N_2(N-N_2)\right] P(N_2;t). \end{aligned} \tag{10.26}$$

To obtain statements for the deterministic case, we define the mean value:

$$\langle N_2(t)\rangle = \sum_{N_2=0}^{\infty} N_2\; P(N_2;t). \tag{10.27}$$

By multiplying the master equation with N_2 and a following summation, we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle N_2(t)\rangle = \frac{E_2 - E_1}{N}\;\langle N_2(N-N_2)\rangle + (E_{21} - E_{12}). \tag{10.28}$$

Using the approximation $\langle (N_2)^2\rangle \approx \langle N_2\rangle^2$ and the abbreviations

$$x_2 = \frac{\langle N_2\rangle}{N}\;;\; \alpha = E_2 - E_1\;;\; \beta = \frac{E_{21} - E_{12}}{N} \tag{10.29}$$

we achieve the corresponding deterministic equations:

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \alpha\; x_2\,(1 - x_2) + \beta. \tag{10.30}$$

Let us compare the stationary behaviour of both the stochastic and the deterministic models. In the case of two technologies, we can explicitly derive the stationary solution of the master equation:

$$P^0(N_2) = \frac{W^+(N_2)\; W^+(N_2-1)\;\cdots\; W^+(N-1)}{W^-(N_2+1)\; W^-(N_2+2)\;\cdots\; W^-(N)}\; P^0(N). \tag{10.31}$$

In the deterministic case, the stationary solution for $E_2 > E_1$ is:

$$x_2 \;= 1, \quad \text{that is} \quad N_2 \;= N. \tag{10.32}$$

The new (better) technology will replace the old one. If a final stable stationary solution is reached by the system, all of the plants will use the new technology.

For the stochastic case, we get quite a different picture. In particular, the probability distribution $P(N_2) \neq 0$ for $N_2 \neq N$. This means that in the stochastic case, the old and new technologies will coexist, if the difference between E_1 and E_2 is not too large. Otherwise, the old technology would only survive in a small niche.

An interesting case occurs, if the spontaneous rates are $E_{21} = E_{12} = 0$. In this case the states $N_1 = 0$, $N_2 = 0$ are called absorber states. This means that the system cannot leave these states. For the stationary solution of the master equation, we get:

$$P^0(N_2) = \sigma_1 \, \delta_{0N_2} + \sigma_2 \, \delta_{NN_2} \, ; \sigma_1 + \sigma_2 = 1 \tag{10.33}$$

σ_i is a real number between zero and one. From the absorber states $N_2 = 0$ and $N_1 = 0$ ($N_2 = N$) the other states cannot be reached.

The initial state determines which stationary solution is occupied by the system. After some calculations we can give for σ_2 in the limit $N \gg 1$ and $N_2(t = 0)$ the following equation:

$$\sigma_2 = \begin{cases} 0 & \text{for} \quad E_2 < E_1 \\ 1 - \left(\frac{E_1}{E_2}\right)^{N_2(0)} & \text{for} \quad E_2 > E_1 \end{cases} \tag{10.34}$$

with $N_2(0) = N_2(t = 0)$ being the initial state of the system. If $N_2(0)$ is the number of users at time $t = 0$ of technology 2, then σ_2 is the probability that for $t \to \infty$ $N_2 = N$ users change to technology 2. σ_1 is the probability that for $t \to \infty$ $N_2 = 0$ (i.e. $N_1 = N$) users changes to technology 2, that is, that the new technology has not survived. In general, σ_i is the survival probability of technology i.

If a small number of plants $N_2(0)$ use a new technology, this technology will disappear if its growth rate E_2 is smaller than that of the old technology. If the growth rate E_2 of the new technology is considerably larger, the new technology will succeed with a probability of

$$1 - \left(\frac{E_1}{E_2}\right)^{N_2(0)} . \tag{10.35}$$

With a small probability of

$$\left(\frac{E_1}{E_2}\right)^{N_2(0)} \tag{10.36}$$

the old technology will still be used within the system.

In the deterministic case a new technology with higher growth rates is always successful. This is the case of pure Darwinian selection where the fittest and only the fittest survives at the end of the process.

What we observe in socio-economic systems is usually a variety of technologies (Saviotti, 1996). This can be either explained by the action of mutation processes

and error reproduction or the presence of stochastic processes. As we showed above in the stochastic situation, the new better technology will only survive with a certain probability. Empirical studies of growth processes in ensembles of technologies might shed light on which growth mechanism is present in a certain system.

10.4.3.2 Technologies with Quadratic Growth Rates

In the following we consider, again, the case with absorber states; this means we neglect the terms E_{21}, E_{12}. As an extension to the model introduced above, instead of the linear terms (E_1, E_2), we now introduce non-linear components (b_1, b_2, V). V is a parameter which stands for the spatial volume of the system in biochemical applications. In general terms it is related to the size of the system. In principle, we could use it as an additional tuning parameter for density effects, but we can also set it to 1, which leaves us with N as the only size parameter.

Remember that we still have a system with a constant total number of elements (plants), so all processes appear as transition processes. The non-linearities might stand for processes of catalytic self-reproduction as well as for transition or mutation processes influenced by other types (technologies). We can write the transition probabilities as:

$$W^+(N_1) = \frac{b_1}{NV}\, N_2\, N_1^2 \tag{10.37}$$

$$W^-(N_1) = \frac{b_2}{NV}\, N_2^2\, N_1 \tag{10.38}$$

(compare Ebeling et al. (1981) in detail for an application to biochemical systems). We assume that the sum of the occupation numbers remains constant:

$$N_1 + N_2 = N = \text{const.} \tag{10.39}$$

The probability distribution of the occupation numbers is:

$$P(N_1, N_2; t). \tag{10.40}$$

Analogous to our formulation in the previous section, we can formulate the master equation as follows:

$$\begin{aligned}\frac{\partial}{\partial t} P(N_1;t) = {} & W^+(N_1-1)\, P(N_1-1;t) + W^-(N_1+1)\, P(N_1+1;t) \\ & - \left[W^+(N_1) + W^-(N_1)\right]\, P(N_1;t).\end{aligned} \tag{10.41}$$

By multiplying the master equation with N_k/V and summing over all occupation numbers, we obtain, after factorization of the mean values, the deterministic equation:

$$\langle N_k \rangle / V = x_k \quad \text{with} \quad x_1 + x_2 = \frac{N}{V} = C = \text{const.}$$

$$\dot{x}_i = b_i \, x_i^2 - \varphi \, x_i \; ; \quad i = 1, 2 \tag{10.42}$$

with

$$\varphi = \frac{b_1 x_1^2 + b_2 x_2^2}{C}. \tag{10.43}$$

Equations of the same form have been derived for so-called hypercyclic systems to describe the evolution of macromolecules (Eigen and Schuster, 1978). Their behaviour is well understood. As a result of the quadratic terms in the growth rates, the phase space is split into two regions separated by a separatrix S_i (Fig. 10.3):

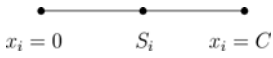


Fig. 10.3 Phase space in the deterministic case

Of note here is that, because of the $N = \text{const.}$ condition, the dimensionality of the system decreases. Due to the nature of the transition process, the occupation numbers in the system change along a diagonal of the two-dimensional phase space. In the deterministic picture, the selection behaviour depends on the initial conditions $x_i(t = 0)$ of the system. A certain technology i can only succeed ($x_i = C$ for $t \to \infty$), if the initial condition places the system beyond the point of the separatrix:

$$x_i(0) \; > \; S_i \; ; \quad S_i = \frac{C \, b_j}{b_i + b_j}. \tag{10.44}$$

Thus it is possible that a new technology i, even if it has a greater growth rate, will not succeed because from the outset the number of users (plants) is too small.

$$x_i(t = 0) \; < \; S_i. \tag{10.45}$$

Such a situation has been called once-forever selection or hyperselection. In the case of macromolecular evolution, this feature was used to explain the uniqueness of the genetic code. For technological evolution we have argued elsewhere that hyperselection is an alternative explanation for so-called lock-in phenomena of technologies (Bruckner et al., 1994, 1996). The situation changes again if we look at the stochastic picture. Here the phase space appears as follows (Fig. 10.4):

The two absorber states for the stochastic case are again:

$$N_2 = 0 \, , \; N_1 = N \, , \tag{10.46}$$

$$N_2 = N \, , \; N_1 = 0 \, . \tag{10.47}$$

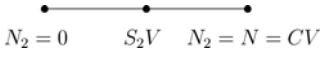


Fig. 10.4 Phase space in the stochastic case

The initial state is $N_2(t) = N_2(t = 0)$:

$$P(N_2; 0) = \delta_{N_2 N_2(t=0)} \; . \tag{10.48}$$

The final state is:

$$P(N_2; t = \infty) = \sigma_1 \; \delta_{0N_2} + \sigma_2 \; \delta_{NN_2} \; . \tag{10.49}$$

After some calculations we obtain for the case $N_2(0) = 1$ (one user of technology 2 occurs at $t = 0$):

$$\sigma_2 = \frac{1}{\left(1 + \dfrac{b_1}{b_2}\right)^{N-1}} \tag{10.50}$$

σ_2 is the probability that the new technology wins the competition process.

In the stochastic case, the separatrix S is penetrated with a certain probability. The once-forever behaviour disappears. The better technology can win the selection process with a certain probability even if it starts with only one user. This probability increases rapidly for small overall numbers of users, as follows from the equation and as can be seen in Fig. 10.5 (Ebeling and Sonntag, 1986). In the stochastic description, the presence of fluctuations is responsible for helping the technology to cross the barrier to entry in the market.

This effect is size-dependent. In smaller systems the probability for survival of new technologies – even in a hyperselection situation – increases. This can also be interpreted in economic terms. Small system size then refers to the size of the market in which the competition between the old and the new technology takes place. If a new technology is protected by certain mechanisms it can win in a market of limited competition. Thus, it can gain a significant number of early adopters and

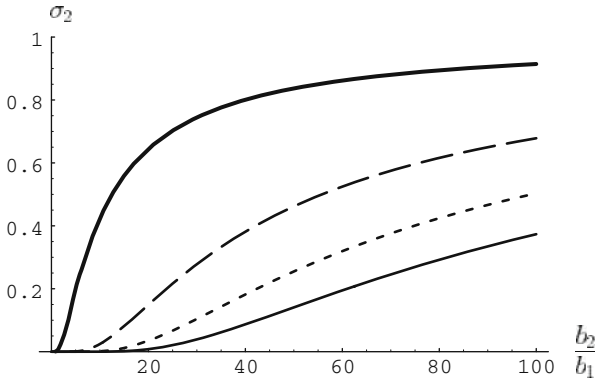


Fig. 10.5 Survival probability of a new technology in systems of different size (*fat line* $N = 10$, *dashed line* $N = 40$, *dotted line* $N = 70$, *fine line* $N = 100$)

move on to a more open, bigger market where it would then have an improved chance of survival. In other words, in small niches, even in a lock-in situation, a new technology can replace an old one. By travelling from one niche to another, or in growing niches (dynamic niches), the new technology would even be able to gain control of the whole market. There is some empirical evidence indicating that such processes take place in technological change (for a discussion see Bruckner et al. (1998) and Aigle et al. (2008)).

10.4.3.3 Technologies with Mixed Growth Rates

In the following, we consider the general case where the growth rates of a certain technology contain both linear as well as non-linear terms. The transition probabilities for this case are:

$$W^+(N_1) = \frac{E_1}{N}\, N_1\, N_2 + \frac{b_1}{NV}\, N_2\, N_1^2, \tag{10.51}$$

$$W^-(N_1) = \frac{E_2}{N}\, N_1\, N_2 + \frac{b_2}{NV}\, N_1\, N_2^2. \tag{10.52}$$

Again, we assume here that $E_{12} = E_{21} = 0$, so we have absorber states. Further, the number of elements (plants) in the system is constant: $N_1 + N_2 = N = \text{const}$. Analogously to the case studies above, the master equation reads:

$$\begin{aligned}\frac{\partial}{\partial t} P(N_1;t) = {} & W^+(N_1-1)\, P(N_1-1;t) + W^-(N_1+1)\, P(N_1+1;t) \\ & - \left[W^+(N_1) + W^-(N_1)\right]\, P(N_1;t).\end{aligned} \tag{10.53}$$

In the same way as described above, we can also derive the deterministic counterpart to the stochastic dynamics (Hartmann-Sonntag et al., 2004).

For the general case, the survival probability of a new technology $\sigma_{N_2(0),N}$ can be calculated as:

$$\sigma_{N_2(0),N} = \frac{1 + \sum_{j=1}^{N_2(0)-1} \prod_{i=1}^{j} \dfrac{E_1 + b_1 \dfrac{N-i}{V}}{E_2 + b_2 \dfrac{i}{V}}}{1 + \sum_{j=1}^{N-1} \prod_{i=1}^{j} \dfrac{E_1 + b_1 \dfrac{N-i}{V}}{E_2 + b_2 \dfrac{i}{V}}}. \tag{10.54}$$

This result was first obtained in Ebeling et al. (1981) and extensively applied to socio-economic problems in Bruckner et al. (1996). As can be seen from the above formula, the survival probability for a new technology σ not only depends on the selection advantage (the relation of the parameters E_i, b_i), but also depends on the size of the system N (the overall user number) and the initial number of users.

10.4.3.4 Summary – General Aspects of the Survival Probability of a New Technology

The stochastic analysis of competition processes of technologies in a market where the number of overall users is restricted shows remarkable differences to the results based on a deterministic description. Let us note here that by restricting ourselves to markets with a fixed number of economic agents, we also restrict ourselves to technological substitution processes. However, using the approach of under-occupied systems, this restriction does not represent a limitation for the innovation process. A variety of future possible innovative types is included in the system all times. The population of agents (fixed in size) can travel unlimited over the field of innovations. Even with this restriction to substitution processes, the stochastic analysis compared with the deterministic one represents an increase in mathematical complexity at the descriptive level. For the multidimensional general case, an analytical solution of the master equation is not available. Some results can be obtained using computer simulations (Schimansky-Geier, 1981; Heinrich and Sonntag, 1981; Bruckner et al., 1990). However, what we have shown so far is how statements about the survival probability in the long run can be derived. Results are available for certain special cases as we presented above. Let us summarize the results for the special case of two-dimensional systems with $N_1 + N_2 = N = \text{const.}$ and with the two absorber states $N_2 = 0$, $N_2 = N$. Absorber states means here that once the states $N_2 = 0$ or $N_2 = N$ have been reached, they cannot be left anymore.

As mentioned above, this case can be treated exactly (Schimansky-Geier, 1981; Ebeling et al., 1981; Ebeling and Feistel, 1982; Ebeling et al., 1986). In accordance with the absorber character, we can assume that the stationary probability which is the target of evolution has a delta character and can be written as:

$$P(N_2; t = \infty) = \sigma\ \delta_{NN_2} + (1 - \sigma)\ \delta_{0N_2}\ . \tag{10.55}$$

Here, σ is the survival probability of the second technology which is supposed to be the new one entering the market. An expression for the survival probability σ can be calculated with the help of one constant of motion for the general case (Ebeling et al., 1981):

$$\sigma_{N_2(0),N} = \frac{1 + \sum_{j=1}^{N_2(0)-1} \prod_{i=1}^{j} \frac{W_i^-}{W_i^+}}{1 + \sum_{j=1}^{N-1} \prod_{i=1}^{j} \frac{W_i^-}{W_i^+}} \quad \text{for} \quad 0 < N_2(0) < N \tag{10.56}$$

and

$$\sigma_{N_2(0),N} = 1 \quad \text{for} \quad N_2(0) = N\ . \tag{10.57}$$

As described in Sect. 10.4.2.1, the transition probabilities W^+ and W^- represent quite a variety of different processes all of which are linked to certain parameters. Further, the initial conditions and the system size N are important for the survival of a new technology (for details see Bruckner et al., 1996).

In particular, we considered the following representations of transition probabilities:

1. For technologies with linear growth rates (linear case)

$$W^+_{N_2} = E_2 \frac{N - N_2}{N} N_2 \; ; W^-_{N_2} = E_1 \frac{N - N_2}{N} N_2. \tag{10.58}$$

2. For technologies with quadratic growth rates (quadratic case)

$$W^+_{N_2} = b_2 \frac{N - N_2}{NV} N_2^2 \; ; W^-_{N_2} = b_1 \frac{(N - N_2)^2}{NV} N_2. \tag{10.59}$$

3. For technologies with certain mixed growth rates (general case)

$$W^+_{N_2} = E_2 \frac{N - N_2}{N} N_2 + b_2 \frac{N - N_2}{NV} N_2^2, \tag{10.60}$$

$$W^-_{N_2} = E_1 \frac{N - N_2}{N} N_2 + b_1 \frac{(N - N_2)^2}{NV} N_2. \tag{10.61}$$

We derived certain expressions for the survival probabilities by solving the absorber problem. These expressions can be treated further, and certain special cases discussed.

1. For the linear case the survival probability has been defined as

$$\sigma_{N_2(0),N} = \frac{1 - \left(\frac{E_1}{E_2}\right)^{N_2(0)}}{1 - \left(\frac{E_1}{E_2}\right)^{N_1)}}. \tag{10.62}$$

We now consider the special case of very large systems (many possible users). In this case, σ can be simplified to:

$$\sigma_{N_2(0),N\to\infty} = 0 \quad \text{for} \quad E_2 < E_1 \tag{10.63}$$

and

$$\sigma_{N_2(0),N\to\infty} = 1 - \left(\frac{E_1}{E_2}\right)^{N_2(0)} \quad \text{for} \quad E_2 > E_1 \, . \tag{10.64}$$

If we then consider a certain initial condition, the case that a technology starts with just one user (plant) $N_2(0) = 1$, we can reduce the expression for the survival probability one step further

$$\sigma_{N_2(0)=1,N\to\infty} = 0 \qquad \text{for} \quad E_2 < E_1 \tag{10.65}$$

$$\sigma_{N_2(0)=1,N\to\infty} = 1 - \frac{E_1}{E_2} \qquad \text{for} \quad E_2 > E_1\,. \tag{10.66}$$

2. For the quadratic case, the survival probability has been defined as

$$\sigma_{N_2(0),N} = \frac{1 + \sum_{j=1}^{N_2(0)-1} \left(\frac{b_1}{b_2}\right)^j \binom{N-1}{j}}{\left(1 + \frac{b_1}{b_2}\right)^{N-1}}. \tag{10.67}$$

This equation can be written down in a more specified manner if we consider a new technology starting with just one user $N_2(0)$:

$$\sigma_{N_2(0),N} = \frac{1}{\left(1 + \frac{b_1}{b_2}\right)^{N-1}}. \tag{10.68}$$

3. For the general case the survival probability has the following form:

$$\sigma_{N_2(0),N} = \frac{1 + \sum_{j=1}^{N_2(0)-1} \prod_{i=1}^{j} \frac{E_1 + b_1 \frac{N-i}{V}}{E_2 + b_2 \frac{i}{V}}}{1 + \sum_{j=1}^{N-1} \prod_{i=1}^{j} \frac{E_1 + b_1 \frac{N-i}{V}}{E_2 + b_2 \frac{i}{V}}}. \tag{10.69}$$

In the special case of a technology starting with just one user $N_2(0) = 1$, we obtain:

$$\sigma_{N_2(0),N} = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{i=1}^{j} \frac{E_1 + b_1 \frac{N-i}{V}}{E_2 + b_2 \frac{i}{V}}}. \tag{10.70}$$

In general, the transition probabilities depend on the system size, the system parameters and the initial conditions. In the stochastic case, we obviously find a

niche effect. In the niche, the sharpness of selection is diminished. In the linear case, good and bad technologies can temporally co-exist. In the quadratic case, the once-forever effect of a competition and selection process is countered by the niche effect. The niche represents a possibility for a better technology to win the competition even when a lock-in effect is present. Locally developed niches may play a constructive role in technological evolution; these can be observed in large complex systems, such as economic systems. The description, with the help of the master equation, includes processes which are important in these small domains (niches). In a niche, the new technology is protected from extinction for a limited time. After winning the competition in this small area, the new technology can "infect" the whole system, and may be established at the end.

10.4.4 Stochastic Analysis of Multiple Decision Processes – The Modelling of Technological Networks

10.4.4.1 Ehrenfest's Urn-Models with Higher Correlations

In Sect. 10.4.1.1, we introduced the urn model of Ehrenfest. The idea of this model was extended to processes which also change the total number of elements in a system. In subsequent sections, we then restricted ourselves to the Ehrenfest approach by considering transition probabilities only, and derived analytical results for the survival probability in the special case of two competing technologies. In this section, we again use the idea of a system with constant size, but extend the dimensionality of the system. Here, we have in mind under-occupied systems where the number of possible different technologies is large. Only the number of plants searching for a new technologies is restricted. Now we will consider the competition between s different technologies. This corresponds to the existence of s urns, which are filled with $N_1, N_2, \ldots, N_s$ spheres. We start a stochastic game where, at random times, spheres are taken out of one urn and put into another. Here, we consider a game with binary decisions, where the transition from one urn j to another urn i is given by the following transition probabilities:

$$W(N_1 \ldots, N_i + 1, \ldots, N_j - 1, \ldots, N_s | N_1, \ldots, N_i, \ldots, N_j, \ldots, N_s) = A_{ij}\, N_j + B_{ij}\, N_i N_j + \sum_k C_{ijk}\, N_i N_j N_k. \tag{10.71}$$

This is a generalization of the transition rates introduced so far in the case of two technologies. In particular, we have in mind the mutual support or hindering of technologies. As indicated in the economic literature, innovations are often the outcome of an innovation network with different actors involved. We can assume that the decision of a certain plant or firm to implement a particular technology depends both on other firms using the same technology as well as on other firms using different but somehow related technologies. The stochastic game we propose here is a hopping process of plants between technologies.

The mean values of the occupation numbers in the thermodynamic limit follows approximately a differential equation. In the sections above, we showed such approximation for the case of two technologies with linear, quadratic and specific non-linear growth rates. The deterministic equations we obtained correspond to Fisher–Eigen–Schuster equations. For the generalization proposed above a set of Lotka-Volterra equation follows in the deterministic limit:

$$\frac{\mathrm{d}}{\mathrm{d}t}x_i = \sum_j \left(A_{ij}\, x_j + B_{ij}\, N x_i x_j + \sum_k C_{ijk}\, N^2 x_i x_j x_k \right) \qquad (10.72)$$

$$\text{with} \quad \frac{\langle N_i \rangle}{N} = x_i \ .$$

The general case can only be handled by computer simulations. The equation above comprises the case of two technologies with non-linear growth rates if we introduce the following correspondence between parameters (Ebeling et al., 2000):

$$A_{21} = 0\,; \;\; B_{12} = \frac{E_1}{N}\,; \; C_{12k} = \frac{b_1}{N}\,\delta_{1k}\,; \; k = \{1;2\}, \qquad (10.73)$$

$$A_{12} = 0\,; \;\; B_{21} = \frac{E_2}{N}\,; \; C_{21k} = \frac{b_2}{N}\,\delta_{2k}, \qquad (10.74)$$

$$W(N_2+1, N_1-1|N_1, N_2) = E_2\,\frac{N_1 N_2}{N} + b_2\,\frac{N_2^2 N_1}{N}. \qquad (10.75)$$

10.4.4.2 Decision Processes and the Dynamics of a Network of Technologies

In this section, we return to the general dynamics of interacting technologies. In Sect. 10.4.2.1, we gave a short economic interpretation of processes like spontaneous generation, self-reproduction, decay and conversion or transition. All of these processes can be interpreted in terms of decision processes by firms or plants related to expansion or shrinking, and the choice of different technologies from a set of technologies available. The case of innovation interpreted as first occupation of a technological possibility unoccupied so far is just one process in a whole set of decision processes made by firms. The substitution case between old and new technologies, which we examined earlier in detail, represents a very specific decision. In the following, we will give a detailed interpretation of different possible decision processes (for further economic interpretations please also consult Bruckner et al. (1996)).

We start again with the set of occupation numbers. N_i is the number of plants using a certain technology i. We consider a network of s different technologies competing with one another $i = 1, \ldots, s$. We will no longer stick to the assumption that the total number of plants in the system will remain constant. This way, we also consider growth and decline processes, not only of certain used technologies in the system, but also of the system itself – the market as a whole.

We can differentiate between changes which occur for just one type of user (where the type or the group that the users belong to is characterized by the technology they use). For instance, the number of users of the technology i may increase or decrease. This is a stepwise process which only changes the occupation number by one:

$$\begin{array}{rcl}(N_i) & \longrightarrow & (N_i + 1) \\ (N_i) & \longrightarrow & (N_i - 1).\end{array} \tag{10.76}$$

Furthermore, changes can occur between different types of users. Mathematically, this is expressed by the simultaneous change of two occupation numbers:

$$\begin{pmatrix} N_i \\ N_j \end{pmatrix} \longrightarrow \begin{pmatrix} N_i - 1 \\ N_j + 1 \end{pmatrix}. \tag{10.77}$$

Plants can take decisions to a certain technology which they have used before; they can develop a new technology, or they can use a technology already established on the market; they can also further develop a technology which they have already used and in this way create a new type of technology. Not only technologies can enter a market; firms can also do so. The creation of a new firm might be connected to an established technology or related to the introduction of a new technology. The problems that technologies face when entering a market are comparable to those faced by firms doing the same thing. Concerning technologies, barriers for entry could hinder the introduction of a specific technology. On the other side, a network of firms might create support for a new firm, or also for a new technology, to enter the market (the system). Coalitions, cooperation and collaboration are examples of synergetic effects in the introduction of new technologies.

We will now formulate some possible changes in the language of transition probabilities:

1. The number of plants using a certain technology i increases:

$$N_i \longrightarrow N_i + 1. \tag{10.78}$$

This change can be the result of different processes which are further differentiable.

- $A_i^0 N_i$ – linear self-reproduction: the number of plants will grow according to the existing number of plants. If these plants belong to one firm, firm growth is modelled. If they belong to different firms, the growth of an industrial sector is considered. The growth rate of this process is linear. Let us remember here that, if only this process takes place in the system, the technology would increase exponentially.
- $A_i^1 N_i^2$ – self-amplification (second order self-reproduction): here, the number of plants already using the technology creates a network effect of higher order which accelerates the growth momentum for this specific technology.

- $B_{ij}\ N_i N_j$ – sponsoring or supporting from other plants: in this case, plants using a different technology j are relevant for the growth of the technology i. One can think of systems of coupled technologies, one supporting the other, or of production chains where one technology relies on others.

For this case, the transition probabilities can be written as:

$$W(N_i + 1, N_j | N_i, N_j) = A_i^0\ N_i + A_i^1\ N_i^2 + B_{ij}\ N_i N_j. \tag{10.79}$$

The different coefficients or parameters describe the strength of a certain effect which acts on the system.

2. Spontaneous formation of a new plant with a certain technology

$$N_i \longrightarrow N_i + 1. \tag{10.80}$$

- Φ_0 – spontaneous formation: here the entry of plants is not connected to the number of those already existing. One can think of a startup. The startup can either begin with an already existing technology or be linked to the development of a new technology. In the latter case, N_i would be zero at the beginning. One would usually assume that the number of new plants created spontaneously is relatively low.

$$W(N_i + 1, N_j | N_i, N_j) = \Phi_0. \tag{10.81}$$

3. Decrease of the number of plants using technology i

$$N_i \longrightarrow N_i - 1. \tag{10.82}$$

- $D_i^0 N_i$ – linear decrease: Given that each technology in the market occupies a certain niche in the market, one can assume that the number of plants using a technology which cannot survive will depend on the size of the population of all plants using that technology.
- $D_i^1 N_i^2$ – non-linear decrease or restricted capacity: this process stands for a network effect related to the number of plants using a certain technology. If too many plants use a particular technology they inhibit each other by occupying market shares. This phenomenon drives plants out of the market. This process is responsible for the existence of a restricted capacity in markets with linearly growing technologies. Without the presence of decreasing terms of higher order, one would be confronted with infinite, exploding markets, which stands in contradiction to empirical observations (Hartmann-Sonntag et al., 2004).

$$W(N_i - 1, N_j | N_i, N_j) = D_i^0\ N_i + D_i^1\ N_i^2. \tag{10.83}$$

4. Origin of a new technology connected to the formation of a new plant (induced innovation)

$$\begin{pmatrix} 0 \\ N_j \end{pmatrix} \longrightarrow \begin{pmatrix} N_i = 1 \\ N_j \end{pmatrix}. \tag{10.84}$$

- $M_{ij}N_j$ – Here we assume that the creation of a startup with a new technology is not a purely spontaneous process, but rather it is related to the number of plants using another technology relevant, in any case, for the new one.

$$W(N_i = 1, N_j | N_i = 0, N_j) = M_{ij}\, N_j. \tag{10.85}$$

5. Change in the use of a technology (conversion, transition)

$$\begin{pmatrix} N_i \\ N_j \end{pmatrix} \longrightarrow \begin{pmatrix} N_i + 1 \\ N_j - 1 \end{pmatrix}. \tag{10.86}$$

- $A_{ij}N_j$ – simple transition from j to i: in this case the decision to take over a new technology is only influenced by the number of plants using a certain technology. One can interpret this process in the following way. If the number of plants using the same technology increases, then the competition between these plants also increases, and thus plants might be motivated to look for another technology in order to increase their chances on the market. In the case where the technology is not yet occupied (has not yet been invented), the transition will also create an innovation for the system. Let us note here that, besides spontaneous generation and induced innovation, this process is important for the exploration of new areas in the technological space.
- $B_{ij}N_iN_j$ – transition from j to i, in addition, is promoted by i: in general one can assume that plants do not act in isolation. Conversely, the information flows between firms and plants about market conditions and technological change are an important part of economic processes. Important in the process we discuss here is the decision to use a certain new technology j (here we use *new* in the sense that the technology is new for the plant); this is related to the number of firms already using the technology. We can further assume that the number of firms using a specific technology can be interpreted as a measure of attractiveness of that particular technology. In this case, the process represents one form in which imitation can be modelled.
- $C_{ijk}N_iN_jN_k$ – transition from j to i, that, in addition, is promoted (sponsored) by j and k: this process represents one possibility to introduce a network effect in the decision process of one firm to use a certain technology. One can imagine that the technologies j and k are related in sense of a production chain or that they complement each other (Hartmann-Sonntag et al., 2004).

$$W(N_i + 1, N_j - 1|N_i, N_j) = A_{ij}\ N_j + B_{ij}\ N_i N_j + \sum_k C_{ijk}\ N_i N_j N_k. \tag{10.87}$$

The model presented is composed in a modular way. Different processes related to the change in the occupation number space were introduced. We have referred to these processes elsewhere as elementary processes. In the presentation above we tried to give examples of processes relevant in the decision behaviour of firms using different technologies. However, alternative definitions and the introduction of further processes are possible within the model framework. The task consists of the definition of processes which can be observed empirically in the economy. The model represents a specific way to operationalize processes of decision making inside firms, the information flows between firms and the interactive pattern between technologies. By relating all of these processes to the occupation number space, a certain reduction of information occurs. On the other hand, the different parameters allow for the possibility of including further economic information. The advantage of the modular structure of the model is that it puts different processes together and places them in an evolutionary framework. Growth of firms, substitution processes, invention and imitation, and startups and firm closings are all part of one model. If we link the model to its deterministic counterpart, the *instrumentarium* of dynamic systems becomes available. In this way, at least, we can hope to gain some insights into the analytic structure of the model and possible stationary states as well as their stability behaviour. Using the stochastic model for simulations we can obtain a lot of statements about a system's behaviour. Interesting investigations can be made by varying the parameters; different kinds of connections can be analysed.

Beyond the economic interpretation chosen for the model, the framework can be also applied in quite different contexts. In any of these new application areas types, elements and the network of connections have to be completely reinterpreted. Some authors applied the model, for example, to biochemical processes (Ebeling and Sonntag, 1986), or to growth, competition and the evolution of scientific specialties (Bruckner et al., 1990), or to the dynamics of values and competences (Scharnhorst, 1999).

10.4.4.3 Further Analysis of the Probability Distribution

Whether we obtain the probability distribution $P(N_1, \ldots, N_s; t)$ analytically or by computer simulation, it is possible to derive the time–behaviour of the moments (mean values, correlation functions), in dependence on the parameters of the system using the generation function (Heinrich and Sonntag, 1981; Ebeling and Sonntag, 1986). For the partial differential equation given in Sect. 10.4.2.1, we can obtain an approximate solution by introducing the transformation (Nicolis and Prigogine, 1977)

$$F = \exp\left[\Psi(\eta_i)N\right]\ ; \quad \eta_i = s_i - 1 \tag{10.88}$$

whereby the function $\psi(\eta_i)$ is expanded in a Taylor series

$$\Psi = \sum_i a_i\ \eta_i + \frac{1}{2!}\sum_{i,j} b_{ij}\ \eta_i\eta_j + \frac{1}{3!}\sum_{i,j,k} c_{ijk}\ \eta_i\eta_j\eta_k + \ldots. \tag{10.89}$$

The coefficients a_i and b_{ij} in terms of the moments of the probability distribution are as follows:

$$a_i = \frac{\langle N_i\rangle}{N}, \tag{10.90}$$

$$b_{ii} = \frac{1}{N}\left[\left\langle N_i^2\right\rangle - \langle N_i\rangle\right], \tag{10.91}$$

$$b_{ij} = \left[\langle N_i N_j\rangle\right]. \tag{10.92}$$

These equations give us information about the deviation from the Poisson distribution (Poisson distribution means that the coefficients $b_{ij} = 0,\ \forall\ i, j$ and all exponential coefficients of higher order are also zero). From the equations we can get a closed set of differential equations for the mean values a_i and the variances b_{ij}. Provided that we can neglect the coefficients of higher than second order, we can achieve an approximate solution of the set of differential equations. Especially for processes which satisfy a multi-Poisson distribution, the coefficients higher than first order are zero in the Taylor expansion. This is the case if we obtain for F or Ψ, respectively, a partial differential equation linear in s_i and at most of first order. If we know the mean values $a_i = \langle N_i\rangle/N$ in this case and start at $t = 0$ with a multi-Poisson distribution, we can obtain an exact solution for the probability distribution. The solution is a multi-Poisson distribution for all times

$$P(N;t) = \prod_i \frac{\langle N_i(t)\rangle^{N_i}}{N_i!}\ \exp\left[-\langle N_i\rangle\right]. \tag{10.93}$$

Now we answer the question: Which processes (transition probabilities from Sect. 10.4.2.1) lead to a multi-Poisson distribution, i.e. which processes have in the generating function at most terms with a first-order derivative with respect to s_i and are linear in s_i? Then, the generating function looks as follows:

$$\dot{F}(s;t) = \sum_j A_i^{(2)}\ (1 - s_i)\ \frac{\partial F}{\partial s_i} + A_{ij}^{(3)}\ (s_i - s_j)\ \frac{\partial F}{\partial s_j}. \tag{10.94}$$

So we can get (Hartmann-Sonntag et al., 2004)

$$\dot{\Psi}(\eta;t) = \sum_i\sum_j \left\{-A_i^{(2)}\ \eta_i a_i + A_{ij}^{(3)}\ (\eta_i - \eta_j)\right\}, \tag{10.95}$$

$$\dot{a}_i = -A_i^{(2)}\ a_i + \sum_{i\neq j} A_{ij}^{(3)}\ a_j - a_i\ \sum_{i\neq j} A_{ij}^{(3)}. \tag{10.96}$$

If we start at $t = 0$ with a multi-Poisson distribution, the probability distribution remains a multi-Poisson distribution for all times

$$P(N;t) = \exp\left[-\langle N(t)\rangle_C\right] \prod_i \frac{\langle N_i(t)\rangle^{N_i}}{N_i!}. \tag{10.97}$$

The extinction probability of the whole component is

$$P(0;t) = \exp\left[-\langle N_C(t)\rangle\right]. \tag{10.98}$$

This extinction probability depends on the overall number of individuals N_C of the component and decreases exponentially with $\langle N_C(t)\rangle$.

A detailed description for the calculation of the moments of the probability distribution $P(N_1, \ldots, N_s; t)$ and the probability distribution was given in the work of Heinrich and Sonntag (1981). The time–behaviour of the moments could be achieved with the help of generating function. Deviations from the Poisson behaviour were investigated. Fluctuations and their influence on the system structure and behaviour could be analysed. The analysis of time-dependent correlation functions could also be interesting for the description of socio-economic systems.

10.5 Summary

This contribution is devoted to the study of innovation processes in socio-economic contexts. In particular, we investigate the influence of stochastic effects on processes of self-organization and evolution. We take a special network perspective. Starting with the newly emergent field of complex networks theory, we develop our own approach of sensitive networks relevant to the description of an innovation.

Innovation is first introduced on a general level as a specific process that changes the composition and dynamic constitution of a system. We use a discrete representation of the system in terms of space of occupation numbers; then innovation can be described as a hopping process between positive cones. Further, we introduce the notion of an under-occupied system. In doing so, we implement a set of possible future paths of developments in our modelling. We relate this abstract notion of an innovation to the discussion of innovation processes in economics.

In the economic literature, innovation has been understood as the outcome of processes running on networks of different actors. We concentrate on firms and technologies in this contribution. We present a network theory of innovation by mapping the dynamic interactions related to the emergence of an innovation to a graph (Hartmann-Sonntag et al., 2004).

Figure 10.1 illustrated our approach. The system is composed of a large set of enumerable types. Each of these types is represented by a node. At a certain point in time, only a small part of these nodes are active. The pattern of interaction between them (including processes of self-influence) determines the dynamic composition

of the system, which is visualized in terms of (active) links between the nodes. We assume that the activated part of the network is embedded in a much larger network of inactive nodes and links. The inactive nodes represent future possibilities in the evolution of the system. An innovation appears when an unoccupied node becomes occupied for the first time. With this initial occupation, the set of links connecting the "new" node with already occupied nodes also becomes activated. It is readily apparent that such an event changes the whole composition of the system.

In another contribution (Hartmann-Sonntag et al., 2004) we have discussed structural (static) properties of such a relational network. In this way, instruments from random graph theory and percolation theory become relevant (see also Sonntag, 1984a,b; Sonntag et al., 1981). In particular, formulae for the probabilities of the occurrence and the distribution of components and cycles in large networks have been obtained by combinatorial considerations. We have shown that connectivity is a central measure in the structure of these kinds of networks. Evolving, dynamic networks reveal different phenomena compared to those derived from random graphs. The results correspond better to findings in real data of empirical networks such as the appearance of power-law distributions of the degree function.

In this contribution, we turn our attention to the discussion of the dynamic properties or the evolution of networks. In particular, we derive descriptions for a network of interacting technologies and interacting firms. We call these networks sensitive because innovation processes described in terms of the removal or appearance of a node might change the dynamic behaviour of the system dramatically. We use a stochastic description of such an evolving network and base this on the theory of birth and death processes. We introduce different forms of transition probabilities for closed systems, as well as for open, growing and declining systems. We define master equations and non-linear differential equations as their deterministic equivalents. Contrary to deterministic models, the stochastic description offers the advantage that, at finite times, new technologies (types) can arise or "die out". Further, the emergence of an innovation can be treated as a singular stepwise event implying the transition from an under-occupied to an occupied state.

In some special cases the master equation can be solved analytically and a stationary survival probability for an innovation can be derived. We show (Hartmann-Sonntag et al., 2004) that the stochastic dynamics differ essentially from the deterministic. Separatrices, which decompose the phase space, cannot be intersected in the deterministic case. In the stochastic case they can be crossed. This way, a "once-forever" selection or hyperselection of technologies that is known in economics as lock-in phenomenon, can be avoided.

The stochastic model of networked dynamic interactions of technologies is further generalized in a multi-dimensional case. Processes representing different non-linearities are discussed in the context of technological change. Although we use technological evolution as the main reference point, we also point to the fact that the modular structure of the model also allows for its application in rather different fields. Applications in fields like biology, population theory and science of science were presented as examples (Feistel and Ebeling, 1989; Bruckner et al., 1990).

References

Aigle, T., Krstacic-Galic, A., Marz, L., Scharnhorst, A. (2008) Busse als Wegbereiter. Zu einem frühen Markt für alternative Antriebe. WZB – discussion paper SP III 2008-102. Wissenschaftszentrum Berlin für Sozialforschung, Berlin

Albert, R., Barabási, A.-L. (2000) Topology of Evolving Networks: Local Events and Universality. Physical Review Letters 85:5234–5237 [arXiv:cond-mat/0005085]

Albert, R., Barabási, A.-L. (2002) Statistical Mechanics of Complex Networks. Reviews of Modern Physics 74:47–97 [arXiv:cond-mat/0106096]

Albert, R., Jeong, H., Barabási, A.-L. (1999) Diameter of the World-Wide Web. Nature (London) 401:130–131 [arXiv:cond-mat/9907038]

Albert, R., Jeong, H., Barabási, A.-L. (2000) Error and Attack Tolerance of Complex Networks. Nature (London) 406:378–382, 409:542 (correction) [arXiv:cond-mat/0008064]

Allen, P. M. (1994) Coherence, Chaos and Evolution in the Social Context. Futures 26:583–597

Barabási, A.-L. (2002) Linked: The New Science of Networks. Perseus Publishing, Cambridge, MA

Barabási, A.-L., Jeong, H., Néda, Z., Ravasz, E., Schubert, A., Vicsek, T. (2002) Evolution of the Social Network of Scientific Collaborations. Physica A 311:590–614 [arXiv:cond-mat/0104162]

Bartholomay, A. F. (1958a) On the Linear Birth and Death Processes of Biology as Markoff Chains. The Bulletin of Mathematical Biophysics 20:97–118

Bartholomay, A. F. (1958b) Stochastic Models for Chemical Reactions: I. Theory of the Unimolecular Reaction Process. The Bulletin of Mathematical Biophysics 20:175–190

Bartholomay, A. F. (1959) Stochastic Models for Chemical Reactions: II. The Unimolecular Rate Constant. The Bulletin of Mathematical Biophysics 21:363–373

Bianconi, G., Barabási, A.-L. (2001a) Competition and Multiscaling in Evolving Networks. Europhysics Letters 54:436–442 [arXiv:cond-mat/0011029]

Bianconi, G., Barabási, A.-L. (2001b) Bose-Einstein Condensation in Complex Networks. Physical Review Letters 86:5632–5635 [arXiv:cond-mat/0011224]

Bornholdt, S., Schuster, H. G. (Eds.) (2003) Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH Verlag, Weinheim

Bruckner, E., Ebeling, W., Jiménez-Montaño, M. A., Scharnhorst, A. (1994) Hyperselection and Innovation Described by a Stochastic Model of Technological Evolution. In: Leydesdorff, L., van den Besselar, P. (Eds.) Evolutionary Economics and Chaos Theory: New Directions in Technology Studies. Pinter Publishers, London, pp. 70–90

Bruckner, E., Ebeling, W., Jiménez-Montaño, M. A., Scharnhorst, A. (1996) Nonlinear Stochastic Effects of Substitution: An Evolutionary Approach. Journal of Evolutionary Economics 6:1–30

Bruckner, E., Ebeling, W., Scharnhorst, A. (1989) Stochastic Dynamics of Instabilities in Evolutionary Systems. System Dynamics Review 5:176–191

Bruckner, E., Ebeling, W., Scharnhorst, A. (1990) The Application of Evolution Models in Scientometrics. Scientometrics 18:21–41

Bruckner, E., Ebeling, W., Scharnhorst, A. (1998) Technologischer Wandel and Innovation – Stochastische Modelle für innovative Veränderungen in der Ökonomie [Technological Change and Innovation – Stochastic Models for Innovative Changes in the Economy]. In: Schweitzer, F., Silverberg, G. (Eds.) Evolution und Selbstorganisation in der Ökonomie. Jahrbuch der Selbstorganisation, Vol. 9. Duncker & Humblot, Berlin, pp. 361–382 [in German]

Buchanan, M. (2002) Nexus: Small Worlds and the Groundbreaking Science of Networks. W. W. Norton & Company, New York, London

Dorogovtsev, S. N., Mendes, J. F. F., Samukhin A. N. (2003) Principles of Statistical Mechanics of Uncorrelated Random Networks. Nuclear Physics B 666:396–416 [arXiv:cond-mat/0204111]

Drossel, B., McKane, A. J. (2003) Modelling Food Webs. In: Bornholdt, S., Schuster, H. G. (Eds.) Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH Verlag, Weinheim, pp. 218–247 [arXiv:nlin.AO/0202034]

Ebeling, W., Engel, A., Feistel, R. (1990a) Physik der Evolutionsprozesse. Akademie-Verlag, Berlin [in German]

Ebeling, W., Engel, H., Herzel, H. (1990b) Selbstorganisation in der Zeit. Wissenschaftliche Taschenbücher, Vol. 309. Akademie-Verlag, Berlin [in German]

Ebeling, W., Feistel, R. (1982, 2nd. ed. 1986) Physik der Selbstorganisation und Evolution. Akademie-Verlag, Berlin [in German]

Ebeling, W., Feistel, R. (1994) Chaos und Kosmos: Prinzipien der Evolution. Spektrum-Verlag, Heidelberg, Berlin, Oxford [in German]

Ebeling, W., Feistel, R., Hartmann-Sonntag, I., Schimansky-Geier, L., Scharnhorst, A. (2006) New species in evolving networks – stochastic theory of sensitive networks and applications on the metaphorical level:BioSystems 85(1):65–71.

Ebeling, W., Freund, J., Schweitzer, F. (1998) Komplexe Strukturen: Entropie und Information. B. G. Teubner-Verlag, Stuttgart, Leipzig [in German]

Ebeling, W., Karmeshu, Scharnhorst, A. (2001) Dynamics of Economic and Technological Search Processes in Complex Adaptive Landscapes. Advances in Complex Systems 4(1):71–88. This article is part of the special issue: Helbing, D., Schweitzer, F. (Guest Eds.) Complex Dynamics in Economics. Advances in Complex Systems 4(1):1–176

Ebeling, W., Molgedey, L., Reimann, A. (2000) Stochastic Urn Models of Innovation and Search Dynamics. Physica A 287:599–612

Ebeling, W., Scharnhorst, A. (1986) Self-organization Models for Field Mobility of Physicists. Czechoslovak Journal of Physics B 36:43–46

Ebeling, W., Scharnhorst, A. (2000) Evolutionary models of innovation dynamics. In: Helbing, D., Herrmann, H. J., Schreckenberg, M., Wolf, D. E. (Eds.) Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics. Springer, Berlin, Heidelberg, New York, pp. 43–56

Ebeling, W., Scharnhorst, A., Jiménez-Montaño, M. A., Karmeshu (1999) Evolutions- und Innovationsdynamik als Suchprozeß in komplexen adaptiven Landschaften [Evolution and Innovation Dynamics as Search Processes in Complex Adaptive Landscapes]. In: Mainzer, K. (Ed.) Komplexe Systeme und Nichlineare Dynamik in Natur und Gesellschaft: Komplexitätsforschung in Deutschland auf dem Weg ins nächste Jahrhundert. Springer-Verlag, Berlin, Heidelberg, New York, pp. 446–473 [in German]

Ebeling, W., Sonntag, I. (1986) A stochastic description of evolutionary processes in underoccupied systems. BioSystems 19:91–100

Ebeling, W., Feistel, R., Hartmann-Sonntag, I., Schimansky-Geier, L., Scharnhorst, A. (1986) New species in evolutionary networks – stochastic theory and applications on the metaphoric level. BioSystems 85:65–71

Ebeling, W., Sonntag, I., Schimansky-Geier, L. (1981) On the evolution of biological macromolecules II: Catalytic networks. Studia Biophysica 84:87–88 and microfiche 1/37.54

Ehrenfest, P., Ehrenfest, T. (1907) Über zwei bekannte Einwände gegen das Boltzmannsche H-Theorem [On two known objections against the H-theorem of Boltzmann]. Physikalische Zeitschrift 8:311–314 [in German]. Reprinted in: Klein, M. J. (Ed.) with an introduction by Casimir, H. B. G. (1959) Paul Ehrenfest: Collected Scientific Papers. North-Holland Publishing, Amsterdam, item 15, pp. 146–149

Eigen, M. (1971) The selforganization of matter and the evolution of biological macromolecules. Die Naturwissenschaften 58:465–523

Eigen, M., Schuster, P. (1977) The hypercycle: A principle of natural self-organization. Part A: Emergence of the hypercycle. Die Naturwissenschaften 64:541–565

Eigen, M., Schuster, P. (1978) The hypercycle: A principle of natural self-organization. Part C: The realistic hypercycle. Die Naturwissenschaften 65:341–369

Faloutsos, M., Faloutsos, P., Faloutsos, C. (1999) On power-law relationships of the internet topology. Computer Communications Review 29:251–262

Feistel, R., Ebeling, W. (1978) Deterministic and stochastic theory of sustained oscillations in autocatalytic reaction systems. Physica A 93:114–137

Feistel, R., Ebeling, W. (1989) Evolution of Complex Systems. Physikalische Monographien. VEB Deutscher Verlag der Wissenschaften, Berlin; Evolution of Complex Systems: Selforganisation,

Entropy and Development. Mathematics and Its Applications, East European Series, Vol. 30. Kluwer Academic Publishers, Dordrecht, Boston, London
Fell, D. (1997) Understanding the Control of Metabolism. Frontiers in Metabolism, Vol. 2. Portland Press, London
Fell, D., Wagner, A. (2000) The small world of metabolism. Nature Biotechnology 18:1121–1122
Feller, W. (1951) Two singular diffusion problems. Annals of Mathematics, 2nd Series 54:173–182
Frenken, K. (2000) A Complexity Approach to Innovation Networks: The Case of the Aircraft Industry (1909-1997). Research Policy – A Journal Devoted to Research Policy, Research Management and Planning 29:257–272
Haken, H. (1978) Synergetics: An Introduction, 2nd. enl. ed. (1983) 3rd. rev. enl. ed.. Springer Series in Synergetics, Vol. 1. Springer-Verlag, Berlin, Heidelberg, New York
Haken, H. (1996) Synergetik und Sozialwissenschaften [Synergetics and Social Sciences]. Ethik und Sozialwissenschaften: EuS – Streitforum für Erwägungskultur 7:587–594 [in German]
Harary, F., Norman, R. Z., Cartwright, D. (1965) Structural Models: An Introduction to the Theory of Directed Graphs. John Wiley & Sons, New York, London, Sydney
Hartmann-Sonntag, I., Scharnhorst, A., Ebeling, W. (2004) Modelling Selforganization and Innovation Processes in Networks. Physics E-Print Archive: arXiv:cond-mat/0406425 (http://arxiv.org/abs/cond-mat/0406425)
Haythornthwaite, C. (1996) Social Network Analysis: An Approach and Technique for the Study of Information Exchange. Library and Information Science Research 18:323–342
Heinrich, R., Sonntag, I. (1981) Analysis of the Selection Equations for a Multivariable Population Model: Deterministic and Stochastic Solutions and Discussion of the Approach for Populations of Self-Reproducing Biochemical Networks. Journal of Theoretical Biology 93:325–361
Huberman, B. A. (2001) The Laws of the Web: Patterns in the Ecology of Information. The MIT Press, Cambridge, MA
Hunter, L., Elias, M. J. (2000) Interracial Friendships, Multicultural Sensitivity, and Social Competence: How Are They Related? Journal of Applied Developmental Psychology 20:551–573
Jain, S., Krishna, S. (2001) A Model for the Emergence of Cooperation, Interdependence and Structure in Evolving Networks. Proceedings of the National Academy of Sciences of the United States of America 98:543-547 [arXiv:nlin.AO/0005039]
Jeong, H., Mason S. P. (2001) Lethality and Centrality in Protein Networks. Nature (London) 411:41–42 [arXiv:cond-mat/0105306]
Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A.-L. (2000) The Large-Scale Organization of Metabolic Networks. Nature (London) 407:651–654 [arXiv:cond-mat/0010278]
Jiménez-Montaño, M. A., Ebeling, W. (1980) A Stochastic Evolutionary Model of Technological Change. Collective Phenomena 3:107–114
Kaniovski, Y. M. (2000) A Comparison of Diffusion Approximations and Actual Limits in Births and Death Processes of Noisy Evolution. Journal of Evolutionary Economics 10:545–555
Katz, J. S. (1999) The Self-Similar Science System. Research Policy – A Journal Devoted to Research Policy, Research Management and Planning 28:501–517
Kirman, A. (2003) Economic Networks. In: Bornholdt, S., Schuster, H. G. (Eds.) Handbook of Graphs and Networks: From the Genome to the Internet. Wiley-VCH Verlag, Weinheim, pp. 273–294
Klimontovich, Yu. L. (1995) Statistical Theory of Open Systems. Vol. 1: A Unified Approach to Kinetic Description of Processes in Active Systems. Fundamental Theories of Physics, Vol. 67. Kluwer Academic Publishers, Dordrecht, Boston, London
Kohlrausch, K. W. F., Schrödinger, E. (1926) Das Ehrenfestsche Modell der H-Kurve [Ehrenfest's model of the H-curve]. Physikalische Zeitschrift 27:306-313 [in German]. Reprinted in: Schrödinger, E. (1984) Collected Papers/Gesammelte Abhandlungen, Volume 1/Band 1, Contributions to Statistical Mechanics/Beiträge zur statistischen Mechanik. Verlag der Österreichischen Akademie der Wissenschaften, Friedr. Vieweg & Sohn, Braunschweig, Wiesbaden, pp. 349–357
Kowol, U., Küppers, G. (2005). Innovation Networks: A New Approach to Innovation Dynamics. In: van Geenhuizen, M., Gibson, D. V., Heitor M. V. (Eds.) Regional Development and

Conditions for Innovation in the Network Society. International Series on Technology Policy and Innovation. Purdue University Press, Indiana, pp. 61–86

Mainzer, K. (Ed.) (1999) Komplexe Systeme und Nichlineare Dynamik in Natur und Gesellschaft: Komplexitätsforschung in Deutschland auf dem Weg ins nächste Jahrhundert. Springer-Verlag, Berlin, Heidelberg, New York [in German]

Maurer, S. M., Huberman, B. A. (2000) Competitive Dynamics of Web Sites. Journal of Economic Dynamics and Control 27:2195–2206 [arXiv:nlin.CD/0003041]. This article is part of the special issue: Juillard, M., Marcet, A. (Guest Eds.) Computing in Economics and Finance. Proceedings of the Sixth International Conference of the Society for Computational Economics, Barcelona, Spain, July 6-8, 2000. Journal of Economic Dynamics and Control 27(11-12):1939–2265

Montoya, J. M., Solé, R. V. (2002) Small World Patterns in Food Webs. Journal of Theoretical Biology 214:405–412 [arXiv:cond-mat/0011195]

Nachtigall, C. (1998) Selbstorganisation und Gewalt. Internationale Hochschulschriften, Vol. 259. Waxmann Verlag, Münster, New York, München, Berlin [in German]

Nagurney, A. (2003) Report: Some Recent Developments in Network Economics. Networks 41:68–72

Newman, M. E. J. (2000) Models of the Small World. Journal of Statistical Physics 101:819–841 [arXiv:cond-mat/0001118]

Nicolis, G., Prigogine, I. (1977) Self-Organization in Non-Equilibrium Systems: From Dissipative Structures to Order Through Fluctuations. John Wiley & Sons, New York

Prigogine, I. (1955) Introduction to Thermodynamics of Irreversible Processes. American Lecture Series, Vol. 185. Charles C. Thomas Publisher, Springfield, IL. (1962) 2nd. rev. ed., (1967) 3rd. ed. Interscience Publishers, New York

Prigogine, I., Sanglier, M. (Eds.) (1987) Laws of Nature and Human Conduct: Specificities and Unifying Themes. Proceedings of the Eighth. International Discoveries Symposium, Brussels, Belgium, October 7–9, 1985. G.O.R.D.E.S./Task Force of Research, Information and Study of Science, Brussels

Pyka, A. Küppers, G. (Eds.) (2002) Innovation Networks: Theory and Practice. New Horizons in the Economics of Innovation Series. Edward Elgar Publishing, Cheltenham, Northampton, MA

Saviotti, P. P. (1996) Technological Evolution, Variety and the Economy. Edward Elgar Publishing, Cheltenham, Northampton, MA

Saviotti, P. P. (2001) Networks, National Innovation Systems and Self-Organization. In: Fischer, M. M., Fröhlich, J. (Eds.) Knowledge, Complexity and Innovation Systems. Advances in Spatial Science. Springer-Verlag, Berlin, Heidelberg, New York, pp. 21–45

Saviotti, P. P., Mani, G. S. (1995) Competition, Variety and Technological Evolution: A Replicator Dynamics Model. Journal of Evolutionary Economics 5:369–392

Saviotti, P. P., Nooteboom B. (Eds.) (2000) Technology and Knowledge: From the Firm to Innovation Systems. Edward Elgar Publishing, Cheltenham, Northampton, MA

Scharnhorst, A. (1999) Modelle von Wertedynamik und Kompetenzentwicklung [Models for the Dynamics of Values and Competence Development]. In: Erpenbeck, J., Heyse, J. (Eds.) Kompetenzbiographie – Kompetenzmileu – Kompetenztransfer. QUEM-Report – Schriften zur beruflichen Weiterbildung, Issue 62. Arbeitsgemeinschaft Betriebliche Weiterbildungsforschung e.V. (ABWF), Berlin, pp. 106-140 [in German]

Scharnhorst, A. (2003) Complex Networks and the Web: Insights from Nonlinear Physics. Journal of Computer-Mediated Communication 8(4): http://www.ascusc.org/jcmc/vol8/issue4/scharnhorst.html. This article is part of the special issue: Beaulieu, A., Park, H. W. (Guest Eds.) (2003) Internet Networks: The Form and the Feel. Journal of Computer-Mediated Communication 8(4): http://www.ascusc.org/jcmc/vol8/issue4

Schimansky-Geier, L. (1981) Stochastische Theorie der Nichtgleichgewichtsphasenübergänge in einkomponentigen bistabilen chemischen Reaktionssystemen. Dissertation A (Ph.D.) Thesis, Humboldt University Berlin, Berlin [in German]

Schweitzer, F. (1997). Self-Organization of Complex Structures: From Individual to Collective Dynamics. Gordon and Breach Science Publishers, Amsterdam

Scott, J. (2000) Social Network Analysis: A Handbook, 2nd. ed. SAGE Publications, London, Thousands Oaks, CA
Solé, R. V., Montoya, J. M. (2001) Complexity and Fragility in Ecological Networks. Proceedings of the Royal Society of London, Series B – Biological Sciences 268:2039–2045 [arXiv:cond-mat/0011196]
Sonntag, I. (1984a) Random Networks of Catalytic Biochemical Reactions. Biometrical Journal 26:799–807
Sonntag, I. (1984b) Application of the Percolation Theory to Random Networks of Biochemical Reactions. Biometrical Journal 26:809–813
Sonntag, I., Feistel, R., Ebeling, W. (1981) Random Networks of Catalytic Biochemical Reactions. Biometrical Journal 23:501–515
Strogatz, S. H. (2001) Exploring Complex Networks. Nature (London) 410:268–276
Temkin O. N., Zeigarnik, A. V., Bonchev, D. (1996) Chemical Reaction Networks: A Graph-Theoretical Approach. CRC Press, Boca Raton, FL
Turán, P. (Ed.) (1976) Rényi Alfréd Válogatott Munkái/Selected Papers of Alfréd Rényi, Vol. 2, 1956–1961. Akadémiai Kiadó, Budapest
Vazquez, A. (2001). Statistics of Citation Networks. Physics E-Print Archive: arXiv:cond-mat/0105031 (http://arxiv.org/abs/cond-mat/0105031)
Vogelstein, B., Lane, D., Levine, A. J. (2000) Surfing the p53 Network. Nature (London) 408: 307–310
Wagner, A., Fell, D. (2000) The Small World Inside Large Metabolic Networks. Proceedings of the Royal Society of London, Series B – Biological Sciences 268:1803–1810
Watts, D. J. (1999) Small Worlds: The Dynamics of Networks between Order and Randomness. Princeton Studies in Complexity. Princeton University Press, Princeton, NJ
Weidlich, W. (2000) Sociodynamics: A Systematic Approach to Mathematical Modelling in the Social Sciences. Harwood Academic Publishers, Amsterdam
Ziman, J. (Ed.) (2000) Technological Innovation as an Evolutionary Process. Cambridge University Press, Cambridge, New York

Index