

Network Formation and Dynamics under Economic Constraints

Dissertation

(Cumulative Thesis)

for the award of the degree
“Doctor rerum naturalium”
of the Georg-August-Universität Göttingen
within the doctoral program
Physics of Biological and Complex Systems (PBCS)
of the Georg-August University School of Science (GAUSS)

submitted by

Malte Schröder

from Oldenburg

Göttingen, 2018

Thesis advisory committee

Prof. Dr. Marc Timme

Institute for Nonlinear Dynamics, Georg-August-Universität Göttingen, Germany
Network Dynamics, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
Chair for Network Dynamics, Center for Advancing Electronics Dresden (cfaed) and
Institute for Theoretical Physics, Technical University of Dresden, Dresden, Germany

Prof. Dr. Reiner Kree

Institute for Theoretical Physics, Georg-August-Universität Göttingen, Germany

Jun.-Prof. Dr. Dirk Witthaut

Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), Forschungszentrum Jülich, Jülich, Germany
Institute for Theoretical Physics, University of Cologne, Köln, Germany

Reviewers and members of the examination board

Prof. Dr. Marc Timme

Institute for Nonlinear Dynamics, Georg-August-Universität Göttingen, Germany
Network Dynamics, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
Chair for Network Dynamics, Center for Advancing Electronics Dresden (cfaed) and
Institute for Theoretical Physics, Technical University of Dresden, Dresden, Germany

Prof. Dr. Reiner Kree

Institute for Theoretical Physics, Georg-August-Universität Göttingen, Germany

Further members of the examination board

Jun.-Prof. Dr. Dirk Witthaut

Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), Forschungszentrum Jülich, Jülich, Germany
Institute for Theoretical Physics, University of Cologne, Köln, Germany

Prof. Dr. Ulrich Parlitz

Biomedical Physics, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
Institute for Nonlinear Dynamics, Georg-August-Universität Göttingen, Germany

Prof. Dr. Theo Geisel

Department of Nonlinear Dynamics, Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany

Prof. Dr. Annette Zippelius

Institute for Theoretical Physics, Georg-August-Universität Göttingen, Germany

Date of oral examination:

27th of February, 2018

Affidavit

I hereby confirm that I have written this thesis independently and with no other sources and aids than quoted.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder

List of publications

- [1] **Schröder, M.**, Rahbari, S. H. E., and Nagler, J. (2013),
Crackling noise in fractional percolation,
Nature Communications **4**, 2222
- [2] Chen, W., **Schröder, M.**, D'Souza, R. M., Sornette, D., and Nagler, J. (2014),
Microtransition cascades to percolation,
Physical Review Letters **112**, 155701
- [3]^{†*} **Schröder, M.**, Mannattil, M., Dutta, D., Chakraborty, S., and
Timme, M. (2015),
Transient uncoupling induces synchronization,
Physical Review Letters **115**, 054101
(see Chapter 3)
- [4][†] **Schröder, M.**, Chen, W., and Nagler, J. (2016),
Discrete scale invariance in supercritical percolation,
New Journal of Physics **18**, 013042
- [5][†] Tandon, A., **Schröder, M.**, Mannattil, M., Timme, M., and
Chakraborty, S. (2016),
Synchronizing noisy nonidentical oscillators by transient uncoupling,
Chaos **26**, 094817
- [6]^{†*} **Schröder, M.**, Chakraborty, S., Witthaut, D., Nagler, J., and Timme, M. (2016),
Interaction control to synchronize non-synchronizable networks,
Scientific Reports **6**, 37142
(see Chapter 4)
- [7]^{†*} **Schröder, M.**, Timme, M., and Witthaut, D. (2017),
A universal order parameter for synchrony in networks of limit cycle oscillators,
Chaos **27**, 073119
(see Chapter 5)
- [8]^{†*} **Schröder, M.**, Araújo, N. A. M., and Nagler, J. (2017),
Controlling percolation with limited resources,
Physical Review E **95**, 062302
(see Chapter 6)
- [9]^{†*} **Schröder, M.**, Nagler, J., Timme, M., and Witthaut, D. (2017),
Hysteretic percolation from locally optimal decisions,
submitted (arXiv:1709.07257)
(see Chapter 7)

- [10][†] Wolters, J., Lünsmann, B., Zhang, X., **Schröder, M.**, and Timme, M. (2017),
Quantifying transient spreading dynamics on networks,
submitted (arXiv:1710.09687)
- [11][†] Molkenthin, N., **Schröder, M.**, and Timme, M. (2017),
Discontinuous transitions in temporal networks induced by asymmetric link ad-
hesion,
submitted

[†] manuscript finished during graduate studies

^{*} manuscript included in the main part of this thesis

Contents

1	Introduction	1
2	Fundamentals	11
2.1	Graphs and networks	11
2.1.1	Representation of graphs	12
2.1.2	Describing the network structure	13
2.1.3	Network formation	18
2.1.4	Percolation	24
2.2	Dynamical systems	33
2.3	Networked dynamical systems	38
2.3.1	Synchronization	39
2.3.2	Chaos synchronization	41
3	Transient uncoupling induces synchronization	45
3.1	Main article	46
3.2	Supplementary material	51
4	Interaction Control to Synchronize Non-synchronizable Networks	63
4.1	Main article	64
4.2	Supplementary material	71
5	A universal order parameter for synchrony in networks of limit cycle oscillators	81
6	Controlling percolation with limited resources	91
6.1	Main article	92
6.2	Supplementary material	102
7	Hysteretic percolation from locally optimal decisions	115
7.1	Main article	116
7.2	Supplementary material	122
8	Discussion	145
	Bibliography	153
	Acknowledgements	167

Chapter 1

Introduction

In the last decades our daily lives have become dominated by increasingly interconnected and interdependent systems [64, 128]. The internet and the large number of online shopping, social, or video sharing platforms are probably the most obvious examples, allowing almost instant communication over large distances with anyone in the world at any time. But also the electricity we use to power our computers and access the internet is transmitted from a distant generator via a power grid spanning more than half of Europe and connecting thousands of power plants, growing in complexity with every wind farm or solar power plant added to the grid [34, 185]. Many products we buy will have seen more of the world than we have, they may have been assembled in China with resources from Africa, running software from America with the technical support situated in India, all together requiring a closely coordinated production and supply chain around the world. Similarly, a complex net of dependencies defines today's financial markets, where individual banks loan money to a large number of different clients and take out loans from other banks themselves [65, 73, 169]. Finally, traveling for work or to visit a friend often relies on long-distance flights, connecting trains, buses, cars, and other, more innovative forms of public transport [33, 69].

While these systems make our lives easier every day, they come with their own, new, and sometimes unexpected problems [85, 86]. If any part of one of these systems fails, such as one transmission line in the power grid, one flight or train is late or canceled, or one bank is unable to pay back its loans, the repercussions of this small disruption in normal operation may be enough to bring down the whole system. A broken transmission line in the power grid leads to redistribution of the power flow, potentially causing other lines to be overloaded and fail, eventually resulting in a large scale blackout [195]. One late train results in tracks being blocked and other trains waiting for connections, causing more and longer delays, even in places very far from the cause of the initial delay [69]. In a similar way, an initial outbreak of a disease will spread through the whole world via the global air traffic network, reaching geographically distant countries in a few days [32, 90]. The situation becomes even more complicated with increasing interdependence among these systems. If the power grid fails, the communication network loses power and may fail as well. If communication fails, power plant operators cannot coordinate their reactions, potentially resulting in an even larger blackout of the power grid and a cascade of failures in both systems [35, 133].

Increasingly faster changes and innovations add new layers of complexity and interdependencies to the already complex connectivity and interactions in all of these

systems, making it difficult to predict or control their behavior and to prevent failures [85, 86]. Due to the wide range of scale and function of these systems, detailed case studies on single examples like specific online platforms or parts of the power grid will often only give results applicable to the specific system. To gain deeper insight into how these systems work and understand the sometimes unintuitive phenomena, a common framework and fundamental theory is required.

Complex networks

All examples mentioned above can be described as complex networks with nodes, such as power plants, airports, or people, connected by links, like transmission lines, flights, or friendships. Historically first used by Euler to solve the well-known Königsberg bridge problem [193] the study of networks was applied in social sciences to describe acquaintances or different types of relations in social groups already in the early 20th century [122]. In the last few decades network science has developed into a field of its own, combining approaches from social sciences, graph theory and theoretical physics and developing new tools to study and understand networked systems. Network science has been used to describe a variety of different systems ranging from online communities over technical system like power grids to biological systems like neural networks in the brain or the interactions between different genes and proteins [5, 64, 127, 128].

Initially, the analysis focused on understanding the complex topology, that means on how the individual elements are connected and if there are common patterns. One of the most famous examples for social networks is Stanley Milgram’s small-world experiment in 1967, where participants were asked to send a letter to a given target by only forwarding it to people they know by their first name [186, 187]. Even though the initial participants had little information about the final target, the letters arrived on average in about 6 steps, leading to the popularized “Six Degrees of Separation”.¹ Derivatives of this concept later appeared in other communities such as the “Six Degrees of Kevin Bacon” connecting all actors and the “Erdős Number” in the scientific community of mathematicians and network scientists [81, 82, 172].

The analysis of the structure of social networks also showed a high tendency of triangles, meaning that two of my friends are very likely also friends [9, 64, 191]. Another example of patterns in these seemingly random networks was found in biological examples of gene interaction and neural networks, where most of the connections make up a few very common “motifs”, small sub-networks that are repeated very often, like the triangles in social networks [119]. Many of these motifs can be linked to a specific function, for example feed-back loops to regulate the expression of specific genes in different situations [7, 117].

¹Of the 96 letters only 18 arrived at the target. Even though this number may seem low, it is surprising in itself that people are able to *find* such short paths in social networks.

The earliest theoretical model to describe the structure of such networks was studied by the mathematicians Erdős and Rényi already in the 1950s, assuming completely random connections, and is still used in many basic models to date [66, 67]. With all the patterns found in real networks, however, most networks are anything but random and people investigated more detailed mechanisms to explain the observed structures [5, 127, 128]. Starting with Watts and Strogatz in 1998 giving a possible explanation for the small world effect in social networks observed by Milgram [191], scientists from various disciplines have proposed many models explaining how networks with specific patterns might form, grow, or restructure [14, 93, 153, 189]. Borrowing tools from statistical physics and percolation theory it is often even possible to analytically describe the resulting ensembles of networks. This interdisciplinary approach revealed a broad range of fundamental implications of the network structure for the function, efficiency, and resilience of networks across different applications. For example, many networks like the internet are naturally surprisingly resistant to failures of random links or nodes. That means most of the nodes remain connected even when a large fraction of random nodes or links fail. At the same time, however, these networks are very susceptible to targeted destruction of the most important nodes and links [6, 52, 53, 60].

Dynamical systems

Dynamical systems theory is a cornerstone of the mathematical description of any (not just physical) process. Historically, the first example dates back to Newton's description of the motion of the planets [130]. Now ubiquitous, the description of the dynamics with differential equations or iterated maps goes far beyond simple mechanical systems and is used to also describe electronics, chemical reactions, the behavior of neurons and even to model the dynamics of populations of animals, the process of opinion formation, or the prevalence of a disease in a population [17, 57]. Even though many of the resulting equations cannot be solved exactly, that means there is no explicit equation that describes the state of the system at any given time, dynamical systems theory provides many insightful qualitative results. Especially, dynamical systems theory describes the long term behavior of a system, the stability of fixed points or periodic orbits, and the response to small changes to these states [178]. In terms of the examples above this may correspond to answering questions like: Will a population of animals recover when a small fraction of the population dies? If not, which other species becomes dominant? How easily can an opinion be influenced and what is the easiest way to do this? How quickly will a disease cause an outbreak?

More complex phenomena than fixed points and closed, periodic orbits were discovered in higher dimensional nonlinear systems. One of the best known examples here is the concept of chaos, that means aperiodic, never repeating behavior where two identical systems starting from slightly different states will behave completely differently after a short time [178]. An often depicted example is Lorenz' simplified model resulting from the study of atmospheric convection [116]. The sensitive dependence on initial conditions of chaotic systems means that they are inherently unpredictable,

no matter how well we understand the system as even a very small error in our measurement its state will quickly grow. However, chaotic systems are not without order. In fact, one of the most famous examples of universality comes from Feigenbaum's discovery that the transition to chaos behaves similarly in different systems [178]. General results like these provide a deeper understanding for more complex systems that cannot be easily analyzed in such detail.

Networked dynamical systems

Clearly, most dynamical systems are far from isolated. Neurons in the brain interact via millions of connections, opinion formation depends on the specific interactions of the individuals and electronic devices regularly send and receive signals [12, 17, 21, 128, 180]. However, the connections in these systems are not regular. In such systems the individual units are dynamical systems that are coupled via a network of interactions between these units. Together these networked dynamical systems describe a broad range of dynamic processes from the activity of neural networks over the dynamics of power grids to the spreading of disease via traffic and transportation networks [32, 68, 96, 152]. Instead of simply analyzing the structure of the networks, networked dynamical systems allow more detailed predictions about the processes over time, for example describing the robustness of a system from a dynamical instead of simply from a topological point of view.

In particular, many phenomena depend on the interplay between the network structure and the dynamics of the units. One example is the Braess' Paradox, originally found in a simple traffic model [64]. Adding a new street to a network intuitively allows faster travel. However, sometimes it may lead to longer travel times instead. A similar effect appears when considering the structure of the power grid and the effect of additional transmission lines on its dynamic stability [194]. Another typical example is the emergence of collective dynamics, most prominently of synchronization, meaning identical behavior of all units [178]. Synchronization is important for example for communication, the stable operation of power grids, and consensus in models of opinion formation [12, 77]. Due to this general importance and broad applicability, it is a widely studied problem how to achieve, enable, or stabilize synchronization. The topology of the interaction network between the individual units plays a large role in determining the stability of the synchronized state [139, 165, 184]. The combination of tools from network science and dynamical systems theory allows a deep understanding of these systems and of the effects of the interaction network structure on the collective dynamics of the units [12, 21, 128, 180].

Analyzing such phenomena in basic models makes it possible to identify common patterns and interactions with implications for more complicated systems with a similar structure. For example, how were the people in Milgram's small world experiment able to deliver the letter to the target so quickly? Given only the name, city and occupation of the target, even though they only know their own friends, how could they successfully navigate the social network of the millions of people living in the

US? Solutions for this problem not only uncover aspects of social networks and our communication, but also help technical applications such as efficient package routing in the internet [4, 107, 192]. Similarly, notions that identify central individuals in social networks are helpful to find the most relevant search results in the world wide web, find out which nodes of a network are most important to control to achieve a desired behavior, or which airports to observe to quickly identify the spreading of diseases [12, 104, 128, 134]. For this reason, these problems are often studied in simple model systems to identify the general mechanisms behind the observed phenomena and to understand how network topology affects the dynamics and vice versa.

Controlling network dynamical systems under economic constraints

Many of these networked systems like the internet, transportation, or the power grid are central to our modern world. We rely on their correct functioning and our thorough understanding of how these systems behave and react to changes. More importantly, however, we rely on controlling these systems in order to maintain the desired function. Control often has different goals depending on the specific system. In some systems it is necessary to avoid an undesired state, such as controlling fluctuations in the power grid to prevent a power outage [157, 158, 185] or vaccination and quarantine to prevent the outbreak of a disease [90, 137]. In other systems we want to force the system into a specific state [55, 74, 115], relevant for example for controlling gene regulation networks and marketing in social networks.

Controlling networked systems is typically difficult and any control action is costly. Often the system is not directly or only partially accessible, like most biological systems, or it is not economically feasible to directly control the entire network, such as in large social networks. For example, it would be extremely difficult to vaccinate every person in a short time to stop the outbreak of a disease. Efficient control strategies are required to enable control under such constraints and with minimal costs. Despite this importance to almost any system [85, 86], explicit economic considerations are largely missing from the theoretical analysis of most network dynamical systems.

Often control schemes are designed without economic constraints, minimizing the time until the control achieves the desired effect or maximizing the effect of the control [10, 20, 157]. In some cases, economic constraints are considered implicitly in the analysis of control strategies. One such example is the identification of optimal targets for vaccination such that the vaccination of only a few central individuals stops a disease from spreading [46, 123]. Similarly, in other systems it may be possible to externally determine the state of the whole system by controlling only a few central units, taking advantage of the internal dynamics of the other units to synchronize automatically [184, 201]. This structural control minimizes the number of targets for the control actions, often also reducing the costs for control, and is an important step to designing efficient control schemes. However, explicit consideration of limited resources and optimization of the control schemes under this constraint may require

different approaches or reveal new interactions that are not apparent otherwise [167].

Similarly, the standard models of network formation usually do not include any economic constraints or only model them implicitly or in weaker form. For example the “explosive” percolation rules that model control of network formation do not include explicit cost for controlling the link addition process. Instead, restrictions are considered implicitly by assuming that only one of every two links may be rejected [3, 147]. While such models provide some insight into the effects of economic constraints, the examples of the collective phenomena given above clearly demonstrate that complex interactions can cause entirely new and qualitatively different phenomena [64, 128, 178]. It is therefore necessary model economic constraints explicitly in fundamental models in order to reveal and study their effects.

Finally, it is not only important to understand the effect of such constraints on control schemes but also their impact in the basic models themselves. Most models of network formation, for example, are based on stochastic processes imitating more complex underlying dynamics. While these models reproduce observed network structures, the effects of underlying economic considerations and optimization driving the formation of many social, technical, and biological networks often cannot be understood with these models [76, 93, 95]. It is therefore of current interest to understand the impact of economic considerations on a fundamental level [85, 86]. Limited resources in particular and optimization principles in general have the potential to dramatically change the dynamics of networks and network formation and the effect of control on these systems.

Synopsis and structure of this thesis

In this cumulative thesis we discuss the effect of economic considerations and optimization on network dynamical processes and network formation. In the main part of this thesis we explore the effects on the control of such processes as well as on the processes themselves in five articles, each summarized below. In particular, we consider how to reduce interaction costs by restricting the interactions between coupled dynamical units. We study how these restricted interactions affect the synchronizability of the network and identify an efficient control strategy to guarantee synchronization. Similarly, we consider control of network formation where control of each link is associated with an explicit cost and identify an efficient control strategy in a standard percolation model. Finally, we study how the network formation process changes when economic considerations are driving the formation of the network, linking a complex optimization problem to a local percolation model.

In chapter 2 we give a brief introduction into the most fundamental concepts of networks and dynamical systems. We specifically focus on methods and results relevant for the following articles. We discuss network formation models and random percolation as well as chaotic systems and synchronization. Finally, we also briefly discuss the typical numerical implementation since a large part of the work is done

with extensive simulations when the systems become too complicated for a comprehensive analytical treatment.

In chapter 3 we discuss the impact of restricted interactions on the synchronizability of two chaotic oscillators. We introduce the concept of “transient uncoupling”, that means switching off the coupling between the oscillators depending on their state. In contrast to many other coupling schemes requiring continuous control of the coupling strength [55, 97, 140], transient uncoupling requires only on-off control over the coupling in the system. We show that, even though we reduce the interactions, transient uncoupling can in fact increase the synchronizability and extend the range of favorable coupling strengths. This allows transient uncoupling to induce synchrony in systems with fixed coupling strengths or limited access to the system parameters.

In chapter 4 we extend the discussion of the previous chapter to arbitrary networks of chaotic oscillators. The dependence of the synchronizability of a network on the network topology means that a wide range of networks cannot be synchronized with standard coupling. While many tools have been discussed to solve this problem, from changing the network structure to adaptively changing the coupling strength, we show that simply uncoupling the units depending on their state can guarantee stable synchronization, independent of the topology of the interaction network. Application of such a synchronization scheme might save a large fraction of interaction costs, for example from signal transmission, making it an efficient tool to control network dynamics by enabling synchronization to a desired target state.

Efficient control of networked systems not only requires efficient control schemes but also accurate descriptions and measurements of the state of the system.

In chapter 5 we present a new order parameter to measure the coherence of coupled phase oscillators in the prototypical Kuramoto model. Applications to networks so far have introduced various adaptations of the original Kuramoto order parameter to account for the disordered interaction structure [77, 91, 148, 149]. However, none of the parameters cover the full range of behavior from initial phase locking of individual oscillators to full synchronization. Our proposed order parameter accurately describes the degree of phase coherence with respect to the network structure in all stages of phase locking and synchronization for all network topologies and sizes. A direct relation to the stability of the synchronous state allows analytical calculations and application to similar models and finite networks such as power grids.

In the last two manuscripts we study the impact of economic considerations on network formation models.

In chapter 6 we consider optimal control of random percolation subject to limited resources. Numerous network percolation models have been designed to model the impact of control on the network formation process, notably leading to very abrupt but still continuous, so called “explosive” transitions [3, 10, 47, 147]. However, all these rules allow for, in principle, unlimited interventions. Here, we consider the effect of a limited budget when interventions in the link addition process are costly

explicitly. In order to efficiently use the limited resources we derive an efficient control scheme. In particular, we show that optimal control of percolation directly leads to a discontinuous transition, an unintended consequence since the network structure becomes effectively uncontrollable when a single link can have a macroscopic impact on the size of the largest connected component.

In chapter 7 we analyze network formation driven by local economic considerations. Instead of a random process in typical percolation models, we develop a deterministic network formation model based on a fundamental network supply problem where each node individually satisfies a fixed demand with minimal cost. The resulting coupled nonlinear optimization problem and similar game-theoretic models of network formation are typically hard to solve for larger networks [76, 94, 95]. We prove that our model exactly maps to a local percolation model that allows an efficient solution. Studying the network formation process, we reveal parameter dependent continuous or discontinuous transitions and hysteresis of the global network structure as a result of the individual optimization of the local network structure.

In chapter 8 we summarize the results of the individual manuscripts. We discuss their importance in the context of current topics in network science and dynamical systems and highlight potential applications and promising extensions.

Chapter 2

Fundamentals

In this chapter we provide an overview over theoretical concepts and the mathematical description of networks and dynamical systems. Due to the sheer amount of content related to these fields we cannot possibly cover every aspect. We focus on the most fundamental concepts and provide the basic information relevant to understanding the manuscripts in chapters 3 to 7. We outline other related concepts and ideas and refer the interested reader to the numerous textbooks and review articles on these topics [5, 12, 127, 128, 142, 178]. Since much of the work on percolation and nonlinear dynamical systems requires numerical simulations, in each part we also give a brief description of the basic aspects of relevant numerical concepts and their implementation.

Specifically, we first discuss the mathematical representation of networks and network formation models as well as the basic concepts of percolation theory. We then introduce the fundamentals of dynamical systems and chaos before finally discussing the most important aspects of networked dynamical systems and synchronization for the following manuscripts.

2.1 Graphs and networks

Networks (graphs) consist of *nodes*, also called *vertices* or simply *units*, connected by *links (edges)* between them [128]. Many examples were already mentioned in the introduction, for example social networks where nodes represent people and connections represent friendships. In the simplest case connections are reciprocal or *undirected*, for example two people being – usually mutual – friends. However, links can also be *directed*, for example the relation of followers and followees on Twitter. An illustration of a small undirected network and a similar directed network is shown in Fig. 2.1. For ease of presentation we will mostly restrict ourselves to undirected networks and mention extensions to directed networks where relevant. We do not cover more complex structures, where there may be multiple links describing different types of relationships between the nodes, like multi-graphs, multi-layer networks and similar generalizations of networks. Instead we refer to the corresponding literature for the extensions of the concepts discussed here [22, 102].

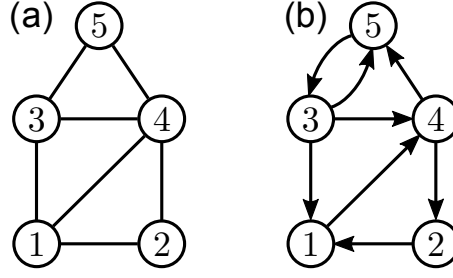


Figure 2.1: (a) A small undirected network with $N = 5$ nodes and $M = 7$ links connecting nodes in both directions. (b) A small directed network with $N = 5$ nodes and $M = 8$ directed links, indicated by the arrows.

2.1.1 Representation of graphs

A graph as a mathematical object consists of a *vertex set* V , typically $V = \{1, 2, \dots, N\}$, describing all nodes and an *edge set* E , for example $E = \{\{1, 2\}, \{1, 4\}, \dots\}$, describing the links between these nodes represented as pairs of vertices [193]. For directed graphs the edges set is given by a set of ordered pairs of vertices such as $E = \{(2, 1), (1, 4), \dots\}$. A graph G is then defined by the combination of these two sets $G = (V, E)$. For example, the vertex- and edge-set of the network shown in Fig. 2.1(a) is

$$\begin{aligned} V &= \{1, 2, 3, 4, 5\} \\ E &= \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 4\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}. \end{aligned} \quad (2.1)$$

The size of a graph usually refers to the number of nodes $|V| = N$ in the network. The number of links in the network is usually denoted by $|E| = M$.

While the representation of a graph via this *edge list* is very compact and sometimes used to encode large networks for digital storage, it is cumbersome for most calculations or algorithms. A more convenient representation is given by the *adjacency matrix* A . Every entry $A_{ij} \in \{0, 1\}$ indicates if the link from node j to i is present, $A_{ij} = 1$, or missing from the network, $A_{ij} = 0$.¹ If the network is undirected and a connection from j to i also implies the reverse connection, the adjacency matrix is symmetric $A_{ij} = A_{ji}$. The adjacency matrices for the two networks in Fig. 2.1 are, respectively,

$$A^{\text{undir}} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} \quad A^{\text{dir}} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}. \quad (2.2)$$

¹Note that there are different conventions depending on the field and the order of indices may be reversed. $A_{ij} = 1$ may imply a (directed) connection from i to j instead of a connection from j to i .

The adjacency matrix naturally generalizes to *weighted* graphs by assigning scalar properties to the links, where typically $A_{ij} > 0$ signifies that a link is present in the network.² For example, in financial networks a directed link may represent a loan and the value A_{ij} may describe the amount of money loaned. For most analytical calculations the adjacency matrix is the most convenient representation of networks.

Numerical representation

The adjacency matrix is ideal to represent small networks numerically due to direct access to every link to check whether a link exists and to potentially modify it. However, for larger networks with thousands of nodes but only few links per node, so called *sparse networks* [128], most of the information in the adjacency matrix refers to absent connections $A_{ij} = 0$ and storing this information explicitly quickly uses too much memory. Similar to the numerical representation of sparse matrices, a numerically very efficient representation of sparse networks is the so called *adjacency list*. As a middle ground between an edge list and an adjacency matrix, the adjacency list stores for each node i individually the nodes j it is connected to. One can think of it as multiple edge lists sorted by nodes, for example for the network shown in Fig. 2.1(a) the adjacency list is

$$\begin{aligned} 1: & \{2, 3, 4\} \\ 2: & \{1, 4\} \\ 3: & \{1, 4, 5\} \\ 4: & \{1, 2, 3, 5\} \\ 5: & \{3, 4\} \end{aligned} \tag{2.3}$$

Depending on the specific requirements of the numerical computations, different data structures are used to represent the adjacency list to allow for fast look-up of specific edges, iteration over all links or insertion and removal of links [98, 128, 145].

2.1.2 Describing the network structure

So far, we only discussed how to represent a network mathematically. However, it is difficult to understand and compare the structure of networks directly from these representations and direct visualizations quickly become unintelligible for larger networks. Unsurprisingly, there are numerous measures to quantify different aspects of connectivity, that means how well a node is connected to other nodes in the networks, or the importance of individual nodes to the structure or function of the network. Here we introduce the most fundamental concepts.

²The meaning of A_{ij} in weighted graphs depends on the context. For example $A_{ij} < 0$ may signify inhibitory connections in neural networks. Similarly, if A_{ij} describes travel times in transportation networks, absent links correspond to $A_{ij} = \infty$.

Degree

The easiest way to quantify the connectivity of a node is simply by counting its links, $k_i = \sum_j A_{ji}$, where sums in this section run over all nodes $j \in \{1, 2, \dots, N\}$ unless explicitly noted otherwise. k_i is called the *degree* of node i [193]. For example, the nodes in the network in Fig. 2.1(a) have degrees 3, 2, 3, 4 and 2, respectively. The notion of degree easily extends to directed networks: simply count the amount of outgoing links as the *out-degree* $k_i^{\text{out}} = \sum_j A_{ji}$ and the amount of incoming links as the *in-degree* $k_i^{\text{in}} = \sum_j A_{ij}$. A single number that summarizes the connectivity of the whole network [5] is the *mean degree*

$$\langle k \rangle = \frac{\sum_i k_i}{N} = \frac{2M}{N}. \quad (2.4)$$

A more detailed measure of connectivity than the mean degree is the list of the degrees of all nodes in a network, called the *degree sequence*. However, the degree sequence does not uniquely determine the network structure and often different networks with identical degree sequence exist. An example for the degree sequence $\{1, 1, 2, 2, 2\}$ is illustrated in Fig. 2.2 [193]. For very large networks the degree sequence is often simply represented as the *degree distribution* $p(k)$, indicating the fraction of nodes with degree k . Alternatively, in models of network formation, where the degree of a vertex is a random variable, the degree distribution $p(k)$ is usually interpreted as the probability of a (uniformly) randomly chosen vertex to have degree k . Many network formation models (see below) are constructed in such a way that they reproduce the observed degree distributions of the networks they are meant to model [50]. However, as with the degree sequence, an identical degree distribution alone is not a sufficient condition that the resulting networks will have similar properties.

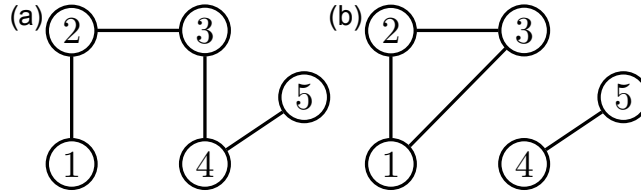


Figure 2.2: Two networks with identical degree sequence $\{1, 1, 2, 2, 2\}$ but different structure.

Paths, distances, and connected components

While the degree of a node describes its local connectivity, the non-local connectivity in the whole network structure is often even more important. For example, communication networks like the internet would not function if not all parts were connected in some way. In this example, there needs to exist a way to reach every node from every other node via a sequence of links for the network to be able to transmit information

between the nodes. Graph theory defines a *path* Π_{v_n, v_1} from node v_1 to v_n as an ordered set of links $\Pi_{v_n, v_1} = (\{v_1, v_2\}, \{v_2, v_3\}, \dots, \{v_{n-1}, v_n\})$ where each link starts at the node where the last link ended [193]. In addition, no edge and no node may be visited more than once when following the path. An exception are *closed paths* (*cycles*) where the first node v_1 is also the last node $v_n = v_1$ in the path. There are multiple generalizations of this notion called *walks* or *trails* allowing, for example, links to be used multiple times. For obvious reasons of similarity, a network without a cycle is called a *tree*, or a *forest* if the network consists of multiple trees, see below [128].

The length of a path is given by the number of links in the path. For example a path $\Pi_{4,1} = (\{1, 2\}, \{2, 3\}, \{3, 5\}, \{5, 4\})$ between nodes 1 and 4 would have length $|\Pi_{4,1}| = 4$ as illustrated in Fig. 2.3. With this definition of path length, paths also allow for an intuitive definition of distance in networks. The (*shortest path*) *distance* d_{ij} from node j to i is simply the length of the shortest path between them, $d_{ij} = \min_{\Pi_{ij}} |\Pi_{ij}|$ where the minimum is taken over all possible paths from j to i [193]. In the network shown in Fig. 2.3 the shortest path between node 1 and 4 has length $d_{4,1} = 3$. If there is no shortest path the distance is usually defined to be $d_{ij} = \infty$. The longest shortest path $d = \max_{i,j} d_{ij}$ is called the *diameter* of the graph [128]. In the example in Fig. 2.3 the diameter is 3, given by the length of the shortest path between nodes 1 and 4 (or 5).

The notion of a path and distances is immediately applicable to directed networks. The sequence of links must now explicitly respect the directionality of the links, $\Pi_{v_n, v_1} = ((v_1, v_2), (v_2, v_3), \dots, (v_{n-1}, v_n))$. Similarly, the length of paths in weighted networks, where the weight of a link describes the distance between the nodes, can be defined as the sum of the weights of all links along a path, instead of simply counting the number of links [128],

$$d_{ji} = \sum_{(k,l) \in \Pi_{ji}} A_{lk}. \quad (2.5)$$

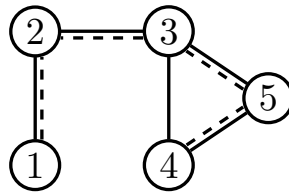


Figure 2.3: A path in a small network. The marked path (dashed lines) between nodes 1 and 4 includes four links and consequently has length 4. The shortest path defining the distance $d_{4,1} = 3$ only consists of three links. Since this is also the longest shortest path of this graph, the graph's diameter is $d = 3$.

Paths define connectivity in terms of which node can be reached from which other node. A *connected component* of a graph, sometimes also called *cluster*, is a set of

nodes where there is a path from every node in the set to every other node in the set [128]. Recalling the example of the internet, to function properly there must be a path between every pair of nodes. The internet must be a single giant connected component, a *connected graph*, like the example network in Fig. 2.2(a), but unlike the network in Fig. 2.2(b). This example illustrates that the concept of connectivity and connectedness is very important for the function of many networks and it is thus the main focus of many network formation models and specifically of percolation theory (see below). In directed networks, where there may be a path from node i to j but not from j to i , the notion of a connected component is more nuanced [128, 183]. In a *strongly connected component* there is a directed path every node to every other node. Otherwise one can define different types of components: the *out-component* of node i comprises all those nodes that can be reached from i and conversely the *in-component* comprises all nodes that can reach node i .

Centrality

Whereas connectivity describes the overall structure of the network, we are often more interested in individual nodes. One of the most widely used concepts is the idea of *centrality* in networks, that means describing which nodes are most important given only the network structure. Different aspects of the structure are more or less important to a the centrality of a node, depending on the specific network, its function and the question. Therefore, these measures are often motivated by functional considerations and their significance depends on the context [64, 93, 128].

One intuitive measure for structural importance is simply the degree of a node, in this context called the *degree centrality* $c_i^{\text{deg}} = k_i$. An example where this measure is actually applied is judging the impact of scientific publications in the network of citations: if a publication has a high (in-)degree that means it is cited often and often considered to be impactful and central to its field. However, the degree centrality is a local measure. It only considers the direct connections of a given node but not which other nodes it is connected to or other aspects of the large scale network structure.

A measure that directly takes into account the neighboring nodes is the *eigenvector centrality* c_i^{ev} in connected networks. Based on the idea that a node is likely to be important if it is connected to other important nodes, it is defined via the self-consistency equations for every node $i \in \{1, 2, \dots, N\}$

$$\lambda c_i^{\text{ev}} = \sum_j A_{ij} c_j^{\text{ev}}, \quad (2.6)$$

with the additional condition that all $c_i^{\text{ev}} \geq 0$ and at least one non-zero entry. Since all entries of the adjacency matrix $A_{ij} \geq 0$ and the network is connected, this is equivalent to the eigenvector \mathbf{c}^{ev} of the adjacency matrix A corresponding to the largest eigenvalue λ_1 ,

$$\lambda_1 \mathbf{c}^{\text{ev}} = A \mathbf{c}^{\text{ev}} \quad (2.7)$$

and is usually normalized to $\sum_i c_i^{\text{ev}} = 1$. Note that this definition naturally extends to

weighted network with non-negative edge weights $A_{ij} \geq 0$. A slightly more involved version of this concept is the basis for ranking the importance of search results of web-sites, though the actual algorithms used are much more complicated [31, 106, 128].

Another concept of centrality is based on a different function of networks. In social, communication, or transport networks, for example, a node may play an important role simply by relaying messages or connecting two large cities [36, 64, 72]. This idea is captured by the concept of *betweenness centrality* c_i^{bet} [128]. Betweenness simply counts the number of shortest paths n_{ab}^i through node i between pairs of nodes b and a . Normalized by the total number of shortest paths n_{ab} from b to a this defines the betweenness centrality

$$c_i^{\text{bet}} = \sum_{a,b} \frac{n_{ab}^i}{n_{ab}}. \quad (2.8)$$

This concept of centrality also extends to the betweenness centrality of links with an analogous calculation, allowing efficient community detection by identifying links “between” the different communities [128].

There are many more centrality measures, for example directly based on distance measures in the network, and corresponding extensions to directed networks, hyper-graphs, multi-layer networks and other generalizations [61, 128]. Many of these measures not only describe the structural importance in a given network but are also related to the dynamics of networked systems (see below).

Clustering

Finally, we discuss one aspect of the network structure that does not describe individual nodes but small local structures, so called *motifs* already mentioned in the introduction. The simplest case are triangles between the nodes, that means cycles of length 3. In many social networks, links often form such triangles, meaning that two friends of mine are likely also friends. This property is called *clustering* and quantified by the *clustering coefficient* cc_i of a node i .³ It is defined as the fraction of triangles including node i relative to all possible triangles given its number of neighbors k_i ,

$$cc_i = \frac{\sum_{j,k} A_{ij} A_{jk} A_{ki}}{k_i(k_i - 1)}, \quad (2.9)$$

where $A_{ij} A_{ik} A_{jk} = 1$ if all links between the three nodes exist and 0 otherwise [64, 128, 191]. Note that both the nominator as well as the denominator count each triangle twice, once for each direction (i, j, k) and (i, k, j) , and this factor cancels out.

³Not to be confused with the concept of a cluster as a connected component in networks. Clustering is also sometimes called *transitivity*, due to its similarity to the mathematical concept: if the links (i, j) and (j, k) exist, then the link (i, k) also (likely) exists.

Similarly, we can define this property for the whole network. The global clustering coefficient is then given by the total number of triangles relative to the number of potential triangles, that means paths of length two where at most one link is missing to form a triangle,

$$cc = \frac{\sum_{i,j,k} A_{ij}A_{jk}A_{ki}}{\sum_{i \neq j \neq k} A_{ij}A_{jk}}. \quad (2.10)$$

Here, the nominator counts each triangle six times, once for each permutation of i , j , and k , and the denominator counts all paths between three mutually distinct nodes. The factor 6 cancels since each triangle contributes six paths of length two.

2.1.3 Network formation

To understand how the topology of a network impacts different processes we cannot solely rely on individual examples of measured network structures. Instead, we need to compare the behavior in large sets of different networks to identify common patterns. Network formation models are used to construct such surrogate networks and act as baseline models to compare processes across network with different properties. More complicated models following observed or expected behavior may even help to identify some of the fundamental forces guiding the formation of real world networks. Most network formation models are based on stochastic processes, randomly adding nodes or links to the network. This results in an *ensemble* of random networks, that means a set of possible structures $\{G_1, G_2, \dots\}$ where each structure is assigned a corresponding probability $p(G_i)$, depending on the network formation model [5, 128].

In this section we will discuss the most common network formation models and the most important properties of the resulting networks. We first discuss the most basic random network formation model before moving on to a model that more closely reproduces the scale-free degree distribution observed in many networks. We then discuss the small world model as an example of a random network formation model designed to explain other structural aspects of real world networks. Finally, we briefly discuss one example of a game theoretic model of network formation.

Poisson random graphs

The simplest model of random networks are *Poisson random graphs*, sometimes referred to simply as random graphs, and named *Erdős-Rényi graphs* after Paul Erdős and Alfréd Rényi who thoroughly analyzed this model already in 1960 [66, 67, 127]. In this model the network consists of a given number of N nodes where each link (i, j) is present with a fixed probability p , independent of all other links. The resulting graph ensemble is usually denoted $G(N, p)$, where the two arguments refer to the two parameters of the model. Two different realizations of a small network from the ensemble $G(50, 0.05)$ are illustrated in Fig. 2.4.

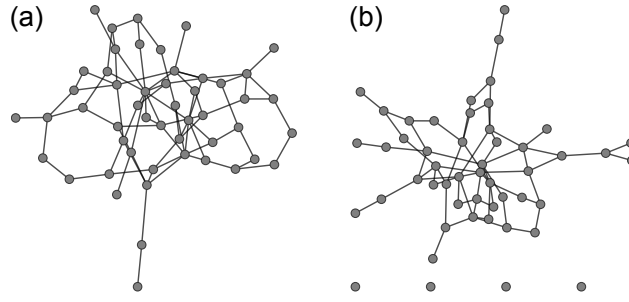


Figure 2.4: **Poisson random networks.** (a,b) Two realizations of the Poisson random graph model for identical parameters $N = 50$ and $p = 0.05$. In fact, every possible network has a non-zero probability, however, some network structures are more likely than others.

The independence of the links simplifies the analysis of most properties of the resulting networks. For example, the number of links $|E| = M$ of graphs in $G(N, p)$ is distributed binomially as

$$p(M) = \binom{\frac{N(N-1)}{2}}{M} p^M (1-p)^{\frac{N(N-1)}{2} - M} \quad (2.11)$$

with mean $\langle M \rangle = pN(N-1)/2$, where $N(N-1)/2$ is the total number of links possible in the network. Similarly, the degree distribution is given by

$$p(k) = \binom{N-1}{k} p^k (1-p)^{N-k-1}, \quad (2.12)$$

with mean degree $\langle k \rangle = (N-1)p$, since each node is connected to at most $N-1$ other nodes. In the limit of large networks, $N \rightarrow \infty$, and if the mean degree $\langle k \rangle = (N-1)p \rightarrow Np$ remains constant, the degree distribution is approximated by a Poisson distribution, giving the name for this network ensemble [5, 127],

$$p(k) \rightarrow \frac{(Np)^k e^{-Np}}{k!}. \quad (2.13)$$

One important difference of a Poisson random networks compared to regular lattices is the small diameter. For example, in a square lattice with $N = L^2$ nodes the longest shortest path connecting two opposite corners has length $d = 2L = \mathcal{O}(N^{1/2})$. In contrast, the diameter of the largest connected component of a Poisson random graph scales at most as $d = \mathcal{O}(\log N)$ for large networks [128].

A closely related model, denoted $G(N, M)$, fixes the number of links $|E| = M$ in the graph and assumes the M links are uniformly randomly distributed among all possible $N(N-1)/2$ links (see section 2.1.4 below). The two ensembles $G(N, p)$ and $G(N, M = pN^2/2 \approx pN(N-1)/2)$ behave identically with respect to most properties

when considering large networks, that means when $N \rightarrow \infty$ and $M \approx pN^2/2 \rightarrow \infty$, since the fluctuations in the number of links in $G(N, p)$ are small compared to the total number of links [27, 28, 127, 128]. In particular, this condition fulfilled in the typical limit of networks with a constant mean degree $\langle k \rangle$, such that $p \sim N^{-1}$ and $M \sim N$.

The Poisson random graph model is often used as a baseline model to illustrate processes and phenomena in network dynamical systems without assuming any additional structural properties of the networks. Consequently, there are many properties of real networks it cannot reproduce. For example, the degree distribution of many real networks is not Poisson and the clustering coefficient of many social networks is much higher than in Poisson random graphs.

Scale-free networks

In contrast to Poisson random networks, a large number of real networks have an (almost) scale-free degree distribution, that means $p(k) \sim k^{-\alpha}$, usually with $2 \leq \alpha \leq 3$ [15, 50]. This means there are some nodes with a much larger degree in the network than one would expect for independently randomly distributed links, so called *hubs*. The *configuration model* [93, 127] creates random networks with any given degree distribution and can therefore also create *scale-free networks*.⁴ However, here we focus on another model that offers a mechanistic explanation for the emergence of a scale-free degree distribution instead of only a way to replicate the observed distribution.

Introduced by Albert-László Barabási and Réka Albert in 1999 the *Barabási-Albert model* creates a network by sequentially adding more and more nodes until a given size N is reached [14]. Starting from a small, completely connected network with $N_0 = m + 1$ nodes, at each step a new node is added to the network and connected to m existing nodes. However, the links are not established uniformly at random. Instead, the links are established following *preferential attachment* [14, 146, 200], meaning a node i is selected for the link with probability proportional to its degree k_i . Fig. 2.5 shows the first few steps of the evolution, one realization of a larger network, and the resulting degree distribution.

This model is supposed to mirror the “rich-get-richer” phenomenon observed in real networks, for example a scientific article with a large number of citations is likely to be cited again by new publications [64]. This preferential attachment mechanism then creates networks where the degree distribution follows a power law

$$p(k) \sim k^{-3} \quad (2.14)$$

for large k when $N \rightarrow \infty$, illustrated in Fig. 2.5(f). This can, for example, be shown by solving the mean-field master equation describing the degree distribution of the network [14].

⁴More specifically, the configuration model creates a network with a given degree sequence, possibly randomly drawn from a specific degree distribution.

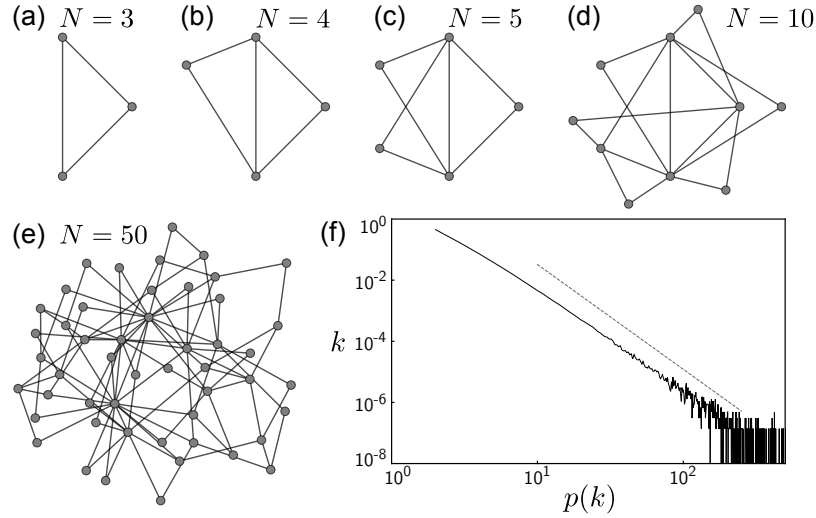


Figure 2.5: **Barabási-Albert scale-free networks.** (a-d) One example of the evolution of the Barabási-Albert model with $m = 2$. (e) Example of a larger network with $N = 50$ nodes. The two nodes from the initial networks are still well visible and have accumulated a large number of connections. (f) The degree distribution of a single network with $N = 10^6$ nodes. The degree distribution approximately follows a power law with exponent -3 (grey dashed line).

Networks created by this model also serve to illustrate the so called *core-periphery structure* observed and quantified in many networks [29, 88]. The idea is that there are a few highly connected nodes (the hubs), which are critical for the connectivity and shortest paths in the network, form the core of the network and many weakly connected, less important nodes form the periphery. One example for this structure is the airline network, where a few large airports are densely connected by long distance flights and form the core of the network. Many smaller, local airports are often only connected to relatively few, close-by airports and form the periphery [189].

Related models based on a very similar premise were already analyzed by Yule in 1925 and Price in 1976, studying the network of citations of scientific publications [146, 200]. Other network formation models built on describing growing networks include generalizations of preferential attachment, leading to power law degree distributions with different exponents, or vertex copying models, used to describe the emergence of scale-free degree distributions in biological contexts, where preferential attachment is an implausible mechanism [105, 173].

Small world networks

Yet another large difference between real world networks and the basic Poisson random graphs is the so called *small world effect*. This means that a network that has a comparatively low mean degree has a high clustering coefficient (many triangles)

but at the same time has a very small diameter. Intuitively, these two properties are at odds with one another since high clustering means that most links are in local, densely connected communities whereas a small diameter means that nodes must also be well connected across their local neighborhood [127].

The *Watts-Strogatz model* solves this apparent paradox by considering a systematic change from a regular network with high clustering to a random network with a low diameter [191]. The basis for this model is a ring with N nodes, where every node is connected to its $\langle k \rangle / 2$ nearest neighbors on each side. The model then interpolates between this regular network and a random network by rewiring each link in the network with probability q_{rew} . Rewiring means the link is disconnected from one of its end nodes and reconnected to another, uniformly randomly chosen node. This creates “shortcuts” in the network. For small q_{rew} there are only few shortcuts and for $q_{\text{rew}} \rightarrow 1$ the network becomes completely random. Examples for the resulting network structures are illustrated in Fig. 2.6.

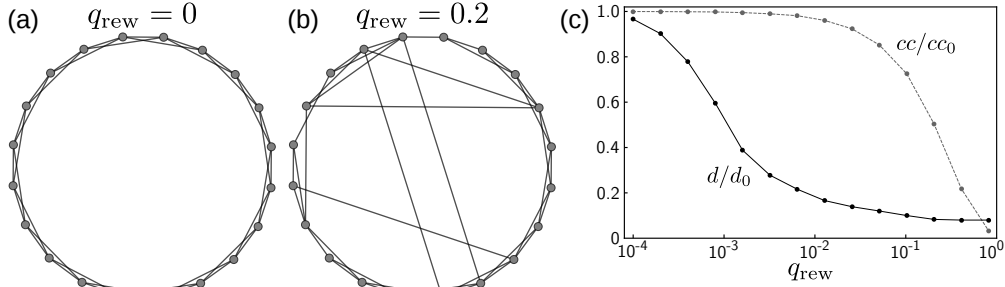


Figure 2.6: **Watts-Strogatz small world networks.** (a,b) Watts-Strogatz model networks with $N = 20$ nodes and $\langle k \rangle = 4$ for different parameters $q_{\text{rew}} = 0$ and $q_{\text{rew}} = 0.2$, respectively. (c) Diameter and clustering coefficient relative to their initial value at $q_{\text{rew}} = 0$ for large networks with $N = 1000$ nodes and $\langle k \rangle = 20$ (results averaged over 20 realizations). For intermediate values of q_{rew} the networks show the small world effect: a high clustering coefficient (grey line) and a small diameter (black line) at the same time.

In the regular network nodes are only connected to their local neighbors and therefore the network has a high clustering coefficient but a large diameter. As the rewiring probability q_{rew} increases some links are rewired and act as shortcuts to strongly decrease the diameter of the network. At the same time most of the triangles are preserved and the clustering coefficient remains large. This is called the small world effect in networks, in reference to structures often found in real social networks. It shows that a small number of shortcuts in the network are sufficient to strongly reduce its diameter. When q_{rew} becomes large almost all links become randomly rewired and, while the diameter remains small, the triangles in the network are destroyed. While this network formation model reproduces the small world effect, the degree distribution does not match those of real networks.

This model easily generalizes to different initial topologies, such as two dimensional lattices, and links may be added instead of rewired without strongly affecting the resulting network characteristics for small q_{rew} [103, 107]. The Watts-Strogatz model and similar models based on hierarchical network construction have been used to study routing and message passing in complex networks and explain the results of Milgram's letter experiment described in the introduction [23, 24, 192]. In general, the Watts-Strogatz model is a very convenient model to illustrate processes or phenomena across different network structures, ranging from regular to random, with a single network formation model with only one parameter q_{rew} .

Game theoretic models

Another approach to network formation is to directly construct the networks from basic principles, often resulting in optimization models [26, 76, 153]. Specifically in social networks it is plausible that people form links not randomly but based on the individual benefit they obtain from the links. This leads to game theoretic models of network formation [13, 64].

One such model is the *distance based utility* model [93–95] where each node obtains a utility u_i from the network given by

$$u_i = \sum_j b(d_{ij}) - ck_i, \quad (2.15)$$

where the function $b(d_{ij}) > 0$ and the constant c are parameters. A node gains utility from all nodes it is connected to via shortest paths of length d_{ij} but each direct link a node establishes is associated with a cost c . The utility a node gains from being connected to other nodes (directly or indirectly) is usually assumed to decay with increasing distance, that means $b(d_{ij})$ is a decreasing function. One example is access to information from different people where it is easier to ask a friend for help than a friend of a friend of a friend. The cost c then represents the time investment to maintain the direct friendships.

Such game theoretic network formation models allow deeper analysis, defining refined concepts such as *efficiency*, that means the optimal network structure that maximizes the total utility. At the same time there is a natural definition of (*pairwise*) *stability*, that means a link is only stable if it is beneficial for both its nodes. Analyzing this model shows, for example, that efficient networks (maximizing the total utility) are not always (pairwise) stable as individual nodes would like to form or cut links [92, 93]. It is possible to solve the distance based utility model for efficient network structures, revealing three possibilities: the network either has no links, forms a star, or is completely connected (see Fig. 2.7).

Other models based on optimization are often applied to model network formation directly from physical, technical or economical constraints. Examples include models for the venation network in plant leafs [99, 153], the structure of the world airline

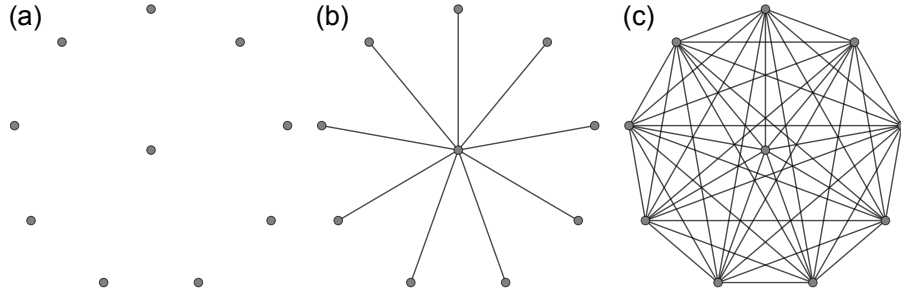


Figure 2.7: **Distance based utility.** (a-c) Efficient network structures resulting from the distance based utility model. If the costs per link are too large the network remains empty. If the cost are very small, all links are formed. For intermediate values of the cost a star-network is the overall efficient solution. However, the network may not be stable when no node wants to become the center due to the costs associated with the direct links.

network or similar distribution networks [76, 189], whose structure depends on trade-offs between the travel distance and the cost of links. Often these models very closely reproduce real network structures of a specific system. However, they are difficult to analyze due to their high complexity and only provide insight in their specific contexts.

2.1.4 Percolation

Percolation is the formation of global connectivity from initially unconnected individual components [80, 156, 177].⁵ Examples of percolation include the aptly named coffee filters (percolators), where water flows through connected spaces between the coffee grounds, similar behavior in porous rocks important for oil extraction, and the gelation of polymers described by Flory, often cited as the first application of percolation [70, 177]. Tools from percolation theory are also applied in the description of epidemic spreading, where global connectivity means that an initial outbreak will affect a non-vanishing fraction of the whole population, or to study the robustness of different networks to random (or targeted) failures [37, 54, 60, 121, 128]. Additionally, many standard percolation problems are analytically solvable [80]. As such the percolation transition is one of the simplest and most studied examples of a phase transition in statistical physics [170].

In this section we first discuss the basic idea of percolation on simple lattices to introduce the concepts of percolation theory. We then consider the emergence of global connectivity in random network percolation (Erdős-Rényi networks) and discuss properties of the phase transition, the critical exponents, and how to measure them in finite size simulations. Finally, we briefly discuss the numerical implementation of simulating percolation in very large networks.

⁵Sometimes the term percolation is also used in the reverse context, meaning the breakdown of global connectivity.

Percolation on a lattice

Chain - discontinuous percolation

The simplest percolation problem is percolation on a one dimensional lattice (chain) with N nodes and, consequently, $M = N - 1$ links [177].⁶ Similar to the Poisson random graph model, we assume each link is present independently of other links with probability p and absent otherwise. This is called *bond percolation*. The same concepts also apply to *site percolation*, where the nodes of the network are present with a given probability and a link exists only if both its nodes are present. Clearly, if $p = 0$ the network is disconnected since no links are present, all nodes are isolated. On the other hand, for $p = 1$ all links are present and the network is globally connected. Between these two extremes, there is a *percolation transition* from individual, disconnected nodes to global connectivity. This is illustrated in Fig. 2.8.

For a lattice it is easy to define global connectivity. The system “percolates” if there is a path from one end of the system to the opposite end, if a so called *spanning cluster* exists [156, 177]. For percolation models on networks without an underlying geometry global connectivity is usually measured via the size S_1 of the largest connected component or cluster in the network relative to the size of the network N [5, 128]. In both cases we are interested in the limit $N \rightarrow \infty$ of infinitely large networks for both the mathematical analysis and to draw parallels with phase transitions known from statistical physics. Percolation is then defined as the transition from $S_1/N \rightarrow 0$ to $S_1/N > 0$ where a macroscopic (*giant*) cluster emerges in the limit $N \rightarrow \infty$.

In our one dimensional example a spanning cluster must contain all links in the network. The probability $P = p^{N-1}$ that such a cluster exists vanishes for any $p < 1$ as $N \rightarrow \infty$. It follows that the *percolation threshold* p_c where a spanning cluster appears is $p_c = 1$, as illustrated in Fig. 2.8(f).

Another quantity of interest is the *cluster size distribution* $n(S)$. It describes the number $n(S)$ of clusters of size S in the network. The quantity $Sn(S)/N$ then describes the probability that a uniformly randomly chosen node belongs to a cluster of size S . For the line graph, this is easy to calculate ignoring as

$$\frac{Sn(S)}{N} = Sp^{S-1}(1-p)^2, \quad (2.16)$$

when ignoring the boundaries of the chain, since a cluster with S nodes must contain all $S - 1$ links in the chain between these nodes but the two links on the ends of the cluster must not be present [177]. This also shows that for any $p < 1$ the number of clusters of size S decreases exponentially with increasing cluster size. Just as with the spanning cluster, there will be no macroscopic cluster $S_1/N > 0$ for any $p < 1$ as $N \rightarrow \infty$.

⁶One may consider a circle with periodic boundary conditions instead of a line. The analysis is identical in both cases. In the limit of $N \rightarrow \infty$ effects of the different boundary conditions disappear.

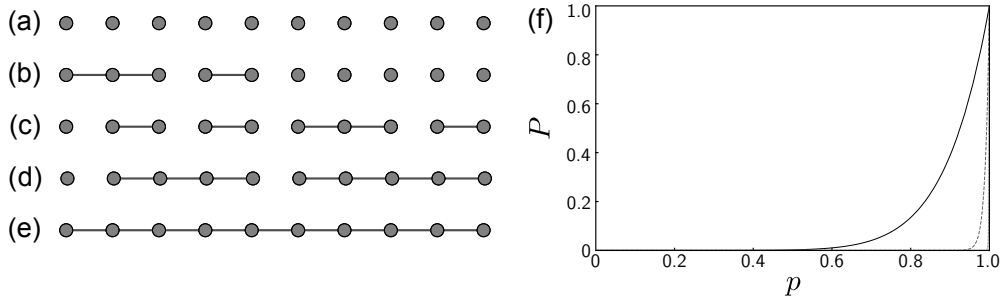


Figure 2.8: **One-dimensional bond percolation.** (a-e) The chain network for different probabilities $p \in \{0, 0.25, 0.5, 0.75, 1\}$. (f) Probability $P = p^{N-1}$ that a spanning cluster exists for different system sizes $N = 10$ (black), $N = 100$ (grey dashed), and $N = 1000$ (light grey dashed). In the limit of large systems $N \rightarrow \infty$ the system only percolates at $p_c = 1$.

This percolation transition is an example of a discontinuous phase transition where the order parameter, the probability that a spanning cluster exists, jumps discontinuously from 0 to 1 as the control parameter, the probability p of a link being present, is changed [63, 156, 170].

Bethe lattice - continuous percolation

A slightly more interesting transition occurs in another analytically solvable example, the *Bethe lattice* [177].⁷ The Bethe lattice is a regular, infinitely branching tree where each node has the same number of neighbors $\langle k \rangle$, illustrated in Fig. 2.9. Here we will only consider $\langle k \rangle = 3$ for ease of presentation but equivalent calculations are possible for larger $\langle k \rangle$. The Bethe lattice is effectively an infinite dimensional structure⁸ and we measure the global connectivity in terms of the relative size of the largest connected component. To understand bond percolation on the Bethe lattice, consider an arbitrary node. Let q be the probability that one of its links does *not* connect this node to the giant cluster. Since all nodes in the lattice are equivalent, the self-consistency equation

$$q = 1 - p + pq^2 \quad (2.17)$$

holds, since either the link is not present with probability $(1 - p)$ or the link is present but none of the other two links on this branch are connected to the giant cluster with probability pq^2 . This equation has the two solutions $q = 1$ (corresponding to a network without a giant cluster) and $q = (1 - p)/p$. With this, the probability that a node belongs to the giant component, which is equal to the fraction of nodes in the

⁷The Bethe lattice is also called *Cayley tree*, which technically should refer to a finite version of the lattice.

⁸The number of nodes on the “surface” at distance d from a central node increases exponentially and is half the total number of nodes up to this distance. This means the “volume” is proportional to the “surface” making the Bethe lattice effectively infinite dimensional [177].

giant component is given by

$$S_1/N = 1 - q^3 = 1 - \left(\frac{1-p}{p}\right)^3 \quad (2.18)$$

above the percolation threshold $p_c = 1/2$. In this case the phase transition is continuous, it corresponds to a critical, second order phase transition similar to the spontaneous magnetization in the Ising model [5, 156, 177]. The order parameter S_1/N varies continuously but is not differentiable at the transition.⁹ As for the transitions known from thermodynamics, one can define critical exponents characterizing the behavior at and around the transition [170]. For example, for the Bethe lattice we find $S_1/N \sim (p - p_c)^\beta$ by expanding equation (2.18), defining the exponent $\beta = 1$. We will discuss these critical exponents and how to measure them in simulations in more detail below.

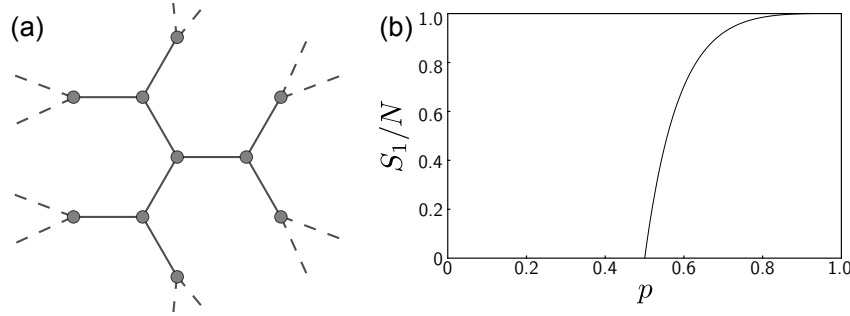


Figure 2.9: **Bond percolation on the Bethe lattice.** (a) Structure of the Bethe lattice for $\langle k \rangle = 3$. (b) Size of the largest cluster S_1 relative to the total number of nodes in the system N in the thermodynamic limit $N \rightarrow \infty$. As derived in the text, a macroscopic cluster emerges continuously after the percolation threshold $p_c = 1/2$.

Random network percolation

Random network percolation follows the same idea as the above examples but without any underlying structure. Instead, links between any pair of nodes may be present or absent. Note that this model is equivalent to the construction of Poisson random graphs discussed above. Using the degree distribution of the resulting network in the limit of large $N \rightarrow \infty$ and $p \sim N^{-1}$

$$p(k) \approx \frac{(Np)^k e^{-Np}}{k!}, \quad (2.19)$$

⁹To be precise, the correct order parameter is the average $\langle S_1/N \rangle$, since S_1/N can only take discrete values in finite systems and will always jump discontinuously in individual realizations. For ease of notation we drop the $\langle \cdot \rangle$ but understand that these statements refer to the corresponding ensemble average.

we can easily derive the size of the largest cluster. Similar to the above argument, assume the probability that with probability q a given node does not belong to the giant component. This is only true if none of the neighbors of the node belongs to the giant component. We again find a self consistency equation, similar to that for the Bethe lattice, by averaging over the possible number of neighbors of that node

$$q = \sum_k p(k) q^k = e^{-Np(1-q)}, \quad (2.20)$$

or in terms of the largest cluster size

$$S_1/N = 1 - q = 1 - e^{-NpS_1/N}. \quad (2.21)$$

Note that we consider the limit where $\langle k \rangle = (N-1)p \rightarrow Np$ is constant such that the exponent is always finite. This equation has a positive solution S_1/N only for $p > p_c = 1/N$ or equivalently $\langle k \rangle > 1$ [5, 127, 128].

Many percolation rules, especially for simulations, are formulated in terms of network growth, where links are added sequentially, one at a time. One then defines the *link density* $l = M/N$ as the control parameter. As discussed above, for the Poisson random graph the two ensembles $G(N, p)$ and $G(N, M)$ behave similarly for large networks. Following the above argument, in this setting random network percolation occurs at $l_c = p_c N/2 = 1/2$, illustrated in Fig. 2.10 [127].

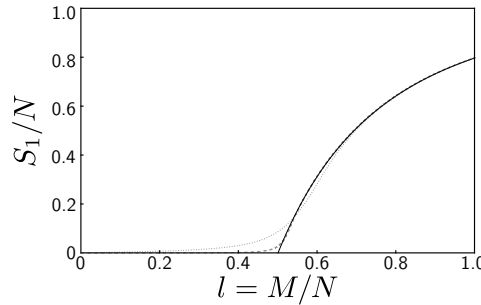


Figure 2.10: **Random network percolation.** Average size of the largest cluster S_1/N relative to the total number of nodes in the system N . The grey lines show numerical averages over 1000 realizations for systems with size $N = 2^{10}$ (dotted) and $N = 2^{15}$ (dashed). The black line shows the analytical results in the thermodynamic limit $N \rightarrow \infty$. As derived in the text, a macroscopic cluster emerges continuously after the percolation threshold $l_c = 1/2$ ($p_c = 1/N$).

As for the Bethe lattice, the percolation transition is a second order critical phase transition. The behavior of the system close to the percolation threshold is defined by power laws with critical exponents characterizing the behavior of the different

observables above, at, or below the critical point [170], for example

$$\begin{aligned} S_1 &\sim (l_c - l)^{-1/\sigma} && \text{for } l \rightarrow l_c^-, \\ S_1/N &\sim (l - l_c)^\beta && \text{for } l \rightarrow l_c^+, \\ \langle S \rangle &\sim |l - l_c|^{-\gamma} && \text{for } l \rightarrow l_c^\pm, \\ n(S) &\sim S^{-\tau} && \text{for } l = l_c. \end{aligned} \quad (2.22)$$

For random network percolation these exponents are

$$\sigma = 1/2, \quad (2.23)$$

$$\beta = 1, \quad (2.24)$$

$$\gamma = 1, \quad (2.25)$$

$$\tau = 5/2. \quad (2.26)$$

With the help of equation (2.21) it is easy to verify that $\beta = 1$, finding $S_1/N \approx 4(l - l_c)$ in random network percolation [5, 80, 177].

Finite size scaling

The discussion of phase transitions and critical (power law) behavior only strictly applies in the limit of large systems $N \rightarrow \infty$. In finite systems the behavior is only approximately described by a power law but modified by effects from the finite system size, such as boundaries of the system. For example, the (average of the) relative size of the largest cluster S_1/N will vary smoothly with p , converging to the critical transition and becoming non-analytic at p_c only as the system size diverges, see Fig. 2.10 [177]. Understanding the transition in finite systems is important since any simulation, experiment, or application is necessarily finite. Finite size scaling provides a framework to describe this behavior at the transition in finite systems and enables us to accurately measure the critical exponents from direct simulations of finite systems.

The basic idea of finite size scaling can be understood by starting from the critical scaling of an observable O , $O(p) \sim |p - p_c|^a$. Assuming there is only a single relevant length scale in the system, we define the *correlation length* $\xi \sim |p - p_c|^{-\nu}$, for example via the typical diameter of clusters in percolation problems [80, 156, 177]. At the transition, when an infinite cluster spans the whole system, this length scale itself diverges and the system becomes self-similar.¹⁰ However, in finite size systems the typical length scale cannot exceed the linear size L of our system and we (naively) set $\xi = L$ at p_c . Then $O(p) \sim |p - p_c|^a \sim \xi^{-a/\nu}$ becomes

$$O(p_c) \sim L^{-a/\nu}, \quad (2.27)$$

relating the scaling of our observable in finite systems of size L at the critical point to the scaling in the infinite system near the critical point with the exponent $-a/\nu$.

¹⁰There is no discernible length scale in the system anymore and no matter how much we zoom in or out, the system will look identical. The system has a *fractal* structure at the transition.

In a more rigorous derivation, one assumes a scaling ansatz of the form

$$O(p) = L^{-a/\nu} F' \left(\frac{L^\nu}{\xi^\nu} \right) = L^{a/\nu} F \left([p - p_c] L^{1/\nu} \right), \quad (2.28)$$

where F is a scaling function depending on the observable and the process. This leads to the same relation with a constant factor $F[0]$.

With the above relation determines the ratios of critical exponents from measurements of the corresponding observables at the critical point in finite size systems. A similar argument also allows us to determine ν via measuring a finite size (pseudo) critical point $p_c(L)$, scaling as $p_c(L) - p_c \sim L^{-1/\nu}$ [147, 177]. $p_c(L)$ is usually defined via the observed maximum of the correlation length or certain observables such as the susceptibility or, correspondingly, the maximum of the mean cluster size (excluding the largest cluster) in percolation systems [80, 177]. In some cases analytic calculations to determine the critical point and some of the critical exponents are possible with the help of *renormalization group* approaches, relating the behavior of systems of different sizes L and L' at different states p and p' [80, 177].

The simple argument given above does not quite work in all dimensions, however. Specifically, the Bethe lattice or random networks are effectively infinite dimensional and thus have no linear extension L . In fact, it is well known that above an *upper critical dimension* d_c the critical exponents characterizing a phase transition become independent of the dimension and are identical to results for the mean field. In the context of percolation, this means the random graph where any two nodes may be connected. This critical dimension for percolation is $d_c = 6$ [84, 156]. Above this dimension, and thus also for random graphs, the system behaves *as if* it had dimension $d = d_c$. Thus, for random network percolation we set $L = N^{1/d_c}$ to obtain the scaling form

$$\begin{aligned} O(p_c) &\sim N^{-a/(d_c\nu)} \\ O(p) &\sim N^{-a/(d_c\nu)} F \left[(p - p_c) N^{1/(d_c\nu)} \right], \end{aligned} \quad (2.29)$$

resulting in the finite size scaling relation $S_1/N \sim N^{-\beta/(d_c\nu)}$ with $-\beta/(d_c\nu) = -1/3$ [197].

In fact, the rigorous mathematical analysis and universality of the scaling ansatz and the (hyper)scaling relations relating the different critical parameters and the dimension of the system are still subject of ongoing research. For example, the scaling above the upper critical dimension may depend on the chosen boundary conditions, resulting in different critical exponents, even if the systems are identical otherwise [79, 100, 196]. Still, for our application to random percolation finite size scaling is a well established and reliable method.

Beyond random percolation

Many extensions of random percolation are possible. Relevant for real world networks is the study of percolation in arbitrary, for example scale-free networks. The critical behavior is different to random network percolation and a macroscopic component already emerges at $p_c = 0$. In this sense, scale-free networks are resilient to random failures since a large connected component always exists. On the other hand, the study of targeted percolation, that means adding nodes or links in a specific order, shows that scale-free networks are less resilient than random networks under targeted attacks. Another famous example is the study of *interdependent* networks or networks of networks, where the function of a node in one network depends on a corresponding node in another network. An often cited example is the interdependence of the power grid and communication network, leading to a large blackout in Italy, analyzed in [35]. In such networks a small failure in one network can cause a cascade of failures in both networks eventually leading to the breakdown of connectivity of both networks in a discontinuous phase transition [22].

Another interesting finding that sparked a large amount of interest in percolation was the observation of *explosive percolation*. In an attempt to model control of the percolation transition Achlioptas et al. introduced a percolation rule where at each step two links are selected randomly. One of these links is then added to the network based on some selection rule [3, 147]. In this specific example, the *product rule* always selects the link that minimizes the product $S(i)S(j)$ of the sizes $S(i)$ and $S(j)$ of the two clusters connected by the link. Achlioptas et al. found a seemingly discontinuous transition, unexpected in random network percolation. Detailed study of the model and even rigorous mathematical analysis finally showed that the transition is continuous but behaves very differently from standard random percolation [58, 78, 150, 151]. The transition is extremely sudden with a critical exponent $\beta \approx 0.08$, compared to $\beta = 1$ for standard random percolation. Even in very large but finite systems the largest cluster S_1 increases in size by a large fraction of the total system size in a single step. This largest gap $\Delta S_1/N \sim N^{-b}$ disappears only very slowly with b around 0.05, depending on the specific model [126].

Based on this research many percolation models have been developed to reveal many more diverse phenomena and unexpected critical and super-critical behavior [63]. In particular, much effort has been focused on finding genuinely discontinuous percolation transitions, for example via direct control of the (largest) clusters in the system, and what properties models must have to allow for a discontinuous transition [10, 25, 42, 44, 49]. It was shown that any finite fraction of randomly added links during the percolation process will always lead to continuous percolation [163]. Similarly, the proof showing the continuity of explosive percolation shows the same for the general class of *Achlioptas processes*, where at each step a finite number of links is evaluated and one of those links must be added [150, 151].

Numerical implementation

The numerical implementation of percolation is not always straightforward, especially for very large systems. Most percolation models are simulated by adding links to the network (or lattice) one by one. By correctly weighting each step these results can be easily transformed to the percolation models discussed above, where each link is present or absent with a given probability p [129]. More complicated percolation models are, however, already formulated as a sequential link addition process and usually directly adopt the link density $l = M/N$ as the control parameter.

In order to track the sizes of the components, and avoid recalculating the connected components after each step, one typically uses a *union-find* structure (*disjoint set* data structure) for percolation on lattices. This algorithm allows fast updating of the cluster sizes for any underlying network structure when links are added [129]. For random networks, where any link between two nodes can be added, it is possible to simplify the simulation considerably with a small approximation. Instead of considering the individual nodes and the detailed network structure, we only consider the cluster size distribution n_S , that is the number n_S of clusters of size S . Effectively, we simulate the master equation of a cluster merging process instead of the detailed network structure [48, 150]. Of course this will not exactly reproduce the percolation process, since we potentially add a link that is already present in the network. However, for sufficiently large systems, this probability quickly disappears and only about 15 errors are made in a simulation with $N = 2^{20}$ nodes.

To take full advantage of this idea, we employ a data structure called a *binary tree* [128, 145] to store the clusters as pairs $(S, n(S))$. This structure allows us to quickly find and update clusters with a specific size. By keeping track of the total number of nodes $\sum_S S n(S)$ in all clusters in the sub-trees of each node of the binary tree, we can also very quickly select a random node and corresponding cluster. This algorithm is slightly improved by using two different structures for ordering the clusters and selecting a random cluster as well as by empirical adaptations depending on the specific percolation model, such as explicitly tracking very large clusters individually. Specifically, we use a *binary heap* [128] to pick a random cluster and a sorted binary tree to find clusters of a specific size [162].

2.2 Dynamical systems

Most systems, including many networks, are not static but change over time, they are *dynamic*. We describe a dynamical system by its *state* $x(t)$ and its evolution in time t [17, 87, 142, 178, 180]. In general, both the state as well as the time may be continuous or discrete variables. For example, when we describe the motion of a car or the motion of planets, where the state is simply the position in space, both time and space are continuous variables. On the other hand, when we describe the daily number of passengers on a bus, time is measured in discrete days and the number of passengers is measured in (hopefully) discrete persons. Dynamical systems theory provides tools to understand both kinds of systems but here we restrict the presentation to continuous time and state variables. This means we have a time $t \in \mathbb{R}$ and a state $x(t) \in \mathbb{R}$ at each time. In general, the state $\mathbf{x}(t) \in \mathbb{R}^n$ may be an n -dimensional vector-valued variable, denoted by bold-face notation, for example position and velocity in classical mechanical systems.

Mathematically, a dynamical system in continuous time is characterized by a *flow* Φ describing the evolution of the state $\mathbf{x}(t)$ of the system in time as $\mathbf{x}(t+T) = \Phi(\mathbf{x}(t), T)$ [87, 178]. The flow is typically given as the solution of a differential equation describing the change of the state $\mathbf{x}(t)$,

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}(t)). \quad (2.30)$$

In the following, we drop the explicit time dependence denoting the state only as \mathbf{x} when there is no possibility of confusion.

In this section we briefly present the basic concepts of attractors and (linear) stability with examples in one and two dimensional systems and then discuss higher dimensional chaotic systems. Finally, we outline the basic ideas of the numerical implementation to measure the exponents characterizing the linear stability.

Attractors and stability

In many cases we are interested in the long term behavior of the system. For example, when modelling a population of animals, we want to know whether the animals die out or the population settles to a fixed level and how this depends on the system parameters. We answer this question by considering states $x^*(t)$ that, once reached, do not change. Rigorously, we are looking for minimal closed *invariant sets* $\mathbb{X}^* \subset \mathbb{R}^n$. That means if a trajectory starts in \mathbb{X}^* , $\mathbf{x}(t) \in \mathbb{X}^*$, it remains in \mathbb{X}^* for all times, $\mathbf{x}(T > t) \in \mathbb{X}^*$. If all trajectories that start arbitrarily close to this set \mathbb{X}^* converge to \mathbb{X}^* for $t \rightarrow \infty$, the set is called an *attractor* [178]. This is quantified by the (*linear*) *stability* of \mathbb{X}^* . For example, we can think of at least two simple invariant states for the animals: (i) if there are no animals left, no new animals are born and the population remains extinct and (ii) if the same number of animals are born and die, the population remains constant. Which of the possibly many attractors a system settles on depends on the specific initial conditions.

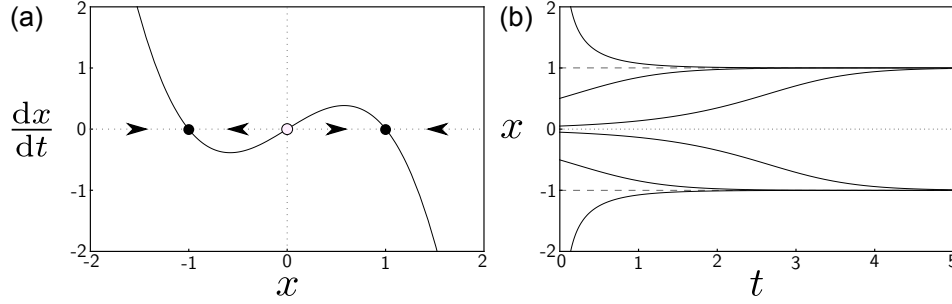


Figure 2.11: **Fixed points.** (a) Phase space plot of the dynamical system $dx/dt = x - x^3$. (b) Evolution of the trajectories for multiple initial conditions. The system has three fixed points given by the solutions to $f(x^*) = 0$, two of which are stable (filled dots, dashed lines) and one which is unstable (open circle, dotted line).

In a one dimensional system $dx/dt = f(x)$ the state x only either increases ($f(x) > 0$), decreases ($f(x) < 0$), or does not change at all ($f(x) = 0$). The points where x does not change are the only possible attractors, called *fixed points*, of the dynamics. As an example, consider the system

$$\frac{dx}{dt} = f(x) = x - x^3 \quad (2.31)$$

with three fixed points $x_1^* = -1$, $x_2^* = 0$ and $x_3^* = 1$. From the illustration in Fig. 2.11 we easily see that the system will move to either $x_1^* = -1$ or $x_3^* = 1$ for large t , these fixed points are called (*asymptotically*) *stable*. In contrast, the system will move away from the *unstable* fixed point $x_2^* = 0$ (unless it starts exactly at this fixed point) [87, 178].

To quantify the stability of the fixed points we consider an infinitesimal change $\delta x \ll 1$ away from the fixed point, $x = x^* + \delta x$. If the system returns to the fixed point from this initial condition, $x \rightarrow x^*$ ($\delta x \rightarrow 0$) for large times $t \rightarrow \infty$, the fixed point is considered linearly stable. The *linear stability* of the fixed point is described by the evolution of this infinitesimal change

$$\frac{d\delta x}{dt} = f(x^* + \delta x) - f(x^*) = \frac{df(x^*)}{dx} \delta x + \mathcal{O}(\delta x^2), \quad (2.32)$$

where we disregard the higher order terms. This system has the solution $\delta x(t) \sim \exp(\lambda t)$. The exponent $\lambda = df(x^*)/dx$ corresponding to this fixed point, called the *Lyapunov exponent*, characterizes the (linear) stability of the fixed point [87, 178]. Here we use the notation $df(x^*)/dx$ to indicate that the derivative should be evaluated at the point x^* . If $\lambda < 0$ a small perturbation decays and the system returns to the fixed point, the fixed point is *linearly stable*. If $\lambda > 0$ the system is unstable. For our example system Eq. (2.31), it is easy to verify the stability of the fixed points with exponents $\lambda_1 = -2$, $\lambda_2 = 1$ and $\lambda_3 = -2$, respectively.

A similar idea applies to fixed points in higher dimensional systems. Here, the linear system becomes

$$\frac{d\delta\mathbf{x}}{dt} = \frac{d\mathbf{f}(\mathbf{x}^*)}{d\mathbf{x}}\delta\mathbf{x} + \mathcal{O}(\delta\mathbf{x}^2) = J(\mathbf{x}^*)\delta\mathbf{x} + \mathcal{O}(\delta\mathbf{x}^2) \quad (2.33)$$

with the *Jacobian matrix* J evaluated at the fixed point \mathbf{x}^* and the solution $\delta\mathbf{x}(t) \sim \exp(J(\mathbf{x}^*)t)$. In this case each fixed point has multiple associated exponents given by the eigenvalues $\lambda_{1,2,\dots}$ of J . The fixed point is only stable if the real part of all eigenvalues is negative, otherwise it is unstable [87, 178].

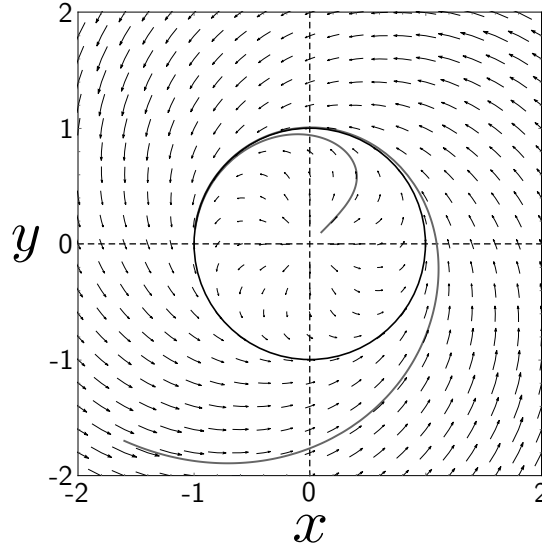


Figure 2.12: **Two dimensional dynamical systems.** Phase space plot of a stable limit cycle in the system $dr/dt = 1 - r^2$, $d\theta/dt = r^2$, given in polar coordinates. Arrows indicate the direction of the flow. The stability of the limit cycle illustrated with two example trajectories.

In two dimensional systems fixed points are not the only possible attractors: the system may move along closed *periodic orbits*, similar to the rhythmic motion of an undamped pendulum. As fixed points, these periodic orbits may be attracting or repelling nearby trajectories and are called *limit cycles* if there is a trajectory that approaches the orbit as $t \rightarrow \pm\infty$. One example of a two dimensional system with a limit cycle and one unstable fixed point is illustrated in Fig. 2.12. We determine the (linear) stability in a similar way as above in equation (2.33) by considering the evolution of a small perturbation. However, now the Jacobian $J(\mathbf{x}^*(t))$ along the periodic

orbit is not constant but changes over time.¹¹ The Poincaré-Bendixon theorem proves that in two dimensions fixed points and limit cycles are the only possible attractors [178]. Higher order systems admit more complicated dynamics, such as chaotic orbits, discussed in the next section.

Chaotic systems

Dynamics in three or more dimensions may be much more complicated than fixed points and limit cycles. Whereas a limit cycle in two dimensions separates the phase space into two disjoint parts, see Fig. 2.12, this is no longer true in three dimensions. In fact, the trajectories of a three dimensional system can settle on aperiodic, fractal attractors that have no volume but are more than surfaces. This is possible since the trajectories can still fold and circle around each other without ever crossing, in contrast to two dimensional systems. Perhaps even more surprising is the fact that nearby trajectories on these attractors diverge exponentially. This means, even though the system settles on an attractor, a small change in initial condition will grow over time, resulting in so called sensitive dependence on initial conditions. Such an attractor is called a *chaotic attractor*, as the dependence of the dynamics on the initial conditions seems random, even though it is defined by a deterministic differential equation. Due to their fractal geometries these attractors are also sometimes referred to as *strange* or *fractal attractors* [87, 178].

One example for chaotic dynamics is the Rössler system [155] given by

$$\begin{aligned}\frac{dx}{dt} &= -y - z, \\ \frac{dy}{dt} &= x + ay, \\ \frac{dz}{dt} &= b + z(x - c)\end{aligned}\tag{2.34}$$

with only one nonlinear equation. For parameters $a = 0.2$, $b = 0.2$ and $c = 5.7$ the system exhibits a chaotic attractor illustrated in Fig. 2.13. Calculating the linear stability of the attractor we find $|\delta\mathbf{x}| \sim \exp(\lambda t)$ with the real part of the largest Lyapunov exponent $\lambda \approx 0.07$, even though the attractor itself is stable. We discuss chaotic dynamics more in the context of synchronization below and in the articles in chapters 3 and 4.

¹¹A perturbation along the periodic orbit will necessarily result in – after one full period – constant displacement and thus an exponent of $\lambda = 0$ associated with any periodic orbit. Therefore, to determine the (linear) stability of the orbit one should consider only *transverse* perturbations perpendicular to the orbit.

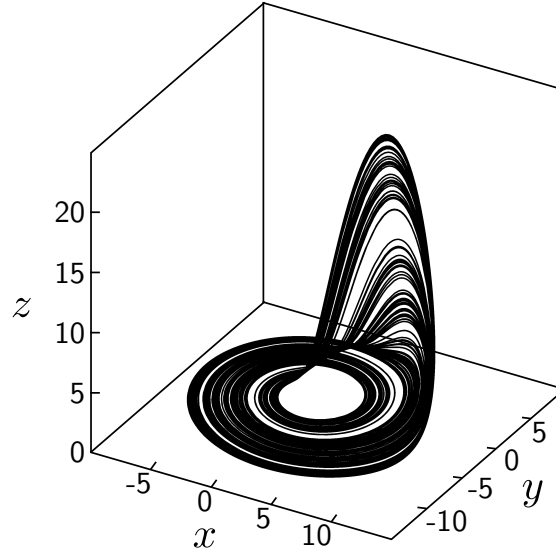


Figure 2.13: **Chaotic systems.** Chaotic attractor of the Rössler system described by equation (2.34).

Numerical implementation

For the numerical solution of the differential equations we use either the standard fourth order Runge-Kutta algorithm or a fifth order Runge-Kutta algorithm with adaptive time steps [145].

In order to calculate the largest Lyapunov exponent from those simulations we directly work in the corresponding linear system [19]. This means we consider a perturbation $\delta \mathbf{x}$ with time evolution given by

$$\frac{d\delta \mathbf{x}}{dt} = J[\mathbf{x}^*(t)] \delta \mathbf{x} \quad (2.35)$$

along the trajectory $\mathbf{x}^*(t)$. Writing $\delta \mathbf{x} = \exp(r) \mathbf{u}$ we separate the perturbation into its amplitude $\exp(r)$ and its direction \mathbf{u} . With the additional constraint $\mathbf{u} \cdot \mathbf{u} = 1$, where \cdot denotes the scalar product of two vectors, the derivative $d\mathbf{u}/dt$ must be orthogonal to \mathbf{u} and we find

$$\begin{aligned} \frac{dr}{dt} &= (J[\mathbf{x}^*(t)] \mathbf{u}) \cdot \mathbf{u} \\ \frac{d\mathbf{u}}{dt} &= J[\mathbf{x}^*(t)] \mathbf{u} - \frac{dr}{dt} \mathbf{u}. \end{aligned} \quad (2.36)$$

From these variables we directly calculate the real part of the largest Lyapunov exponent as $\lambda = \lim_{t \rightarrow \infty} r(t)/t$ when starting with $r(0) = 0$ as initial condition.

2.3 Networked dynamical systems

So far we have discussed the structure of networked systems and the dynamics of individual units. Often, these two aspects are not separate. Many networked systems are not static but change their state or structure over time. Similarly, most dynamical systems are not isolated but interact with many other systems. *Networked dynamical systems* refer to such systems where the nodes are dynamical systems and the links in the network describe the interactions between the units [17, 180].¹² For example, we could describe the population of animals in different geographical areas with identical population dynamics in each area and coupling via a network of migration between these areas [18, 143]. Other examples include the dynamics of epidemics, where coupling is given by the transportation network [32, 38], the dynamics of power grids, where units are individual generators coupled by transmission lines [152, 157, 195], or the activity of neural networks, where the interactions describe signal transmission between neurons via synapses [59, 96].

In general, we consider a network of N coupled units with adjacency matrix A , $A_{ij} \in \{0, 1\}$. The dynamics of an individual unit i are then given by

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{f}_i(\mathbf{x}_i) + \sum_{j=1}^N A_{ij} \mathbf{g}_{ij}(\mathbf{x}_i, \mathbf{x}_j), \quad (2.37)$$

for $i \in \{1, 2, \dots, N\}$ and $\mathbf{x}_i \in \mathbb{R}^n$. The function \mathbf{f}_i describes the internal dynamics of the unit and the function \mathbf{g}_{ij} describes the coupling between the units, summed over all nodes with a connection to i [128]. Typically, the internal dynamics and the coupling functions are identical for all nodes. In particular, we only consider systems with identical coupling $\mathbf{g}_{ij} = \mathbf{g}$ and $\mathbf{f}_i(\mathbf{x}_i) = \mathbf{f}(\mathbf{x}_i, \omega_i)$, where the functional form of the dynamics is identical for all nodes but may depend on different parameters ω_i .

There are also many examples where the structure of the network changes dynamically, sometimes depending on the dynamics of individual units. The most prominent examples are the brain, where individual neurons form new synapses based on their mutual activity, and our continuously changing social networks [9, 59, 89, 109]. However, in this section we consider only networks with a fixed topology and mainly discuss synchronization phenomena in networks of dynamical units described by Eq. (2.37). First, we discuss phase locking and synchronization in the Kuramoto model, then we consider synchronization of chaotic units and stability in these systems.

¹²There is no fundamental difference between a high dimensional dynamical system and a network formed by many small, coupled dynamical systems. However, often the individual units have very similar or even identical dynamics, giving these systems additional structure. This allows us to apply tools from both network science as well as dynamical systems theory to study these systems.

2.3.1 Synchronization

One of the most common collective dynamical phenomena is *synchronization*, where different units align their behavior such that they are in the same state at the same time [12, 142]. There are many examples such as clapping, singing, or dancing to a song, marching to a beat, and even our circadian rhythm synchronizing to the day-night cycle. In these cases the behavior is determined by an external driving force. Interestingly, in some systems synchronization also emerges without an external rhythm, seen in the famous example of synchronization of the walking of pedestrians on London's Millennium Bridge [181]. Many more biological systems show synchronization, including cell activity and the blinking of fireflies synchronizing to a common on-off frequency [12, 178]. Synchronization also plays an important role in many technical systems, for example for the function of communication and positioning system such as GPS [120]. Due to the strongly interconnected financial markets, even the prices of goods, stocks, or different currencies are related and sometimes exhibit weaker forms of synchronization [174]. In more abstract models, synchronization is interpreted as the emergence of consensus in opinion formation [17].

Mathematically, synchronization between two units i and j means that their state is identical $\mathbf{x}_i(t) = \mathbf{x}_j(t)$ for all times t . This is called *identical* or *complete synchronization*. Often this condition is not exactly fulfilled and there are extensions to the definition of synchrony that also cover weaker forms of synchronization [142].

Phase oscillators, phase locking and the Kuramoto model

The examples above show that synchronization is a ubiquitous phenomenon in a broad range of systems. For a general understanding of synchronization, we need to understand what all of these systems have in common. Individually, many of the systems mentioned above are *limit cycle oscillators*. They exhibit a stable limit cycle, such as the system shown in Fig. 2.12. If the limit cycle is strongly attractive, the fluctuations around the stable periodic orbit are small. We can then describe the state \mathbf{x} of such a system with a single scalar variable, its *phase* $\phi_i \in [0, 2\pi)$ along the periodic orbit, ignoring the amplitude of the oscillations. Such a system is called a *phase oscillator*. Consequently, for sufficiently weak coupling the system will remain close to its limit cycle and interactions between two units are also limited to functions of their phases ϕ_i and ϕ_j along the periodic orbit [2].

Synchronization between two phase oscillators i and j then simply means $\phi_i(t) = \phi_j(t)$ for all times t . A weaker condition does not require identical phases but only identical, typically constant frequencies $d\phi_i(t)/dt = d\phi_j(t)/dt$, such that the difference $\phi_i(t) - \phi_j(t)$ is constant. In this state the two oscillators are called (*fully*) *phase locked* [2, 178]. This is relevant, for example, for power grids, where all generators should run with identical frequency to guarantee stable operation [68].

The *Kuramoto model* is one of the most basic models to describe coupled phase oscillators [2, 110, 111, 179]. It consists of a collection of N oscillators, each moving with an intrinsic frequency ω_i and interacting with attractive sinusoidal coupling with coupling strength K . In the original model, all units are coupled to all other units

$$\frac{d\phi_i}{dt} = \omega_i + \frac{K}{N} \sum_{j=1}^N \sin(\phi_j - \phi_i) . \quad (2.38)$$

In terms of our definition of networked dynamical systems in equation (2.37) this means $f_i(\phi_i) = \omega_i$ and $g(\phi_i, \phi_j) = K/N \sin(\phi_j - \phi_i)$. For weak coupling the oscillators move with their individual frequency. As the coupling becomes stronger individual oscillators become *partially phase locked*, that means their phase differences become bounded and the oscillators move with the same average frequency. For sufficiently large coupling all oscillators become fully phase locked with constant phase differences and finally synchronize completely as $K \rightarrow \infty$.

In order to study the phase locking and synchronization, Kuramoto introduced an order parameter to quantify the phase coherence of the oscillators. Interpreting the phases of the oscillators as phases of complex variables with absolute value 1, we define

$$r(t) \exp[i\Phi(t)] = \frac{1}{N} \sum_{j=1}^N \exp(i\phi_j(t)) \quad (2.39)$$

with the collective phase $\Phi(t)$ and amplitude $r(t)$. If the phases ϕ_i are disordered their contributions cancel out and $r(t)$ is small. If all phases are similar $r(t)$ becomes larger and approaches $r(t) \rightarrow 1$ in the limit of complete synchronization. The order parameter is then defined as the long time average $r = \langle r(t) \rangle_t$, measuring the average phase coherence of the oscillators. In the limit of $N \rightarrow \infty$ mean field calculations show that there exists a critical coupling strength K_c below which no phase locking takes place, $r(K \leq K_c) = 0$. At K_c some oscillators become partially phase locked and the order parameter becomes positive [2, 12, 179]. The system undergoes a second order phase transition, similar to the percolation transition in random networks, illustrated in Fig. 2.14.

Kuramoto oscillators will also synchronize if the coupling network is not completely connected. However, synchronization emerges in a different way, depending on the network structure. As in percolation, individual clusters of connected nodes with similar phases become (partially) phase locked first. These clusters slowly grow and merge as the coupling strength increases [77, 91, 148, 149]. Again, it is possible to estimate the critical coupling depending on the degree distribution with mean field calculations and the definition of the order parameter can be extended to take into account the specific network topology. An overview of the Kuramoto model, the order parameter, and its extensions to complex coupling networks is given in the introduction of the article in chapter 5.

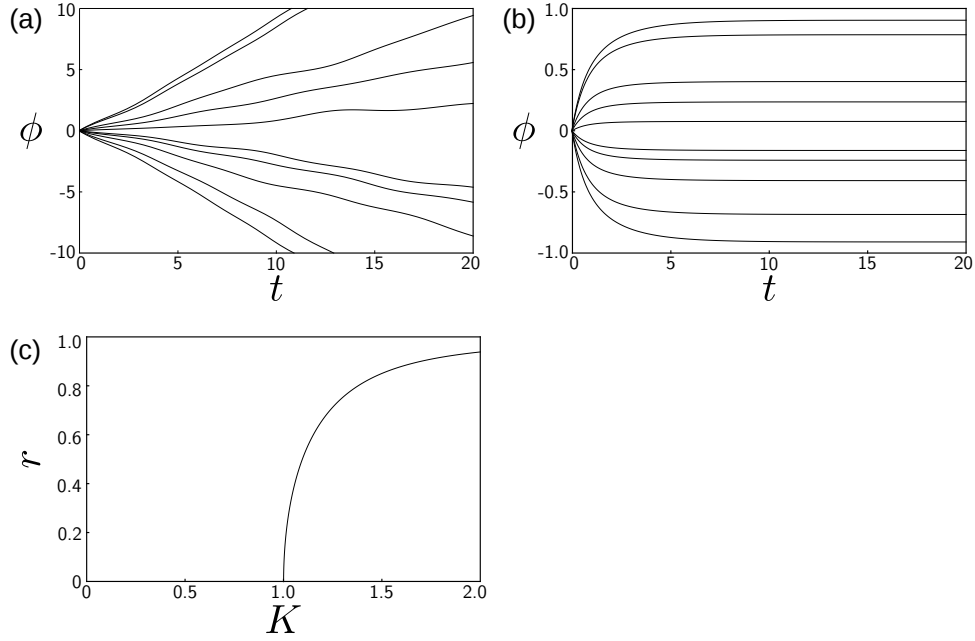


Figure 2.14: **Kuramoto model of coupled limit cycle oscillators.** (a,b) Trajectories of $N = 10$ coupled Kuramoto oscillators for weak ($K = 0.5$) and strong ($K = 1.5$) coupling, respectively. (c) Order parameter r in the limit of $N \rightarrow \infty$ with frequencies distributed normally with mean 0 and variance $\pi/8$ [2, 12, 179].

2.3.2 Chaos synchronization

In principle it is easy to imagine that any two identical, coupled systems synchronize eventually, even if the coupling strength is small. However, chaotic systems discussed above challenge this intuition. Due to their chaotic dynamics, even a small initial difference in the state of two identical chaotic systems grows exponentially. Perhaps surprisingly, it is still possible to synchronize chaotic systems [138, 140, 141]. In fact, depending on the system and the coupling, chaotic systems exhibit a variety of different types of synchronization. These relations between the units' states range from complete synchronization over *lag synchronization*, where one unit follows another unit with a time lag $\mathbf{x}_i(t) = \mathbf{x}_j(t + \tau)$ even when there is no explicit delay in the coupling itself to arbitrary functional relations $\mathbf{x}_i = \mathbf{h}(\mathbf{x}_j)$ in *generalized synchronization* of possibly non-identical units [56, 140, 142, 154, 159, 178].¹³ Here, however, we focus only on complete synchronization of identical chaotic systems.

¹³The phase locking of limit cycle oscillators is a form of lag synchronization, as one oscillators will lag behind the other by a fixed phase, which means by a fixed time if they rotate with constant, non-zero frequency.

To illustrate chaotic synchronization we consider two coupled Rössler oscillators i and j coupled only in their x -coordinate

$$\begin{aligned}\frac{dx_i}{dt} &= -y_i - z_i + K(x_j - x_i), \\ \frac{dy_i}{dt} &= x_i + ay_i, \\ \frac{dz_i}{dt} &= b + z_i(x_i - c),\end{aligned}\tag{2.40}$$

and equivalently for the other oscillator j . The two systems become completely synchronized for sufficiently large coupling K but desynchronize again when K becomes too large. Fig. 2.15 shows the largest Lyapunov exponent of the synchronized state. This non-trivial behavior depends strongly on the system and its parameters as well as the topology of the interaction network and the type of coupling between the units. However, in the next section we will see how it is possible to determine the stability of the synchronized state separating the network topology from the specifics of the dynamical system. This allows us to understand the stability of the synchronized state in any network in relation to a system of only two coupled units.

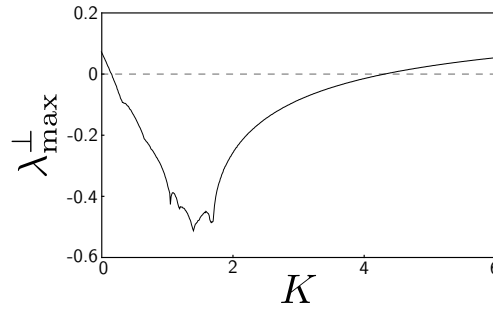


Figure 2.15: **Synchronization of chaotic oscillators.** Stability of the completely synchronized state of two x -coupled Rössler oscillators [Eq. (2.40)] characterized by the largest transverse Lyapunov exponent λ_{\max}^{\perp} depending on the coupling strength K .

Master stability function

A big question when dealing with large networks of oscillators is how to control or synchronize them efficiently. Depending on the task, one can find many connections between the structure of the interaction network and the synchronizability of the units [16, 71, 115, 124]. The *master stability function* relates the stability of the synchronized state in any network to the topology of the interaction network and the stability of the synchronized state of two coupled units [128, 139].¹⁴ In the following

¹⁴The idea of the master stability function is also described in the articles in chapter 3 and 4 for the specific model studied.

we assume a network of identical coupled dynamical systems described by

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{f}(\mathbf{x}_i) + \sum_{j=1}^N A_{ij} \mathbf{g}(\mathbf{x}_i, \mathbf{x}_j), \quad (2.41)$$

where $\mathbf{x}_i \in \mathbb{R}^n$. We assume diffusive coupling with coupling strength K and constant coupling matrix $C \in \mathbb{R}^{n \times n}$ with $C_{kl} \in \{0, 1\}$ such that

$$\mathbf{g}(\mathbf{x}_i, \mathbf{x}_j) = KC(\mathbf{x}_j - \mathbf{x}_i). \quad (2.42)$$

We write $\tilde{\mathbf{X}} = (\mathbf{x}_1, \dots, \mathbf{x}_N)^T$ and similarly $\tilde{\mathbf{F}}(\tilde{\mathbf{X}}) = (\mathbf{f}(\mathbf{x}_1), \dots, \mathbf{f}(\mathbf{x}_N))^T$ such that $\tilde{\mathbf{X}}_{ik} = (\mathbf{x}_i)_k$ and $\tilde{\mathbf{F}}(\tilde{\mathbf{X}})_{ik} = (\mathbf{f}_i(\mathbf{x}_i))_k$ for $i \in \{1, 2, \dots, N\}$ and $k \in \{1, 2, \dots, n\}$. Equation (2.41) then becomes

$$\frac{d\tilde{\mathbf{X}}}{dt} = \tilde{\mathbf{F}}(\tilde{\mathbf{X}}) + K(L \otimes C)\tilde{\mathbf{X}}, \quad (2.43)$$

where $L = A - \text{diag}(k_1, \dots, k_N) \in \mathbb{R}^{N \times N}$ is the *Laplacian matrix* of the coupling network and \otimes denotes the direct product

$$(L \otimes C) = \begin{pmatrix} -k_1 C & A_{12} C & \dots \\ A_{21} C & \ddots & \\ \vdots & & \end{pmatrix} \quad (2.44)$$

such that $(L \otimes C)_{ijkl} = L_{ij} C_{kl}$ with $i, j \in \{1, 2, \dots, N\}$ and $k, l \in \{1, 2, \dots, n\}$.¹⁵ Given a synchronous state $\tilde{\mathbf{X}}^* = (\mathbf{x}^*, \mathbf{x}^*, \dots)^T$, small perturbations around this state are subject to the linear dynamics

$$\frac{d\delta\tilde{\mathbf{X}}}{dt} = [\mathbf{D}\tilde{\mathbf{F}}(\tilde{\mathbf{X}}^*) + K(L \otimes C)]\delta\tilde{\mathbf{X}} = \tilde{\mathbf{J}}(\tilde{\mathbf{X}}^*)\delta\tilde{\mathbf{X}}. \quad (2.45)$$

The important observation here is that both parts of $\tilde{\mathbf{J}}(\tilde{\mathbf{X}}^*)$ have an $N \times N$ block structure and $\mathbf{D}\tilde{\mathbf{F}}(\tilde{\mathbf{X}}^*) = \mathbf{1}_{N \times N} \otimes \mathbf{D}\mathbf{f}(\mathbf{x}^*)$ is diagonal in this block structure. Thus, it is possible to diagonalize the system using the diagonalization of L without affecting the block-diagonal first part. This separates the perturbation into the eigenvectors ξ_k to the eigenvalues γ_k of L with dynamics

$$\frac{d\xi_k}{dt} = [\mathbf{D}\mathbf{f}(\mathbf{x}^*) + K\gamma_k C]\xi_k, \quad (2.46)$$

for $k \in \{1, 2, \dots, N\}$. This is the same as the variational equation of a system of two coupled units with coupling strength $K\gamma_k$. This means the stability of a synchronous state in an arbitrary network of coupled identical units with coupling strength K is defined by the largest Lyapunov exponents $\lambda(K\gamma_k)$ of a system of two coupled units with coupling strengths $K\gamma_k$ given by the eigenvalues of the Laplacian of the interaction network.

¹⁵The product $(L \otimes C)\tilde{\mathbf{X}}$ is then given by $((L \otimes C)\tilde{\mathbf{X}})_{ik} = \sum_{j=1}^N \sum_{l=1}^n L_{ij} C_{kl} (\mathbf{x}_j)_l$

Chapter 3

Original manuscript: Transient uncoupling induces synchronization

Citation

Schröder, M., Mannattil, M., Dutta, D., Chakraborty, S., and Timme, M. (2015), Transient uncoupling induces synchronization, *Phys. Rev. Lett.* **115**, 054101

DOI: 10.1103/PhysRevLett.115.054101

©2015 by the American Physical Society. Permission to reprint this article in the authors' thesis is explicitly granted by the American Physical Society.

Original contribution

Design of the research was done together with M. Timme and S. Chakraborty. I performed most simulations and calculations, analyzed the data and generated all figures in the main manuscript (Fig. 2 and 5 based on initial contributions of M. Mannattil). I wrote main parts of the manuscript with M. Timme. I designed and wrote the supplemental information and performed all simulations and created all figures. I revised the manuscript during the review process.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder

Transient Uncoupling Induces Synchronization

Malte Schröder,^{1,†} Manu Mannattil,^{2,‡} Debabrata Dutta,^{3,*§} Sagar Chakraborty,^{2,4,||} and Marc Timme^{1,¶}

¹*Network Dynamics, Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany*

²*Department of Physics, Indian Institute of Technology Kanpur, Uttar Pradesh 208016, India*

³*S.N. Bose National Centre for Basic Sciences, Saltlake, Kolkata 700098, India*

⁴*Mechanics and Applied Mathematics Group, Indian Institute of Technology Kanpur, Uttar Pradesh 208016, India*

(Received 26 March 2015; published 29 July 2015)

Finding conditions that support synchronization is a fertile and active area of research with applications across multiple disciplines. Here we present and analyze a scheme for synchronizing chaotic dynamical systems by transiently *uncoupling* them. Specifically, systems coupled only in a fraction of their state space may synchronize even if fully coupled they do not. While for many standard systems coupling strengths need to be bounded to ensure synchrony, transient uncoupling removes this bound and thus enables synchronization in an infinite range of effective coupling strengths. The presented coupling scheme therefore opens up the possibility to induce synchrony in (biological or technical) systems whose parameters are fixed and cannot be modified continuously.

DOI: 10.1103/PhysRevLett.115.054101

PACS numbers: 05.45.Xt, 05.45.Gg

Synchronization is one of the most prevalent collective phenomena in coupled dynamical systems [1]. Synchronization and related consensus phenomena have been frequently found in biological, ecological, physical, engineering, and social systems such as in predator-prey dynamics, the spread of epidemics, the migration of large populations, systems of self-driven particles, and systems of social or technical dynamics [2–12]. For chaotic systems, synchronization typically emerges only within a specific range of coupling strengths and is impossible otherwise [1,13–15].

In this Letter, we propose and analyze a way of inducing synchronization between coupled chaotic oscillators by transient uncoupling: If the system is in a certain predefined subset of its state space, coupling is active; otherwise it is inactive. We systematically study the dependence of successful synchronization on the fraction of state space where coupling is active. Synchronization may emerge even for systems that do not synchronize when coupled continuously in time (i.e., standard coupling). Furthermore, the system may synchronize for an infinite range of coupling strengths, even though this is often not possible for ordinarily coupled chaotic systems. A systematic numerical analysis reveals how transverse stability properties vary across the attractor with the location of active coupling, not only between more or less stable synchrony, but all the way from stability to instability for the same system. This demonstrates that transient uncoupling modifies the collective dynamics in a nontrivial way. These results may find applications in inducing synchrony in systems whose local coupling parameters cannot be continuously varied with ease, but only switched on or off.

Standard coupling.—To start, consider a system of two unidirectionally coupled chaotic oscillators

$$\frac{d\mathbf{x}_1}{dt} = \mathbf{F}(\mathbf{x}_1), \quad (1)$$

$$\frac{d\mathbf{x}_2}{dt} = \mathbf{F}(\mathbf{x}_2) + \alpha \mathbf{C} \times (\mathbf{x}_1 - \mathbf{x}_2), \quad (2)$$

where $\mathbf{x}_1(t), \mathbf{x}_2(t) \in \mathbb{R}^d$ denote the states of the driving and driven unit, respectively, \mathbf{C} is a square coupling matrix, and α is the coupling constant that determines the overall strength of coupling [13]. As an explicit example throughout this Letter we consider identical x -coupled Rössler oscillators defined by $\mathbf{F}(\mathbf{x}) = (- (y + z), x + ay, b + z(x - c))^T$ [16] and $\mathbf{C} \in \mathbb{R}^{3 \times 3}$, where $\mathbf{C}_{ij} = 1$ for $i = j = 1$ and $\mathbf{C}_{ij} = 0$ otherwise. Further, $a = b = 0.2$, $c = 5.7$, and we take $\mathbf{x}_i = (x_i, y_i, z_i)^T$ as a convenient notation. Other chaotic systems exhibit qualitatively the same phenomena as those presented below [17].

Depending on the coupling strength α , such systems do or do not synchronize towards $\mathbf{x}_1(t) = \mathbf{x}_2(t) =: \mathbf{x}_S(t)$. In particular, like many other coupled chaotic systems, Rössler oscillators are known to typically synchronize for intermediate coupling strengths α , but not if coupled too strongly or too weakly [Figs. 1(a)–1(c)].

These qualitative synchronization properties depend on the (“transverse”) dynamics of the difference $\mathbf{x}_\perp = \mathbf{x}_1 - \mathbf{x}_2$. A Taylor expansion to first order in the $(\mathbf{x}_\perp)_i$ yields

$$\begin{aligned} \dot{\mathbf{x}}_\perp &= \mathbf{F}(\mathbf{x}_1) - \mathbf{F}(\mathbf{x}_2) - \alpha \mathbf{C} \times (\mathbf{x}_1 - \mathbf{x}_2) \\ &\approx [\mathbf{J}(\mathbf{x}_S(t)) - \alpha \mathbf{C}] \mathbf{x}_\perp, \end{aligned} \quad (3)$$

where $\mathbf{J}(\mathbf{x}) = \partial_{\mathbf{x}} \mathbf{F}(\mathbf{x})$ is the local Jacobian of \mathbf{F} . For the system (3) to relax to $\mathbf{x}_\perp(t) \rightarrow 0$, its maximum transverse Lyapunov exponent

$$\lambda_{\max}^\perp = \lim_{t \rightarrow \infty} \frac{1}{t} \ln \frac{|\mathbf{x}_\perp(t)|}{|\mathbf{x}_\perp(0)|} \quad (4)$$

needs to be negative [1]. Figure 1(d) illustrates λ_{\max}^\perp as a function of the coupling constant α . This clearly links,

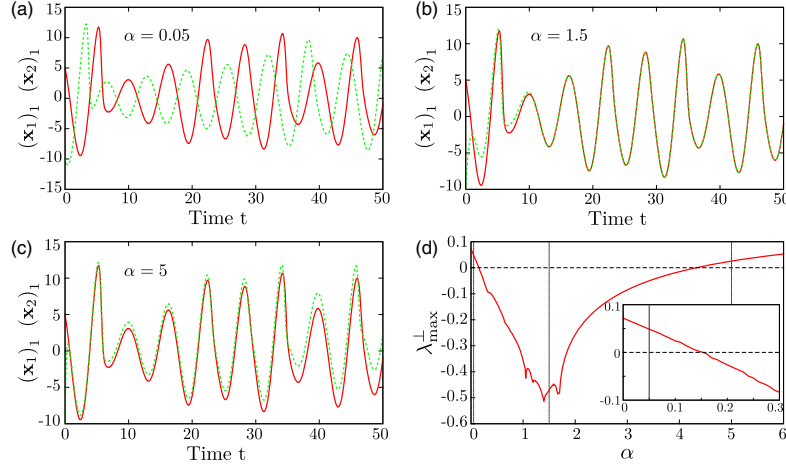


FIG. 1 (color online). Synchronization depends on coupling strength. Trajectories of the driving (solid line) and driven (dashed line) unit of two coupled chaotic oscillators for (a) $\alpha = 0.05$, (b) $\alpha = 1.5$, and (c) $\alpha = 5$ as indicated in panel (d). (d) The maximum transverse Lyapunov exponent λ_{\max}^{\perp} indicates synchronization for intermediate coupling only.

in a quantitative way, the coupling strength and the qualitative changes in collective dynamics observed before [Figs. 1(a)–1(c)].

Transient uncoupling.—We now introduce transient uncoupling via a factor

$$\chi_A(\mathbf{x}_2) = \begin{cases} 1 & \text{for } \mathbf{x}_2 \in A; \\ 0 & \text{for } \mathbf{x}_2 \notin A \end{cases} \quad (5)$$

in the coupling term,

$$\frac{d\mathbf{x}_2}{dt} = \mathbf{F}(\mathbf{x}_2) + \alpha \chi_A(\mathbf{x}_2) \mathbf{C} \times (\mathbf{x}_1 - \mathbf{x}_2). \quad (6)$$

Here $A \subseteq \mathbb{R}^d$ is a subset of the driven unit's state space where coupling is active. The two units are thus effectively coupled only within a subset $\mathbb{R}^d \times A$ of their common state space. For $A = \mathbb{R}^d$, the units are ordinarily coupled continuously in time.

Practically relevant subsets A are defined by clipping a region of state space along the direction of a particular coordinate axis,

$$A_{\Delta} = \{\mathbf{x}_2 \in \mathbb{R}^d : |(\mathbf{x}_2)_1 - (\mathbf{x}_2^*)_1| \leq \Delta\}, \quad (7)$$

where \mathbf{x}_2^* is a suitable point and the subscript “1” refers to the first coordinate of \mathbf{x}_2 and \mathbf{x}_2^* . Thus, coupling is only active within a column of width 2Δ centered around $(\mathbf{x}_2^*)_1$. Here, $(\mathbf{x}_2^*)_1 = 1.2$ was chosen as the center of the attractor in the x direction. An example realization for Rössler oscillators is illustrated in Fig. 2.

Such transient uncoupling modifies the collective dynamics of the coupled system in a nontrivial way (Figs. 3 and 4). Specifically, for a fixed coupling strength α , for which standard coupling would not lead to synchronization, clipping in an intermediate interval Δ induces synchronization.

Obviously, for $\Delta \rightarrow 0$ the units become completely uncoupled and cannot synchronize. Similarly, for no clipping $\Delta \rightarrow \Omega/2$ (where Ω is the width of the attractor along the clipping direction) we reobtain the original system with standard coupling that does not synchronize. For intermediate clipping, however, we find stable synchronization. As the clipping fraction Δ becomes just one additional parameter of the system, we expect the Lyapunov exponent to vary continuously with respect to Δ . An analysis of the transverse Lyapunov exponent as a function of the clipping fraction $\Delta' = 2\Delta/\Omega$ confirms this (Fig. 3).

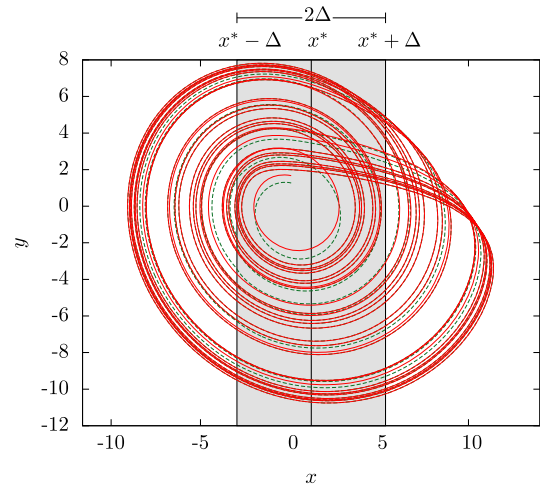


FIG. 2 (color online). Transient uncoupling through state-space clipping. The dynamics of two synchronized chaotic oscillators in the x - y plane with $x^* = (\mathbf{x}_2^*)_1 = 1.20$, $\Delta = 4.16$, and $\alpha = 7.0$ (driving: solid curve; driven: dashed curve). Coupling is only active in the interval $x_2 \in [x^* - \Delta, x^* + \Delta]$ (shaded in gray).

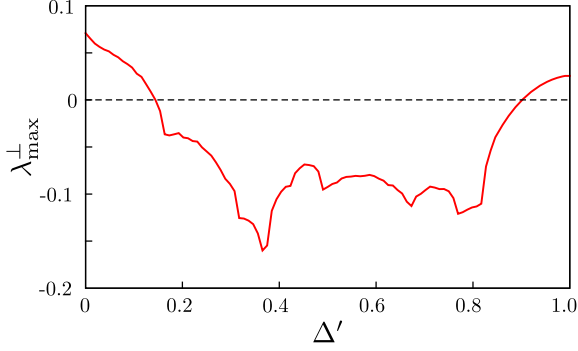


FIG. 3 (color online). Synchronization induced by transient uncoupling. Maximum transverse Lyapunov exponent for two transiently uncoupled chaotic oscillators (for parameters see text) for $\alpha = 5$ and clipping with $(\mathbf{x}_2^*)_1 = 1.20$. Synchronization emerges for moderate clipping, i.e., for intermediate values of Δ' , though not without clipping ($\Delta' = 1$).

Intriguingly, we find that for a fixed clipping interval Δ , the dependence on the coupling strength α is changed not only quantitatively but also qualitatively (compare Fig. 4 to Fig. 1). In particular, for intermediate transient uncoupling (intermediate values of Δ), synchrony emerges in an infinite range of coupling strengths α , thus in particular for arbitrarily large coupling [Fig. 4(d)]. This is in contrast to many chaotic oscillators which, when ordinarily coupled, exhibit an upper bound above which synchronization fails [13]. In fact we explicitly checked that the same phenomenon also emerges in Rössler oscillators with other parameters, and in pairs of coupled Lorenz and coupled Chen oscillators as well as for larger networks [17].

Optimal uncoupling.—We now analyze direction dependencies of transient uncoupling. Interestingly, the range of coupling strengths α for which the system synchronizes increases when the clipping fraction decreases from $\Delta' = 1$, as Fig. 5(a) illustrates. Moreover, the range of clipping fractions for which synchronization emerges depends on the exact direction in state space along which clipping is applied. For instance, clipping along the x axis seems more synchronizing in this sense than clipping along the y axis [compare Fig. 5(a) to Fig. 5(b)]. Oblique directions exhibit even broader ranges of clipping fractions where synchrony emerges [Fig. 5(c)].

In fact, certain directions of clipping are optimal. Because of the shape of the attractor, excursions of trajectories that substantially vary z are rare compared to those that vary the other two coordinates. Thus, clipping is desirable in the x - y plane. To quantify the effectiveness of clipping depending on its direction in the x - y plane, we measure the fraction of clipping

$$S(\theta) = \int_0^1 s(f, \theta) df \quad (8)$$

for which the system synchronizes when α is fixed. Here, we have measured the angle θ counterclockwise from the x axis and have defined the synchrony indicator

$$s(f, \theta) = \begin{cases} 1 & \text{for } \lambda_{\max}^{\perp} < 0; \\ 0 & \text{for } \lambda_{\max}^{\perp} \geq 0 \end{cases} \quad (9)$$

and the temporal clipping fraction

$$f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \chi_A(\mathbf{x}_2(t)) dt \quad (10)$$

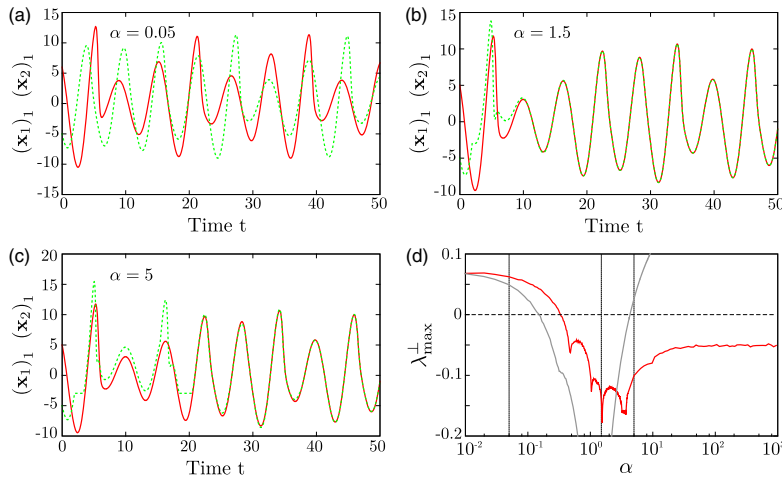


FIG. 4 (color online). Transient uncoupling induces synchronization in an infinite range of coupling strengths. Trajectories of the driving (solid line) and driven (dashed line) units for (a) $\alpha = 0.05$, (b) $\alpha = 1.5$, and (c) $\alpha = 5$, the same as in Fig. 1. The clipping is given by Eqs. (5) and (7) with $(\mathbf{x}_2^*)_1 = 1.20$ and $\Delta = 4.16$ as in Fig. 2. (d) Maximum transverse Lyapunov exponent λ_{\max}^{\perp} as a function of the coupling strength α ; note the logarithmic scale. The gray line shows λ_{\max}^{\perp} for normal, unclipped coupling. With transient uncoupling, synchronization is stable for arbitrarily large coupling strengths.

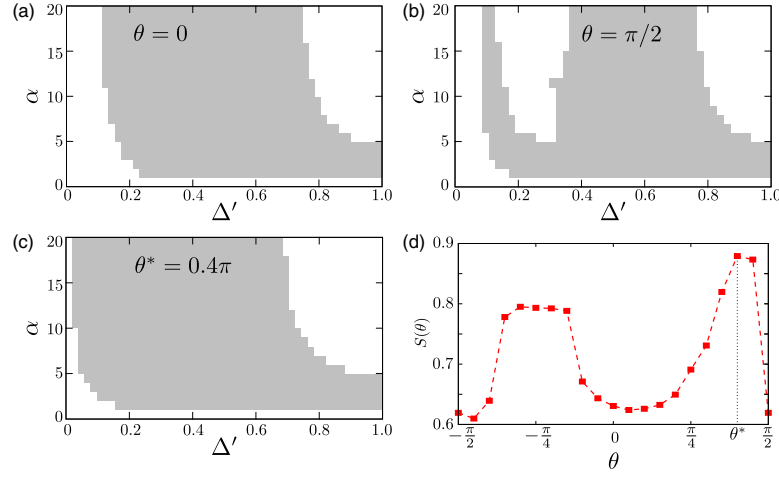


FIG. 5 (color online). Extended synchronization range by transient uncoupling and optimal clipping. (a)–(c) Depending on the coupling strength α and the percentage ($\Delta' = 2\Delta/\Omega$) of the state space where coupling is active, the system may or may not synchronize. The dark area marks the parameters where the synchronized state is stable, i.e., $\lambda_{\max}^{\perp} < 0$. Clipping is (a) in x direction [$x^* = (\mathbf{x}_2^*)_1 = 1.2$], (b) in y direction [$y^* = (\mathbf{x}_2^*)_2 = -1.5$], and (c) in the direction $y \approx 3.1x$ ($\theta^* = 0.4\pi$). Accordingly, the direction of clipping can be optimized to achieve the largest possible clipping range. (d) Effectiveness $S(\theta)$ [Eq. (8)] of clipping along the direction θ for fixed $\alpha = 10$.

such that larger values of $S(\theta)$ indicate that synchronization emerges in a larger range of clipping fractions.

The curve $S(\theta)$ has two local maxima [Fig. 5(d)], indicating two locally optimal clipping directions, one of which is globally optimal (at $\theta^* \approx 0.4\pi$). Why is there such a complicated dependence on direction?

Transverse stability depends on uncoupling location.—We can better understand the synchronization of the system by characterizing the stability for a family of clipping functions $\chi(\mathbf{x}) = \chi_A(\mathbf{x})$ where

$$A = A_{\mathbf{x}_2^*, r} := \{\mathbf{x}_2 \in \mathbb{R}^d : |\mathbf{x}_2 - \mathbf{x}_2^*| \leq r(\mathbf{x}_2^*)\}. \quad (11)$$

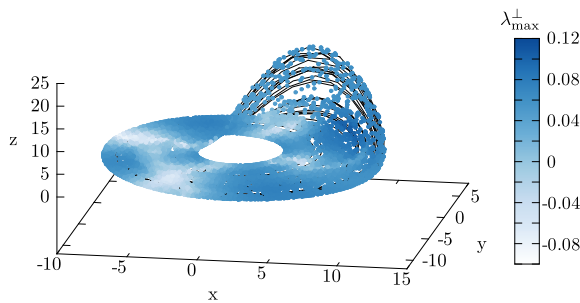


FIG. 6 (color online). Multiple switches between stability and instability depending on the coupling location. The color of the points indicates the maximum transverse Lyapunov exponent of the system with clipping to sets $A = A_{\mathbf{x}_2^*, r}$ [Eq. (11)], indicating stable synchronization (light) and no synchronization (dark) depending on the coupling location on the attractor. Parameters are $\alpha = 5$ and $f = 0.05$.

Coupling is thus active if and only if \mathbf{x}_2 is in a sphere of radius $r(\mathbf{x}_2^*)$ around \mathbf{x}_2^* . We sample the center points randomly from the attractor (i.e., the invariant measure) of the uncoupled system and choose the size $r(\mathbf{x}_2^*)$ such that the coupling is active during a fraction f of the time.

The results show that the impact of the uncoupling strongly depends on the position where clipping is applied. In particular, at identical system parameters, synchrony can be either stable or unstable, depending on where the coupling is active (Fig. 6). This holds even though the coupling is active for the same fraction f of time. The attractor regions of positive and negative transverse Lyapunov exponents alternate depending on the direction from the origin. As these different regions of stability and instability each occur two times on the 2π phase cycle (circulating the origin) and at roughly equal phase distance, this explains the two maxima (and the two minima) of the curve $S(\theta)$ found above [Fig. 5(d)]. This heterogeneous dependence on the exact location indicates that transient uncoupling, despite being represented by a linear reduction of the coupling term, modifies the collective dynamics of the system in a strongly nonlinear way. As a consequence, the clipping sets A need to be determined individually for each given system to be synchronized.

Conclusion.—In summary, we have proposed transient uncoupling to modify whether a system of coupled chaotic oscillators synchronizes. Most generally, these results demonstrate that continuous-time coupling is not required for synchronization, even for very simple coupling schemes [21]. Interestingly, uncoupling can synchronize systems that would fail to synchronize if ordinarily coupled. Furthermore, it can even remove any upper bound on

the coupling strengths that enable synchronization. As a natural extension, it would be challenging to explore how systems capable of weaker forms of collective dynamical coordination, such as phase synchronization, lag synchronization, or generalized synchronization, would behave if transiently uncoupled [1]. Additionally, our scheme may extend synchronization regimes not only in continuous-time systems (described by differential equations and discussed throughout the Letter), but also for chaotic maps and systems temporally switching between different continuous dynamics, cf. e.g., [22,23].

Stability properties of chaotic systems are known to vary locally with the system's state as quantified by the local Lyapunov exponent [24–26]. For transverse systems, studied above, local stability depends on the direction of the difference vector $\mathbf{x}_\perp = \mathbf{x}_1 - \mathbf{x}_2$. For small coupling strengths, the direction of this vector in the uncoupled transverse system accurately indicates the regions of state space where coupling will be most effective. However, when the coupling is stronger or active in an extended region of state space, the trajectories are more strongly modified by the coupling. In particular, whether coupling at one point is effective or not in general depends nonlinearly on the coupling in the rest of state space. Optimizing the regions of active coupling in this respect might enhance synchronizability even further.

As experimental chaotic systems often exhibit intrinsically fixed, or at least restricted, internal and coupling settings, the question emerges how to synchronize them. Transient uncoupling by state-space clipping may help to induce synchronization for a wider range of coupling strengths, with potential applications to chaotic lasers, electric and electronic circuits, communication systems, and chaos-based cryptography [27–37].

S. C. gratefully acknowledges financial support from the INSPIRE faculty fellowship awarded by the Department of Science and Technology, Government of India. This work was partially supported by the Federal Ministry of Education and Research (BMBF) Germany under Grant No. 03SF0472E and by a grant from the Max Planck Society, both to M. T.

*Present address: CGG Services UK, Crompton Way, Crawley, West Sussex RH10 9QN, United Kingdom.

†malte@nld.ds.mpg.de

‡mmanu@iitk.ac.in

§debabrata.dutta@cgg.com

||sagarc@iitk.ac.in

*timme@nld.ds.mpg.de

- [1] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization* (Cambridge University Press, Cambridge, England, 2003).
- [2] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Phys. Rev. Lett.* **75**, 1226 (1995).
- [3] B. Blasius, A. Huppert, and L. Stone, *Nature (London)* **399**, 354 (1999).

- [4] M. Kuperman and G. Abramson, *Phys. Rev. Lett.* **86**, 2909 (2001).
- [5] D. H. Zanette and M. Kuperman, *Physica (Amsterdam)* **309A**, 445 (2002).
- [6] P. Wieland, Ph.D. thesis, Universität Stuttgart, 2010.
- [7] M. Aguiar, P. Ashwin, A. Dias, and M. Field, *J. Nonlinear Sci.* **21**, 271 (2011).
- [8] S. H. Strogatz, D. M. Abrams, A. McRobie, B. Eckhardt, and E. Ott, *Nature (London)* **438**, 43 (2005).
- [9] M. H. Matheny, M. Grau, L. G. Villanueva, R. B. Karabalin, M. C. Cross, and M. L. Roukes, *Phys. Rev. Lett.* **112**, 014101 (2014).
- [10] H.-A. Tanaka, *J. Phys. A* **47**, 402002 (2014).
- [11] V. Flunkert, S. Yanchuk, T. Dahms, and E. Schöll, *Phys. Rev. Lett.* **105**, 254101 (2010).
- [12] J. Klinglmayr, C. Kirst, C. Bettstetter, and M. Timme, *New J. Phys.* **14**, 073031 (2012).
- [13] L. M. Pecora, T. L. Carroll, G. A. Johnson, D. J. Mar, and J. F. Heagy, *Chaos* **7**, 520 (1997).
- [14] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, *Phys. Rev. Lett.* **76**, 1804 (1996).
- [15] M. Timme, *Europhys. Lett.* **76**, 367 (2006).
- [16] O. E. Rössler, *Phys. Lett.* **57A**, 397 (1976).
- [17] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.115.054101> for results for Rössler oscillators with different parameters and coupled Lorenz and Chen systems, as well as larger networks, including Refs. [18–20].
- [18] L. M. Pecora and T. L. Carroll, *Phys. Rev. Lett.* **80**, 2109 (1998).
- [19] E. N. Lorenz, *J. Atmos. Sci.* **20**, 130 (1963).
- [20] G. Chen and T. Ueta, *Int. J. Bifurcation Chaos Appl. Sci. Eng.* **09**, 1465 (1999).
- [21] L. Junge and U. Parlitz, *Phys. Rev. E* **64**, 055204 (2001).
- [22] J. F. Heagy and T. L. Carroll, *Chaos* **4**, 385 (1994).
- [23] R. E. Amritkar and N. Gupte, *Phys. Rev. E* **47**, 3889 (1993).
- [24] H. Abarbanel, R. Brown, and M. Kennel, *J. Nonlinear Sci.* **1**, 175 (1991).
- [25] B. Eckhardt and D. Yao, *Physica (Amsterdam)* **65D**, 100 (1993).
- [26] U. Parlitz, J. Schumann-Bischoff, and S. Luther, *Chaos* **24**, 024411 (2014).
- [27] R. Roy and K. S. Thornburg, Jr., *Phys. Rev. Lett.* **72**, 2009 (1994).
- [28] T. Sugawara, M. Tachikawa, T. Tsukamoto, and T. Shimizu, *Phys. Rev. Lett.* **72**, 3502 (1994).
- [29] M. Sciamanna and K. A. Shore, *Nat. Photonics* **9**, 151 (2015).
- [30] K. M. Cuomo and A. V. Oppenheim, *Phys. Rev. Lett.* **71**, 65 (1993).
- [31] J. C. Sprott, *Phys. Lett. A* **266**, 19 (2000).
- [32] E.-W. Bai, K. E. Lonngren, and J. C. Sprott, *Chaos Solitons Fractals* **13**, 1515 (2002).
- [33] A. Uchida, *Optical Communication with Chaotic Lasers* (John Wiley & Sons, New York, 2012).
- [34] V. Annovazzi-Lodi, S. Donati, and A. Scire, *IEEE J. Quantum Electron.* **32**, 953 (1996).
- [35] L. Kocarev and U. Parlitz, *Phys. Rev. Lett.* **74**, 5028 (1995).
- [36] R. Mislavaty, E. Klein, I. Kanter, and W. Kinzel, *Phys. Rev. Lett.* **91**, 118701 (2003).
- [37] N. Jiang, W. Pan, B. Luo, L. Yan, S. Xiang, L. Yang, D. Zheng, and N. Li, *Phys. Rev. E* **81**, 066217 (2010).

Supplementary Material

accompanying the manuscript

Transient Uncoupling Induces Synchronization

Malte Schröder, Manu Mannattil, Debabrata Dutta, Sagar Chakraborty, Marc Timme

In the main manuscript, we introduce the idea of transient uncoupling and demonstrate that it can not only extend the range of coupling strengths enabling synchronization. Transient uncoupling does also induce qualitative changes. One key example is that in systems allowing synchronization only in some bounded interval, transient uncoupling can induce synchronization for an infinite range of coupling strengths. Thus, in particular it can stabilize synchronization for arbitrarily large coupling strengths, where the normally coupled system would fail to synchronize.

In the main manuscript the ideas and novel phenomena are presented in a general way and illustrated for the well-known system of two coupled Rössler oscillators. Yet a wide variety of different dynamical systems, from other oscillatory units to larger networks, exhibit qualitatively the same behavior.

This Supplementary Material repeats the analysis of the manuscript for several additional examples, in particular illustrating the phenomenon for unidirectionally coupled master-slave systems of Rössler oscillators at different parameters, for Lorenz as well as for Chen oscillators and finally for larger networks (extending the master stability formalism [1] to include transient uncoupling, with explicit applications to two exemplary networks).

Furthermore, as explicitly requested by a referee, we present our simple probabilistic argument describing the extension of the interval of coupling strengths that lead to stable synchronization. The results indicate that this argument does capture rough trends describing the synchronizability regime, but can neither quantitatively nor qualitatively capture the novel phenomena we observed. As explained in the manuscript (in particular see Fig. 6) this is due to the complex dependence of stability on the position in state space where coupling is active.

Throughout this supplement we use the general notation for our coupled chaotic systems

$$\frac{d\mathbf{x}_1}{dt} = \mathbf{F}(\mathbf{x}_1), \quad (1)$$

$$\frac{d\mathbf{x}_2}{dt} = \mathbf{F}(\mathbf{x}_2) + \alpha \chi_A(\mathbf{x}_2) \mathbf{C} \times (\mathbf{x}_1 - \mathbf{x}_2), \quad (2)$$

where α denotes the coupling strength, \mathbf{C} is a square coupling matrix and $\mathbf{x}_i =: (x_i, y_i, z_i)^\top$. $\chi_A(\mathbf{x}_2)$ is the indicator function describing the transient uncoupling

$$\chi_A(\mathbf{x}_2) = \begin{cases} 1 & \text{for } \mathbf{x}_2 \in A; \\ 0 & \text{for } \mathbf{x}_2 \notin A. \end{cases} \quad (3)$$

Rössler oscillator for different parameters. We again consider two x -coupled Rössler oscillators [2] with dynamics [Eq. (1)] given by

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} -(y + z) \\ x + ay \\ b + z(x - c) \end{pmatrix} \quad (4)$$

and coupled only in the x -coordinate, i.e., $\mathbf{C}_{ij} = 1$ for $i = j = 1$ and $\mathbf{C}_{ij} = 0$ otherwise. However, we use a different set of parameters: $a = b = 0.1$, $c = 14$. Transient uncoupling is again realized by clipping a region of state space in x -direction with

$$A_\Delta = \{\mathbf{x}_2 \in \mathbb{R}^d : |(\mathbf{x}_2)_1 - (\mathbf{x}_2^*)_1| \leq \Delta\}, \quad (5)$$

where $(\mathbf{x}_2)_1$ denotes the first component of (\mathbf{x}_2) and we chose $(\mathbf{x}_2^*)_1 = 1.325$ as the center of the attractor in x -direction.

As in the main manuscript, for standard coupling the synchronized state is only stable for α in a bounded interval, illustrated in Fig. 1. Again, intermediate clipping $\Delta' \approx 0.48$ ($\Delta = 10$) removes the upper bound on the coupling strength α and stabilizes the synchronized state for arbitrarily large coupling strengths (see Fig. 2). We calculate the maximal transverse Lyapunov exponent of this system for a fixed coupling strength $\alpha = 20$ and different clipping fractions $\Delta' = 2\Delta/\Omega$, where Ω is the width of the attractor in x -direction. We find again that transient uncoupling stabilizes the synchronized state (see Fig. 3) while it is unstable for standard coupling ($\Delta' = 1$).

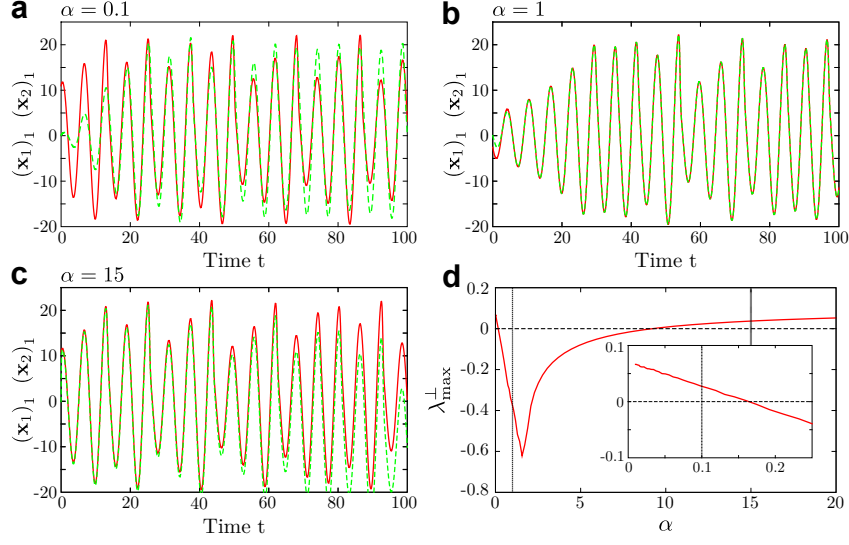


FIG. 1. Stable synchronization only for a bounded interval of coupling strengths. Trajectories of the driving (red solid line) and driven (green dashed line) units of two coupled Rössler systems for (a) $\alpha = 0.1$, (b) $\alpha = 1$, and (c) $\alpha = 15$ as marked in panel d). (d) Maximum transverse Lyapunov exponent indicates synchronization only for intermediate coupling strengths.

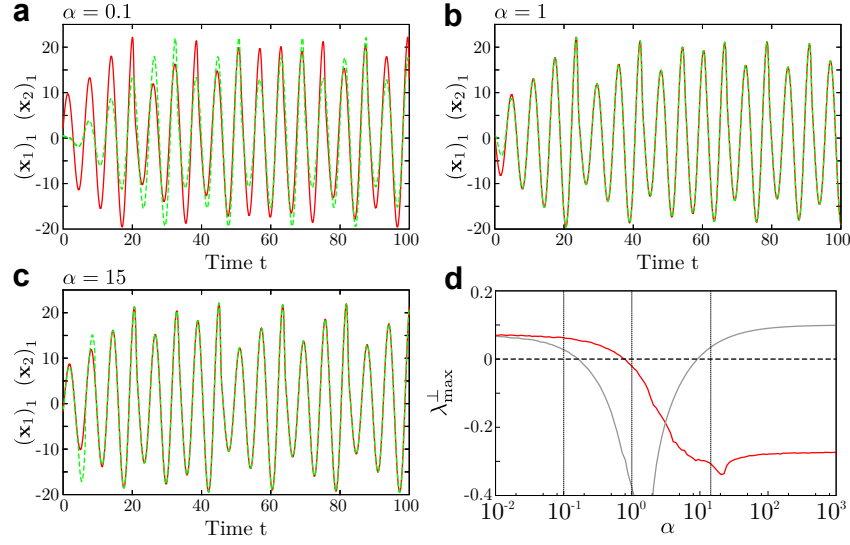


FIG. 2. Synchronization for arbitrarily large coupling strengths by transient uncoupling. Trajectories of the driving (red solid line) and driven (green dashed line) units of two coupled Rössler systems for (a) $\alpha = 0.1$, (b) $\alpha = 1$, and (c) $\alpha = 15$, the same as in Fig. 1 and marked in panel d). Transient uncoupling is realized by the clipping given by Eq. (3) and (5) with $(\mathbf{x}_2^*)_1 = 1.325$ and $\Delta' \approx 0.48$ ($\Delta = 10$). (d) Maximum transverse Lyapunov exponent as a function of the coupling strength α , note the logarithmic scale. The light grey line shows the result of unclipped coupling for comparison. With transient uncoupling synchronization is stable for arbitrarily large coupling strengths.

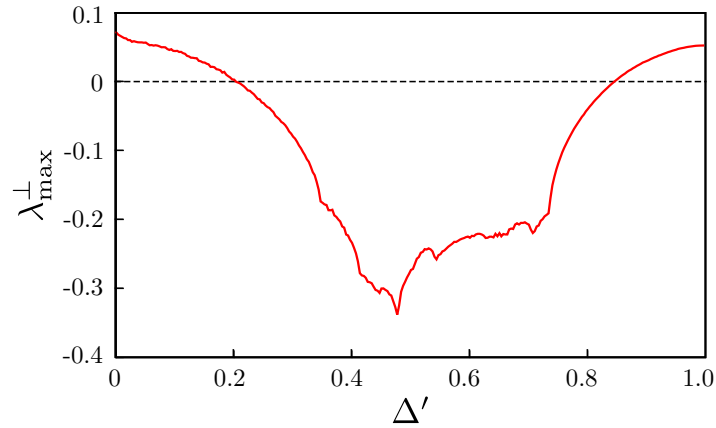


FIG. 3. Synchronization induced by transient uncoupling. Maximum transverse Lyapunov exponent for two transiently coupled Rössler oscillators [Eq. (4)] for $\alpha = 20$ and clipping [Eq. (3) and (5)] with $(\mathbf{x}_2^*)_1 = 1.325$. Synchronization emerges for moderate clipping, i.e., intermediate values of Δ' , although not without clipping ($\Delta' = 1$).

Lorenz system. As another example we consider two z -coupled Lorenz systems [3] with dynamics [Eq. (1)] given by

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} \sigma(y - x) \\ x(\rho - z) - y \\ xy - \beta z \end{pmatrix} \quad (6)$$

with $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$ and coupled with $\mathbf{C}_{ij} = 1$ for $i = j = 3$ and $\mathbf{C}_{ij} = 0$ otherwise. Transient uncoupling is realized by clipping a region of state space in z -direction with

$$A_\Delta = \{\mathbf{x}_2 \in \mathbb{R}^d : |(\mathbf{x}_2)_3 - (\mathbf{x}_2^*)_3| \leq \Delta\}, \quad (7)$$

where we chose $(\mathbf{x}_2^*)_3 = 25$ approximately in the center of the attractor in z -direction.

For this system transient uncoupling also enhances the stability of the synchronized state for arbitrarily large coupling strengths α as shown in Fig. 4 for strong clipping $\Delta' \approx 0.04$. However, we find that transient uncoupling is not always favorable for synchronization. Intermediate clipping can strongly destabilize the system compared to normal coupling as illustrated in Fig. 5(a) ($\Delta' \approx 0.5$) and Fig. 5(b) for fixed coupling strength $\alpha = 1000$.

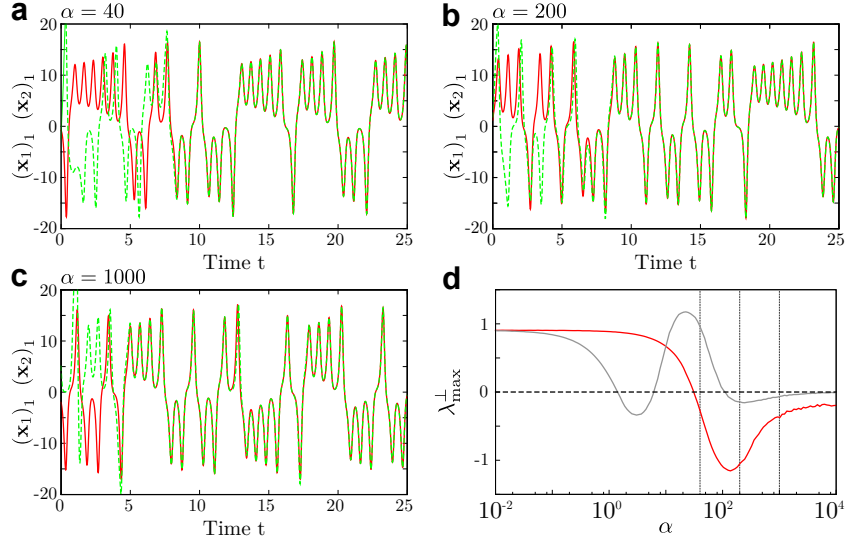


FIG. 4. Trajectories of the driving (red solid line) and driven (green dashed line) units of two coupled Lorenz systems for (a) $\alpha = 40$, (b) $\alpha = 200$, and (c) $\alpha = 1000$. The clipping is given by Eq. (3) and (7) with $(\mathbf{x}_2^*)_3 = 25$ and $\Delta' \approx 0.04$. (d) Maximum transverse Lyapunov exponent as a function of the coupling strength α , note the logarithmic scale. Synchronization is stabilized for strong coupling compared to standard coupling without clipping (light grey line).

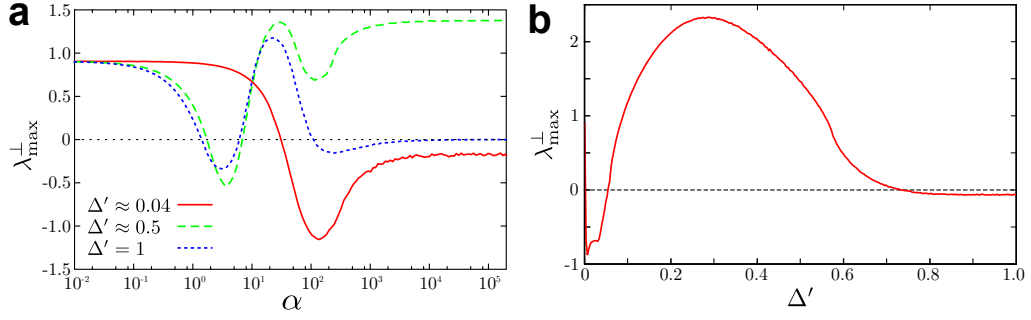


FIG. 5. Transient uncoupling induces synchronization for arbitrarily large coupling strengths. (a) Maximum transverse Lyapunov exponent for two transiently coupled Lorenz oscillators [Eq. (6)]. The clipping is given by Eq. (3) and (7) with $(\mathbf{x}_2^*)_3 = 25$. Depending on the extent of transient uncoupling synchronization can be enhanced (strong clipping $\Delta' \approx 0.04$ ($\Delta = 1$), red solid line) or hindered (intermediate clipping $\Delta' \approx 0.5$ ($\Delta = 12$), green dashed line) compared to standard coupling (no clipping $\Delta' = 1$, blue dotted line). (b) Maximum transverse Lyapunov exponent of the same system for fixed $\alpha = 1000$. Synchronization is most stable for strong clipping, i.e., small values of Δ' , while intermediate clipping destabilizes the system.

Chen system. As a further example we consider two z -coupled Chen systems [4] with dynamics [Eq. (1)] given by

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} a(y - x) \\ (c - a)x - zx + cy \\ xy - bz \end{pmatrix} \quad (8)$$

with $a = 35$, $b = 3$, $c = 28$ and coupled with $\mathbf{C}_{ij} = 1$ for $i = j = 3$ and $\mathbf{C}_{ij} = 0$ otherwise. Transient uncoupling is realized by clipping a region of state space in z -direction with

$$A_\Delta = \{\mathbf{x}_2 \in \mathbb{R}^d : |(\mathbf{x}_2)_3 - (\mathbf{x}_2^*)_3| \leq \Delta\}, \quad (9)$$

where we chose $(\mathbf{x}_2^*)_3 = 26.5$ approximately in the center of the attractor in z -direction.

The Chen systems coupled in this way synchronize only for coupling strengths in a bounded interval as was the case for the Rössler oscillators. Again, transient uncoupling allows for synchronization for arbitrarily large coupling strengths as shown in Fig. 6, depending on the extent of transient uncoupling [Fig. 7(a)]. Similarly, transient uncoupling induces synchronization for intermediate clipping but not when clipping is too weak $\Delta' \rightarrow 1$ or too strong $\Delta' \rightarrow 0$ [see Fig. 7(b)].

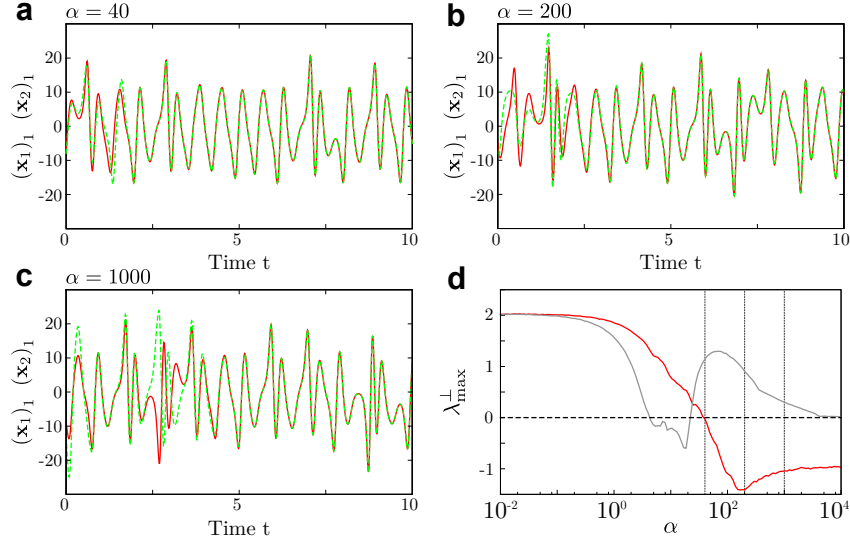


FIG. 6. Transient uncoupling induces synchronization for arbitrarily large coupling strengths. Trajectories of the driving (red solid line) and driven (green dashed line) units of two coupled Chen systems for (a) $\alpha = 40$, (b) $\alpha = 200$, and (c) $\alpha = 1000$. The clipping is given by Eq. (3) and (9) with $(\mathbf{x}_2^*)_3 = 26.5$ and $\Delta' \approx 0.2$ ($\Delta = 5$). (d) Maximum transverse Lyapunov exponent as a function of the coupling strength α , note the logarithmic scale. Transient uncoupling stabilizes synchronization for arbitrarily large coupling strengths compared to standard coupling without clipping (light grey line).

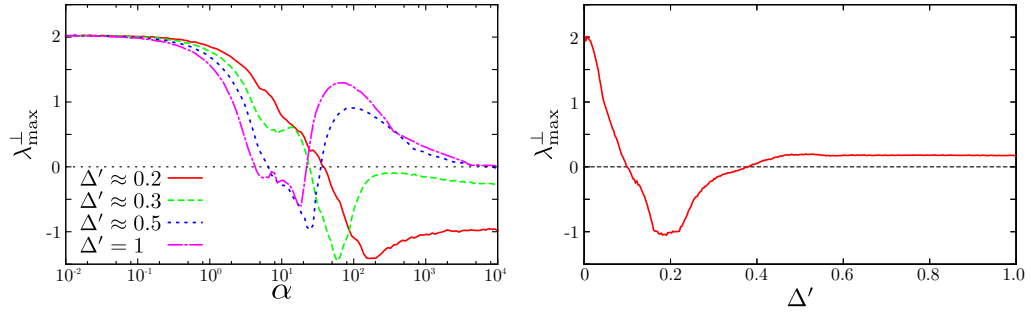


FIG. 7. Transient uncoupling induces synchronization for arbitrarily large coupling strengths. (a) Maximum transverse Lyapunov exponent for two transiently coupled Chen oscillators [Eq. (8)]. The clipping is given by Eq. (3) and (9) with $(\mathbf{x}_2^*)_3 = 26.5$. Depending on the extent of transient uncoupling the range of coupling strengths that lead to stable synchronization can be extended ($\Delta = 5$ ($\Delta' \approx 0.2$), red solid line and $\Delta' \approx 0.3$ ($\Delta = 7$), green dashed line) or only shifted ($\Delta = 12$ ($\Delta' \approx 0.5$), blue dotted line) compared to standard coupling (no clipping $\Delta' = 1$, purple dashed-dotted line). (b) Maximum transverse Lyapunov exponent of the same system for fixed $\alpha = 2000$. Synchronization emerges for moderate clipping, i.e., intermediate values of Δ' , although not without clipping ($\Delta' = 1$)

Extension to networks. The effects of transient uncoupling on larger networks of chaotic oscillators can be described easily by following the master stability function formalism introduced in [1]. We realize transient uncoupling homogeneously across all interactions in the network, i.e., the set A where coupling is active is the same for all interactions.

The equation describing our network of N coupled, identical oscillators is then given by

$$\frac{d\mathbf{x}}{dt} = \tilde{\mathbf{F}}(\mathbf{x}) + \alpha\chi_A(\mathbf{x})(\mathbf{G} \otimes \mathbf{C})\mathbf{x}, \quad (10)$$

where $\mathbf{x} \in \mathbb{R}^{dN}$ describes the vector of states of all oscillators, $\mathbf{G} \in \mathbb{R}^{N \times N}$ describes the coupling network, $\mathbf{C} \in \mathbb{R}^{d \times d}$ defines the coupling as above, $\mathbf{G} \otimes \mathbf{C}$ is the direct product and we use $\tilde{\mathbf{F}}(\mathbf{x}) = (\mathbf{F}(\mathbf{x}_1), \mathbf{F}(\mathbf{x}_2), \dots)^\top$. For the asymmetrically coupled master-slave systems used above the coupling network is given by

$$G = \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix}. \quad (11)$$

The variational equations around the synchronous orbit \mathbf{x}_S then read

$$\begin{aligned} \frac{d(\delta\mathbf{x})}{dt} &= D\tilde{\mathbf{F}}(\mathbf{x}_S)\delta\mathbf{x} + \alpha\chi_A(\mathbf{x}_S)(\mathbf{G} \otimes \mathbf{C})\delta\mathbf{x} + \alpha[D\chi_A(\mathbf{x}_S)\delta\mathbf{x}](\mathbf{G} \otimes \mathbf{C})\mathbf{x}_S \\ &= [D\tilde{\mathbf{F}}(\mathbf{x}_S) + \alpha\chi_A(\mathbf{x}_S)(\mathbf{G} \otimes \mathbf{C})]\delta\mathbf{x}, \end{aligned} \quad (12)$$

since $(\mathbf{G} \otimes \mathbf{C})\mathbf{x}_S = 0$. These equations can then be described in terms of eigenvectors of the coupling matrix

$$\dot{\xi}_k = [D\mathbf{F}(\mathbf{x}_S) + \alpha\gamma_k\chi_A(\mathbf{x}_S)\mathbf{C}]\xi_k \quad (13)$$

for $k \in \{1, 2, \dots, N\}$ where γ_k are the eigenvalues of \mathbf{G} , the $\xi_k \in \mathbb{R}^d$ are small variations with respect to the synchronous orbit and $D\mathbf{F}(\mathbf{x}_S)$ is the Jacobian matrix of a single system. It is thus sufficient to study the master stability function, the largest Lyapunov exponent of the system

$$\dot{\xi}_k = [D\mathbf{F}(\mathbf{x}_S) + \sigma\chi_A(\mathbf{x}_S)\mathbf{C}]\xi_k, \quad (14)$$

to determine the stability of arbitrary networks with transient uncoupling.

We now present two examples of networks of coupled Rössler oscillators [Eq. (4)] with parameters as in the main manuscript ($a = b = 0.2$, $c = 5.7$). We realize transient uncoupling via

$$A_\Delta = \{\mathbf{x}_2 \in \mathbb{R}^d : |(\mathbf{x}_2)_1 - (\mathbf{x}_2^*)_1| \leq \Delta\}, \quad (15)$$

where we chose $(\mathbf{x}_2^*)_1 = 1.2$ in the center of the attractor in x -direction and $\Delta = 4.16$ as in the main manuscript.

As expected by the applicability of the master stability function formalism, results identical to those for the two-node case can be found in larger networks: first, we consider an undirected chain of $N = 9$ oscillators [Fig. 8(a)]. In this case the eigenvalues of the coupling network G are such that the synchronous state is unstable for all coupling strengths α when coupled normally. Applying transient uncoupling stabilizes the synchronous state for coupling strengths larger than some lower bound $\alpha_{\min} \approx 2.75$ [Fig. 8(c)]. Secondly, we show the effects of transient uncoupling on an undirected Erdős-Rényi-random network with $N = 50$ nodes and $M = 100$ edges [Fig. 8(b)]. This network has a stable synchronized state only for a small range of coupling strengths $0.3 \lesssim \alpha \lesssim 0.4$ when coupled normally. With transient uncoupling, however, the oscillators in the network synchronize for the open interval $\alpha \gtrsim 0.7$ as shown in Fig. 8(d).

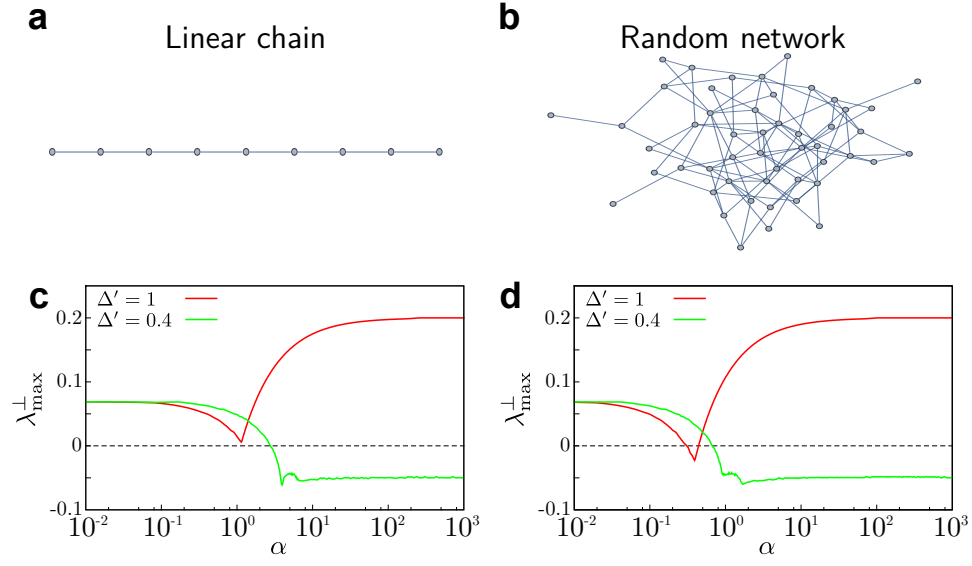


FIG. 8. Extension of transient uncoupling to networks. a) and b) show the (undirected) coupling networks, a linear chain with $N = 9$ nodes and an Erdős–Rényi random network with $N = 50$ nodes and $M = 100$ edges. c) and d) show the resulting largest transverse Lyapunov exponent for the networks a) and c), respectively. The red (dark) line shows the result for standard coupling (no clipping $\Delta' = 1$), the green (light) line the results with transient uncoupling $\Delta' = 0.4$ ($\Delta = 4.16$). As in the case of the master–slave system, transient uncoupling allows for synchronization in an infinite range of coupling strengths, thus even for arbitrarily large coupling strengths.

Why a simple probabilistic argument fails. A simple analytical argument that can be made in an attempt to explain the effects of transient uncoupling ignores all structure of the attractor and considers only the fraction of time the coupling is active

$$f = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \chi_A(\mathbf{x}_2(t)) dt. \quad (16)$$

This leads to the description of the transiently uncoupled system as a system normally coupled with an effective coupling strength of

$$\alpha_{\text{eff}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \alpha \chi_A(\mathbf{x}_2(t)) dt = f \alpha. \quad (17)$$

A system that normally coupled synchronizes for $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ should then synchronize for a range $\alpha_{\text{eff}} \in [\alpha_{\min}, \alpha_{\max}]$, or $\alpha \in [\alpha_{\min}/f, \alpha_{\max}/f]$. We therefore expect a wider, but by construction always bounded, range of coupling strengths for stable synchronization.

This simple probabilistic argument, however, does not describe the observed phenomenon of stable synchronization in an open interval $\alpha > \alpha_{\min}$. Indeed, while it describes the trend of an extended range of coupling strengths leading to stable synchronization, it fails to capture both the qualitative as well as quantitative change in the dynamics with transient uncoupling. Fig. 9 shows the largest transverse Lyapunov exponent (color-coded) of the system used in the main manuscript and the fraction $S(\alpha)$ of clipping ranges leading to stable synchronization together with the predictions by the probabilistic argument (black lines). $S(\alpha)$ is calculated as in Eq. 8 in the main manuscript:

$$S(\alpha) = \int_0^1 s(f, \alpha) df \quad (18)$$

with the synchronization indicator

$$s(f, \alpha) = \begin{cases} 1 & \text{for } \lambda_{\max}^{\perp} < 0; \\ 0 & \text{for } \lambda_{\max}^{\perp} \geq 0. \end{cases} \quad (19)$$

Evidently, there is a large discrepancy between the prediction and results of the numerical simulations. As we showed in the main manuscript, the effect of coupling depends strongly on the position where it is applied (Fig. 6 in the main manuscript). The chaotic evolution of the state of the system then leads to complex, non-linear interactions between coupling applied at different positions. For example, applying coupling at positions A and B can destabilize the system compared to applying coupling individually at either A or B . Due to this complex dependence of transverse stability on the location of active coupling on the attractor, the simple probabilistic argument, that ignores all of this structure of the attractor, can neither quantitatively nor qualitatively capture the novel phenomena we observed.

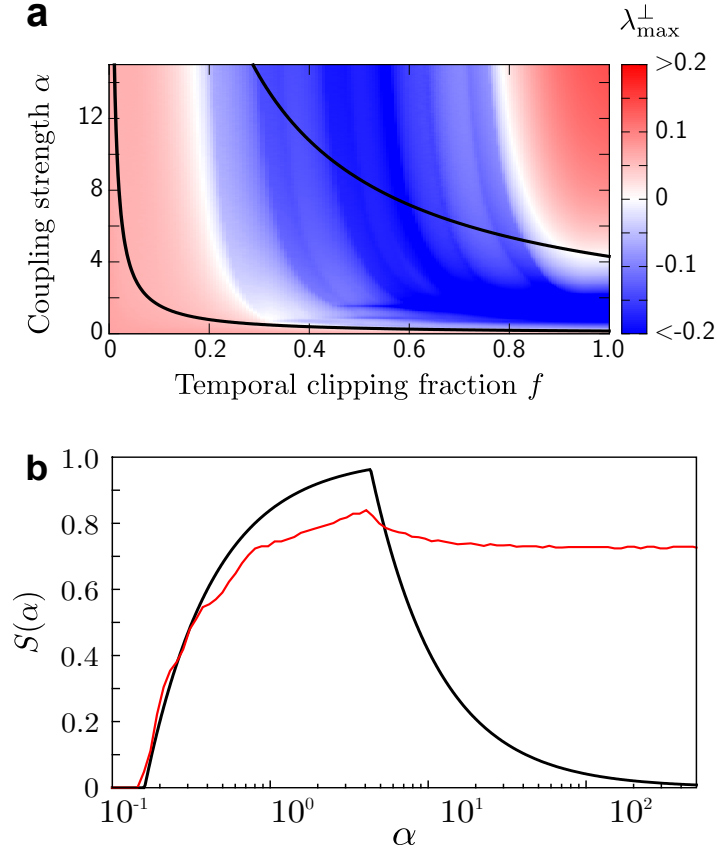


FIG. 9. Probabilistic description does not predict synchronization in an infinite range of coupling strengths. (a) Largest transverse Lyapunov exponent (color-coded) for two x -coupled Rössler oscillators [Eq. (4), $a = b = 0.2$, $c = 5.7$] with clipping in x -direction [Eq. (5), $(\mathbf{x}_2^*)_1 = 1.2$] for different coupling strengths and clipping fractions. The temporal clipping fraction f increases monotonically, though not linearly, with the clipping parameter Δ' used in the other figures (i.e., $f = 0$ corresponds to $\Delta' = 0$ and $f = 1$ corresponds to $\Delta' = 1$ but values in between differ). Synchronization is stable in the blue region (dark region in the middle). The black lines show the prediction of the probabilistic argument for the boundaries of the region of stable synchronization. (b) Synchronization region $S(\alpha)$ [Eq. (18)] showing a finite, non-zero range of temporal clipping fractions f leading to stable synchronization, even for large coupling strengths. The black line shows the prediction of the probabilistic argument: a decreasing synchronization range for large coupling. Evidently, the predictions neither quantitatively nor qualitatively capture the actual dynamics of the system under transient uncoupling.

-
- [1] L. M. Pecora and T. L. Carroll, Phys. Rev. Lett., **80**, 2109 (1998).
 - [2] O. E. Rössler, Phys. Lett. A, **57**, 397 (1976).
 - [3] E. N. Lorenz, J. Atmos. Sci., **20**, 130 (1963).
 - [4] G. Chen and T. Ueta, Int. J. Bifurcation Chaos Appl. Sci. Eng., **09**, 1465 (1999).

Chapter 4

Original manuscript: Interaction Control to Synchronize Non-synchronizable Networks

Citation

Schröder, M., Chakraborty, S., Witthaut, D., Nagler, J., and Timme, M. (2016), Interaction Control to Synchronize Non-synchronizable Networks, *Sci. Rep.* **6**, 37142

DOI: 10.1038/srep37142

©2016 Schröder, Chakraborty, Witthaut, Nagler, and Timme. This article was published open-access under the Creative Commons Attribution 4.0 International License (CC-BY) [1]. The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Original contribution

Conception and design of the research was done together with M. Timme. I performed all simulations and calculations, analyzed the data and generated all figures for the main manuscript. I designed the contents and wrote main parts of the manuscript together with M. Timme. I designed and wrote the supplemental information and performed all simulations and created all figures. I revised the manuscript during the review process.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder

SCIENTIFIC REPORTS

OPEN

Interaction Control to Synchronize Non-synchronizable Networks

Malte Schröder¹, Sagar Chakraborty², Dirk Witthaut^{3,4}, Jan Nagler⁵ & Marc Timme^{1,6}

Received: 30 September 2016

Accepted: 24 October 2016

Published: 17 November 2016

Synchronization constitutes one of the most fundamental collective dynamics across networked systems and often underlies their function. Whether a system may synchronize depends on the internal unit dynamics as well as the topology and strength of their interactions. For chaotic units with certain interaction topologies synchronization might be impossible across all interaction strengths, meaning that these networks are non-synchronizable. Here we propose the concept of *interaction control*, generalizing transient uncoupling, to induce desired collective dynamics in complex networks and apply it to synchronize even such non-synchronizable systems. After highlighting that non-synchronizability prevails for a wide range of networks of arbitrary size, we explain how a simple binary control may localize interactions in state space and thereby synchronize networks. Intriguingly, localizing interactions by a fixed control scheme enables stable synchronization across all connected networks regardless of topological constraints. Interaction control may thus ease the design of desired collective dynamics even without knowledge of the networks' exact interaction topology and consequently have implications for biological and self-organizing technical systems.

One of the simplest and most common types of collective dynamics of a networked system is synchrony, the state in which all units behave identically^{1,2}. Synchrony emerges, and is often essential, in natural and artificial systems alike, e.g. in the dynamics of circadian oscillators and neural circuits as well as in communication networks and power grids^{3–10}. More than 25 years ago, Pecora and Carroll^{11–13} uncovered that even chaotic units may synchronize; under certain conditions they coordinate their dynamics even though individually the units generate dynamics that are sensitive to small variations in the initial conditions.

The types of chaotic units jointly with their interaction topology and strength determine whether synchronization is possible at all¹⁴. Some combinations of system types and interaction topologies do not enable synchronization of the units for *any* coupling strength, rendering those systems non-synchronizable¹³. Yet, several technical systems demand synchronization of their units^{10,15–19}, requiring generic methods to achieve synchronization, ideally despite such obstacles. In fact, chaos synchronization has attracted a broad range of applications from secure communication to new paradigms of network analysis^{17,20–24}.

In this article, we investigate how a simple control of network interactions guarantees reliable synchronization independent of the specific interaction topology. We first highlight that a wide range of systems with sparse connectivity are non-synchronizable, even if they exhibit at least indirect connections (paths) between any two units. We then systematically extend a method of transient uncoupling that has been studied for two coupled oscillators²⁵ to propose a general scheme of *interaction control* applicable to any network. We show that localizing the interactions among the units to small regions of state space not only extends the synchronization range but newly creates synchrony, even for non-synchronizable networks. We further show that interaction control in fact enables synchronization regardless of the underlying interaction topology. The proposed scheme of interaction control leaves the system entirely non-interacting in most of state space, potentially saving interaction costs. Interaction control may thus help establishing collective dynamical states desired for network function in a simple and efficient way.

¹Network Dynamics, Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany.

²Department of Physics, Indian Institute of Technology Kanpur, U.P. 208016, India. ³Forschungszentrum Jülich, Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), 52428 Jülich, Germany. ⁴Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany. ⁵Computational Physics, IfB, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland. ⁶Department of Physics, Technical University of Darmstadt, 64289 Darmstadt, Germany. Correspondence and requests for materials should be addressed to M.S. (email: malte@nld.ds.mpg.de)

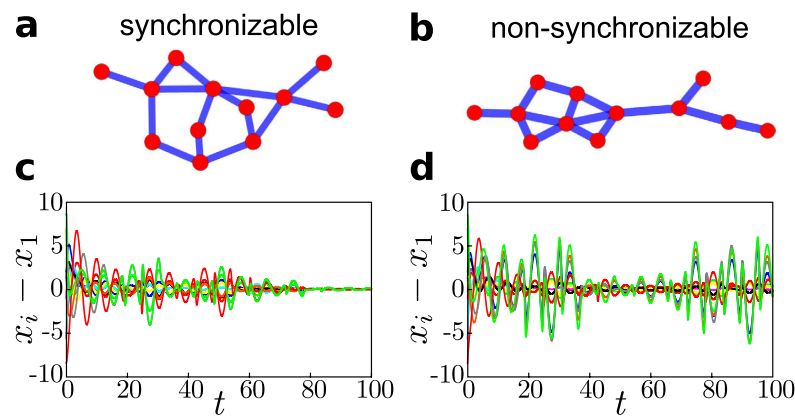


Figure 1. Synchronizable or not? Similar networks may exhibit different synchronization properties. (a,b) Two networks, despite having the same number of units, identical units, identical coupling strength and identical degree sequences exhibit different synchronizability (here for $N = 12$ coupled Rössler oscillators with identical parameters (see text), and identical coupling strength $\alpha = 0.5$). (c) Network (a) enables stable synchronization. (d) For network (b) the synchronized state is unstable. More generally, no choice of α results in stable synchronization: in this sense, the network is entirely non-synchronizable.

Results

Problem setting. Consider networks of N units with dynamics given by

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{f}(\mathbf{x}_i) + \mathbf{C}_i(\mathbf{x}_i, \mathbf{x}), \quad (1)$$

where $\mathbf{x}_i \in \mathbb{R}^d$ is the state of unit i , $\mathbf{f}(\mathbf{x}_i)$ describes the internal dynamics and $\mathbf{C}_i(\mathbf{x}_i, \mathbf{x})$ represents the pairwise interactions between the unit's state variable \mathbf{x}_i with the full network's state $\mathbf{x} \in \mathbb{R}^{Nd}$. The interactions are given by

$$\mathbf{C}_i(\mathbf{x}_i, \mathbf{x}) = c(\mathbf{x}_i) \sum_{j=1}^N A_{ij} \mathbf{h}(\mathbf{x}_j - \mathbf{x}_i), \quad (2)$$

where $A_{ij} \in \{0, 1\}$ denotes the adjacency matrix of the undirected interaction network and \mathbf{h} is the interaction function. For the control scheme introduced below, we write $c(\mathbf{x}_i)$ to be a general control function that localizes interactions in state space (see below). For a system without control we have constant $c(\mathbf{x}_i) = \alpha$.

For the numerical examples presented throughout this article we consider the units as Rössler systems²⁶ given by $\mathbf{f}(\mathbf{x}) = (-y - z, x + uy, v + z(x - w))^T$ with parameters $u = v = 0.2$ and $w = 5.7$ and diffusive coupling with $\mathbf{h}(\mathbf{x}_j - \mathbf{x}_i) = [(x_j - x_i), 0, 0]^T$ for $\mathbf{x} = (x, y, z)^T$. Interaction control is equally applicable in other settings, specifically for Rössler oscillators with different parameters, networks of other chaotic units and systems with various types of limited observability (see Supplementary Information for details).

Prevalence of non-synchronizable networks. Synchronizability of such networks of chaotic units depends on the interaction topology. Trivially, if the network is not connected synchronization is impossible. Yet, even connected networks exhibiting (at least) indirect interaction paths among every pair of units may be non-synchronizable, compare also^{27,28}. Indeed, whereas some networks may be synchronizable, similar networks with similar statistics of their topologies are non-synchronizable: Fig. 1 illustrates topology and dynamics of two networks, where one is synchronizable and the other is not, despite both having identical dynamical units and identical degree sequence. More generally, we highlight that a large fraction of sparse networks with heterogeneous degree sequence is indeed non-synchronizable (Fig. 2).

Control to localize interactions. The most obvious way to change the synchronizability of a non-synchronizable network is modifying the network topology such that the synchronized state becomes stable, an approach followed previously^{29,30}. However, changing the topology is often costly, if not impossible, in particular if the exact network topology is unknown. Can stable synchronization be achieved for these non-synchronizable networks at all?

Let us control the interactions to a small, local part of state space and now take $c(\mathbf{x}_i)$ in Eq. (2) to be a binary switch as in ref. 25: the control function then regulates whether the units are coupled at strength $c(\mathbf{x}_i) = \alpha$, if the local state is in some small region of state space where $\|\mathbf{x}_i - \mathbf{s}\| < r$ (for some offset point \mathbf{s}), or whether the units do not interact, $c(\mathbf{x}_i) = 0$ otherwise. In the limit of $r \rightarrow \infty$, the units interact for all \mathbf{x}_i in state space such that $c(\mathbf{x}_i) \equiv \alpha$ and we recover the original network of coupled chaotic units.

For small r the control strongly localizes the interactions (e.g., in the following examples with $r = 2.75$ interaction is active only about 5% of the time), thereby vastly reducing the information exchanged across the network.

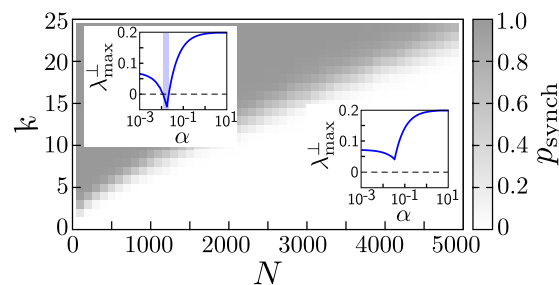


Figure 2. Prevalence of non-synchronizable networks. Synchronizability of ensembles of Barabási-Albert networks⁴⁹ of N units with $M = Nk$ links. The main panel shows the probability p_{synch} that these networks are synchronizable (measured as the fraction of 100 networks that enable stable synchronization for some value of the coupling strength α). The panel shows a clear transition to non-synchronizable networks with increasing sparseness (decreasing k). Insets: examples of the largest transverse Lyapunov exponent as a function of the coupling strength α for networks of $N = 1000$ units. Top left: synchronizable network ($k = 16$); the range of coupling strengths α enabling synchronization is shaded. Bottom right: Non-synchronizable network ($k = 4$).

For communication systems, for instance, where information exchange comes with energetic or other costs, such interaction control might reduce these costs by limiting the interactions. At the same time, localizing the interactions in this way stabilizes synchrony for a range of choices of s and r . In the following examples, we employ $s \approx (-8.7, 2.3, 0.01)^T$ (see Supplementary Information “Choice of the coupling region” and Fig. S1) and systematically vary the localization radius r of the coupling as well as the coupling strength α .

Enabling synchronization by interaction control. Intriguingly, we find that interaction control may reliably enable synchronization of networks if the interactions become strongly localized. In particular, synchronization is achieved even for systems that are non-synchronizable without control. We first illustrate the effects of interaction control for the small non-synchronizable network displayed in Fig. 1(b) in dependence of coupling strength α and localization radius r , see Fig. 3. For small r (highly localized coupling), synchronization becomes stable as long as the coupling strength is sufficiently large. For moderate r , synchronization is still possible, but only in some interval of coupling strengths. Without control ($r = \infty$), however, the network is entirely non-synchronizable.

To understand how interaction control is successful in enabling synchronization, consider the following intuitive argument: compare the local Lyapunov exponents (expansion or contraction) of the system with and without coupling. In some regions of state space the coupled system will be less expanding (more contracting) than the uncoupled system while in other regions the coupled system is more expanding (less contracting). Intuitively, applying interaction control and activating coupling only at the former, more contracting regions will lead to overall stronger contraction and will thus be beneficial for synchronization. This is the basic mechanism of interaction control. Note, however, that this argument is only approximate as it neglects the impact of interaction control on the local Lyapunov exponents: for instance, activating coupling only in one region, A or B , might be beneficial for synchronization, whereas activating coupling in both, A and B , might destabilize the synchronized state due to the effect of coupling in A on the effectiveness of coupling in B (see also Supplementary Information “Choice of the coupling region” and Fig. S1). Nevertheless, the general mechanism is applicable for a wide range of network structures and chaotic systems:

In fact, the qualitative behavior is robust for larger networks and, intriguingly, generalizes to all connected network topologies: Fig. 4 illustrates the typical characteristics for a network of $N = 1000$ units taken from the regime of non-synchronizability displayed in Fig. 2, now with interaction control localizing the coupling up to a parameter r . Systematically varying both r and α shows a common pattern: (i) For large r , i.e. without control or only weak localization, the system remains non-synchronizable (ii) For moderate r , the system becomes synchronizable for a finite interval of coupling strengths. (iii) For some sufficiently small r , even non-synchronizable networks become synchronizable for an infinite range of coupling strengths. Combined with the fact that every finite network has a finite and thus bounded spectrum the theory of master stability¹⁴ implies that interaction control may enable stable synchronization for *all* connected network topologies by strongly localizing where the units interact: for sufficiently large coupling strengths all eigenvalues of any finite, undirected graph fall within the range of negative transverse Lyapunov exponents, see Fig. 5.

Furthermore, interaction control can enable synchronization not only across network topologies but is successful for a range of different dynamical units and under various observability conditions. For instance, for different parameters in the chaotic regime of the Rössler oscillator we find qualitatively the same results to those presented above [see Supplementary Information “Rössler oscillator for different parameters” and Fig. S2 (a,b)]. Additionally, we find similar effectiveness of interaction control for different other dynamical systems, e.g., Lorenz³¹ and Chen³² oscillators (see Supplementary Information “Lorenz System”, “Chen System” and Fig. S3 and S4 respectively). Moreover, interaction control is applicable in networks with limited observability or limited controllability: All of the above examples already demonstrate successful interaction control with only one of the dynamical variables (for example only x) observed for each unit. We also find that interaction control can enable synchronization when measurements are possible at only a few discrete points in time [see Supplementary

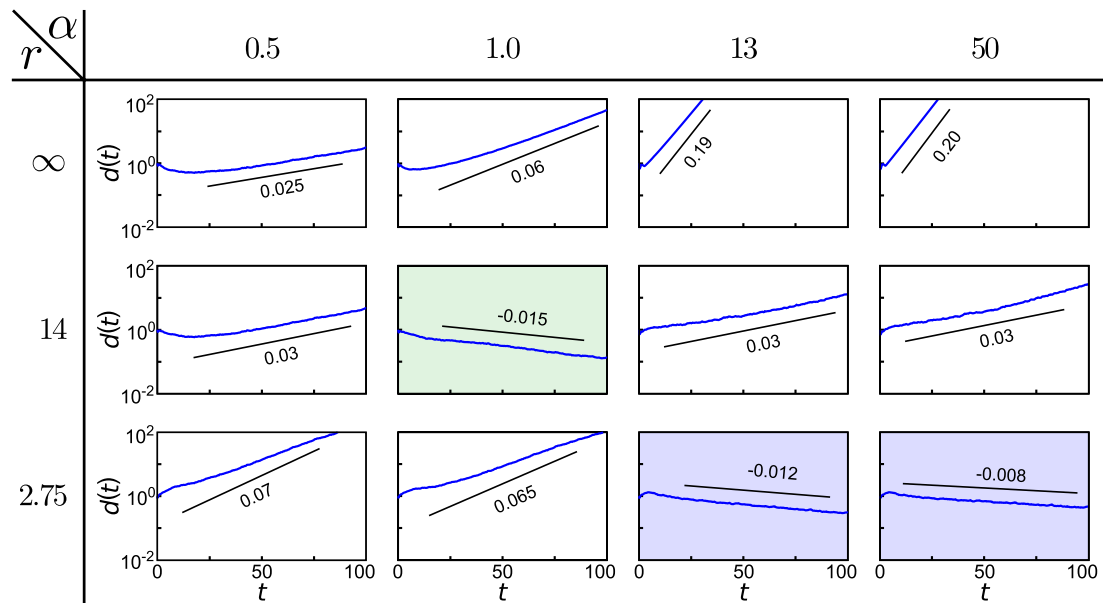


Figure 3. Interaction control to synchronize non-synchronizable networks. Panels display the average convergence or divergence of trajectories for the non-synchronizable interaction network displayed in Fig. 1(b). Averages are taken over $R = 500$ initial conditions randomly drawn from boxes of linear size 0.01 around points drawn randomly on the attractor. Each panel displays the relative average divergence of the states of the units $d(t) = \exp\left[\frac{1}{R} \sum_{k=1}^R \log(\delta(t)/\delta(0))\right]$ where $\delta(t) = \sum_{i,j=1}^N ||\mathbf{x}_i - \mathbf{x}_j||$. The black lines illustrate the scaling expected from the maximum transverse Lyapunov exponent. Without control ($r = \infty$) the network is non-synchronizable, independent of the coupling strength α . For moderate control, the system is synchronizable (negative exponent, highlighted by green shading) for some intermediate range of α . For strong control where interactions are highly localized, stable synchrony prevails at sufficiently large coupling strengths (highlighted by blue shading).

Information “Rössler oscillator for different parameters” and Fig. S2(c)]. Finally, we considered interaction control in networks where a fraction of units is not observable and thus not directly controllable. Depending on the interaction topology of the entire network, all units may become synchronizable through interaction control under some conditions on the coupling strengths. Generally, the controlled part of the network stays or even becomes synchronizable in the presence of interaction control, irrespective of its topology (see Supplementary Information “Partially controlled networks” and Fig. S5).

Conclusion

Many networks are non-synchronizable for various types of coupled units and across all interaction strengths because synchronizability is intrinsically limited by the topology of the interaction network. Here we propose interaction control to synchronize arbitrary networks, even if they are entirely non-synchronizable without control. Generalizing the idea of transient uncoupling previously suggested for two coupled oscillators²⁵ to arbitrary networks is thereby generically successful and operates by localizing interactions in state space. The interaction control scheme requires no changes to the network topology and exploits only a binary switch to strongly localize interactions to a small region in state space. As it works across all network topologies, the topology of any given network even need not be known.

Previous studies discussing time or state dependent uncoupling enhanced the stability of the synchronized state and extended synchronizability of systems that are already synchronizable^{25,33–35}. Related works aimed at enabling synchronization in non-synchronizable networks focused on topological constraints and permanent changes to the network topology^{36–38} or adaptive coupling strengths requiring permanently active interactions and detailed control over the coupling of the individual units^{39,40}. In contrast, in this article we demonstrate how interaction control may synchronize previously non-synchronizable networks. Specifically, interaction control induces a *qualitative* rather than a quantitative change of the synchronizability interval that cannot be explained by extending existing ranges of synchronizability or modifying effective coupling strengths. As we report, non-synchronizability prevails among sparse networks with heterogeneous degree sequence, emphasizing the range of systems for which interaction control may be valuable.

Intuitively, interaction control increases synchronizability by disabling coupling in regions of state space where the trajectories of the coupled system diverge more than those of the uncoupled system. Thus, interaction control enables synchronization with little information transmission between the units, thereby providing a potentially efficient control for engineered systems where interaction generates costs in terms of energy or other resources^{33,41}, for example for communication with chaos synchronization^{17,42}. At the same time, interaction

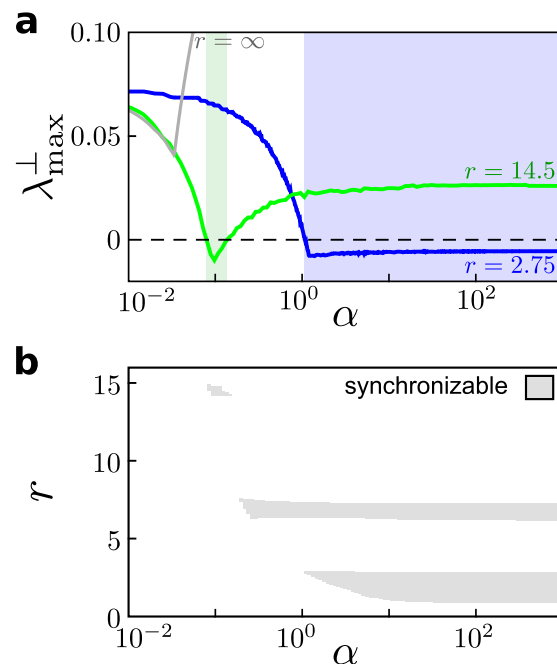


Figure 4. Control generically enables synchronization for large networks. Localized interactions induce stable synchronization even in large networks with arbitrary topology. Panel (a) shows the largest transverse Lyapunov exponent for a non-synchronizable Barabási-Albert network of $N = 1000$ units and $k = 4$ without control (light gray line, compare Fig. 2) and with interaction control. Moderately localized interactions (control parameter $r = 14.5$) enable stable synchronization in a small range of coupling strengths only (shaded in green). Stronger localization ($r = 2.75$) enables stable synchronization for all sufficiently large coupling strengths (shaded in blue). Panel (b) illustrates the synchronizability for all combinations of coupling strength α and localization radius r .

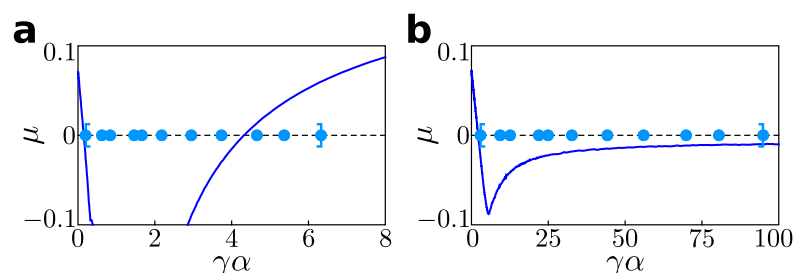


Figure 5. Stable synchronization for arbitrary network topologies. Master stability function μ for real values $\gamma\alpha$, where γ are the eigenvalues of the Laplacian of the interaction network and α is the coupling strength (see also Supplementary Information “Extension of the master stability formalism”). The light blue points mark $\gamma\alpha$ for the eigenvalues of the non-synchronizable network displayed in Fig. 1(b). (a) Without control the network is non-synchronizable because some transverse modes are unstable regardless of the choice of α (shown here for $\alpha = 1$). (b) With interaction control ($r = 2.75$) all transverse modes are stable (negative $\lambda_{\perp\max}$) if the coupling strength is large enough ($\alpha = 15$), since the master stability function is negative for large $\gamma\alpha$. Similarly, interaction control can be used to synchronize any connected, undirected network independent of its topology, since for sufficiently large coupling strengths all transverse modes will become stable.

control can be successfully applied to induce synchronization even in systems with limited observability. For instance access may only be possible to some of the dynamical variables of each node, measurements at discrete points in time or in the presence of unobservable and thus uncontrollable units (see Supplementary Information), opening up potential perspectives, e.g., also for natural and synthetic biological systems^{43,44}. Following the intuitive mechanism, from a general dynamical systems perspective interaction control might be applicable to any collective state that exhibits instabilities due to coupling among variables. Specifically, this includes potentially inducing different collective dynamics, for instance phase synchronization, or enabling coordinated dynamics also for delayed or pulsed interactions^{36,37,45}.

In summary, interaction control offers a generic control scheme for collectively coordinated networks. Although requiring interaction only in a small region of state space, interaction control enables synchronization in all connected networks independent of their specific topology, even if the network would normally be non-synchronizable. Generally, interaction control may functionally help beyond enabling synchrony, for instance to create consensus among interacting agents⁴⁶. Interaction control may thus offer a complementary network control method^{47,48} and thereby a valuable paradigm for enabling a number of different collective dynamical phenomena in a range of networked systems.

References

- Pikovsky, A., Rosenblum, M. & Kurths, J. *Synchronization: a universal concept in nonlinear sciences*, (Cambridge University Press, 2003).
- Strogatz, S. H. *Sync: How Order Emerges From Chaos In the Universe, Nature, and Daily Life* (Hachette Books, New York, 2004), reprint edn.
- Gonze, D., Bernard, S., Waltermann, C., Kramer, A. & Herzl, H. Spontaneous synchronization of coupled circadian oscillators. *Biophys. J.* **89**, 120–129 (2005).
- Mirollo, R. E. & Strogatz, S. H. Synchronization of pulse-coupled biological oscillators. *SIAM J. Appl. Math.* **50**, 1645–1662 (1990).
- Buzsáki, G. & Chrobak, J. Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. *Curr. Opin. Neurobiol.* **5**, 504–510 (1995).
- Kirst, C., Geisel, T. & Timme, M. Sequential desynchronization in networks of spiking neurons with partial reset. *Phys. Rev. Lett.* **102**, 068101 (2009).
- Kirst, C., Timme, M. & Battaglia, D. Dynamic information routing in complex networks. *Nat Commun* **7**, 11061 (2016).
- Bregni, S. *Synchronization of Digital Telecommunications Networks* (Wiley, New York, 2002).
- Blaabjerg, F., Teodorescu, R., Liserre, M. & Timbus, A. V. Overview of control and grid synchronization for distributed power generation systems. *IEEE Trans. Ind. Electron.* **53**, 1398–1409 (2006).
- Rohden, M., Sorge, A., Timme, M. & Witthaut, D. Self-organized synchronization in decentralized power grids. *Phys. Rev. Lett.* **109**, 064101 (2012).
- Pecora, L. M. & Carroll, T. L. Synchronization in chaotic systems. *Phys. Rev. Lett.* **64**, 821–824 (1990).
- Pecora, L. M. & Carroll, T. L. Driving systems with chaotic signals. *Phys. Rev. A* **44**, 2374–2383 (1991).
- Pecora, L. M. & Carroll, T. L. Synchronization of chaotic systems. *Chaos* **25**, 097611 (2015).
- Pecora, L. M. & Carroll, T. L. Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.* **80**, 2109–2112 (1998).
- Tyrrell, A., Auer, G. & Bettstetter, C. Emergent slot synchronization in wireless networks. *IEEE Trans. Mobile Comput.* **9**, 719–732 (2010).
- Klinglmayr, J., Kirst, C., Bettstetter, C. & Timme, M. Guaranteeing global synchronization in networks with stochastic interactions. *New J. Phys.* **14**, 073031 (2012).
- Kinzel, W., Englert, A. & Kanter, I. On chaos synchronization and secure communication. *Phil. Trans. R. Soc. A* **368**, 379–389 (2010).
- Rohden, M., Sorge, A., Witthaut, D. & Timme, M. Impact of network topology on synchrony of oscillatory power grids. *Chaos* **24**, 013123 (2014).
- Nishimura, J. & Friedman, E. J. Robust convergence in pulse-coupled oscillators with delays. *Phys. Rev. Lett.* **106**, 194101 (2011).
- Parlitz, U. Estimating model parameters from time series by autosynchronization. *Phys. Rev. Lett.* **76**, 1232–1235 (1996).
- Reidler, I. *et al.* Coupled lasers: phase versus chaos synchronization. *Opt. Lett.* **38**, 4174–4177 (2013).
- Englert, A. *et al.* Zero lag synchronization of chaotic systems with time delayed couplings. *Phys. Rev. Lett.* **104**, 114102 (2010).
- Yu, D., Righero, M. & Kocarev, L. Estimating topology of networks. *Phys. Rev. Lett.* **97**, 188701 (2006).
- Timme, M. & Casadiego, J. Revealing networks from dynamics: an introduction. *J. Phys. A* **47**, 343001 (2014).
- Schröder, M., Mannatil, M., Dutta, D., Chakraborty, S. & Timme, M. Transient uncoupling induces synchronization. *Phys. Rev. Lett.* **115**, 054101 (2015).
- Rössler, O. E. An equation for continuous chaos. *Phys. Lett. A* **57**, 397–398 (1976).
- Timme, M. Does dynamics reflect topology in directed networks? *Europhys. Lett.* **76**, 367 (2006).
- Timme, M. & Wolf, F. The simplest problem in the collective dynamics of neural networks: is synchrony stable? *Nonlinearity* **21**, 1579 (2008).
- Nishikawa, T. & Motter, A. E. Synchronization is optimal in nondiagonalizable networks. *Phys. Rev. E* **73**, 065106 (2006).
- Sevilla-Escoboza, R. *et al.* Experimental implementation of maximally synchronizable networks. *Physica A* **448**, 113–121 (2016).
- Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).
- Chen, G. & Ueta, T. Yet another chaotic attractor. *Int. J. Bifurcation Chaos* **9**, 1465–1466 (1999).
- Junge, L. & Parlitz, U. Synchronization using dynamic coupling. *Phys. Rev. E* **64**, 055204 (2001).
- Chen, L., Qiu, C. & Huang, H. Synchronization with on-off coupling: Role of time scales in network dynamics. *Phys. Rev. E* **79**, 045101 (2009).
- Chen, L., Qiu, C., Huang, H., Qi, G. & Wang, H. Facilitated synchronization of complex networks through a discontinuous coupling strategy. *Eur. Phys. J. B* **76**, 625–635 (2010).
- Flunkert, V., Yanchuk, S., Dahms, T. & Schöll, E. Synchronizing distant nodes: A universal classification of networks. *Phys. Rev. Lett.* **105**, 254101 (2010).
- Heiligenthal, S. *et al.* Strong and weak chaos in nonlinear networks with time-delayed couplings. *Phys. Rev. Lett.* **107**, 234102 (2011).
- Nishikawa, T. & Motter, A. E. Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions. *Proc. Natl. Acad. Sci.* **107**, 10342 (2010).
- Zhou, C. & Kurths, J. Dynamical weights and enhanced synchronization in adaptive complex networks. *Phys. Rev. Lett.* **96**, 164102 (2006).
- DeLellis, P., di Bernardo, M. & Turci, L. F. R. Fully adaptive pinning control of complex networks. In *Proceedings of 2010 IEEE international symposium on circuits and systems* (2010).
- Fradkov, A. L., Andrievsky, B. & Evans, R. J. Chaotic observer-based synchronization under information constraints. *Phys. Rev. E* **73**, 066209 (2006).
- Klein, E., Mislavaty, R., Kanter, I. & Kinzel, W. Public-channel cryptography using chaos synchronization. *Phys. Rev. E* **72**, 016214 (2005).
- Karzbrun, E., Tayar, A. M., Noireaux, V. & Bar-Ziv, R. Programmable on-chip dna compartments as artificial cells. *Science* **345**, 829–832 (2014).
- Tayar, A. M., Karzbrun, E., Noireaux, V. & Bar-Ziv, R. H. Propagating gene expression fronts in a one-dimensional coupled system of artificial cells. *Nat. Phys.* **11**, 1037–1041 (2015).
- Jahnke, S., Memmesheimer, R.-M. & Timme, M. Stable irregular dynamics in complex neural networks. *Phys. Rev. Lett.* **100**, 048102 (2008).
- Tang, C.-B., Wu, B., Wang, J.-B. & Li, X. Evolutionary origin of asymptotically stable consensus. *Sci. Rep.* **4**, 4590 (2014).
- Liu, Y.-Y., Slotine, J.-J. & Barabási, A.-L. Controllability of complex networks. *Nature* **473**, 167–173 (2011).

48. Cornelius, S. P., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. *Nat. Commun.* **4**, 1942 (2013).
49. Barabási, L.-A. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509 (1999).

Acknowledgements

We gratefully acknowledge support by the Göttingen Graduate School for Neurosciences and Molecular Biosciences (DFG Grant GSC 226/2 [MS]), the INSPIRE faculty fellowship awarded by the Department of Science and Technology, Government of India (DST/INSPIRE/04/2013/000365 [SC]), the Helmholtz association (grant no. VH-NG-1025 [DW]), the ETH Risk Center (SP RC 08-15 [JN]), the Federal Ministry of Education and Research (BMBF) Germany under grant no. 03SF0472E [DW,MT] and a grant from the Max Planck Society [MT].

Author Contributions

M.S. and M.T. conceived and designed the research. M.S. worked out the numerical simulations and prepared the figures. M.S., S.C. and M.T. analysed the data. All authors provided ideas and analysis tools, interpreted the results and wrote the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Schröder, M. *et al.* Interaction Control to Synchronize Non-synchronizable Networks. *Sci. Rep.* **6**, 37142; doi: 10.1038/srep37142 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

Supplementary Material
 accompanying the manuscript
Interaction Control to Synchronize Non-synchronizable Networks

Malte Schröder, Sagar Chakraborty, Dirk Witthaut, Jan Nagler, Marc Timme

In the main manuscript, we discuss how interaction control creates synchronizability for networks of coupled chaotic units. Specifically, we demonstrated that interaction control enables to synchronize networks that are non-synchronizable without control, irrespective of their network topology. Here, we first discuss the choice of the offset point \mathbf{s} used for the examples in the main manuscript. Second, we formally extend the master stability formalism¹ to include interaction control. Third, we illustrate the universality of this approach by applying interaction control to networks of Rössler systems at different parameters, to networks of Lorenz and to networks of Chen systems with qualitatively the same results as those presented in the main manuscript. Finally, we consider three aspects of interaction control in systems with limited observability, specifically, when we only have access to a single variable of each unit, to measurements at discrete points in time and in the presence of unobservable (and thus uncontrollable) units in the network.

Throughout this supplement we use the same notation as in the main manuscript

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{f}(\mathbf{x}_i) + \mathbf{C}_i(\mathbf{x}_i, \mathbf{x}), \quad (1)$$

where $\mathbf{x}_i \in \mathbb{R}^d$ is the state of unit i , $\mathbf{f}(\mathbf{x}_i)$ describes the internal dynamics and $\mathbf{C}_i(\mathbf{x}_i, \mathbf{x})$ represents the pairwise interactions between local state variable \mathbf{x}_i and the remaining network's state $\mathbf{x} \in \mathbb{R}^{Nd}$ for a network of N units. Again, the interactions are defined as

$$\mathbf{C}_i(\mathbf{x}_i, \mathbf{x}) = c(\mathbf{x}_i) \sum_{j=1}^N A_{ij} \mathbf{h}(\mathbf{x}_j - \mathbf{x}_i), \quad (2)$$

where $A_{ij} \in \{0, 1\}$ denotes the adjacency matrix of the undirected interaction network, \mathbf{h} is the interaction function and $c(\mathbf{x}_i)$ is a general control function that localizes interactions in state space.

Supplementary Note 1

Choice of the coupling region.

In the examples in the main manuscript we employed interaction control via

$$c(\mathbf{x}_i) = \begin{cases} \alpha & \text{if } \|\mathbf{x} - \mathbf{s}\| < r, \\ 0 & \text{else} \end{cases}, \quad (3)$$

where $\mathbf{s} \approx (-8.7, 2.8, 0.01)^T$ and r controls the size of the coupling region. Here we explain how to choose \mathbf{s} . To find suitable parameters for interaction control we computed the stability of the synchronized state of two coupled Rössler systems for various choices of potential offset points \mathbf{s}' and a suitable distance $r(\mathbf{s}')$. We chose $R = 10000$ points randomly from the attractor (invariant measure) as potential offset points. For each point we calculate $r(\mathbf{s}')$ such that $c(\mathbf{x}_i) = \alpha$ for a fraction of 5% of points on the invariant measure. With these parameters and $\alpha = 5$ we calculated the maximum transverse Lyapunov exponent λ_{\max}^{\perp} . The point \mathbf{s} for which the maximum transverse Lyapunov exponent is minimal is chosen as the offset point. Results of the simulations are shown in Fig. S1. Depending on choice of \mathbf{s} , interaction control will be more or less efficient. If one effectively optimizes the function $c(\mathbf{x}_i)$ to increase stability one might expect even better results, allowing for example stable synchronization with minimal coupling effort.

A faster way to determine a feasible, though probably less efficient, coupling region can be understood by considering a simple argument that qualitatively explains how interaction control works: comparing the local Lyapunov exponents for the uncoupled and coupled system provides a measure of how effective coupling is at any given point in state space. A suitable coupling region restricts coupling to efficient points and, more importantly, disables it at points where coupling is detrimental to synchronization. One would then naturally expect more stable synchronization. This method requires only derivatives of the individual dynamics $\mathbf{f}(\mathbf{x}_i)$ which are either known or can be estimated from measurements and will provide a feasible coupling region for general systems that is expected to enhance synchronizability. Note, however, that this method uses only local indicators but ignores global effects, such as coupling at one point changing the effectiveness of coupling at another point. Thus, while this method might serve as an efficient way to determine a feasible coupling region, it is not guaranteed to result in an efficient one.

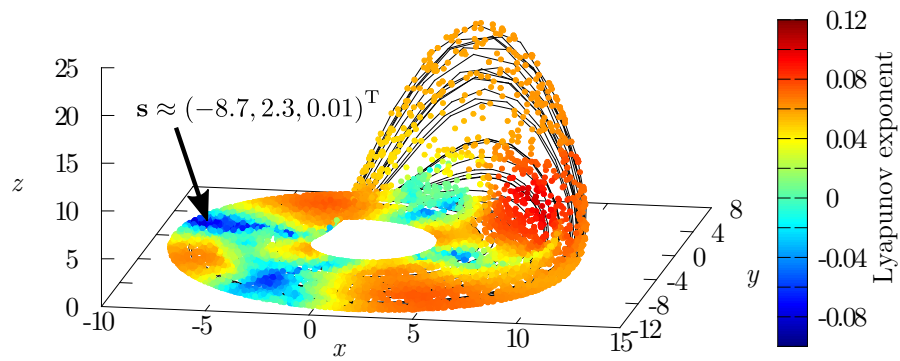


Figure S1. Largest transverse Lyapunov exponent of two coupled Rössler systems, figure adapted from². Simulations were done with parameters s' (marked by the location of the points), $r(s')$ such that $c(\mathbf{x}_i) = \alpha$ for a fraction of 5% of points on the invariant measure and $\alpha = 5$ (see text). The color indicates the resulting largest transverse Lyapunov exponent. The most stable synchronized state (minimum largest transverse Lyapunov exponent) is achieved for $s \approx (-8.7, 2.8, 0.01)^T$ (indicated by the arrow) and marks our choice for the offset point.

Supplementary Note 2

Extension of the master stability formalism. Here we derive the extension of the master stability function formalism introduced in¹ for interaction control. We assume the same dynamics as described in Eq. 1.

The equation describing our network of N coupled, identical units with dynamical variables $\mathbf{x}_i \in \mathbb{R}^d$ in d dimensions can then be written as

$$\frac{d\mathbf{x}}{dt} = \tilde{\mathbf{F}}(\mathbf{x}) + \tilde{\mathbf{C}}(\mathbf{x}) (\mathbf{G} \otimes \mathbf{H}) \mathbf{x}, \quad (4)$$

where $\mathbf{x} \in \mathbb{R}^{dN}$ describes the vector of states of all oscillators and we write $\tilde{\mathbf{F}}(\mathbf{x}) = (\mathbf{f}(\mathbf{x}_1), \mathbf{f}(\mathbf{x}_2), \dots)^T$ [similarly for $\tilde{\mathbf{C}}(\mathbf{x})$ combining the individual $c(\mathbf{x}_i)$]. The coupling is defined by $\mathbf{G} \in \mathbb{R}^{N \times N}$ describing the Laplacian of the coupling network, $\mathbf{H} \in \mathbb{R}^{d \times d}$ defining the coupling between the coordinates and $\mathbf{G} \otimes \mathbf{H}$ representing the direct product. As an example, two bidirectionally coupled units with coupling between the x -coordinates would be described by

$$\mathbf{G} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (5)$$

The variational equations around the synchronous orbit \mathbf{x}_S then read

$$\begin{aligned} \frac{d(\delta\mathbf{x})}{dt} &= \mathbf{D}\tilde{\mathbf{F}}(\mathbf{x}_S)\delta\mathbf{x} + \tilde{\mathbf{C}}(\mathbf{x}_S)(\mathbf{G} \otimes \mathbf{H})\delta\mathbf{x} + [\mathbf{D}\tilde{\mathbf{C}}(\mathbf{x}_S)\delta\mathbf{x}](\mathbf{G} \otimes \mathbf{H})\mathbf{x}_S \\ &= [\mathbf{D}\tilde{\mathbf{F}}(\mathbf{x}_S) + \tilde{\mathbf{C}}(\mathbf{x}_S)(\mathbf{G} \otimes \mathbf{H})]\delta\mathbf{x}, \end{aligned} \quad (6)$$

since $(\mathbf{G} \otimes \mathbf{H})\mathbf{x}_S = 0$. These equations can then be described in terms of eigenvectors of the coupling network

$$\frac{d\xi_k}{dt} = [\mathbf{D}\mathbf{f}(\mathbf{x}_S) + \gamma_k c(\mathbf{x}_S) \mathbf{H}] \xi_k \quad (7)$$

for $k \in \{1, 2, \dots, N\}$ where γ_k are the eigenvalues of \mathbf{G} , the $\xi_k \in \mathbb{R}^d$ are small variations with respect to the synchronous orbit and $\mathbf{D}\mathbf{f}(\mathbf{x}_S)$ is the Jacobian matrix of a single unit. It is thus sufficient to study the master stability function $\mu(\gamma\alpha)$, defined as the largest Lyapunov exponent of the system

$$\frac{d\xi}{dt} = [\mathbf{D}\mathbf{f}(\mathbf{x}_S) + \gamma c(\mathbf{x}_S) \mathbf{H}] \xi, \quad (8)$$

to determine the stability of arbitrary networks with interaction control.

In the following we use $c(\mathbf{x}_i) \in \{0, \alpha\}$ as in the main manuscript. We compute the master stability function for real values of the parameter $\gamma\alpha$ (undirected networks) both with and without control. To illustrate the effect of interaction control we show results for eigenvalues of a non-synchronizable network, illustrating how it becomes synchronizable with interaction control.

Supplementary Note 3

Rössler oscillator for different parameters. We again consider a network of Rössler units³ with dynamics [Eq. (1)] given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} -(y+z) \\ x+uy \\ v+z(x-w) \end{pmatrix} \quad (9)$$

and coupled only in the x -coordinate, i.e., $\mathbf{h}(\mathbf{x}_j - \mathbf{x}_i) = [(x_j - x_i), 0, 0]^T$ for $\mathbf{x} = (x, y, z)^T$. However, we use a different set of parameters: $u = v = 0.1$, $w = 14$. Here, we choose the control function for interaction control as

$$c(\mathbf{x}_i) = \begin{cases} \alpha & \text{if } |x_i - x^*| < d, \\ 0 & \text{else} \end{cases}, \quad (10)$$

where $x^* = 1.325$ is the center of the attractor in x -direction, i.e. we localize control to a box of width $2d$ in x -direction. We illustrate results for the master stability function without and with control for $d = 10$ in Fig. S2 (a,b). Note that both the coupling as well as the coupling region only depend on the x -coordinate, thus showing that interaction control can be successfully applied without access to the y and z variable of the system. Furthermore, in Fig. S2 (c), we assumed limited observability in the sense that measurements of the unit's states are only possible at discrete time points with only about five measurements per full oscillation. Consequently, we can only adjust the control function at these times: at the time of a measurement, the state of the coupling is fixed as active ($c = \alpha$) or inactive ($c = 0$) depending on the current state of the unit for a time Δt_{meas} until the next measurement. Note that the coupling *input* [i.e. $\mathbf{x}_j - \mathbf{x}_i$] is still continuous in time. As in the main manuscript, all networks become synchronizable regardless of their specific topology due to the interaction control, even if they were non-synchronizable without control.

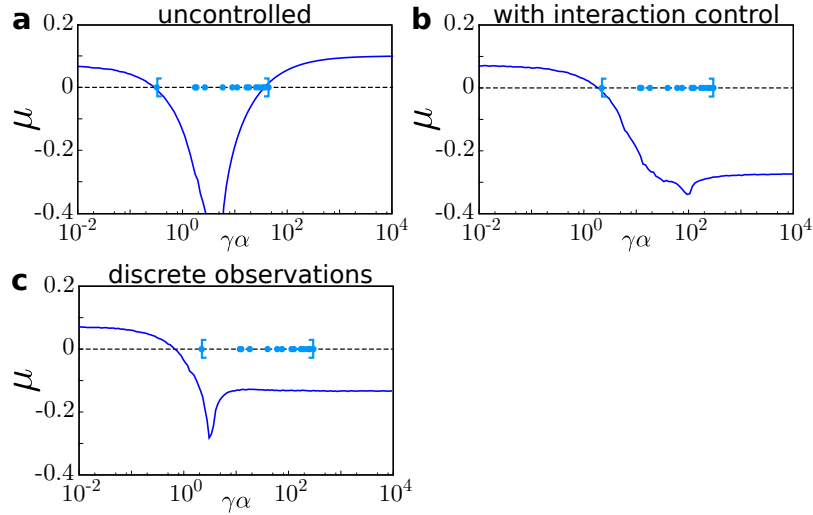


Figure S2. Master stability function of coupled Rössler units for real values $\gamma\alpha$, where γ are the eigenvalues of the Laplacian of the coupling network and α is the coupling strength. The light blue points illustrate eigenvalues of a non-synchronizable network with $\alpha = 2$ [panel (a)] and $\alpha = 10$ [panel (b)]. a) Without control the network is non-synchronizable, some transverse modes are unstable. b) With interaction control (Eq. (10), $d = 10$) all transverse modes are stable if the coupling strength is sufficiently large, since the master stability function is negative for all sufficiently large $\gamma\alpha$. Similarly, interaction control can be used to synchronize any undirected network independent of its topology, since for large coupling strengths all transverse modes will be stable. c) Even with limited observability, i.e., only about five measurements per full oscillation ($\Delta t_{\text{meas}} = 1$), interaction control is still successful in enabling stable synchronization regardless of network topology.

Supplementary Note 5

Lorenz system. As another example we consider Lorenz units⁴ with dynamics [Eq. (1)] given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} \sigma(y-x) \\ x(\rho-z)-y \\ xy-\beta z \end{pmatrix} \quad (11)$$

with $\sigma = 10$, $\rho = 28$ and $\beta = 8/3$ and coupled with $\mathbf{h}(\mathbf{x}_j - \mathbf{x}_i) = [0, 0, (z_j - z_i)]^T$. Interaction control is realized with

$$c(\mathbf{x}_i) = \begin{cases} \alpha & \text{if } |z_i - z^*| < d \\ 0 & \text{else} \end{cases}, \quad (12)$$

where we chose $z^* = 25$ approximately in the center of the attractor in z -direction. We illustrate results for the master stability function without and with control for $d = 1$ in Fig. S3. While all networks of coupled Lorenz oscillators are synchronizable for very large coupling strengths, interaction control both decreases the coupling strength necessary to induce stable synchronization and increases the stability of the synchronized state, enhancing synchronizability of all networks.

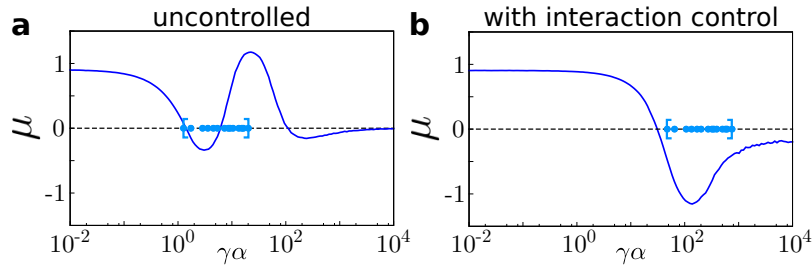


Figure S3. Master stability function of coupled Lorenz units for real values $\gamma\alpha$, where γ are the eigenvalues of the Laplacian of the coupling network and α is the coupling strength. The light blue points illustrate eigenvalues for an example network with $\alpha = 3$ [panel (a)] and $\alpha = 100$ [panel (b)]. a) Without control the synchronized state is unstable since some transverse modes are unstable. Synchronization would be (weakly) stable only for very large coupling strengths. b) With interaction control (Eq. (12), $d = 1$) all transverse modes are stable if the coupling strength is sufficiently large, since the master stability function is negative for all sufficiently large $\gamma\alpha$ and the stability of the synchronized state is enhanced. Similarly, interaction control can be used to synchronize any undirected network independent of its topology, since for large coupling strengths all transverse modes will become stable.

Supplementary Note 6

Chen system. As a final example we consider Chen units⁵ with dynamics [Eq. (1)] given by

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} u(y-x) \\ (w-u)x - zx + wy \\ xy - vz \end{pmatrix} \quad (13)$$

with $u = 35$, $v = 3$, $w = 28$ and coupled with $\mathbf{h}(\mathbf{x}_j - \mathbf{x}_i) = [0, 0, (z_j - z_i)]^T$. Interaction control is realized with

$$c(\mathbf{x}_i) = \begin{cases} \alpha & \text{if } |z_i - z^*| < d \\ 0 & \text{else} \end{cases}, \quad (14)$$

where we chose $z^* = 26.5$ approximately in the center of the attractor in z -direction. We illustrate results for the master stability function without and with control for $d = 5$ in Fig. S4. As in the main manuscript, all networks become synchronizable regardless of their specific topology with interaction control, even if they were non-synchronizable without control.

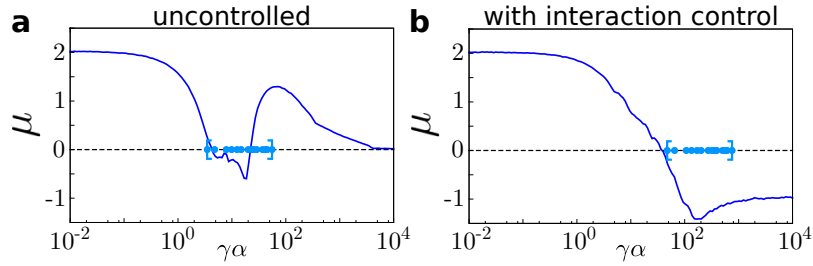


Figure S4. Master stability function of coupled Chen units for real values $\gamma\alpha$, where γ are the eigenvalues of the Laplacian of the coupling network and α is the coupling strength. The light blue points illustrate eigenvalues of a non-synchronizable network with $\alpha = 8$ [panel (a)] and $\alpha = 100$ [panel (b)]. a) Without interaction control the network is non-synchronizable, some transverse modes are unstable. b) With interaction control (Eq. (14), $d = 5$) all transverse modes are stable if the coupling strength is sufficiently large, since the master stability function is negative for all sufficiently large $\gamma\alpha$. Similarly, interaction control can be used to synchronize any undirected network independent of its topology, since for large coupling strengths all transverse modes will become stable.

Supplementary Note 7

Partially controlled networks. We consider a network where some units are not observable and thus not affected by interaction control, these uncontrolled nodes are instead continuously coupled to their neighbors [$c(\mathbf{x}_i) = \alpha$]. In such partially controlled networks, success of the method, by construction, depends on the set of controlled units and the network structure. The general effect can be readily understood considering the simplified case of a single uncontrolled unit: we split the network into two disjoint sets, a (connected) set of controlled units A and a set of the single uncontrolled unit B , as sketched in Fig. S5(a). Consider now the two parts separately: the controlled part A will synchronize as any other network under the effect of interaction control for coupling strength $\alpha \geq \alpha_{\min,A}$. Considering input from B to A as a small outside perturbation, interaction control still enables stable synchronization of the controlled part A .

Assuming A is synchronized, unit B receives input in form of the synchronized state from all its connections to A . Synchronization will typically be stable only in a finite range of coupling strengths, $\alpha \in [\alpha_{\min,B}, \alpha_{\max,B}]$ (e.g. for Rössler or Chen systems). Thus, synchronizability of the complete network A and B is either possible in a finite interval of coupling strengths if $\alpha_{\min,A} < \alpha_{\max,B}$ or synchronization is only stable in part A of the network if $\alpha_{\min,A} \geq \alpha_{\max,B}$.

For larger sets of uncontrolled units the structure of the individual sets and their interaction becomes more important. The general idea, however, holds: both parts of the network must be synchronizable for the same coupling strength in order to allow synchronization of the complete network [illustrated in Fig. S5(b-e)].

In summary, success of interaction control in partially controlled networks depends on the network structure. Whereas synchronization of the whole network possible under the (necessary) condition $\alpha_{\min,A} < \alpha_{\max,B}$, the controlled part A of the network will always be synchronizable for large enough coupling strengths (disregarding the perturbation by part B). This potentially enables selective control over specific parts of a given network as long as the outside disturbance of the uncontrolled units is not too large.

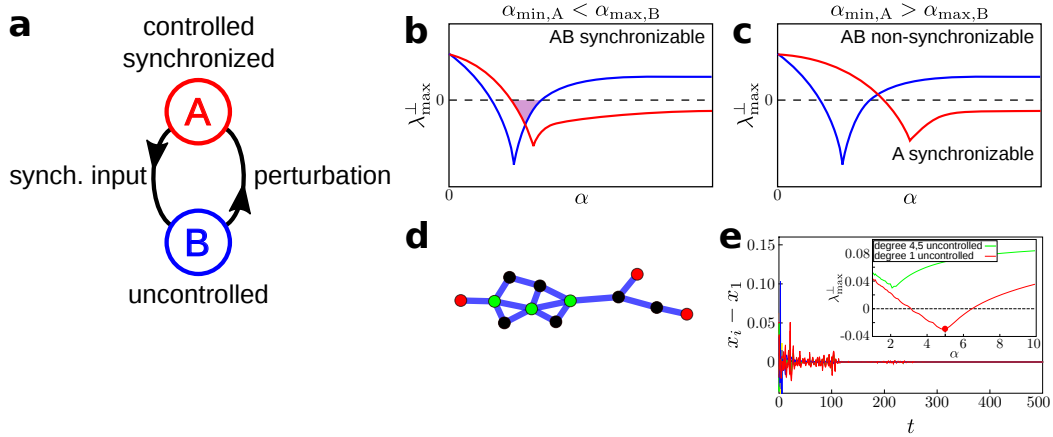


Figure S5. Panel (a) shows a schematic of a partially controlled network: we consider two sets of units, controlled (A) and uncontrolled (B). Panel (b) and (c) show a sketch of the maximum transverse Lyapunov exponent for the two parts A and B in two different cases. While interaction control guarantees synchronizability of the controlled part A as long as it is connected (disregarding the perturbation from B), synchronization of part B is typically only possible in a small range of coupling strengths, if at all. Consequently, synchronization of the complete network is only possible in the finite range of coupling strengths illustrated in panel (c) if both A and B are synchronizable for the same coupling strength. Panel (d) shows an example of a partially controlled network of Rössler units (see above: "Rössler oscillator for different parameters"). Two sets of uncontrolled nodes are marked in green and red, the corresponding maximum Lyapunov exponent is shown in the inset of panel (e). If nodes with large degree [green in panel (d)] are uncontrolled, complete synchronization of the network is not possible and the system diverges. For another set of uncontrolled nodes [red in panel (d)] synchronization is possible in a finite range of coupling strengths. Panel (e) shows one set of trajectories for the second case with $\alpha = 5$, the differences between the units disappear and all units synchronize.

References

1. Pecora, L. M. & Carroll, T. L. Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.* **80**, 2109–2112 (1998).
2. Schröder, M., Mannattil, M., Dutta, D., Chakraborty, S. & Timme, M. Transient uncoupling induces synchronization. *Phys. Rev. Lett.* **115**, 054101 (2015).
3. Rössler, O. E. An equation for continuous chaos. *Phys. Lett. A* **57**, 397–398 (1976).
4. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).
5. Chen, G. & Ueta, T. Yet another chaotic attractor. *Int. J. Bifur. Chaos* **09**, 1465–1466 (1999).

Chapter 5

Original manuscript: A universal order parameter for synchrony in networks of limit cycle oscillators

Citation

Schröder, M., Timme, M., and Witthaut, D. (2017),
A universal order parameter for synchrony in networks of limit cycle oscillators,
Chaos **27**, 073119

DOI: 10.1063/1.4995963

©2017 Schröder, Timme, and Witthaut. Reprinting the Version of Record for use in the authors thesis is explicitly permitted by AIP Publishing's License to Publish Agreement.

Original contribution

Design of the research was done together with D. Witthaut. I performed all simulations and parts of the analytical calculations, analyzed the data and generated all figures. I designed the contents and wrote the manuscript together with D. Witthaut. I revised the manuscript during the review process.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder



A universal order parameter for synchrony in networks of limit cycle oscillators

Malte Schröder,¹ Marc Timme,^{1,2} and Dirk Witthaut^{3,4}

¹Network Dynamics, Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany

²Chair for Network Dynamics, Institute for Theoretical Physics and Center for Advancing Electronics Dresden (cfaed), Technical University of Dresden, 01069 Dresden, Germany

³Forschungszentrum Jülich, Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), 52428 Jülich, Germany

⁴Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany

(Received 12 April 2017; accepted 13 July 2017; published online 27 July 2017)

We analyze the properties of order parameters measuring synchronization and phase locking in complex oscillator networks. First, we review network order parameters previously introduced and reveal several shortcomings: none of the introduced order parameters capture all transitions from incoherence over phase locking to full synchrony for arbitrary, finite networks. We then introduce an alternative, universal order parameter that accurately tracks the degree of partial phase locking and synchronization, adapting the traditional definition to account for the network topology and its influence on the phase coherence of the oscillators. We rigorously prove that this order parameter is strictly monotonously increasing with the coupling strength in the phase locked state, directly reflecting the dynamic stability of the network. Furthermore, it indicates the onset of full phase locking by a diverging slope at the critical coupling strength. The order parameter may find applications across systems where different types of synchrony are possible, including biological networks and power grids. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4995963>]

Many dynamical systems in physics, biology, or engineering can be described as coupled phase oscillators, often in a network with a complex interaction topology. The prototypical model considered in this context is networks of Kuramoto oscillators. To study the synchronization in such systems, several order parameters have been introduced, adapting the original Kuramoto order parameter, defined for all-to-all coupled oscillators, to complex interaction networks. However, none of the order parameters manages to fully track the transition from oscillators moving at their individual frequencies to full synchronization of the network. Here, we propose a universal order parameter to study synchronization in finite networks of phase oscillators, tracking all stages of synchronization. This order parameter may be used to study systems where different stages of synchrony are relevant. Additionally, we rigorously prove several helpful qualities, relating the order parameter not only to the synchrony, but also to the dynamical stability of the network.

I. INTRODUCTION

Many oscillatory systems enter stable limit cycles as their dynamic steady state. If such systems are coupled, they often interact only through their positions along their periodic orbit, their phases. The simplest prototypical model to describe such coupled phase oscillators is the celebrated Kuramoto model.^{1,2} It characterizes the collective dynamics of a variety of phase oscillator systems ranging from chemical reactions³ and neural

networks^{4,5} to coupled Josephson junctions,⁶ laser arrays,⁷ optomechanical systems,⁸ and mean-field quantum systems.^{9,10}

Studies of the Kuramoto model and more general phase oscillator networks typically focus on the onset of synchronization between the individual oscillators.^{1-3,11,12} Starting from the analytical results for the mean field behavior in the all-to-all coupled Kuramoto model, correctly predicting the emergence of partial phase locking, extensions of this result to various network topologies were developed.¹³⁻¹⁶ These extensions often use a similar methodology and define an adapted order parameter to analyze the transition to synchrony. Interestingly, none of these order parameters captures all transitions from the incoherent to the completely synchronized state for arbitrary, finite networks.

Depending on the application, different states of phase ordering are relevant and a different order parameter is appropriate. Commonly, the onset of partial phase locking has received most interest.¹⁻³ For example, partial phase locking indicates the growth of number fluctuations in quantum mean-field models.^{9,10} In contrast, in technical systems such as power grids, a fully phase locked state is required for stable operation.¹⁷⁻²⁰

We propose a universal order parameter that accurately reflects the phase coherence of phase oscillators in any network, describing the initial growth of partially phase locked clusters as well as the convergence to full synchrony. This order parameter is particularly suited to study the fully phase locked state as it directly reflects the dynamic stability of this steady state. It increases monotonically with the coupling strength, in contrast to previously defined mean field order parameters.

II. PHASE OSCILLATORS AND THE KURAMOTO MODEL

Limit cycles are ubiquitous as dynamically stable states in a wide range of systems. When such systems are coupled, interactions can typically be approximated as interactions between their phases θ_i . The Kuramoto model

$$\frac{d\theta_i}{dt} = \omega_i + K \sum_{j=1}^N A_{ij} \sin(\theta_j - \theta_i), \quad (1)$$

is one of the simplest models for such coupled phase oscillators. It describes the dynamics of N oscillators with natural frequencies ω_i and sinusoidal coupling. The parameter K denotes the coupling strength of the interactions and $A_{ij} \in \{0, 1\}$ is the adjacency matrix of the interaction network, describing which nodes interact with which other nodes. The results easily extend to inhomogeneous coupling strengths with $A_{ij} \in \mathbb{R}$. In many applications, interactions between individual oscillators are reciprocal, and in the following, we assume an undirected network, i.e., a symmetric adjacency matrix $A_{ij} = A_{ji}$. Similarly, we can without loss of generality consider a co-rotating frame such that the natural frequencies of the oscillators are centered around 0 and we have $\sum_i \omega_i = 0$, where the sum runs from 1 to N . In the following, we only consider connected networks, as otherwise we can treat the connected sub-systems individually.

The dynamics of coupled Kuramoto oscillators depends strongly on the strength K of the interactions. For small coupling K , all oscillators rotate (almost) independently with their natural frequencies ω_j . In this state, the phases are *incoherent*. Above some critical coupling strength $K \geq K_{c1}$, a subset of the oscillators starts to synchronize such that their time averaged frequencies $\langle \frac{d\theta_i}{dt} \rangle_t$ become identical. The phases of these oscillators then move together in a *partially phase locked* state and their phase differences $\theta_i - \theta_j$ are bounded. When the coupling becomes even stronger, $K \geq K_{c2}$, a *fully phase locked* state appears. All oscillators synchronize to a common frequency $\frac{d\theta_i}{dt} = \text{const.} = 0$ and the phase differences between all nodes become constant $\theta_i - \theta_j = \text{const.}$ Further increasing the coupling reduces the phase differences until *complete synchronization* of the oscillators, defined by $\theta_i - \theta_j = 0$, is achieved as $K \rightarrow \infty$. The details of this evolution and specific transitions between the different stable steady states depend on the structure of the interaction network and the distribution of the natural frequencies.^{2,16,21,22} An example illustrating the dynamics of a small random network of oscillators for various coupling strengths is shown in Fig. 1.

Most studies focus on the transition from incoherent oscillators moving at their individual frequencies to a partially phase locked state.^{1-3,11,12} In a variety of technical systems, however, partial phase coherence is not sufficient for stable function. For instance, Kuramoto-like dynamics appear in a second order model describing the frequency dynamics of power grids^{17-20,22-25}

$$M_i \frac{d^2 \theta_i}{dt^2} + D_i \frac{d\theta_i}{dt} = P_i + \sum_{j=1}^N K A_{ij} \sin(\theta_j - \theta_i). \quad (2)$$

Here, M_i is the inertia, D_i the damping coefficient, and P_i the power injection at node i . The phases $\theta_i(t)$ describe the state

of rotating machines (generators or motors) and the coupling their interactions via power transmission lines. In the steady state $\frac{d\theta_i}{dt} = 0$, required for stable operation of the power grid, all machines work at the same frequency. This state is characterized by the same equations that describe a fully phase locked state in the Kuramoto model. The stability of this state and how the phase cohesiveness in the network scales with the coupling strength is an important question.²⁶

Ideally, a universal order parameter would be able to characterize both the transition to partial and to full phase locking and the properties of a phase locked state in arbitrary, especially finite networks.

III. KURAMOTO ORDER PARAMETERS

To quantitatively study the transitions from an incoherent state to a fully synchronous state, one typically introduces an order parameter to measure the phase coherence. For the original all-to-all coupling model, Kuramoto introduced the complex order parameter^{2,3}

$$r(t) e^{i\psi(t)} = \frac{1}{N} \sum_{i=1}^N e^{i\theta_i}, \quad (3)$$

where $\psi(t)$ describes the collective phase of all oscillators and $r(t)$ the degree of phase coherence. A single measure for the phase ordering is then given by the long time average of the absolute value of the order parameter

$$\begin{aligned} r_{\text{Kuramoto}}^2 &= \langle |r(t) e^{i\psi(t)}|^2 \rangle_t = \langle r(t)^2 e^{i\psi(t)} e^{-i\psi(t)} \rangle_t \\ &= \left\langle \frac{1}{N^2} \sum_{i,j=1}^N e^{i(\theta_i - \theta_j)} \right\rangle_t \\ &= \frac{1}{N^2} \sum_{i,j=1}^N \langle \cos(\theta_i - \theta_j) \rangle_t. \end{aligned} \quad (4)$$

This order parameter measures the average of the phase differences of all pairs of oscillators. If the oscillators are incoherent, the time average vanishes and the order parameter is 0. When a fraction of the oscillators are partially phase locked, the cosine of their phase differences becomes positive and does not disappear in the time average; the order parameter becomes positive.

In the original case for N all-to-all coupled oscillators with natural frequencies ω_i following a distribution $g(\omega)$, the mean-field theory correctly predicts the transition to partial phase coherence at the critical coupling $K_{c1} = 2/[\pi g(0)]$ if the frequency distribution g is unimodal and symmetric around zero. For larger coupling strengths $K > K_{c1}$, the order parameter then grows continuously as $r(K) \propto \sqrt{1 - K_{c1}/K}$.² As such, this order parameter characterizes the transition from an *incoherent* state to a *partially* phase locked state.

This original order parameter is clearly unsuited when studying more general interaction networks. One would compare the phases of two oscillators in the network that are only interacting indirectly via a (possibly very long) chain of intermediate oscillators. As such, several adaptations of the

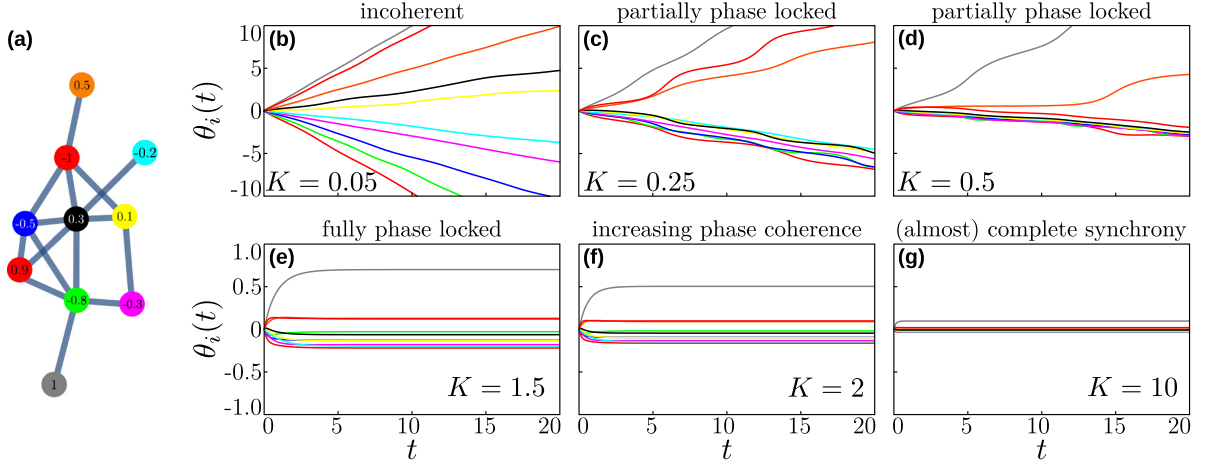


FIG. 1. Synchronization in the Kuramoto model. Dynamics of the Kuramoto model for $N = 10$ oscillators with a random interaction network. The phase coherence between neighboring oscillators increases with the coupling strength, eventually leading to full synchrony of all oscillators. (a) Topology of the interaction network, the numbers denote ω_i of the respective oscillator. (b) For small coupling, the oscillators move (almost) independently with their individual frequencies (slope). (c) and (d) As the coupling strength increases beyond $K_{c1} = 0.1$ some oscillators enter a partially phase locked state and their phases evolve with the same time-averaged frequency. (e) If the coupling strength becomes larger than $K_{c2} = 1$, all nodes are phase locked and move with the same constant frequency $\frac{d\theta_i}{dt} = 0$. (f) and (g) Further increasing the coupling strength reduces the phase differences of the oscillators until complete synchrony $\theta_i - \theta_j = 0$ is achieved for $K \rightarrow \infty$.

order parameter have been introduced to study the effect of the network topology on the synchronization of Kuramoto oscillators:

The first definition used by Restrepo *et al.*^{16,27,28} considers an intuitively defined local order parameter

$$r_i = \left| \sum_{j=1}^N A_{ij} \langle e^{i\theta_j} \rangle_t \right|, \quad (5)$$

for oscillator i , measuring the phase coherence of all neighboring oscillators. A global order parameter is then easily defined as the average of the local order parameters

$$r_{\text{net}} = \frac{\sum_{i=1}^N r_i}{\sum_{i=1}^N k_i}, \quad (6)$$

where $k_i = \sum_j A_{ij}$ is the degree of node i .

A second definition^{14,29} adapts the original order parameter Eq. (3) weighting each node with its degree

$$r_{\text{mf}} = \left\langle \left| \frac{\sum_{i=1}^N k_i e^{i\theta_i}}{\sum_{i=1}^N k_i} \right| \right\rangle_t. \quad (7)$$

This order parameter ignores the specific network topology in favor of a mean-field view of network ensembles to simplify analytical calculations.

Finally, a definition of an order parameter to study local synchronization used in Ref. 15 derives from the original order parameter Eq. (4), restricting it to the network topology and only averaging over the phase differences between directly connected nodes

$$r_{\text{link}} = \frac{1}{\sum_{i=1}^N k_i} \sum_{i,j=1}^N A_{ij} |\langle e^{i(\theta_i - \theta_j)} \rangle_t|. \quad (8)$$

The above order parameters work well for their respective use, for example, to study synchronization analytically in

mean-field network models. However, none of them accurately captures the whole transition to synchronization, especially in smaller networks. We illustrate this in Fig. 2 for a small random network: While r_{net} clearly captures the transition to full phase locking at $K_{c2} = 1$, it is effectively 0 before full phase locking becomes stable and does not indicate where individual nodes enter the partially phase locked state for $K < 1$. Conversely, r_{link} describes these transitions but cannot cover the convergence to full synchrony as $r_{\text{link}} = 1$ in the fully phase locked state, regardless of the network topology. Finally, r_{mf} works well to describe the behavior for a large ensemble of networks but is clearly unsuited for use with specific, particularly small, networks as it ignores the specific network structure and is large already for weak coupling. It is easy to construct further examples where, for instance, the mean field order parameter r_{mf} is non-monotonous with respect to the coupling strength K , even in the fully phase locked state.

IV. A UNIVERSAL ORDER PARAMETER FOR COMPLEX NETWORKS

In order to have both a practically applicable and relevant order parameter as well as describe the whole evolution from an incoherent state to complete synchronization, we propose a universal network order parameter:

Definition 1. Given a network of coupled Kuramoto oscillators Eq. (1), phase ordering is measured by

$$\begin{aligned} r_{\text{uni}} &= \frac{1}{\sum_{i=1}^N k_i} \sum_{i,j=1}^N A_{ij} \langle \Re(e^{i(\theta_i - \theta_j)}) \rangle_t \\ &= \frac{1}{\sum_{i=1}^N k_i} \sum_{i,j=1}^N A_{ij} \langle \cos(\theta_i - \theta_j) \rangle_t. \end{aligned} \quad (9)$$

As r_{link} , this definition respects the topology of the interaction network and considers only phase differences between

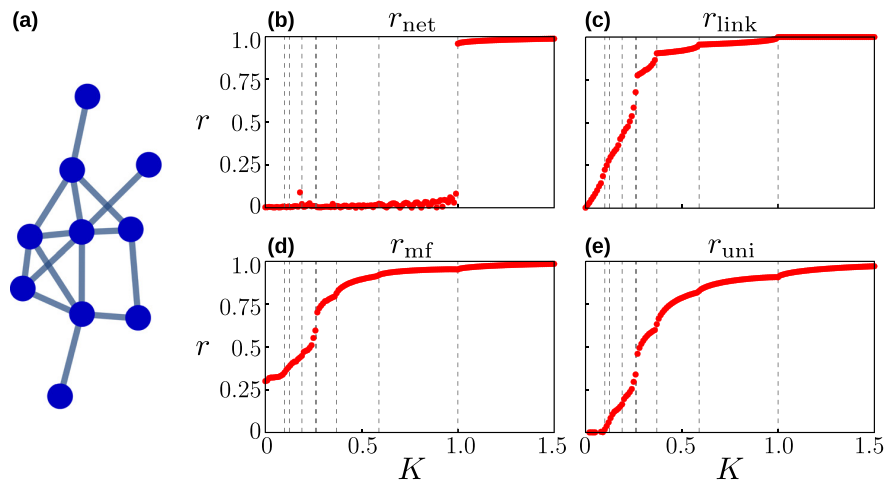


FIG. 2. Order parameters to measure phase coherence in networks. Different order parameters measuring the phase coherence in complex networks of Kuramoto oscillators, describing the transition from a completely incoherent state [$K=0$, cf. Fig. 1(b)] to full synchrony [$K \rightarrow \infty$, cf. Fig. 1(g)]. None of the order parameters used in the literature r_{net} , r_{mf} , and r_{link} captures all transitions. (a) Topology of the interaction network, cf. Fig. 1(a). (b) r_{net} is almost 0 until the fully phase locked state becomes stable at $K=1$. It fails to capture transitions in the partially phase locked regime. (c) In contrast, r_{link} captures the transitions in the partially phase locked regime well. However, $r_{\text{link}}=1$ in the fully phase locked state for $K \geq 1$ and does not capture the convergence to complete synchrony. (d) r_{mf} measures globally averaged phase coherence. It fails to accurately represent the incoherent and partially phase locked state with respect to the actual network topology, especially for small networks. (e) Our universal order parameter r_{uni} accurately reflects the degree of phase coherence in all stages of synchronization. All results show the long time limit of the order parameter starting from identical initial conditions $\theta_i=0$; the black dashed lines mark transitions where single nodes enter a (partially) phase locked state.

neighboring nodes. In contrast to r_{link} , the definition of r_{uni} reduces to the original Kuramoto order parameter Eq. (4) for a completely connected network as desired. Figure 2(d) illustrates the behavior in comparison to the other network order parameters, showing that it accurately captures the transitions in all stages of phase locking (cf. Fig. 3).

A. Synchronization and stability

The order parameter r_{uni} gives a full account of the emergence of synchrony. It accurately follows both the transitions to partially and fully phase locked states as well as the convergence to complete synchrony.

We illustrate this central result in Fig. 2 for a small random network. Whenever one of the nodes enters a partially phase locked state, we observe a strong kink in $r_{\text{uni}}(K)$. Hence, we can directly track the growth of phase locked clusters. In fact, the slope dr_{uni}/dK diverges when approaching these transition points from the right. We rigorously

prove this result for the transition to full phase locking below (cf. Theorem 1).

The universal order parameter has further advantages compared to the alternatives discussed above. First, r_{uni} quantifies the dynamical stability of a phase-locked steady state (cf. Lemma 1). This becomes most apparent in a ring of N oscillators with identical natural frequencies, $\omega_i=0$ for all $i \in \{1, 2, \dots, N\}$, where all interactions have identical coupling strength $K=1$. Clearly, in a fully phase locked state all phase differences between neighboring nodes need to be identical while the cumulative phase difference around the ring must be a multiple of 2π .^{30,31} Under these conditions, we can characterize the phase locked states by a mode m describing the total phase change around the ring $2\pi m$. The individual phases are then given by

$$\theta_i^* = \frac{2\pi im}{N}, \quad (10)$$

	partial phase locking	full phase locking	complete synchrony	classifies stability	finite networks
r_{net}	✗	✓	✓	✓	✗
r_{link}	✓	✓	✗	✗	✓
r_{mf}	✗	✓	✓	✗	✗
r_{uni}	✓	✓	✓	✓	✓

FIG. 3. A universal order parameter. None of the order parameters used in the literature r_{net} , r_{mf} , and r_{link} capture all transitions. Following the observations in Fig. 2, r_{net} fails to capture transitions in the partially phase locked regime. It also fails to describe phase coherence for some small networks, most easily seen for just two connected oscillators. r_{link} does not capture the transition to complete synchrony and, since $r_{\text{link}}=1$ in the fully phase locked state, it does not classify stability. r_{mf} does not reflect the phase ordering in networks for partially or fully phase locked states, since it measures global phase coherence. As such, it does not represent stability of the phase locked steady states which depends on local phase differences and is not suited for small networks. The order parameter r_{uni} accurately reflects the transitions for all stages of synchronization and correctly classifies stability of different phase locked states in arbitrary, even small networks.

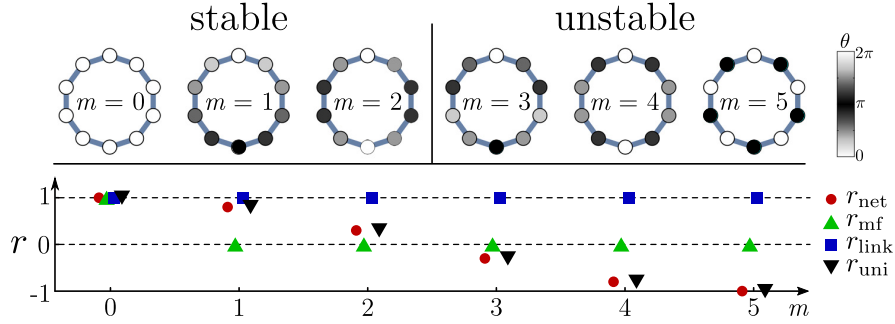


FIG. 4. Order parameters and stability. Steady states in a ring network with $N=10$ nodes and the corresponding values of the different order parameters (shifted horizontally for better visibility). The state $m=0$ is the most stable as the phase differences between neighboring nodes are the smallest. The phase locked states become more unstable with increasing m . r_{mf} and r_{link} do not provide information about the stability of the steady state, being either zero for most of the states or identical to one for all phase locked states, respectively. Our universal order parameter r_{uni} accurately reflects the stability of the different states.

with $m \in \{-N/2, -N/2 + 1, \dots, N/2\}$, illustrated for $m \geq 0$ in Fig. 4. Here and in the following, we use an asterisk $*$ to denote a phase locked steady state θ_i^* of the Kuramoto model Eq. (1).

The phase locked states with $|\theta_i^* - \theta_{i-1}^*| < \pi/2$, that means $m \in \{-N/4, N/4\}$, are linearly stable, and the remaining states are unstable. Our order parameter r_{uni} reflects the linear stability of these different steady states - the state with perfectly aligned phases ($m=0$) is most stable and has $r_{uni} = 1$. All other states have larger phase differences, which impede dynamical stability, and consequently lower values of r_{uni} . This information is completely lost for the alternatives r_{link} and r_{mf} , the first one being identically one for all phase-locked states and the second one being one for the fully aligned state and zero otherwise.

The classification of stability is due to the fact that r_{uni} Eq. (9) counts only the phase differences in the stable region as positive contributions, i.e., when $|\theta_i^* - \theta_j^*| < \pi/2$. As the stability of a phase locked state is directly related to these phase differences, with phase differences close to 0 corresponding to more stable states, the order parameter directly reflects the systems stability of any phase locked state, relevant, for example, for applications to power grids.

A further advantage of r_{uni} for the analysis of phase-locked states is monotonicity (cf. Theorem 2). Intuitively, we expect that an increase in the coupling K leads to a stronger alignment of the phases and thus to an increase in the order parameter. This expectation can be violated for the mean-field order parameter r_{mf} , as it measures global alignment, but an increase in the coupling acts only locally on the links. In contrast, we rigorously prove below that the order parameter r_{uni} is monotonic in the coupling strength K for a phase-locked steady state.

B. Analytical results

To formalize these observations, first consider the linear stability of a phase locked state θ^* for $K \geq K_{c2}$: A small perturbation ξ around the steady state, $\theta_j = \theta_j^* + \xi_j$, evolves as

$$\frac{d}{dt} \xi = \mathbf{J} \xi + \mathcal{O}(\xi^2), \quad (11)$$

where we make use of vector notation $\xi = (\xi_1, \dots, \xi_N)^T$. The Jacobian matrix \mathbf{J} quantifies the linear stability of a phase-

locked steady state. It always has one trivial eigenvalue $\lambda_1 = 0$ with eigenvector $\mathbf{v}_1 = (1, 1, \dots, 1)^T$, representing a global uniform shift of all phases which does not affect the phase-locking of the nodes. In a stable phase locked state, all other eigenvalues are negative $0 > \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$. We denote the associated eigenvectors as $\mathbf{v}_2, \dots, \mathbf{v}_N$.

We can then formalize the above observations about r_{uni} in the following theorems:

Theorem 1. *Given a network of coupled Kuramoto oscillators Eq. (1) with $\sum_i \omega_i = 0$ and $\omega \cdot \mathbf{v}_2 \neq 0$, the derivative of the order parameter r_{uni} Eq. (9) diverges when the fully phase locked state becomes unstable at the critical coupling K_{c2}*

$$dr_{uni}/dK \rightarrow \infty \quad \text{for } K \rightarrow K_{c2}^+.$$

Theorem 2. *Given a network of coupled Kuramoto oscillators Eq. (1) with $\sum_i \omega_i = 0$, in a fully phase locked regime $K > K_{c2}$ the order parameter r_{uni} Eq. (9) is strictly larger than zero for every stable phase locked state and increases monotonically with increasing K .*

In the remainder of this section, we provide the proof for these theorems with the help of two lemmas, relating the order parameter to the eigenvalues of the Jacobian:

Lemma 1. *Given a network of coupled Kuramoto oscillators Eq. (1) with $\sum_i \omega_i = 0$ and $K \geq K_{c2}$ in the stable phase locked state, the order parameter r_{uni} Eq. (9) is given by the negative trace of the Jacobian \mathbf{J}*

$$\begin{aligned} r_{uni} &= -\frac{1}{K \sum_{i=1}^N k_i} \text{tr}(\mathbf{J}) \\ &= -\frac{1}{K \sum_{i=1}^N k_i} \sum_{j=2}^N \lambda_j. \end{aligned} \quad (12)$$

Proof. Explicit calculation of the Jacobian matrix \mathbf{J} in Eq. (11) yields

$$\begin{aligned} J_{i,j} &= K A_{i,j} \cos(\theta_i^* - \theta_j^*) \quad \text{for } i \neq j, \\ J_{i,i} &= -K \sum_{j=1}^N A_{i,j} \cos(\theta_i^* - \theta_j^*). \end{aligned} \quad (13)$$

The lemma then follows directly by calculating the trace. The second equality follows from the fact that the largest eigenvalue of \mathbf{J} is $\lambda_1 = 0$. \square

Given that the eigenvalues of the Jacobian $\lambda_2, \dots, \lambda_N < 0$ are all negative for a stable phase locked state, $K > K_{c2}$, it immediately follows that the order parameter r_{uni} must be positive.

To finish proving the theorems above, we now also relate the derivative dr_{uni}/dK to the eigenvalues $\lambda_1, \dots, \lambda_N$ of the Jacobian matrix and their corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_N$:

Lemma 2. *Given a network of coupled Kuramoto oscillators Eq. (1) with $\sum_i \omega_i = 0$ and $K \geq K_{c2}$, the derivative of the order parameter with respect to the coupling strength is given by*

$$\frac{dr_{\text{uni}}}{dK} = \frac{2}{K^2} \sum_{i=1}^N \frac{1}{k_i} \frac{1}{-\lambda_n} (\mathbf{v}_n \cdot \boldsymbol{\omega})^2 \geq 0. \quad (14)$$

Proof. Consider a global change of the coupling strength $K' = K + \kappa$. This perturbation induces a small change of the steady state phases of the network, $\theta_m^* \rightarrow \theta'_m = \theta_m^* + \zeta_m$. Expanding the steady state condition

$$0 = \omega_i + (K + \kappa) \sum_{m=1}^N A_{i,m} \sin(\theta_m^* + \zeta_m - \theta_i^* - \zeta_i),$$

to leading order in κ and the ζ_m yields

$$\begin{aligned} 0 &= \kappa \sum_{m=1}^N A_{i,m} \sin(\theta_m^* - \theta_i^*) + \sum_{m=1}^N J_{i,m} \zeta_m \\ \Rightarrow \sum_{m=1}^N J_{i,m} \zeta_m &= -\frac{\kappa}{2} \sum_{\ell,m=1}^N A_{\ell,m} \sin(\theta_m^* - \theta_\ell^*) (\delta_{i,\ell} - \delta_{i,m}), \end{aligned}$$

for all $i \in \{1, \dots, N\}$ using the definition of the Jacobian Eq. (13) and the Kronecker δ symbol. In vectorial notation, this set of equations can be written as

$$\mathbf{J} \boldsymbol{\zeta} = -\frac{\kappa}{2} \sum_{\ell,m=1}^N A_{\ell,m} \sin(\theta_m^* - \theta_\ell^*) \mathbf{q}_{(\ell,m)}, \quad (15)$$

where we define the vector $\mathbf{q}_{(\ell,m)}$, whose i th component is given by $q_{(\ell,m),i} = \delta_{i,\ell} - \delta_{i,m}$. The matrix \mathbf{J} is singular, but the vectors $\mathbf{q}_{(\ell,m)}$ are orthogonal to its kernel $[\mathbf{v}_1 = (1, 1, \dots, 1)^T]$ such that we can solve Eq. (15) using the Moore-Penrose pseudo-inverse \mathbf{J}^+ . Decomposing \mathbf{J} into eigenvalues and eigenstates, we thus obtain

$$\begin{aligned} \boldsymbol{\zeta} &= -\frac{\kappa}{2} \sum_{\ell,m=1}^N A_{\ell,m} \sin(\theta_m^* - \theta_\ell^*) \mathbf{J}^+ \mathbf{q}_{(\ell,m)} \\ &= -\frac{\kappa}{2} \sum_{\ell,m=1}^N \sum_{n=2}^N A_{\ell,m} \sin(\theta_m^* - \theta_\ell^*) \frac{1}{\lambda_n} (\mathbf{v}_n \cdot \mathbf{q}_{(\ell,m)}) \mathbf{v}_n. \end{aligned}$$

We then find for the change of the phases

$$\begin{aligned} \frac{d(\theta_j - \theta_i)}{dK} &= \mathbf{q}_{(j,i)} \cdot \lim_{\kappa \rightarrow 0} \underbrace{\frac{\theta(K + \kappa) - \theta(K)}{\kappa}}_{=\boldsymbol{\zeta}/\kappa} \\ &= -\frac{1}{2} \sum_{\ell,m=1}^N A_{\ell,m} \sin(\theta_m^* - \theta_\ell^*) \\ &\quad \times \sum_{n=2}^N \frac{1}{\lambda_n} (\mathbf{q}_{(\ell,m)} \cdot \mathbf{v}_n) (\mathbf{q}_{(j,i)} \cdot \mathbf{v}_n). \end{aligned}$$

Hence, the derivative of the order parameter is given by

$$\begin{aligned} \frac{dr_{\text{uni}}}{dK} &= \frac{1}{\sum_{i=1}^N k_i} \sum_{i,j=1}^N A_{i,j} \frac{d \cos(\theta_i^* - \theta_j^*)}{dK} \\ &= \frac{1}{\sum_{i=1}^N k_i} \sum_{i,j=1}^N A_{i,j} \sin(\theta_i^* - \theta_j^*) \frac{d(\theta_j - \theta_i)}{dK} \\ &= \frac{1}{2 \sum_{i=1}^N k_i} \sum_{n=2}^N \frac{1}{-\lambda_n} \left[\sum_{i,j=1}^N A_{i,j} \sin(\theta_i^* - \theta_j^*) (\mathbf{q}_{(j,i)} \cdot \mathbf{v}_n) \right]^2. \end{aligned}$$

Now, we use the steady state condition to simplify this expression. We write $\mathbf{q}_{(j,i)} \cdot \mathbf{v}_n = \mathbf{v}_{n,j} - \mathbf{v}_{n,i}$, where $\mathbf{v}_{n,j}$ denotes the j th component of the vector \mathbf{v}_n and we obtain

$$\begin{aligned} &\sum_{i,j=1}^N A_{i,j} \sin(\theta_i^* - \theta_j^*) (\mathbf{q}_{(j,i)} \cdot \mathbf{v}_n) \\ &= \sum_{j=1}^N \mathbf{v}_{n,j} \underbrace{\sum_{i=1}^N A_{i,j} \sin(\theta_i^* - \theta_j^*)}_{=-\omega_j/K} \\ &\quad - \sum_{i=1}^N \mathbf{v}_{n,i} \underbrace{\sum_{j=1}^N A_{i,j} \sin(\theta_i^* - \theta_j^*)}_{=+\omega_i/K} \\ &= -\frac{2}{K} \mathbf{v}_n \cdot \boldsymbol{\omega}. \end{aligned} \quad (16)$$

The derivative of the order parameter then becomes

$$\frac{dr_{\text{uni}}}{dK} = \frac{2}{K^2} \sum_{i=1}^N \frac{1}{k_i} \frac{1}{-\lambda_n} (\mathbf{v}_n \cdot \boldsymbol{\omega})^2,$$

finishing the proof of Lemma 2. \square

For any stable steady state, we have $\lambda_n < 0$ for all $n \in \{2, \dots, N\}$ such that the slope is non-negative. It can become zero only if $\mathbf{v}_n \cdot \boldsymbol{\omega} = 0$ for all $n \in \{2, \dots, N\}$. As the eigenvectors form an orthonormal basis, this would imply that $\boldsymbol{\omega}$ is parallel to \mathbf{v}_1 . As we assume $\sum_j \omega_j = 0$, this is only possible if $\boldsymbol{\omega} = \mathbf{0}$ and we have $dr_{\text{uni}}/dK > 0$ for $K > K_{c2}$.

Finally, as $K \rightarrow K_{c2}^+$ from above the phase locked state becomes unstable with $\lambda_2 \rightarrow 0$. With the assumption $\boldsymbol{\omega} \cdot \mathbf{v}_2 \neq 0$, it follows that the derivative diverges, concluding the proofs for both theorems.

C. Extension to general coupling functions

The proposed order parameter provides a measure for the phase coherence in any network of phase oscillators, independent of the coupling function. We have derived the above analytical results specifically for the most common model, Kuramoto phase oscillators. The proofs relied on the fact that the order parameter directly involves $\cos(\theta_j - \theta_i)$

that also appears in the Jacobian matrix of the system. Thus, while the results are not, in general, valid for arbitrary coupling functions, the above arguments can in many cases be extended to other common models of phase oscillators.

In particular, we consider the second order Kuramoto model Eq. (2) as an application of a related system describing the frequency dynamics of power grids. Since the condition describing the steady operating state [Eq. (2) with derivatives set to 0] is exactly identical to that describing a fully phase locked state in the first order Kuramoto model,²² the arguments given in the proof of Lemma 2 hold without modifications. The new order parameter r_{uni} describes the phase coherence and increases with the coupling strength in the phase locked state also in the second order model. Similarly, the order parameter will always be positive. While the order parameter is not given directly by the trace of the Jacobian of the second order model, it is still closely related to its Eigenvalues, since the Jacobian of the first order system is a sub-block of the second order Jacobian. For example, it is under certain conditions possible to easily express the Eigenvalues of the second order system in terms of the first order Eigenvalues.²² In these cases, this relation directly translates to the order parameter as well.

Similarly, the Kuramoto-Sakaguchi-model³² is a common modification, introducing an off-set in the coupling function $\sin(\theta_j - \theta_i - \alpha)$. In this case, Lemma 1 can be adapted with an additional factor $1/\cos(\alpha)$, easily seen by direct calculation. However, depending on the topology of the interaction network, the existence of frustrated states³³ means that an increase in the coupling strength does not necessarily result in stronger phase coherence. As such, $dr_{\text{uni}}/dK > 0$ cannot be expected to hold. The new order parameter r_{uni} nonetheless provides a way to quantify the phase coherence of such systems, even if they cannot reach full synchrony.

V. CONCLUSION

Kuramoto oscillators are the prototypical systems used to study the synchronization behavior of limit cycle oscillators. The order parameters introduced to study this synchronization capture different aspects of the transition to synchrony. None of the order parameters previously suggested for Kuramoto oscillators on complex networks describes all transitions to partial and full phase locking as well as the convergence to full synchrony in arbitrary networks.

Here, we have proposed a universal order parameter accurately describing the phase coherence in networks of phase oscillators. This order parameter recovers the original Kuramoto order parameter for a fully connected network of oscillators. We have analytically shown that the slope of the order parameter diverges when the fully phase locked state becomes stable, accurately marking this transition even in small networks. For larger coupling strengths, a monotonic increase reflects the slow convergence to complete synchrony and directly relates to the stability of the phase locked state, important, for example, for applications to

power grid models where a fully phase locked state is required for stable operation.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Göttingen Graduate School for Neurosciences and Molecular Biosciences (DFG Grant No. GSC 226/2 to M.S.), the Helmholtz Association (Grant No. VH-NG-1025 to D.W.), the German Federal Ministry of Education and Research (BMBF Grant Nos. 03SF0472B and 03SF0472E to M.T. and D.W.), and the Max Planck Society to M.T. Supported by a grant by the German Research Foundation (DFG) through the Center of Excellence ‘Center for Advancing Electronics Dresden (cfaed)’.

¹Y. Kuramoto, in *International Symposium on Mathematical Problems in Theoretical Physics*, Lecture Notes in Physics Vol. 39, edited by H. Araki (Springer, New York, 1975), p. 420.

²S. H. Strogatz, *Physica D* **143**, 1 (2000).

³Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Berlin, 1984).

⁴H. Sompolinsky, D. Golomb, and D. Kleinfeld, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 7200 (1990).

⁵C. Kist, M. Timme, and D. Battaglia, *Nat. Commun.* **7**, 11061 (2016).

⁶K. Wiesenfeld, P. Colet, and S. H. Strogatz, *Phys. Rev. Lett.* **76**, 404 (1996).

⁷A. G. Vladimirov, G. Kozireff, and P. Mandel, *Europhys. Lett.* **61**, 613 (2003).

⁸G. Heinrich, M. Ludwig, J. Qian, B. Kubala, and F. Marquardt, *Phys. Rev. Lett.* **107**, 043603 (2011).

⁹D. Witthaut and M. Timme, *Phys. Rev. E* **90**, 032917 (2014).

¹⁰D. Witthaut, S. Wimberger, R. Burioni, and M. Timme, *Nat. Commun.* **8**, 14829 (2017).

¹¹J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, and R. Spigler, *Rev. Mod. Phys.* **77**, 137 (2005).

¹²F. Dörfler and F. Bullo, *Automatica* **50**, 1539 (2014).

¹³M. Timme, *Europhys. Lett.* **76**, 367 (2006).

¹⁴S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).

¹⁵J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, *Phys. Rev. Lett.* **98**, 034101 (2007).

¹⁶A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Phys. Rep.* **469**, 93 (2008).

¹⁷M. Rohden, A. Sorge, M. Timme, and D. Witthaut, *Phys. Rev. Lett.* **109**, 064101 (2012).

¹⁸F. Dörfler, M. Chertkov, and F. Bullo, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2005 (2013).

¹⁹A. E. Motter, S. A. Myers, M. Anghel, and T. Nishikawa, *Nat. Phys.* **9**, 191 (2013).

²⁰D. Witthaut, M. Rohden, X. Zhang, S. Hallerberg, and M. Timme, *Phys. Rev. Lett.* **116**, 138701 (2016).

²¹E. A. Martens, E. Barreto, S. H. Strogatz, E. Ott, P. So, and T. M. Antonsen, *Phys. Rev. E* **79**, 026204 (2009).

²²D. Manik *et al.*, *Eur. Phys. J.: Spec. Top.* **223**, 2527 (2014).

²³D. Witthaut and M. Timme, *New J. Phys.* **14**, 083036 (2012).

²⁴D. Witthaut and M. Timme, *Eur. Phys. J. B* **86**, 377 (2013).

²⁵B. Schäfer, M. Matthiae, M. Timme, and D. Witthaut, *New J. Phys.* **17**, 015002 (2015).

²⁶F. Dörfler and F. Bullo, *SIAM J. Control Optim.* **50**, 1616 (2010).

²⁷J. G. Restrepo, E. Ott, and B. R. Hunt, *Phys. Rev. E* **71**, 036151 (2005).

²⁸J. G. Restrepo, E. Ott, and B. R. Hunt, *Chaos* **16**, 015107 (2006).

²⁹T. Ichinomiya, *Phys. Rev. E* **70**, 026116 (2004).

³⁰D. Manik, M. Timme, and D. Witthaut, ‘‘Cycle flows and multistability in oscillatory networks: An overview,’’ preprint arXiv:1611.09825 (2016).

³¹R. Delabays, T. Coletta, and P. Jacquod, *J. Math. Phys.* **57**, 032701 (2016).

³²H. Sakaguchi and Y. Kuramoto, *Prog. Theor. Phys.* **76**, 576 (1986).

³³P. S. Skardal, D. Taylor, J. Sun, and A. Arenas, *Phys. Rev. E* **91**, 010802 (2015).

Chapter 6

Original manuscript: Controlling percolation with limited resources

Citation

Schröder, M., Araújo, Nuno A. M., Sornette, D. and Nagler, J. (2017)
Controlling percolation with limited resources,
Phys. Rev. E **96**, 062302

DOI: 10.1103/PhysRevE.96.062302

©2017 by the American Physical Society. Permission to reprint this article in the authors' thesis is explicitly granted by the American Physical Society.

Original contribution

Conception and design of the research was done together with J. Nagler and N. Araújo. I performed all simulations and calculations, analyzed the data and generated all figures for the main manuscript with the exception of figure 1 (by J. Nagler). I designed and wrote the supplemental information and performed all simulations and created all figures. I designed the contents and wrote the manuscript with J. Nagler and revised it during the review process.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder

PHYSICAL REVIEW E **96**, 062302 (2017)**Controlling percolation with limited resources**Malte Schröder,^{1,*} Nuno A. M. Araújo,^{2,3,†} Didier Sornette,^{4,‡} and Jan Nagler^{5,6,§}¹*Network Dynamics, Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany*²*Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, P-1749-016 Lisboa, Portugal*³*Centro de Física Teórica e Computacional, Universidade de Lisboa, 1749-016 Lisboa, Portugal*⁴*Department of Management, Technology and Economics (D-MTEC), ETH Zürich, Scheuchzerstrasse 7, CH-8092 Zurich, Switzerland*⁵*Computational Physics for Engineering Materials, Institute for Building Materials, ETH Zürich, Wolfgang-Pauli-Strasse 27, HIT, CH-8093 Zurich, Switzerland*⁶*Computational Social Science, Department of Humanities, Social and Political Sciences, ETH Zürich, Clausiusstrasse 50, CH-8092 Zurich, Switzerland*

(Received 25 April 2017; revised manuscript received 20 November 2017; published 4 December 2017)

Connectivity, or the lack thereof, is crucial for the function of many man-made systems, from financial and economic networks over epidemic spreading in social networks to technical infrastructure. Often, connections are deliberately established or removed to induce, maintain, or destroy global connectivity. Thus, there has been a great interest in understanding how to control percolation, the transition to large-scale connectivity. Previous work, however, studied control strategies assuming unlimited resources. Here, we depart from this unrealistic assumption and consider the effect of limited resources on the effectiveness of control. We show that, even for scarce resources, percolation can be controlled with an efficient intervention strategy. We derive such an efficient strategy and study its implications, revealing a discontinuous transition as an unintended side effect of optimal control.

DOI: 10.1103/PhysRevE.96.062302

We are living in a globalized world. Large-scale connectivity, in particular, is essential for the proper functioning of many socioeconomic and technical systems. Examples include technical networks like the internet [1–3] or the world aviation network [4] and a wide range of socioeconomic and financial systems [5–7]. In other cases connectivity may be a liability, allowing the spreading of diseases and other contagion processes [8–10]. Ideally, control of connectivity has the goal to prevent widespread failure, for example, by immunizing a subset of the population to prevent an epidemic. Identifying efficient strategies that use minimal resources is an ongoing problem [11–13]. In many cases, however, one cannot completely prevent an undesirable transition, such as a recession or financial crisis, and tries to delay it as long as possible, often resulting in more severe consequences when the transition inevitably occurs [6,14,15]. Thus, it is essential to understand how to control and delay the emergence of connectivity under the constraint of limited resources and what such unintended consequences may be.

Percolation theory describes the emergence or breakdown of global connectivity depending on the structure of the underlying network with stochastic link addition processes [16–20]. A large body of work has studied the impact of an unlimited number of small interventions in modified models of network growth with the goal to delay the percolation transition. Most of these processes are based on a specific link addition rule. Typically, two (or more) possible candidate links are evaluated at each step and the link is added that delays (or enhances) the percolation transition the most [21]. This “competitive” percolation [22] leads to an extremely

sudden, but still continuous transition, sometimes referred to as “explosive” [22–24]. Other models introduce explicit control over the largest cluster, which further delays the transition and can result in a genuine discontinuous percolation transition [25–28]. Many more models with similar motivation have been studied, leading to a surprising diversity of phenomena [18,21–24,29–37].

In all these examples control is inherent to the link addition process, implicitly assuming unlimited resources and allowing indefinite control. Control in realistic settings, however, will be restricted by limited resources. Here, we derive an efficient resource limited control strategy to delay percolation and discuss the consequences for the resulting percolation transition. In particular, while the delayed transition remains smooth for suboptimal interventions, optimizing the control parameters to maximize the delay results in a discontinuous transition.

I. RESULTS**A. Model**

We develop our framework to efficiently delay the percolation transition based on the prototypical model of classical network formation, percolation of a random graph: new links e_{ij} between nodes i and j are chosen uniformly at random and sequentially added to a set of N initially unconnected nodes [38]. We implement control of link addition by preventing the chosen link from being added (see Fig. 1). This control is costly and preventing a link incurs a cost $c[S(i), S(j)]$, where $S(i)$ and $S(j)$ are the sizes of the respective connected components (clusters) that include the nodes i and j . Once a total budget B is spent, we can no longer control the link addition process. We track the evolution of the relative size of the largest connected component S_1/N as a function of the link density $p = L/N$, where L is the number of links added to the network. For the results presented here, the cost of an intervention is kept constant $c[S(i), S(j)] = 1$ and we assume a budget that scales

*malte.schroeder@ds.mpg.de

†nmaraujo@fc.ul.pt

‡dsornette@ethz.ch

§jnagler@ethz.ch

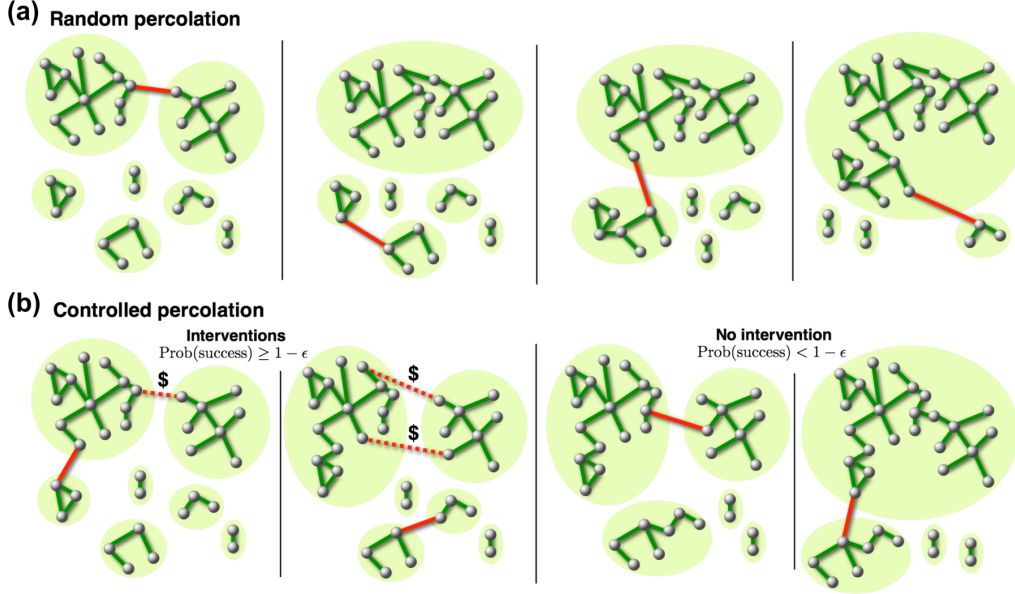


FIG. 1. Controlling the percolation transition. (a) Random percolation: in each step a link is selected uniformly at random and added to the network. (b) Controlled percolation: In each step, we can prevent the chosen link from being added to the network, paying a cost $c[S(i), S(j)]$ from a limited budget B . The constraint of a limited budget requires an efficient control strategy. As described in the text, we only prevent links when the probability that the intervention is successful is sufficiently large, $\text{Prob}(\text{success}) \geq 1 - \epsilon$. We consider an intervention unsuccessful if a similarly large cluster is likely to appear again with the next link e_{kl} . Consequently, we intervene when the probability of such a failure $\text{Prob}(\text{failure}) \approx \text{Prob}[S(k) + S(l) \geq S_{ij}] < \epsilon$ is small (the expected time until a similarly large cluster appears is large). When this failure probability is too large or the budget is exhausted, we do not intervene. As illustrated, this control delays the creation of large clusters and the onset of percolation.

linearly with the number of nodes, $B = bN$, where b is a (finite) constant. Corresponding results are obtained for other cost functions that scale with the size of the clusters, such as $c[S(i), S(j)] = S(i) + S(j)$ (see Supplemental Material [39]). In this case, avoiding the transition completely would clearly require preventing most of the links, which is impossible with limited resources.

In order to efficiently utilize the available resources and decide which links to prevent, we derive a control protocol based on the effect of a single intervention. Consider preventing a link e_{ij} that, when added to the network, would create a cluster of size $S_{ij} = S(i) + S(j)$. If the next link e_{kl} creates a cluster of size $S_{kl} = S(k) + S(l) \geq S_{ij}$, we spent some of our budget in vain, since we did not delay the emergence of a large cluster. Conversely, we can consider the intervention effective, when the next links e_{kl} only create smaller clusters $S_{kl} < S_{ij}$ and the emergence of a large cluster was delayed. Based on this idea we propose a control protocol where we prevent a link e_{ij} only if the expected impact is sufficiently large. We measure this impact by the (expected) number of links $\Delta L_{S_{ij}}$ until a cluster of size at least S_{ij} appears again. Clearly, if $\Delta L_{S_{ij}}$ is large, the intervention is more likely to delay the growth of a large cluster. If this delay is larger than some threshold ΔL_{thres} , we consider the intervention effective and prevent the link, otherwise we do not intervene. In practice, we estimate the expected $\Delta L_{S_{ij}}$ from the current cluster-size distribution n_S as the inverse of the probability that a new link e_{kl} creates

a cluster $S_{kl} \geq S_{ij}$,

$$\begin{aligned} \frac{1}{\langle \Delta L_{S_{ij}} \rangle} &\approx \text{Prob}[S_{kl} = S(k) + S(l) \geq S_{ij}] \\ &= \sum_{\substack{S(l) \neq S(k) \\ S(k) + S(l) \geq S_{ij}}} \frac{S(k)n_{S(k)}}{N} \frac{S(l)n_{S(l)}}{N-1} \\ &\quad + \sum_{2S(k) \geq S_{ij}} \frac{S(k)n_{S(k)}}{N} \frac{S(k)(n_{S(k)} - 1)}{N-1}, \quad (1) \end{aligned}$$

where the first sum describes the probability of a merger of clusters of different size resulting in a cluster at least as large as S_{ij} and the second sum describes similar mergers between clusters with equal size. For simplicity, we ignore that a link already present cannot be added again. Hence, we prevent a link from being added if $\text{Prob}[S(k) + S(l) \geq S_{ij}] < 1/\Delta L_{\text{thres}} := \epsilon$, where ϵ denotes the “intervention intensity,” which is the expected link rejection rate. This protocol is equivalent to stopping the ϵ -fraction most extreme events during the percolation process given sufficient budget. Other control strategies based, for example, on constraining the variance of the cluster size distribution are less efficient but give qualitatively similar results (see Supplemental Material [39]).

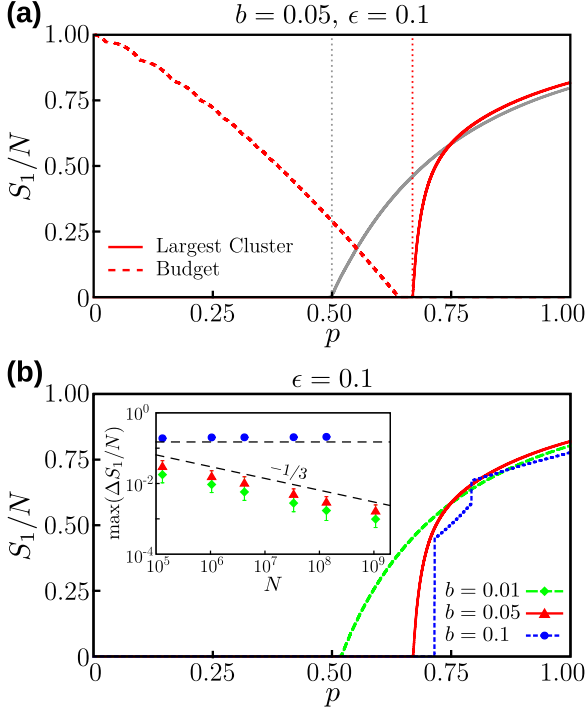


FIG. 2. Effects of resource limited control of percolation. (a) Single realization of the evolution of the relative size of the largest cluster for $N = 2^{25}$ (red solid line) and remaining fraction of the budget (red dashed line) for budget parameter $b = 0.05$ and intervention intensity $\epsilon = 0.1$. Compared to zero budget, the percolation threshold is shifted from $p_c = 0.5$ (gray line, showing random percolation without control) to $p_c \approx 0.67$. Interestingly, the transition remains continuous and in the same universality class. (b) Single realizations of the evolution of the relative size of the largest cluster for $N = 2^{25}$, $\epsilon = 0.1$, and different values of b . Surprisingly, when b becomes large enough, the transition becomes discontinuous. Inset: the largest gap $\max(\Delta S_1/N)$, averaged over 2^{10} to 2^6 realizations; error bars indicate the standard deviation. For small b , the scaling is the same as expected for random percolation, $\max(\Delta S_1/N) \sim N^{-1/3}$. However, for a sufficiently large budget, the largest gap is independent of the network size and the transition is discontinuous.

B. Efficient control of percolation

How much and how efficiently can the percolation transition be delayed with limited resources? As shown in Fig. 2, even with a small budget $B = bN = 0.05N$, meaning less than one intervention in ten link additions until $p_c \geq 1/2$, we can significantly delay the percolation transition compared to random percolation. Compared to the sudden transitions in the models of explosive percolation [21–24,31,37], our control protocol is more effective in delaying the transition (see Appendix A). Interestingly, the transition remains smooth and still belongs to the same universality class as random percolation when the budget is exhausted before the transition (see Table I and Appendix B for results of a finite-size scaling analysis).

Note that in Fig. 2(a) the budget runs out at $p =: p_{\text{last}} < p_c$, before the percolation threshold p_c , and the transition itself

TABLE I. Finite-size scaling. Exponents $-\beta/\nu$ (top) and γ/ν (bottom) found by finite-size scaling analysis. The corresponding fits are shown in Fig. 7 in Appendix B. The values agree with the exponents expected for random percolation $-\beta/\nu = -1/3$ and $\gamma/\nu = 1/3$ when the interventions end before the transition. For $\epsilon = 0.1$ and $b = 0.1$ the result is consistent with the expected $\beta = 0$ of a discontinuous transition.

$b \backslash \epsilon$	$-\beta/\nu$		
	0.1	0.2	0.5
0.01	−0.325(3)	−0.338(8)	−0.337(9)
0.05	−0.35(1)	−0.336(3)	−0.333(8)
0.10	−0.03(5)	−0.338(6)	−0.337(5)
	γ/ν		
	0.1	0.2	0.5
	0.331(3)	0.338(5)	0.331(2)
	0.347(5)	0.343(7)	0.333(7)
0.10	0.40(5)	0.334(7)	0.339(5)

is uncontrolled. We can estimate how long the budget lasts: With a constant intervention rate ϵ we would expect $\Delta L_{\text{int}} = \epsilon \Delta L_{\text{total}}$ interventions to occur during the sampling of ΔL_{total} links. During this period, we add only $N \Delta p = \Delta L = (1 - \epsilon) \Delta L_{\text{total}}$ links. Taking $\Delta L_{\text{int}} = \Delta B = N \Delta b$, we find the budget used in this interval $\Delta b = \frac{\epsilon}{1 - \epsilon} \Delta p$.

However, the budget decays nonlinearly, as seen in Fig. 2(a), which means the true intervention rate also varies with p . This nonlinear dependency results from the behavior of the intervention rate oscillating around an effective linear increase $\epsilon_{\text{eff}}(p) = \epsilon(1 + p/p_c^{\text{max}})/2$ for $p \leq p_c$, where p_c^{max} is the position of the critical point of controlled percolation with intervention intensity ϵ and infinite budget (see Appendix C for details). This observation, together with integration over p , then yields the closed expression defining p_{last}

$$b = \int_0^{p_{\text{last}}} \frac{\epsilon_{\text{eff}}(p)}{1 - \epsilon_{\text{eff}}(p)} dp. \quad (2)$$

As expected, a larger (effective) intervention rate requires a larger budget. Consequently, for a small budget, (i) the budget runs out before the onset of percolation at $p_{\text{last}} < p_c$ (interventions stop), (ii) the process is uncontrolled in a short but extensive window prior to the transition point, and (iii) one observes a continuous transition in the same universality class as random percolation. In contrast to previous percolation rules where delaying the transition changes its universality class, the limited resources in our model are exhausted before the transition. At this point the largest cluster has a fixed finite size and uncontrolled random percolation takes over, resulting in a continuous transition similar to random percolation for different initial cluster-size distributions [40].

C. Optimal control leads to discontinuity

Increasing the budget also increases the delay of the transition. Interestingly, too large a budget also leads to a discontinuous transition [see Fig. 2(b)]. At the same time, increasing the budget further no longer increases the delay of the transition and p_c becomes constant. Clearly, when the budget survives the percolation threshold, additional

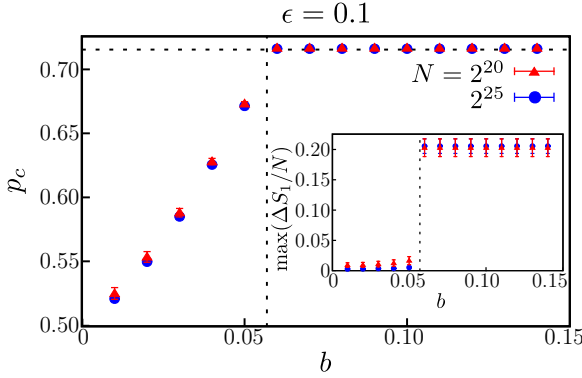


FIG. 3. Discontinuous transition above a critical budget. Percolation threshold p_c measured by the position of the largest gap of S_1 for different values of b . Results are averaged over 1024 and 256 realizations for networks of size $N = 2^{20}$ and 2^{25} , respectively. Error bars indicate the standard deviation. The delay increases with an increasing budget until it becomes constant above a critical budget $b_c \approx 0.058$. At the same time, the transition changes from continuous to discontinuous at $b = b_c$. Inset: The size of the largest gap $\max(\Delta S_1/N)$ for different b .

interventions have no effect on the transition. This suggests that the optimal delay is achieved for an optimal budget lasting exactly until the percolation threshold, $p_{\text{last}} = p_c$. At this point no uncontrolled window exists before the transition and the transition becomes discontinuous.

A similar logic defines the optimal parameters for speeding up the percolation transition (see Appendix D): interventions taken after the transition have no effect, while intervention-free uncontrolled link addition will reduce the effect of previous interventions. Optimal interventions necessarily end exactly at the percolation threshold, regardless of the intended result of the control.

Substituting $p_{\text{last}} = p_c^{\text{max}} \approx 0.72$ in Eq. (2) as the largest observed value of the critical point, we predict the critical budget required for a discontinuous transition for $\epsilon = 0.1$ to be $b_c^{\text{est}} \approx 0.058$. Indeed, this is confirmed by the numerical results shown in Fig. 3: the transition is continuous for $b \leq 0.05$, while the transition for $b \geq 0.06$ is already discontinuous.

But how can the transition become discontinuous for $b > b_c$? Stopping the ϵ -fraction most extreme events prevents any cluster above a certain size C_{thresh} to appear in the network. As more links are added, this threshold slowly increases. This is similar to the dynamics of the Bohman-Frieze-Wormald (BFW) model [41]. In fact, we observe comparable behavior in the subcritical regime: there is a hierarchy of thresholds $p_k > 0$, $k = 3, 4, \dots$ where a new largest cluster of size $S_1 = k$ first appears. As in the BFW model, these p_k converge to constant, finite values $0 < p_k < p_c$ for large systems and announce the critical transition as $p_k \rightarrow p_c$ for $k \rightarrow \infty$ (see Appendix C). Thus, the same mechanism that leads to a discontinuous transition in the BFW model causes a discontinuous transition for optimal resource limited control of percolation [27, 41, 42].

We have studied other control strategies and cost functions, for example, cost proportional to the size of the clusters involved in the link, $c[S(i), S(j)] = S(i) + S(j)$ (see

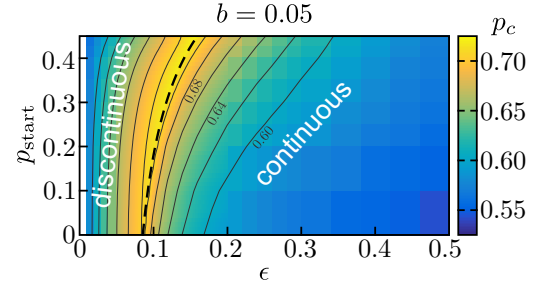


FIG. 4. “Phase diagram” and discontinuous transition as a result of optimal control. Position of the transition p_c for budget parameter $b = 0.05$ as a function of intervention intensity ϵ and intervention start p_{start} . Results for system size $N = 2^{20}$, averaged over $R = 128$ realizations. The largest delay with $p_c \approx 0.72$ is achieved for a set of optimal intervention parameters (bright yellow) that separate the continuous from the discontinuous transition regime. The transition becomes discontinuous as a result of the optimal resource limited control. The black dashed line represents our estimate for this optimal parameter set in $(\epsilon, p_{\text{start}})$ space (see text). The thin lines indicate lines of constant p_c .

Supplemental Material [39]). We find for all of the studied cost functions that a small budget leads to a continuous transition, whereas a larger budget further delays the transition and eventually leads to a discontinuous transition. However, when the cost scales with the size of the clusters, the transition only becomes discontinuous when the budget scales superlinearly $B \sim O(N^a)$ with $a > 1$.

D. Limited observability

One realistic limitation to the control of connectivity is observability. In particular, we might not be aware of problems, such as emerging large clusters, early on in the process and only begin interventions after some time p_{start} . Under these conditions, how do we best utilize a limited budget? Adapting Eq. (2) to include p_{start} leads to the relation $b = \int_{p_{\text{start}}}^{p_c^{\text{max}}} \frac{\epsilon_{\text{eff}}(p)}{1 - \epsilon_{\text{eff}}(p)} dp$ describing the optimal intervention parameters (see also Appendix C). Calculating the optimal start and intensity of the interventions with $b = 0.05$ and the observed $p_c^{\text{max}} = 0.72$, we obtain good agreement with the numerical results in Fig. 4. As explained above, the line of optimal control parameters separates the regimes of continuous and discontinuous transitions. As required by the constraint of limited resources, our control scheme is much more efficient than explosive percolation models at controlling percolation: We achieve $p_c = 0.72$ with only about one intervention per 15 added links; much less than comparable competitive percolation models, which reject one link for each link added (see Appendix A).

Interestingly, we find that for fixed intervention cost interventions close to the percolation threshold are slightly more effective than early interventions (p_c slowly increases as a function of p_{start} along the critical line). This result, however, is specific to constant intervention costs as other cost functions can lead to a different behavior: interventions as early as possible, $p_{\text{start}} = 0$, are optimal for intervention costs that grow with the size of the connected clusters (see Supplemental Material [39]).

II. DISCUSSION

We have derived a control strategy to efficiently delay percolation with limited resources. In contrast to previous models constructed to delay the percolation transition [18,21–24,29–37], we find that the transition remains smooth and in the same universality class as random percolation for nonoptimal control when the resources are exhausted before the transition. Given a fixed budget, maximal delay of the percolation transition is achieved by optimizing the control protocol such that the budget is exhausted exactly at the percolation threshold. While the percolation transition can be delayed by control interventions, this resource optimal delay inevitably results in a discontinuous percolation transition that becomes effectively *uncontrollable*, since the addition of a single link induces a macroscopic change in the connectivity.

It is commonly believed that interventions taken as early as possible can have the biggest impact to avoid large-scale connectivity [6]. We have shown that this is not always the case: a strong effort to intervene right at the beginning can diminish the budget to such an extent that more timely interventions become impossible in crucial stages.

The framework we developed on the basis of random network growth highlights the unintended consequences of trying to control the percolation transition by delaying it [6,14,15]. Likely, similar effects will occur for other control schemes as well. This work may thus help to design control schemes in other networks, specific to the underlying network dynamics and its constraints, in particular when resources are scarce.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Göttingen Graduate School for Neurosciences and Molecular Biosciences [DFG Grant No. GSC 226/2 (M.S.)], the

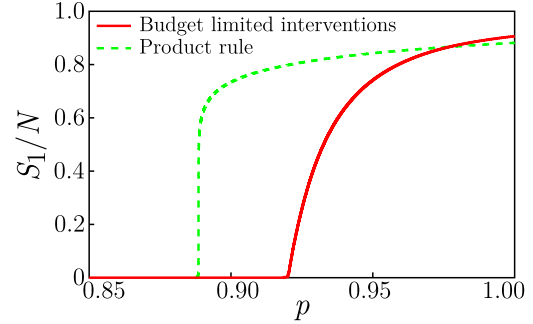


FIG. 5. Effectiveness of resource limited control. Single realizations of the largest cluster size for the budget limited control (red solid line) and the product rule (green dashed line) [21] resulting in explosive percolation ($N = 2^{25}$). The parameters are $b = 0.88$ and $\epsilon = 0.62$. In both models $L_{\text{rej}} \approx 0.88N$ links are rejected until the phase transition occurs. This illustrates that the intervention rule defined in the main text is more effective in delaying the transition and at the same time keeps the transition smoother.

Portuguese Foundation for Science and Technology (FCT) under Contracts No. UID/FIS/00618/2013 and No. IF/00255/2013 (N.A.), the SNF [“The anatomy of systemic financial risk,” Grant No. 162776 (J.N.)], and the ETH Risk Center [RC SP 08-15 (J.N.)].

APPENDIX A: COMPARISON TO THE PRODUCT RULE

To illustrate the effectiveness of the proposed control protocol, we explicitly compare it to the product rule of explosive percolation [21]. The product rule is defined as follows: in each step choose two links uniformly at random

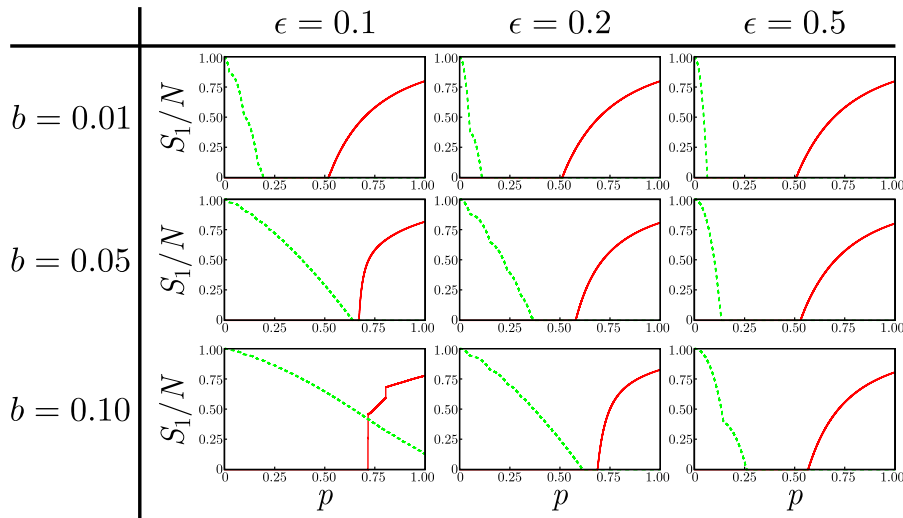


FIG. 6. Controlled percolation. Single realizations of the largest cluster size (red solid lines) and the remaining fraction of the budget (green dashed lines) for various parameter combinations and $N = 2^{25}$. Depending on the parameters the delay between the last interventions (budget reaching 0) and the percolation transition changes. The transition is smoothest when this gap is large. When the budget lasts until after the percolation transition, the transition becomes discontinuous.

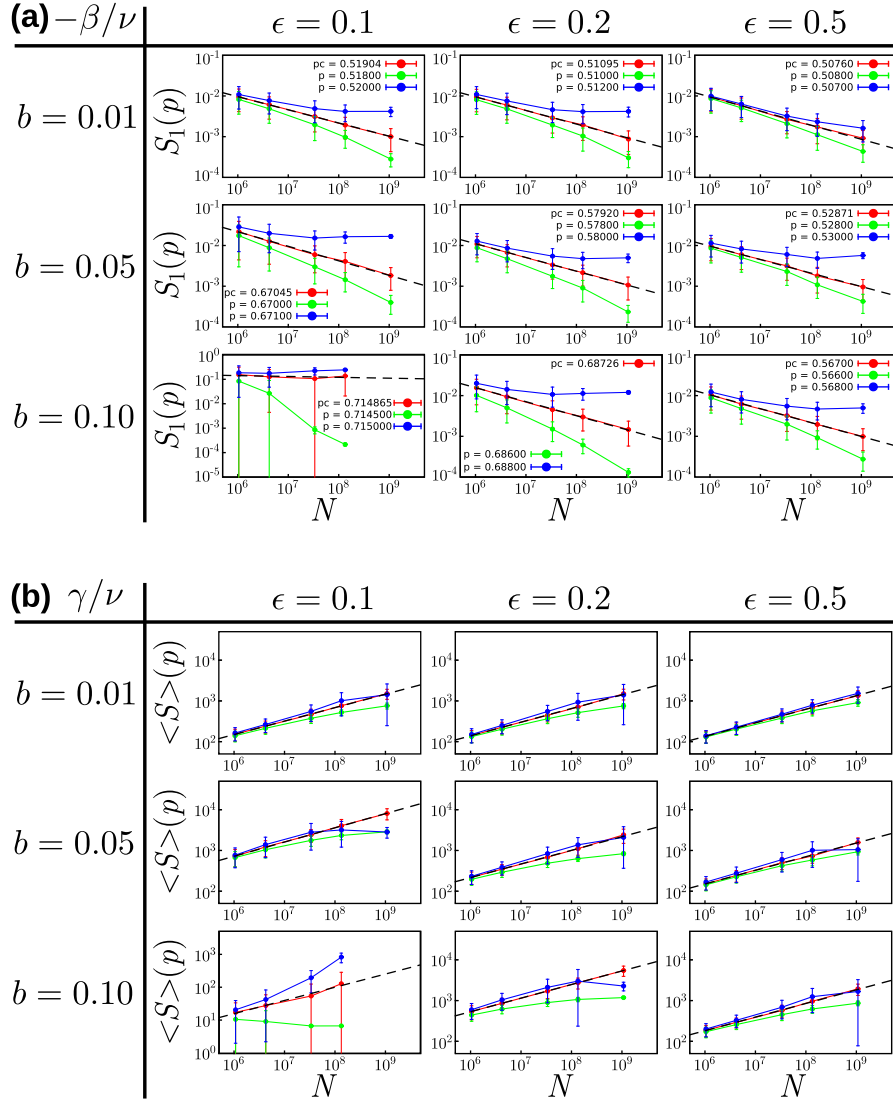


FIG. 7. Finite-size scaling for controlled percolation. Results for finite-size scaling for the estimated critical point and two values of p slightly below and above. The error bars indicate the standard deviation; lines are guides to the eye. Averages are taken over 1024 to 64 realizations for system sizes from $N = 2^{20}$ to $N = 2^{30}$. The black dashed lines indicate the best fits; the resulting exponents are listed in Table I above. (a) Results for the exponent $-\beta/\nu \approx -1/3$, showing the same behavior as expected for random percolation for all continuous transitions. For $b = 0.1$, $\epsilon = 0.1$ we find $\beta \approx 0$, corresponding to a discontinuous transition. (b) Results for the exponent $\gamma/\nu \approx 1/3$, showing the same behavior as expected for random percolation for all continuous transitions.

and add the link that minimizes the product $S(i)S(j)$. This significantly delays the percolation transition, but results in a very abrupt, explosive transition that is continuous but almost indistinguishable from a discontinuous transition even in very large systems. To compare the models, consider the “budget” required for the product rule: in each step one link is rejected, thus for constant cost, $c[S(i), S(j)] = 1$, the product rule requires a budget $B(p) = pN$ up until p . Therefore, until the phase transition at $p_c^{\text{PR}} \approx 0.889$ it uses a budget $B = 0.88N$.

We use the same budget, $b = 0.88$, for our intervention rule and choose a good (although not optimal) intervention intensity $\epsilon = 0.62$. As shown in Fig. 5 our intervention rule delays the percolation transition more efficiently while also keeping the transition in the same universality class as random percolation.

APPENDIX B: FINITE-SIZE SCALING ANALYSIS

Compared to explosive percolation our control scheme does not change the universality class of the transition when the

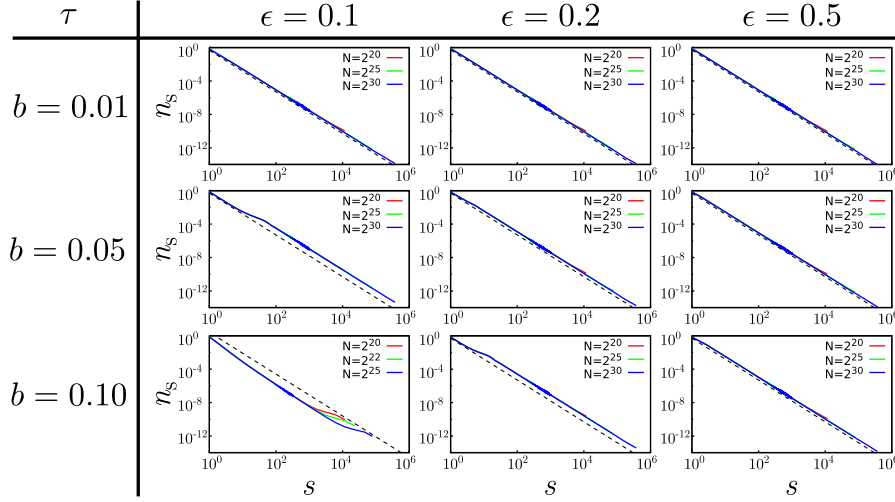


FIG. 8. Critical cluster size distribution. Cluster size distribution n_S for various system sizes N [$N = 2^{20}$, 2^{25} , and 2^{30} averaged over 1024, 256, and 64 realizations, respectively ($N = 2^{20}$, 2^{22} , and 2^{25} for $b = 0.1$, $\epsilon = 0.1$)]. Here, n_S describes the *relative* frequency of clusters of size S . The results are aggregated into logarithmic bins for clusters with size $S > N^{1/3}$ up to $S = 16N^{1/2}$. The scaling is expected to follow a power law $n_S \sim S^{-\tau}$ for large S ; the dashed black lines show the scaling expected for random percolation with exponent $\tau = 5/2$ (not normalized). The peak in the cluster size distribution for small S is a signature of the finite size of clusters in the system when the interventions stop. Larger budgets allow for more interventions shifting the peak to larger S and making it more pronounced. Higher intervention intensities use the budget earlier, shifting the peak to lower S .

budget is exhausted before the transition. To demonstrate this, we conduct a finite-size scaling analysis for various values of the parameters b and ϵ . Figure 6 shows single realizations for these parameters. The transition is continuous in all cases where the budget is exhausted early. Only for $b = 0.1$ and $\epsilon = 0.1$ does the budget last until after the transition and the transition becomes discontinuous.

With the standard assumption of the critical scaling $S_1(p) \sim |p - p_c|^\beta$ and $\langle S \rangle(p) \sim |p - p_c|^{-\gamma}$ for the size of the largest cluster and the mean cluster size, respectively, as well as for the correlation length $\xi(p) \sim |p - p_c|^{-\nu}$, we find the exponents $-\beta/\nu$ and γ/ν from the finite-size scaling fits shown in Fig. 7. The results for the exponents are shown in Table I above. All exponents of the continuous transitions agree well with those expected for random percolation. For $b = 0.1$ and $\epsilon = 0.1$ we obtain exponents expected for a discontinuous transition, $-\beta/\nu = 0$. In particular, the transition never becomes weakly discontinuous or explosive. Similarly, we find the exponent $\tau = 5/2$ for the cluster size distribution at the critical point in all cases, the same as for random percolation. The corresponding results are shown in Fig. 8.

APPENDIX C: OPTIMAL CONTROL PARAMETERS

In order to estimate optimal intervention parameters, we need to predict the point p_{last} when the budget is exhausted.

We first illustrate that the intervention rate is not constant as one might have expected from the definition of the intervention rule. Instead it fluctuates, dropping to small values immediately after the largest cluster size in the system grew (see Fig. 9). This is easiest to understand by considering the first link: we never prevent the first link since the probability to create a cluster of size 2 is $\text{Prob}[S(k) + S(l) \geq 2] = 1 > \epsilon$. Thus the probability of an intervention $\epsilon(p = 0) = 0$. Similarly, the first few links are unlikely to be prevented, since a link creating a cluster of size 3 or larger is chosen with vanishing probability.

We can think about the intervention rule in the following way: We always prevent the most extreme links. This is equivalent to preventing all clusters above a certain size (until these links become too likely). This means, when the size of the largest cluster just changed to S_1 , the probability to create a larger cluster is usually smaller than ϵ . However, the links creating a cluster of size S_1 are not prevented as the probability to create a cluster larger or equal to S_1 is larger than ϵ . Thus, after these microtransitions of the largest cluster size, the intervention probability drops. In fact, we find that these transitions to a new largest cluster size happen at well defined times, constant across different system sizes (see Fig. 9). This behavior is similar to the subcritical evolution of the BFW model [41,42]. This observation also supports the discontinuity of the transition.

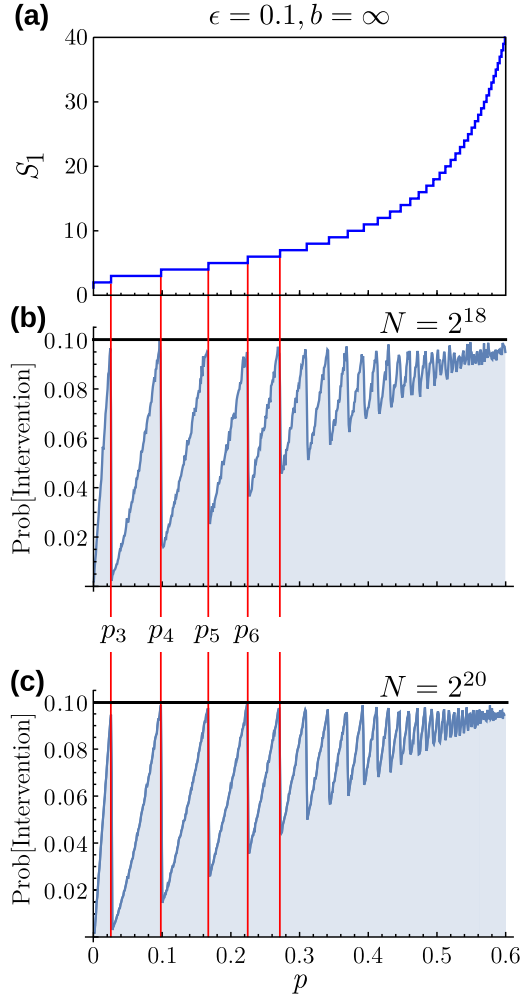


FIG. 9. Distribution of interventions. (a) Single realization of the largest cluster size in the subcritical regime for $N = 2^{25}$ with unlimited budget and $\epsilon = 0.1$. (b), (c) Probability of an intervention for a single link chosen at p for two different system sizes averaged over 100 realizations each. When a new cluster size appears in the network the intervention probability “resets.” This causes the transitions to $S_1 = 3$ at p_3 , $S_1 = 4$ at p_4 , and so on to occur at fixed positions. This behavior is similar to the subcritical evolution of the largest cluster in the BFW model, leading to a discontinuous transition at $p_k \rightarrow p_c$ for $k \rightarrow \infty$.

$$b = \int_{p_{\text{start}}}^{p_{\text{last}}} \frac{\epsilon_{\text{eff}}(p)}{1 - \epsilon_{\text{eff}}(p)} dp$$

$$= p_{\text{last}} - p_{\text{start}} - \frac{2(p_c^{\text{max}} - p_{\text{start}}) \log \left(\frac{-2(p_c^{\text{max}} - p_{\text{start}}) + \epsilon[p_{\text{last}} + p_c^{\text{max}} - 2p_{\text{start}}(1 + p_{\text{last}} - p_c^{\text{max}})]}{(p_c^{\text{max}} - p_{\text{start}})(\epsilon + 2p_{\text{start}}\epsilon - 2)} \right)}{(1 - 2p_{\text{start}})\epsilon}, \quad (\text{C2})$$

assuming again $p_c^{\text{max}} \geq p_{\text{last}}$ is the critical point of the process with parameters ϵ and p_{start} given unlimited budget.

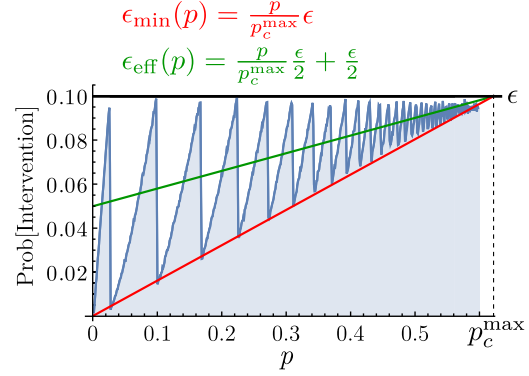


FIG. 10. Estimating the intervention rate. Probability of an intervention for a single link chosen at p for system size $N = 2^{20}$ averaged over 100 realizations. The red and green lines illustrate the approximation used to define the effective intervention rate ϵ_{eff} (green), describing a local average of the true intervention rate (here for $p_{\text{start}} = 0$).

We can use the observed intervention rate to derive an estimate for the budget used for interventions up to p . Since we do not know the exact form of $\epsilon(p)$, we use an empirically determined “effective intervention rate” $\epsilon_{\text{eff}}(p)$, describing a local average of $\epsilon(p)$ (illustrated in Fig. 10 for $p_{\text{start}} = 0$). This intervention rate depends on the intervention parameter ϵ and the position p_c^{max} of the critical point of the process with unlimited budget. When $p_{\text{start}} > 0$, the uncontrolled evolution before the control starts will cause the intervention rate to be larger than in the fully controlled process. We assume that the effective intervention rate at p_{start} is $\epsilon_{\text{eff}}(p_{\text{start}}) = p_{\text{start}}\epsilon + \epsilon/2$ (the value obtained by setting $p_c^{\text{max}} = 1/2$). Directly at and after p_c the effective intervention rate is $\epsilon_{\text{eff}}(p \geq p_c) = \epsilon$. Together this gives

$$\epsilon_{\text{eff}}(p) = \frac{p - p_{\text{start}}}{p_c^{\text{max}} - p_{\text{start}}} (\epsilon/2 - p_{\text{start}}\epsilon) + \epsilon/2 + p_{\text{start}}\epsilon, \quad (\text{C1})$$

for $p_{\text{start}} \leq p \leq p_c$. Obviously, before p_{start} the intervention rate is $\epsilon_{\text{eff}} = 0$ and above p_c the intervention rate is $\epsilon_{\text{eff}} = \epsilon$.

We can now use the argument we gave above and integrate Eq. (C1) over all interventions to find the total budget used. We arrive at the approximate relation

Substituting $p_{\text{last}} = p_c^{\text{max}}$ gives the condition for optimal intervention parameters, which can be solved numerically to

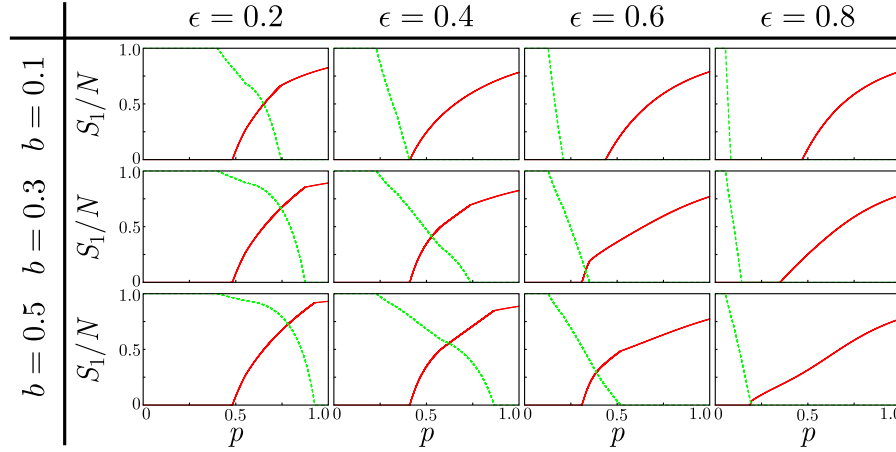


FIG. 11. Enhancing percolation. Single realizations of the largest cluster size (red solid lines) and the remaining fraction of the total budget (green dashed lines) for various parameter combinations for interventions enhancing percolation ($N = 2^{25}$). Depending on the parameters the transition is enhanced more or less strongly. As for delaying the transition, interventions are most efficient when the interventions last exactly until the transition.

find the optimal budget or intervention rate (see Fig. 4). The estimate becomes worse for large values of ϵ and p_{start} and very small values of b , where interventions occur only in a small interval and averaging to ϵ_{eff} becomes inaccurate. For the same reason, the effective intervention rate is a good approximation when estimating the optimal intervention parameters, where interventions last until p_c and the error from averaging is small.

APPENDIX D: ENHANCING PERCOLATION

We have illustrated our results for interventions that are designed to delay the percolation transition. Interestingly, the same logic describing the optimal intervention strategy also applies to enhance percolation. Instead of stopping the ϵ -fraction most extreme events, we simply reverse the protocol, $\text{Prob}[S(k) + S(l) \leq S_{ij}] < 1/\Delta L_{\text{thres}} := \epsilon$, and stop

the ϵ -fraction least extreme events, where we specifically include links connecting nodes in the same cluster as creating a new cluster of size 0. Additionally, we always prevent such intracluster links as there are no less extreme events. However, this is only relevant for nonoptimal interventions after the transition.

Also in this case optimal interventions necessarily end at the percolation threshold. Interventions lasting longer have no additional effect on the threshold and interventions ending earlier create an extensive interval of uncontrolled percolation before the transition, partially negating the effect of the interventions. In Fig. 11 we show examples for single realizations of percolation enhancing interventions. The results confirm that the effect is largest (p_c is smallest) when the budget runs out exactly at $p_{\text{last}} = p_c$. The budget used also shows that it is much more difficult to enhance the percolation than to delay it.

- [1] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [2] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, *Phys. Rev. Lett.* **86**, 3682 (2001).
- [3] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, *Nature (London)* **464**, 1025 (2010).
- [4] T. Verma, F. Russmann, N. A. M. Araújo, J. Nagler, and H. J. Herrmann, *Nat. Commun.* **7**, 10441 (2016).
- [5] P. Gai and S. Kapadia, *Proc. R. Soc. London, Ser. A* **466**, 2401 (2010).
- [6] D. Helbing, *Nature (London)* **497**, 51 (2013).
- [7] M. Elliott, B. Golub, and M. O. Jackson, *Am. Econ. Rev.* **104**, 3115 (2014).
- [8] C. Moore and M. E. J. Newman, *Phys. Rev. E* **61**, 5678 (2000).
- [9] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.* **86**, 3200 (2001).
- [10] R. Cohen, S. Havlin, and D. ben-Avraham, *Phys. Rev. Lett.* **91**, 247901 (2003).
- [11] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. E* **65**, 036104 (2002).
- [12] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley, *Phys. Rev. Lett.* **101**, 058701 (2008).
- [13] F. Morone and H. A. Makse, *Nature (London)* **524**, 65 (2015).
- [14] D. Lee, W. Choi, J. Kertész, and B. Kahng, *Sci. Rep.* **7**, 5723 (2017).
- [15] D. Sornette and P. Cauwels, *Risks* **2**, 103 (2014).
- [16] D. Stauffer and A. Aharony, *Introduction To Percolation Theory* (Taylor & Francis, London, 1992).
- [17] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Phys. Rev. Lett.* **85**, 5468 (2000).
- [18] Y. S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim, *Phys. Rev. Lett.* **103**, 135702 (2009).

SCHRÖDER, ARAÚJO, SORNETTE, AND NAGLER

PHYSICAL REVIEW E **96**, 062302 (2017)

- [19] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge, 2010).
- [20] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
- [21] D. Achlioptas, R. M. D'Souza, and J. Spencer, *Science* **323**, 1453 (2009).
- [22] J. Nagler, A. Levina, and M. Timme, *Nat. Phys.* **7**, 265 (2011).
- [23] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Phys. Rev. Lett.* **105**, 255701 (2010).
- [24] O. Riordan and L. Warnke, *Science* **333**, 322 (2011).
- [25] N. A. M. Araújo and H. J. Herrmann, *Phys. Rev. Lett.* **105**, 035701 (2010).
- [26] K. J. Schrenk, N. A. M. Araújo, and H. J. Herrmann, *Phys. Rev. E* **84**, 041136 (2011).
- [27] W. Chen, J. Nagler, X. Cheng, X. Jin, H. Shen, Z. Zheng, and R. M. D'Souza, *Phys. Rev. E* **87**, 052130 (2013).
- [28] Y. S. Cho, S. Hwang, H. J. Herrmann, and B. Kahng, *Science* **339**, 1185 (2013).
- [29] F. Radicchi and S. Fortunato, *Phys. Rev. Lett.* **103**, 168701 (2009).
- [30] E. J. Friedman and A. S. Landsberg, *Phys. Rev. Lett.* **103**, 255701 (2009).
- [31] P. Grassberger, C. Christensen, G. Bizhani, S.-W. Son, and M. Paczuski, *Phys. Rev. Lett.* **106**, 225701 (2011).
- [32] O. Riordan and L. Warnke, *Phys. Rev. E* **86**, 011129 (2012).
- [33] K. J. Schrenk, A. Felder, S. Deflorin, N. A. M. Araújo, R. M. D'Souza, and H. J. Herrmann, *Phys. Rev. E* **85**, 031103 (2012).
- [34] W. Chen, X. Cheng, Z. Zheng, N. N. Chung, R. M. D'Souza, and J. Nagler, *Phys. Rev. E* **88**, 042152 (2013).
- [35] M. Schröder, S. H. E. Rahbari, and J. Nagler, *Nat. Commun.* **4**, 2222 (2013).
- [36] M. Schröder, W. Chen, and J. Nagler, *New J. Phys.* **18**, 013042 (2016).
- [37] R. M. D'Souza and J. Nagler, *Nat. Phys.* **11**, 531 (2015).
- [38] P. Erdős and A. Rényi, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17 (1960).
- [39] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.96.062302> for results for other cost functions and a comparison to other intervention rules.
- [40] Y. S. Cho, B. Kahng, and D. Kim, *Phys. Rev. E* **81**, 030103 (2010).
- [41] T. Bohman, A. Frieze, and N. C. Wormald, *Random Struct. Alg.* **25**, 432 (2004).
- [42] W. Chen, M. Schröder, R. M. D'Souza, D. Sornette, and J. Nagler, *Phys. Rev. Lett.* **112**, 155701 (2014).

Supplemental Material
 accompanying the manuscript
Controlling percolation with limited resources

Malte Schröder, Nuno Araujo, Didier Sornette, Jan Nagler

In the main manuscript we discussed the control of percolation with limited resources for interventions. In particular, we considered a growth model where, starting from an empty network with N nodes and no links, at each step we choose a link e_{ij} uniformly at random to add to the network. Control is implemented by a choice to prevent this link, paying a cost $c[S(i), S(j)]$ from a limited budget B . To decide whether we intervene in the link addition process we consider the ϵ fraction of links that, when added to the network, would create the largest clusters and prevent those links. Interventions are only possible as long as we have sufficient budget $B \geq c[S(i), S(j)]$, starting from $B = bN$. In the main manuscript we discussed the main features of this model for constant cost $c[S(i), S(j)] = 1$, showing that even a relatively small number of interventions can significantly delay the percolation transition. For small budgets this transition is continuous, but becomes discontinuous at some critical values b_c and ϵ_c . We also showed that the transition is maximally delayed for a given budget when the intervention intensity ϵ is exactly equal to this critical value.

Here, we provide further evidence that the transition indeed becomes discontinuous for a sufficiently large budget. Specifically, we consider unlimited budget in line with the (implicit) assumption in many other percolation rules. Finally, we also illustrate the robustness of our results with respect to different cost functions, among others $c[S(i), S(j)] = S(i) + S(j)$. These results are qualitatively similar and we can establish a direct mapping between results for the different cost functions. A given budget “fixes” the number of interventions up to small fluctuations and there is a direct correspondence between the budget and intervention parameters for both cost functions. Contrary to the results from the main manuscript, however, early interventions are more effective in this case.

UNLIMITED BUDGET — DISCONTINUOUS TRANSITION

To clearly show the discontinuity of the transition when the budget survives until the transition we now consider interventions with an unlimited budget in more detail. We study the largest gap of the largest cluster, following a method from [4] to resolve multiple jumps of the size of the largest cluster. We divide the region around the transition into intervals of width $\Delta p = 4 \cdot 10^{-5}$ and record the largest jump in each of these intervals (Fig. S2). While the transition is blurred out for small systems, a double transition is revealed for larger system sizes. The same behavior can also be seen for averages over the multiple realizations shown in Fig. S1.

We find that the largest gap of the first transition does not decay for increasing system size (even taking into account the smaller spread and thus expected larger averages). In the case of $\epsilon = 0.2$, we can quantify this by assuming a single large jump of size ΔS_1 and negligible contributions of all other changes for a given realization (this approximation becomes better the larger the system). If this assumption is correct, the resulting average should simply be the product of the size of the jump ΔS_1 and the probability that the jump occurs in a given interval. Consequently, we can (approximately) determine the size of the jump ΔS_1 by fitting a Gaussian distribution multiplied by ΔS_1 to the measured average jump size. We find $\Delta S_1^{\text{est}} \approx 0.3$ for the smallest system $N = 2^{20}$ decaying to only $\Delta S_1^{\text{est}} \approx 0.12$ for the larger systems $N = 2^{25}$ and 2^{27} . The fact that ΔS_1^{est} does not decay to zero shows that the jump is indeed macroscopic and the transition is discontinuous (Fig. S2, inset).

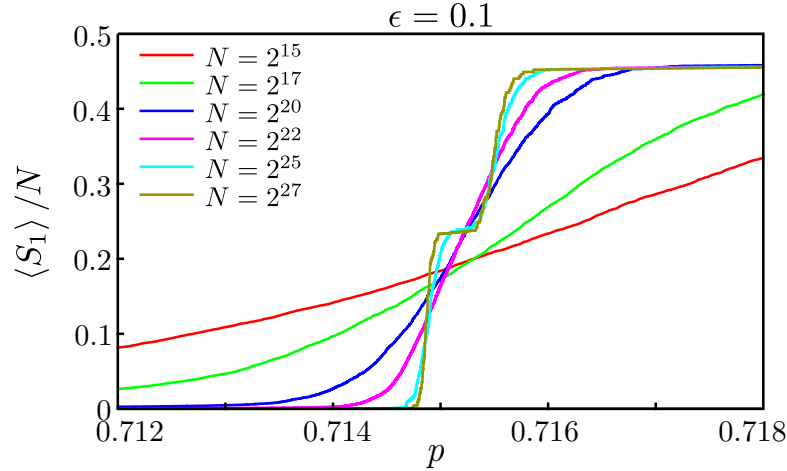


FIG. S1. Average size of the largest cluster during the transition for interventions with unlimited budget and $\epsilon = 0.1$. Averages are taken over 1024 to 64 realizations. This figure illustrates the difficulty in studying properties of the transition across different realizations: the transition is blurred out even for large finite systems, and the double transition is only revealed for very large systems.

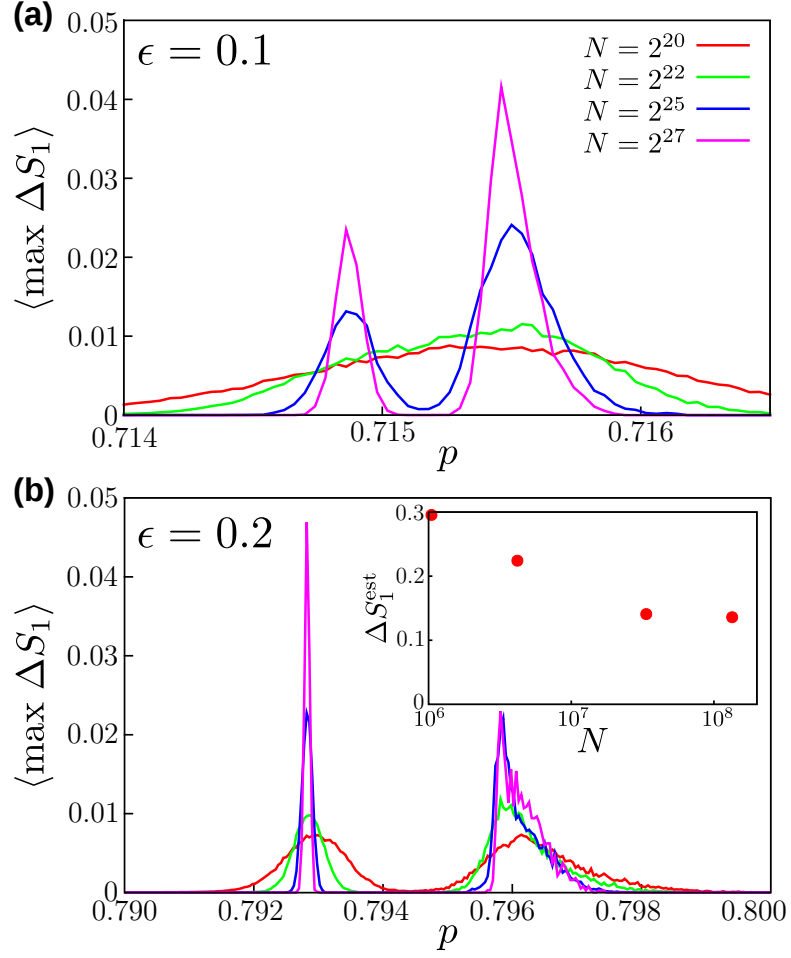


FIG. S2. Average maximum gap in the size of the largest cluster over p . The figures show the maximum size of the largest cluster in intervals $\Delta = 4 \cdot 10^{-5}$, averaged over 16384 to 1024 realizations for $\epsilon = 0.1$ [panel (a)] and $\epsilon = 0.2$ [panel (b)]. Results show that even though the jumps are initially indistinguishable, only two distinct jumps appear for large systems. An estimation of the expected size ΔS_1 of the first jump for $\epsilon = 0.2$ shows that it is becoming constant for large systems (inset, see text for more details), further evidencing that the transition is discontinuous.

RESOURCE-LIMITED CONTROL OF PERCOLATION UNDER VARIOUS COST FUNCTIONS

In order to illustrate the universality of our results, in the following we consider control of percolation using different cost functions as well as intervention rules based on other observables. First, we discuss a theoretical argument, why our results naturally extend to different cost functions.

Due to the self-averaging behavior of the percolation model, the number of interventions for a given cost function $c[S(i), S(j)]$ with given parameters ϵ and b is fixed with a negligible variance (relative to the system size). For large systems we thus find a direct correspondence to a system with constant cost $c' = 1$ and parameters ϵ' and $b' = b'(b, \epsilon)$, where $b'N$ is simply given as the average number of interventions.

We mostly discuss results using the cost function $c[S(i), S(j)] = S(i) + S(j)$, where the cost of an intervention scales with the size of the clusters involved. In this case, for any budget $B = bN$ with constant b we always observe a continuous transition (this means the corresponding $b' < b'_c$). This is illustrated in an overview of single realizations in Fig. S3. This can be understood with the following (rough) argument: consider the average budget used for interventions up to a point p

$$\begin{aligned} \langle b(p) \rangle &= 1/N \left\langle \sum_{\text{Interventions}} S(i) + S(j) \right\rangle \\ &\leq 1/N \sum_{\text{Interventions}} 2 \langle S \rangle \lesssim 2\epsilon \int_0^p \langle S \rangle(p') dp'. \end{aligned} \quad (1)$$

With the standard assumption for the critical scaling $\langle S \rangle(p') \sim |p_c - p'|^{-\gamma}$ this integral is finite for all $p < p_c$ but diverges at p_c . Thus any constant, finite budget b will run out at some point $p_{\text{last}} < p_c$, regardless of the value of ϵ , and the transition will be continuous. Conversely, we can reach any $p_{\text{last}} < p_c$ with a finite budget $B = \mathcal{O}(N)$. We can thus establish a direct mapping between the two cost functions with $b \in [0, \infty)$ for $c[S(i), S(j)] = S(i) + S(j)$ and $b' \in [0, b'_c)$ for constant cost (main manuscript). A delayed start of the interventions at p_{start} does not qualitatively change this mapping.

Due to the different cost of interventions, it becomes much more important to intervene early, when interventions are cheap. Finding the optimal ϵ for a given budget now does not mean keeping the interventions up the longest: if the intensity is too large we prevent relatively unimportant links. If the intensity is too small, some interventions are executed close to the critical point and are very costly, reducing the total number of interventions. The optimal delay is obtained for intermediate values of ϵ , balancing the observed effectiveness of interventions close to the critical point (see main manuscript) with the increasing costs.

The resulting delay of the percolation transition for various parameters, shown in detail in Fig. S4 and S5, illustrates the findings summarized above: (i) a larger budget will always increase the delay of the percolation transition (Fig. S4a), (ii) starting the interventions early and (iii) using an intermediate intensity results in the largest delay of p_c (Fig. S4b, S5).

However, we recover the discontinuous transition observed in the main manuscript for superlinear budget scaling $B \sim \mathcal{O}(N^a)$ with $a > 1$ (Fig. S6). This is required for interventions to last until (after) the percolation transition, where a single intervention will (likely) cost an extensive amount $\Delta B = \mathcal{O}(N)$. Specifically, considering the scaling of the cluster sizes before the transition, we can expect a critical budget on the order of $B_c \sim \mathcal{O}[N \log(N)]$. However, strong finite size effects make this prediction impossible to verify numerically.

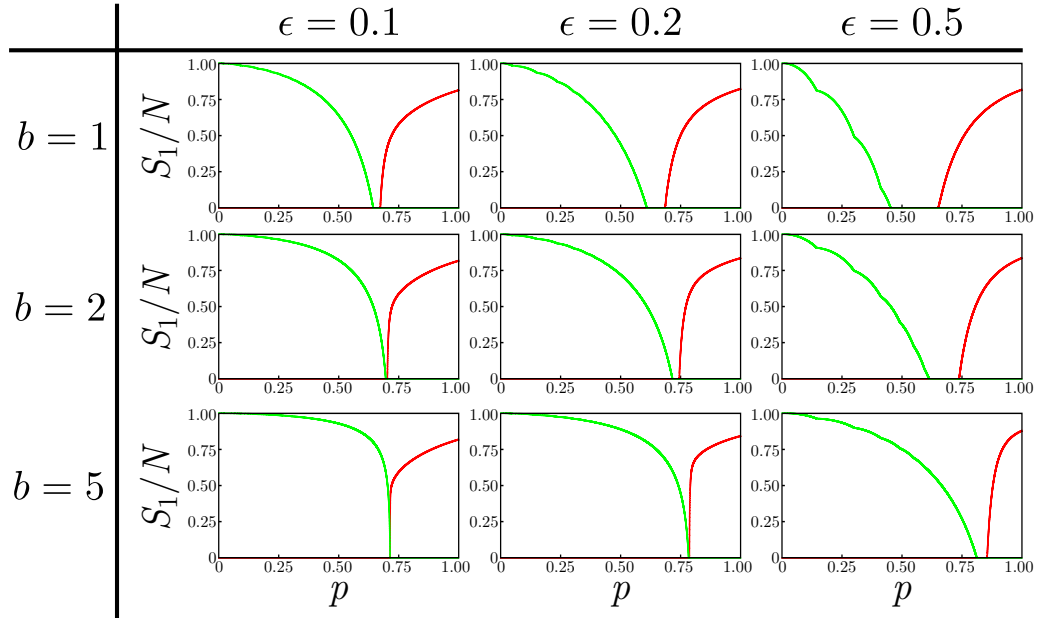


FIG. S3. Single realizations of the largest cluster size and the budget for various parameter combinations with $c[S(i), S(j)] = S(i) + S(j)$ for system size $N = 2^{25}$. The panels show the relative size of the largest cluster (red lines) and the remaining *fraction* of the total budget (green line) for different initial values $B = bN$ and intervention intensities ϵ . Depending on the parameters the delay between the last interventions (budget reaching 0) and the percolation transition changes. The transition is smoothest when this gap is large. As discussed in the text, the transition is always continuous, since the budget runs out before the transition.

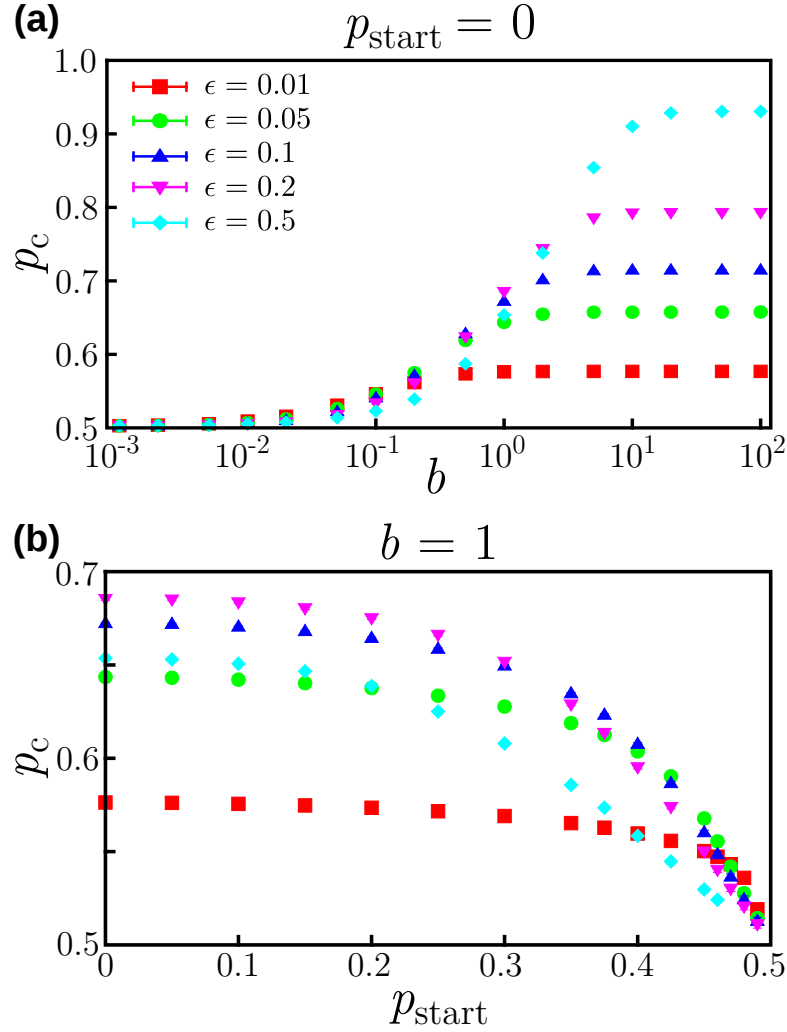


FIG. S4. Position of the percolation transition for various parameter combinations and cost function $c[S(i), S(j)] = S(i) + S(j)$, averaged over 256 realizations of $N = 2^{25}$. Error bars indicating the standard deviation are smaller than the symbol size. (a) p_c when interventions are possible for all p ($p_{\text{start}} = 0$). Obviously, a larger budget allows more interventions and leads to larger p_c . Considering a fixed budget, it is clearly visible that different values of the intervention intensity ϵ are optimal, e.g., large ϵ are feasible for large budgets, while they are sub-optimal for smaller budgets. (b) Resulting p_c for the same ϵ versus p_{start} , now for fixed $b = 1$. Clearly, starting the interventions earlier always results in a larger p_c . Considering a fixed p_{start} , the optimal intervention intensity ϵ changes: small intensities are sub-optimal when interventions are possible early but become optimal as p_{start} approaches $p_c^{\text{ER}} = 1/2$.

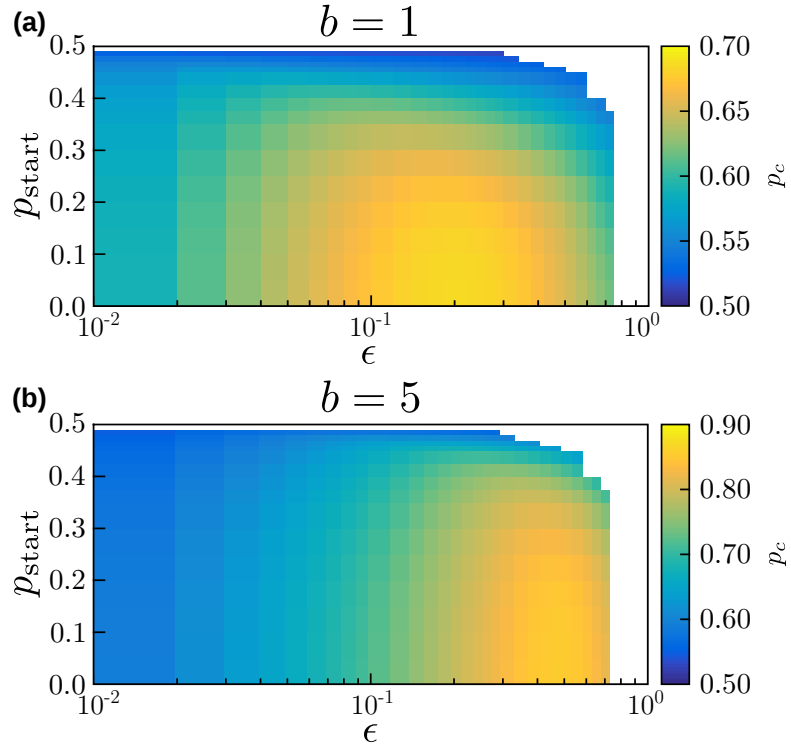


FIG. S5. Position of the phase transition p_c for cost function $c[S(i), S(j)] = S(i) + S(j)$ and interventions with $b = 1$ [panel (a)] and $b = 5$ [panel (b)] and parameters ϵ and p_{start} ($N = 2^5$), averaged over 256 realizations. The qualitative behavior is identical in both cases. Early interventions with intermediate intensity are optimal. Since more budget is available for $b = 5$, the optimal intervention intensity as well as the possible delay is larger than for $b = 1$.

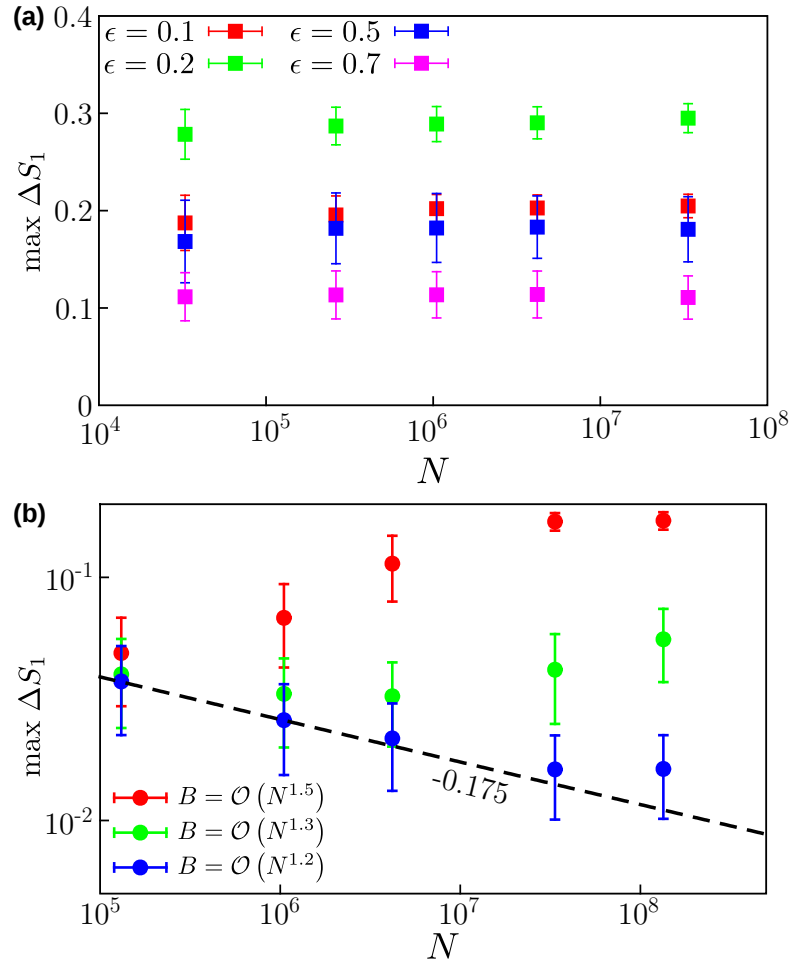


FIG. S6. Discontinuous transition with super-linear budget scaling for cost function $c[S(i), S(j)] = S(i) + S(j)$. Error bars indicate the standard deviation, averages are taken over 1024 to 64 realizations. (a) The largest gap is constant with increasing system size for unlimited budget and different values of the intervention intensity, evidencing a discontinuous transition. (b) The largest gap does not disappear with increasing system size for superlinear budget scaling. The black line shows a power law scaling expected for a continuous transition (slope chosen by eye to approximate the data for small systems). While finite size effects make it difficult to study the behavior for arbitrary super-linear scaling, it is clear that for all $B \geq \mathcal{O}(N^{1.2})$ the largest gap in the size of the largest cluster does not disappear, evidencing a discontinuous transition.

To further illustrate the universality of these results, depending only on the scaling but not the specific choice of the cost function, we considered other intervention cost functions, specifically $c[S(i), S(j)] = \min[S(i), S(j)]$ and $c[S(i), S(j)] = S(i) + S(j)$ as above (Fig. S7), as well as $c[S(i), S(j)] = \max[S(i), S(j)]$ (not shown) with equivalent results. Note that cost scale linearly with the size of the clusters in all cases. In all cases we can map the parameters to corresponding parameters for constant intervention cost with a budget $b' < b'_c$ as discussed above.

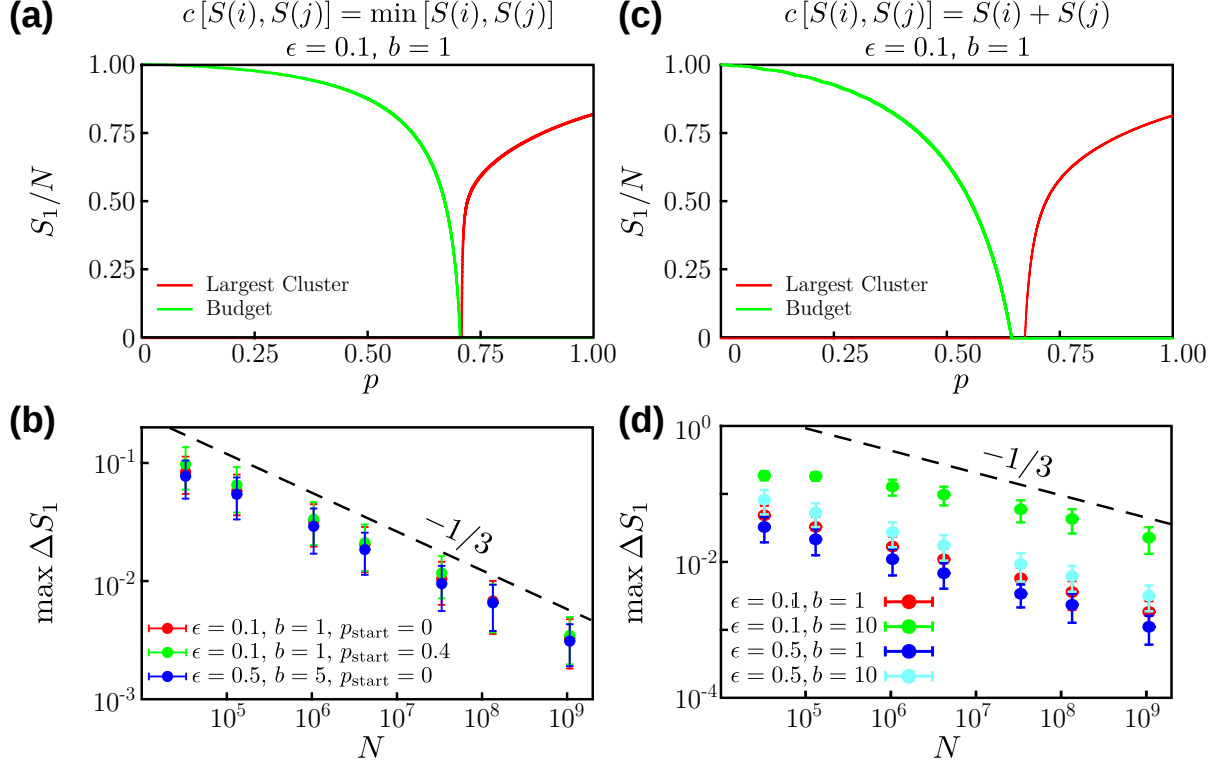


FIG. S7. Results for the intervention rule as used in the main manuscript, but for different intervention costs as noted above the two columns ($N = 2^{25}$). Error bars indicate the standard deviation, averages are taken over 1024 to 64 realizations, depending on the system size. (a,c) Single realizations of the size of the largest cluster as well as the remaining fraction of the total budget. (b,d) The largest gap of the size of the largest cluster for various parameter combinations. The transition is continuous and behaves as expected for random percolation in all cases.

Similarly, we consider different intervention rules. In the main manuscript we derived the intervention rule using an intuitive argument to achieve at least a given effectiveness of the intervention. Here we explicitly demonstrate other intervention rules (Fig. S8 and S9), showing that the qualitative behavior is similar. We again consider intervention cost proportional to the size of the clusters $c[S(i), S(j)] = S(i) + S(j)$.

Specifically, we consider an intervention rule based on the variance of the cluster size distribution: we prevent a link if the change ΔV of the variance is larger than a certain threshold ϵ/N , in order to keep cluster sizes in the network similar (and thus prevent large clusters). While this rule is less complex numerically, as we can track the variance as the network grows, it is also less efficient than the protocol derived in the main manuscript. Additionally, the threshold does not easily scale with the system size: the scaling changes depending on the shape of the cluster size distribution at any given time.

Similarly, we consider the entropy $E = \sum_S n_S \log(n_S)$ instead of the variance, where n_S is the probability that a random cluster has size S . In both cases we find qualitatively similar results as above: interventions should be applied early and with an intermediate threshold (intensity) adjusted to the budget.

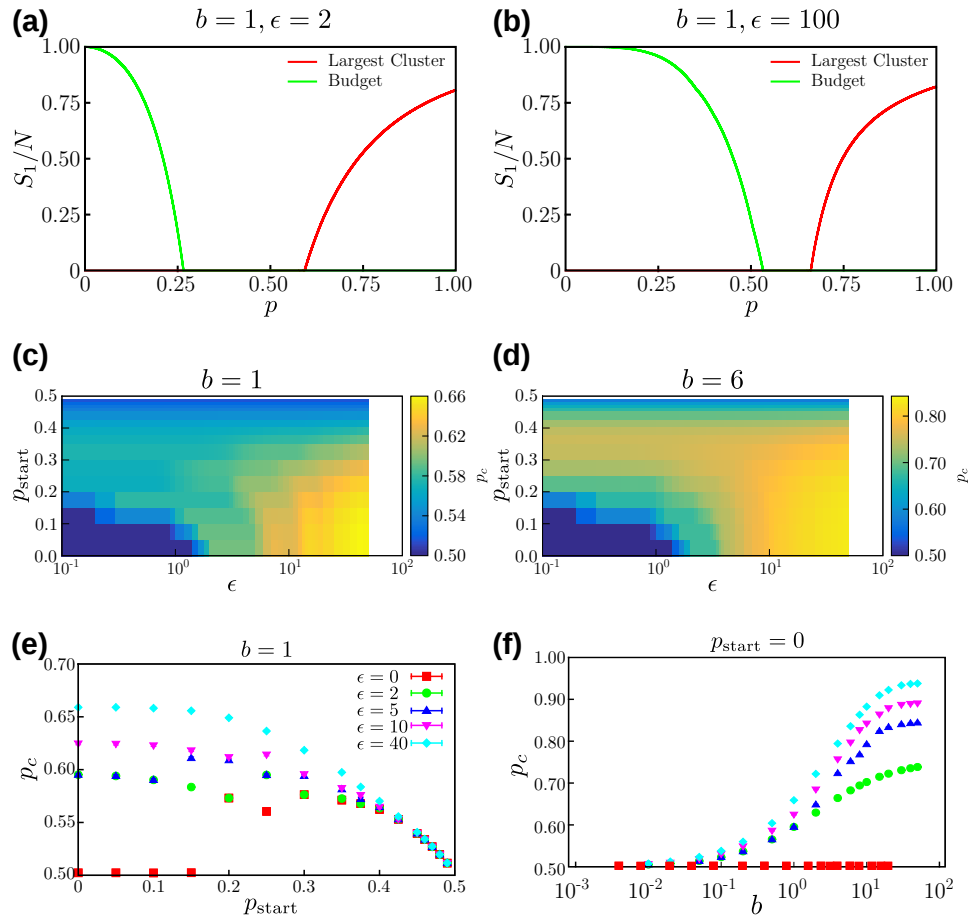


FIG. S8. Results for the variance intervention rule (see text) for $N = 2^{25}$. Error bars indicate the standard deviation, averages are taken over 256 realizations. (a,b) show two examples of single realizations for different intervention thresholds.

(c,d) show the same colormaps plots as in Fig. S5, illustrating the position of the phase transition versus different parameters: due to the cost function, early interventions are preferable.

(e,f) show the resulting position of the percolation transition versus the start of the interventions and the budget, respectively (compare Fig. S4). We find the same qualitative behavior: early interventions are optimal and a larger budget obviously allows for a larger delay. Non-monotonocities in the resulting curves for p_c are due to the non-monotonous scaling of the changes in the variance. The simplest example is the following: for $\epsilon = 0$ we prevent the first link when $p_{\text{start}} = 0$ and thus use all budget at $p = 0$. However, when $p_{\text{start}} > 0$ there are links which actually reduce the variance and our interventions are not useless.

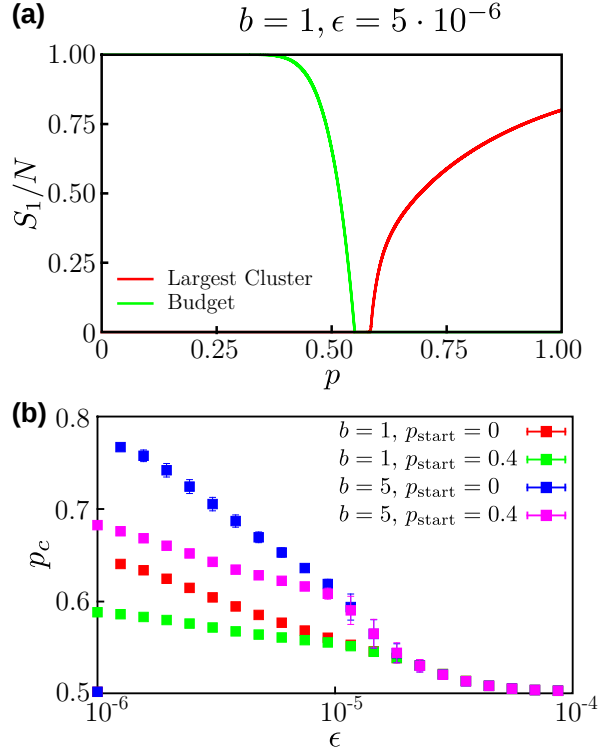


FIG. S9. Results for the entropy intervention rule (see text) for $N = 2^{25}$. Error bars indicate the standard deviation, averages are taken over 1024 realizations. Results are qualitatively similar to the other interventions rules considered: a larger budget will always delay the transition and, due to the cost function, early interventions are preferable. As for the variance interventions, it is possible to choose a parameter ϵ that always stops the first merger at $p = 0$ (datapoints at $\epsilon = 10^{-6}$, $p_c \approx 0.5$).

-
- [1] D. Achlioptas, R. M. D'Souza, and J. Spencer, *Science*, **323**, 1453 (2009).
 - [2] T. Bohman, A. Frieze, and N. C. Wormald, *Random Struct. Algorithms*, **25**, 432 (2004).
 - [3] W. Chen, M. Schröder, R. M. D'Souza, D. Sornette, and J. Nagler, *Phys. Rev. Lett.*, **112**, 155701 (2014).
 - [4] K. J. Schrenk, A. Felder, S. Deflorin, N. A. M. Araújo, R. M. D'Souza, and H. J. Herrmann, *Phys. Rev. E*, **85**, 031103 (2012).

Chapter 7

Original manuscript: Hysteretic percolation from locally optimal decisions

Citation

Schröder, M., Nagler, J., Timme, M., and Witthaut, D. (2017),
Hysteretic percolation from locally optimal decisions,
submitted

arXiv:1709.07257

©2017 Schröder, Nagler, Timme, and Witthaut.

Original contribution

Conception and design of the research with D. Witthaut and M. Timme. I performed all simulations, partially with world map data provided by D. Witthaut. I analyzed the data and generated figures 3-5 and figure 6 based on world maps by D. Witthaut. I designed and wrote the supplemental information and performed all simulations and created most figures (world map figures by D. Witthaut). I performed all analytical calculations, partially together with D. Witthaut. I designed the contents, wrote the manuscript, and revised the manuscript with all authors.

Göttingen, 17.03.2018

Place, Date

Malte Schröder

Malte Schröder

Hysteretic percolation from locally optimal decisions

Malte Schröder,^{1,*} Jan Nagler,^{2,†} Marc Timme,^{1,3,‡} and Dirk Witthaut^{4,5,§}

¹*Network Dynamics, Max Planck Institute for Dynamics and Self-Organization (MPIDS), 37077 Göttingen, Germany*

²*ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland*

³*Chair for Network Dynamics, Center for Advancing Electronics Dresden (cfaed) and Institute for Theoretical Physics, Technical University of Dresden 01069 Dresden, Germany*

⁴*Forschungszentrum Jülich, Institute for Energy and Climate Research - Systems Analysis and Technology Evaluation (IEK-STE), 52428 Jülich, Germany*

⁵*Institute for Theoretical Physics, University of Cologne, 50937 Köln, Germany*

The emergence of large-scale connectivity underlies the proper functioning of many networked systems, ranging from social networks and technological infrastructure to global trade networks. Percolation theory characterizes network formation following stochastic local rules, while optimization models of network formation assume a single controlling authority or one global objective function. In socio-economic networks, however, network formation is often driven by individual, locally optimal decisions. How such decisions impact connectivity is only poorly understood to date. We study how large-scale connectivity emerges from decisions made by rational agents that individually minimize costs for satisfying their demand. We establish an exact mapping of the resulting nonlinear optimization problem to a *local* percolation model and analyze how locally optimal decisions on the micro-level define the structure of networks on the macroscopic scale.

The proper functioning of networked systems fundamentally relies on their established large-scale connectivity. The global connectivity of social, economic and technological networks, such as the internet, trade and transportation networks, enables global communication and exchange, but also the rapid spreading of diseases [1–7]. The loss of connectivity, or even of a single connection, may cause catastrophic effects such as the collapse of ecological networks, blackouts of power grids and other infrastructures, or even a global economic crisis [8–15]. Understanding how global connectivity emerges thus constitutes a key challenge in the field of network science.

Two major theoretical approaches have been established for revealing core properties of the emergence of large-scale connectivity. First, the theory of percolation provides fundamental insights about network formation processes by assuming that new links are established stochastically according to some local rule [16, 17]. For such percolation models, a variety of distinct structure-forming phenomena have been observed, where diverse network topologies emerge even for simple link formation rules [7, 18–24]. Second, global optimization models explain network formation controlled by a central authority or driven by a single global objective function. These models have been studied to construct and understand aspects of the structure of various man-made or biological networks [25–32]. The formation of many socio-economic networks, however, is driven by local agents making individual decisions based on optimizing their own goals. Such settings result in networks constrained by many individual, yet interacting optimization problems. A similar motivation underlies game-theoretic models of network formation [33–38]. These models allow a more detailed analysis of the formation process and the stability of the resulting network. Unsurprisingly, however, they

are often hard, if not impossible, to solve, especially for larger networks, which limits mechanistic insights.

In this Letter we study network formation processes based on rational agents that individually optimize their own local objective function. Given costs for production and transaction (including transport) in an underlying transport network, each agent satisfies its own demand at minimal costs [39, 40]. For the resulting class of nonlinear optimization problems, we establish an exact mapping to a *local* percolation model, enabling us to efficiently investigate the collective network formation. The proposed framework bridges the gap between global optimization models of network formation and stochastic local percolation models.

From optimization to percolation — We analyze a network formation model based on the following fundamental network supply problem. Consider an underlying network of N nodes and M links, describing agents and *potential* transportation routes, where each agent must satisfy its demand. We study the network of trades that actually evolves between the nodes, similar to bond percolation on an underlying network or random graph [16].

Specifically, we assume each node $i \in \{1, \dots, N\}$ is an agent with a fixed demand D_i . The agent satisfies this demand by purchases S_{ki} from any nodes k , including its own node, under the constraint that $\sum_k S_{ki} = D_i$. (Throughout the manuscript all sums run over all nodes, here $k \in \{1, \dots, N\}$, unless otherwise noted.) Each agent will try to achieve this with minimal cost

$$K_i = \sum_k K_{ki}^P + K_{ki}^T, \quad (1)$$

including both production costs K_{ki}^P at node k as well as transaction costs K_{ki}^T between the two nodes (see Fig. 1). The production costs depend on the total amount $S_k =$

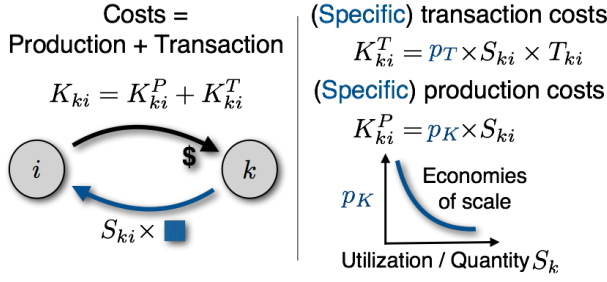


Figure 1. **Cost structure of the supply problem.** Each agent i satisfies its demand by purchases S_{ki} from nodes k at minimum costs, including production and transaction costs [Eq. 1]. The specific production cost at node k decrease with the total production S_k at k , describing economic scale effects. The transaction costs are proportional to the shortest path distance T_{ki} between k and i in the underlying transport network and to the specific transaction costs p_T .

$\sum_j S_{kj}$ produced at node k and are proportional to the amount of purchased goods S_{ki} , resulting in the nonlinear dependence

$$K_{ki}^P = p_k(S_k) S_{ki} \quad (2)$$

on the purchases S_{ki} . The specific production cost p_k describe scale effects, for example increasing efficiency with increasing production. The transaction costs are proportional to the amount of transported goods S_{ki} , the distance T_{ki} between the two nodes in the underlying transportation network and the specific transaction costs p_T ,

$$K_{ki}^T = p_T S_{ki} T_{ki}, \quad (3)$$

where $T_{ki} = \sum_e t_e$ is given as the sum of the distances t_e of all links e along the (shortest) path between k and i in the underlying transport network. All agents solve their individual nonlinear optimization problem [Eq.(1)] simultaneously, defining the network of optimal purchases S_{ki} . The resulting state of this network then corresponds to a Nash-equilibrium [41], where no agent can reduce its cost by changing its purchases given that all other purchases remain constant.

Results — A simple, yet efficient solution to this problem can be found for non-increasing specific production costs p_k . In this case, we find that any agent i chooses a single supplier i^* , such that $S_{i^*i} = D_i$ and $S_{ki} = 0$ for $k \neq i^*$. In general each agent would have to check each node in the network to find its optimal supplier. Interestingly, if the demand of all agents is identical, $D_i \equiv D$, this optimal supplier can be found *locally*: An agent i just has to query its direct neighbors about their current suppliers to find the optimal supplier i^* .

Here, we provide a brief intuitive argument: Any purchase of agent i has to be transported via one of its nei-

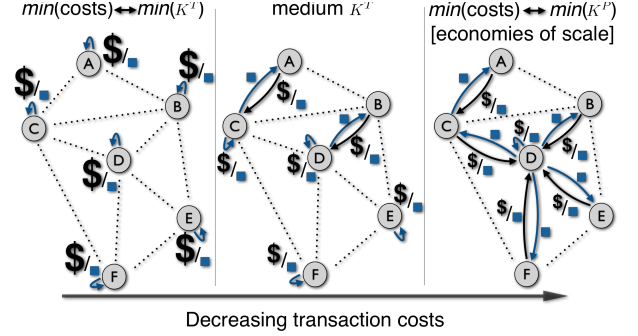


Figure 2. **Local percolation induced by optimization.** The multipartite optimization problem illustrated in Fig. 1 is solved by a local percolation algorithm. Dashed lines show potential transportation routes. As the specific transaction costs p_T decrease, agents start to purchase from other nodes and a network of trades emerges (arrows). When a cluster is established, the specific production costs decrease due to scale effects, enabling discontinuous network growth.

ghors j . Note that the specific production costs only depend on the supplier and transaction costs are additive over the shortest path. If j is buying optimally from j^* , this supplier must also be optimal for i as any other supplier transporting via j would be more expensive. Thus, i will learn about its optimal supplier locally from one of its neighbors (see Supplemental Material Sec. I and II for a rigorous proof and details of the simulation [42]).

We investigate this local percolation model starting with large transaction costs $p_T = \infty$ and, correspondingly, only internal production $i^* = i$ and $S_{ii} = D_i$. As the specific transaction costs p_T decrease, agents minimize their costs by establishing external purchases S_{ki} . Finally, transaction costs disappear at $p_T = 0$ and all agents will have the same supplier (Fig. 2). We study the size $C(i^*)$ of the connected components (clusters) in the network defined by these purchases, this means the number of agents $\{i_1, i_2, \dots\}$ with the same supplier i^* . As for standard percolation we record the size $C_1(p_T)$ of the currently largest cluster (and C_2 for the second largest cluster and so on). In the following examples we consider linearly decreasing specific costs $p_k(S_k) = b_k - aS_k$, where $a \geq 0$ determines the strength of the scale effects.

Discontinuous percolation and hysteresis — We illustrate the emergence of connectivity in a random spatially embedded network in Fig. 3, revealing the importance of scale effects. Weak scale effects (small a) lead to a continuous growth of the largest cluster. Strong scale effects (large a) lead to a discontinuous evolution of the size of the largest cluster in the network. A microscopic decrease of the transaction costs triggers a cascade of decisions: As a suppliers' production costs decrease, a large

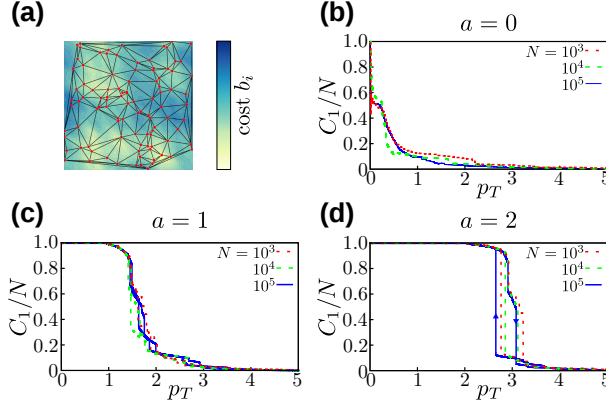


Figure 3. From continuous transitions to discontinuous transitions and hysteresis. The type of the transition to a single globally connected component changes depending on the size of the scale effects in the cost function. (a) Example of a small random network embedded in the unit square (see Supplemental Material Sec. III [42]). Parameters b_i are taken as values of a random function $b(x, y)$. (b)-(d) Single realizations of the evolution of the relative size of the largest cluster for different system sizes N . Small scale effects lead to a continuous growth of the largest cluster. Sufficiently large scale effects lead to a discontinuous transition. Reversing the process, i.e., increasing the transaction costs, leads to a direct reversal for small scale effects, but hysteresis is observed for large scale effects.

fraction of agents join a connected component with only a single supplier. In the language of percolation, a giant connected component emerges in a continuous (small scale effects) or discontinuous (strong scale effects) phase transition.

Moreover, multiple stable states exist for strong scale effects. In an intermediate interval of transaction costs p_T the network settles on one of the possible structures, depending on the previous state of the network: hysteresis emerges. Thus, a large cluster may remain stable after it has emerged for decreasing p_T , even when p_T is increased again [Fig. 3(d)].

Underlying mechanism — To understand the mechanism underlying these different transitions, we analyze a mathematically tractable system in detail: consider a complete graph as underlying network in which $t_e = 1$ for all links, $b_i \in [0, 1]$ uniformly randomly distributed and $D_i = 1/N$, illustrated in Fig. 4(a). We consider the limit of large N and assume a mean field distribution of the b_i , allowing us to re-order the nodes such that $b_i \approx i/N$ for $i \in \{1, \dots, N\}$.

We now track individual decisions by considering the cost $K_i(k)$ agent i pays at a supplier k . Since transaction costs across all links are identical, clearly the first link to be established will be between the node with the highest

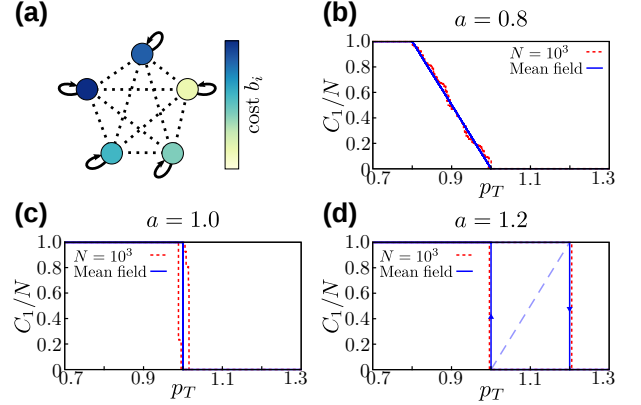


Figure 4. Theoretical analysis of the mechanism of discontinuous transitions and hysteresis. (a) A small example of a complete graph: all nodes are connected with $T_{ij} = t_{ij} = 1$ and random $b_i \in [0, 1]$ (see text). (b)-(d) A single realization of the relative size of the largest cluster for different scale effects a for a small system of $N = 10^3$ nodes and the predicted behavior in the mean field limit $N \rightarrow \infty$ (see text). For small scale effects the largest cluster grows continuously. For sufficiently large scale effects the transition becomes discontinuous as a single agent changing its supplier changes the cost enough to induce a cascade. For even larger scale effects, we find hysteresis for increasing transaction costs. This is the same qualitative behavior as observed for the spatially embedded network (Fig. 3) and will be qualitatively unchanged for different distributions of the b_i .

specific production cost (node N) and the one with the smallest (node 1). This happens when the cost per unit $K_N(1)$ for agent N to import from node 1 become smaller than the cost $K_N(N)$ to buy internally: $K_N(1) = 1/N - 2a/N + p_T < 1 - a/N = K_N(N)$, that is for $p_T < p_T^N = 1 - 1/N + a/N$. Similarly, we can calculate when the next link between agent $N-1$ and node 1 appears: $K_{N-1}(1) = 1/N - 3a/N + p_T < (N-1)/N - a/N = K_{N-1}(N-1)$, that is for $p_T < p_T^{N-1} = 1 - 2/N + 2a/N$. The other agents follow the same pattern.

Considering the two links, we now have to distinguish two cases: if $a < 1$, then $p_T^{N-1} < p_T^N$ and the agents N and $N-1$ will establish their links sequentially at different values of p_T . The largest cluster will grow continuously with a slope of $(1/N)/(p_T^{N-1} - p_T^N) = 1/(a-1)$. However, if the scale effects are strong enough ($a \geq 1$), the cost at node 1 decrease sufficiently for the next link to be established immediately since $p_T^{N-1} \geq p_T^N$. The cluster grows discontinuously in a single cascade [Fig. 4(c)]. If the scale effects are even stronger ($a > 1$), the cluster is stable with respect to single agents changing their supplier for larger values of p_T , causing hysteresis when increasing p_T [Fig. 4(d)]. In this case the line with slope $1/(a-1)$ describes an unstable equilibrium state (see also Supplemental Material Sec. VI [42]).

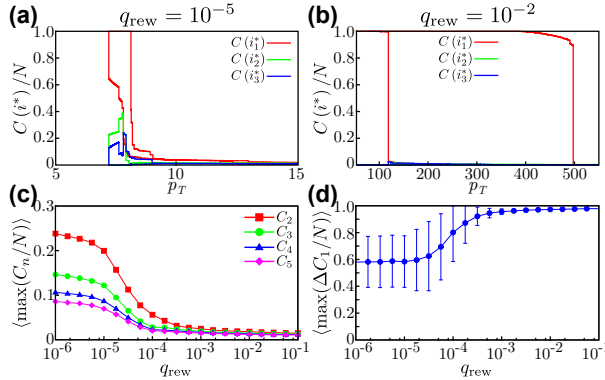


Figure 5. Impact of the network structure on the percolation transition. Size and discontinuities of large connected components for an underlying Watts-Strogatz small world network ($N = 10^4$, $k = 8$, see Supplemental Material Sec. III and V [42]). (a),(b) Single realizations for the size of the components $C(i^*)$ of three specific large suppliers $i \in \{i_1^*, i_2^*, i_3^*\}$. In a network with large diameter [panel (a), $q_{\text{rew}} = 10^{-5}$] multiple clusters grow simultaneously and merge for small p_T . In a network with small diameter [panel (b), $q_{\text{rew}} = 10^{-2}$] one large cluster emerges in a single cascade. (c) Maximum size of the n -th largest cluster as a function of the topological randomness q_{rew} (error bars omitted for visibility). (d) Largest change of the size of the largest cluster (error bars indicate the standard deviation). For large network diameters (small q_{rew}) multiple large clusters emerge independently in different parts of the network. For small network diameters only one large cluster emerges in a single large cascade, characterized by a large value of the size of the largest jump $\max(\Delta C_1/N)$. Averages are taken over $R = 100$ realizations.

Impact of network topology — Besides the scale effects, the network growth is further determined by the underlying physical transportation network. For constant scale effects we find different routes of network formation depending on the structure of the network. If the network diameter (the longest shortest path between any two nodes) is small, only one cluster emerges. If the diameter is large, multiple large clusters appear. This difference is already evident when comparing the spatially embedded (large diameter) and complete network (small diameter) for $a = 1$ [compare Fig. 3(c) and 4(c)].

To systematically study this effect we consider a network class introduced by Watts and Strogatz [23]. Starting from a regular ring network where each node is connected to its k neighbors (large diameter) each link is randomly rewired with probability q_{rew} , which introduces shortcuts and reduces the diameter of the network. If the diameter is large [q_{rew} small, Fig. 5(a)], different suppliers can attract large clusters of agents from their local part of the network when p_T decreases. Fig. 5(c) illustrates the maximum size of the n -th largest cluster, showing

that multiple large clusters emerge for $q_{\text{rew}} \leq 10^{-5}$ (less than one shortcut per node). Only for small values of p_T do these clusters interact and finally merge in small cascades to a single giant cluster. If the diameter is small [q_{rew} large, Fig. 5(b)], only a single cluster emerges, attracting nodes from all parts of the network. The largest cluster then grows in a single cascade until it fills the entire network [Fig. 5(c),(d)].

As an example of a realistic transportation network we consider an elementary model of a world transport network (Fig. 6). We explore different network structures by varying the costs for different modes of transportation, modifying the effective distances of transport via land and sea. If transport via sea is expensive the network has a large diameter and multiple large clusters appear in different regions of the world, merging when p_T becomes small [Fig. 6(a)-(d)]. Conversely, if transport via sea is cheap, the network becomes densely connected with a small diameter and a single largest cluster grows in a sudden cascade [Fig. 6(e)-(g)].

Discussion — We have studied a network formation model based on fundamental economic considerations, describing a network of trading agents minimizing their costs to satisfy a fixed demand. Link addition in this model is driven by local decisions of the individual agents instead of random chance. We showed that under simple conditions the resulting complex, nonlinear optimization problem can be solved with a local percolation model, bridging the gap between optimization-based approaches of network formation [25–32] and prototypical percolation models [7, 18–24]. We revealed how local optimization affects the phase transition to global connectivity. In particular, we illustrated how scale effects and the topology of the underlying transport network determine the emergence of a discontinuous transition and hysteresis in the network formation process.

The observed effects are not only of interest to understand explosive phase transitions but also have important consequences in political economics. First, globalization is not necessarily a continuous process. Very small changes in the transportation processes can change the equilibrium from local to centralized production. Such discontinuous transitions are associated with hysteresis so that large common markets are hard to disentangle. Our results suggest that a small increase of transportation costs through taxes or customs will not undo the formation of a giant connected component.

Furthermore, our model highlights the importance of the transportation network for global economics [39, 40]. Both, economic parameters as well as the current structure of the network, determine whether there may be only one, several, or no giant connected component. Expressed in economic terms, whether there emerges a monopoly, oligopoly, or polypoly is not only a matter of the micro-level dynamics al-

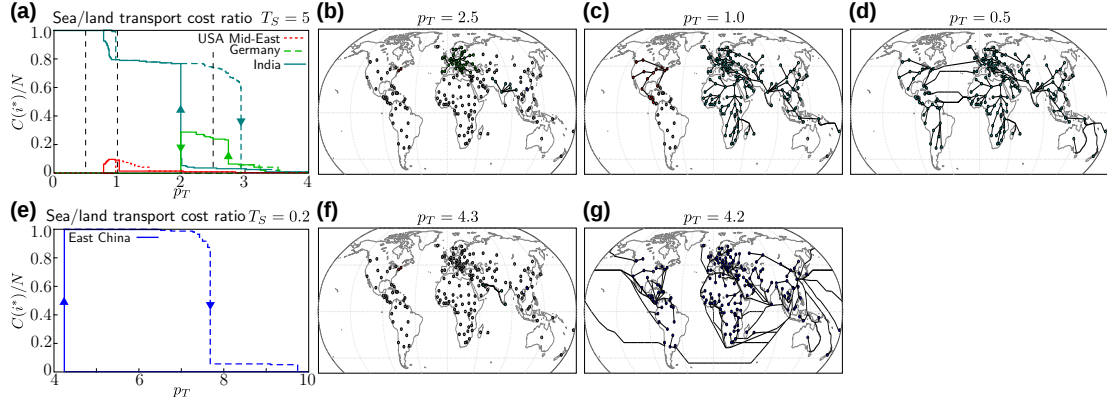


Figure 6. **Preferred modes of transport change network evolution.** Global connectivity induced by locally optimal decisions in a model of a world transport network (see Supplemental Material Sec. VII for details [42]). (a) Evolution of the size $C(i^*)/N$ of specific clusters identified by a supplier i^* when transport via sea is more expensive than transport via land (by a factor $T_S = 5$). Due to the high costs of sea transport, the network diameter is large and multiple large clusters can emerge in different parts of the world. (b)-(d) Network structure and active transport links for different values of the transactions costs p_T . Land routes are preferred to transport via sea. (e) Evolution of the size of the emerging cluster for small sea transport costs ($T_S = 0.2$). Due to sea routes connecting most countries cheaply, only one large component emerges in a single large cascade. This transition happens for larger p_T as the overall diameter of the network is much smaller. (f),(g) The state of the network immediately before and after the transition.

one, but also of the actual structure of the trade network.

We thank J. Többen and S. Klipp for helpful discussions. We gratefully acknowledge support from the Göttingen Graduate School for Neurosciences and Molecular Biosciences (DFG Grant GSC 226/2), from the Helmholtz association (grant no. VH-NG-1025), the German Ministry for Education and Research (BMBF grants no. 03SF0472A and 03SF0472E), the German Science Foundation (DFG) by a grant towards the Center of Excellence *Center for Advancing Electronics Dresden* (cfaed), the ETH Risk Center (RC SP 08-15) and SNF Grant *The Anatomy of systemic financial risk*, No. 162776.

* malte.schroeder@ds.mpg.de

† jnagler@ethz.ch

‡ marc.timme@ds.mpg.de

§ d.witthaut@fz-juelich.de

- [1] R. Albert, H. Jeong, and A.-L. Barabási, *Nature*, **406**, 378 (2000).
- [2] R. Pastor-Satorras and A. Vespignani, *Phys. Rev. Lett.*, **86**, 3200 (2001).
- [3] L. Hufnagel, D. Brockmann, and T. Geisel, *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 15124 (2004).
- [4] D. Brockmann, L. Hufnagel, and T. Geisel, *Nature*, **439**, 462 (2006).
- [5] D. Brockmann and D. Helbing, *Science*, **342**, 1337 (2013).

- [6] R. Albert and A.-L. Barabási, *Rev. Mod. Phys.*, **74**, 47 (2002).
- [7] M. Newman, *SIAM Review*, **45**, 167 (2003).
- [8] R. V. Sole and M. Montoya, *Proc. Roy. Soc. London Ser. B*, **268**, 2039 (2001).
- [9] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White, *Science*, **325**, 422 (2009).
- [10] S. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, *Nature*, **464**, 1025 (2010).
- [11] S. Havlin, D. Y. Kenett, E. Ben-Jacob, A. Bunde, R. Cohen, H. Hermann, J. Kantelhardt, J. Kertész, S. Kirkpatrick, J. Kurths, J. Portugali, and S. Solomon, *Eur. Phys. J. Special Topics*, **214**, 273 (2012).
- [12] M. Elliott, B. Golub, and M. O. Jackson, *Am. Econ. Rev.*, **104**, 3115 (2014).
- [13] D. Witthaut, M. Rohden, X. Zhang, S. Hallerberg, and M. Timme, *Phys. Rev. Lett.*, **116**, 138701 (2016).
- [14] H. Ronellenfitsch, D. Manik, J. Horsch, T. Brown, and D. Witthaut, *IEEE Trans. Power Syst.*, **32**, 4060 (2017).
- [15] J. Nagler, A. Levina, and M. Timme, *Nat. Phys.*, **7**, 265 (2011).
- [16] D. Stauffer and A. Aharony, *Introduction To Percolation Theory* (Taylor & Francis, London, 1992).
- [17] G. Grimmett, *Percolation* (Springer, Berlin, 1999).
- [18] D. Achlioptas, R. D'Souza, and J. Spencer, *Science*, **323**, 1453 (2009).
- [19] O. Riordan and L. Warnke, *Science*, **333**, 322 (2011).
- [20] M. Schröder, S. E. Rahbari, and J. Nagler, *Nat. Commun.*, **4** (2013).
- [21] M. Schröder, W. Chen, and J. Nagler, *New J. Phys.*, **18**, 013042 (2016).
- [22] R. M. D'Souza and J. Nagler, *Nat. Phys.*, **11**, 531 (2015).
- [23] D. J. Watts and S. H. Strogatz, *Nature*, **393**, 440 (1998).
- [24] T. Verma, F. Russmann, N. Araújo, J. Nagler, and H. Herrmann, *Nat. Commun.*, **7** (2016).

- [25] D. P. Bertsekas, *Network optimization: continuous and discrete models* (Athena Scientific, Belmont, 1998).
- [26] M. T. Gastner and M. E. J. Newman, Phys. Rev. E, **74**, 016117 (2006).
- [27] S. Bohn and M. O. Magnasco, Phys. Rev. Lett., **98**, 088702 (2007).
- [28] E. Katifori, G. J. Szöllősi, and M. O. Magnasco, Phys. Rev. Lett., **104**, 048704 (2010).
- [29] H. Ronellenfitsch and E. Katifori, Phys. Rev. Lett., **117**, 138301 (2016).
- [30] D. B. Chklovskii, T. Schikorski, and C. F. Stevens, Neuron, **34**, 341 (2002), ISSN 0896-6273.
- [31] R.-M. Memmesheimer and M. Timme, Physica D, **224**, 182 (2006), ISSN 0167-2789, dynamics on Complex Networks and Applications.
- [32] R.-M. Memmesheimer and M. Timme, Phys. Rev. Lett., **97**, 188101 (2006).
- [33] M. Jackson and A. Wolinsky, J. Econ. Theory, **71**, 44 (1996).
- [34] M. Jackson and A. Watts, J. Econ. Theory, **106**, 265 (2002).
- [35] M. O. Jackson, *Social and economic networks*, Vol. 3 (Princeton University Press, Princeton, 2008).
- [36] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world* (Cambridge University Press, 2010).
- [37] V. Bala and S. Goyal, Econometrica, **68**, 1181 (2000).
- [38] M. D. König, S. Battiston, M. Napoletano, and F. Schweitzer, Games Econ. Behav., **75**, 694 (2012).
- [39] P. R. Krugman, *Geography and trade* (MIT press, 1991).
- [40] P. R. Krugman, J. Pol. Econ., **99**, 483 (1991).
- [41] M. J. Osborne and A. Rubinstein, *A course in game theory* (MIT Press, Cambridge, MA, 1994).
- [42] See Supplemental Material at [URL will be inserted by publisher].
- [43] D. Saupe, in *The science of fractal images* (Springer, New York, 1988) pp. 71–136.

Supplemental Material

accompanying the manuscript

Hysteretic percolation from locally optimal decisions

Malte Schröder, Jan Nagler, Marc Timme and Dirk Witthaut

In the main manuscript we introduced a network formation model where link addition is based on individual decisions of local agents in a fundamental network supply problem. In general the resulting individual optimization problems are too complex to be solved efficiently for large systems. We discussed how these optimization problems map to a local percolation model with an efficient solution, bridging the gap between global optimization models of network formation and stochastic local percolation models. This mapping allowed us to efficiently study the phenomena emerging in this network formation model.

In this supplementary information we give the rigorous proofs for the mapping as well as additional details and examples for the model. First, we consider the mapping of the optimization problem to a percolation model. In particular, we discuss in detail the assumptions and requirements on the cost functions and other parameters for this mapping to be valid and give a rigorous proof. We also describe the simulation procedure resulting from these conditions. Second, we discuss in more detail the parameter choices made for the examples shown in the main manuscript in relation to properties of the model, describing how and why we take the limit of an increasing number of nodes $N \rightarrow \infty$ as a fine-graining rather than an expansion of the system in relation to the underlying network topology (for example random and small world networks). Third, we provide an additional example for an analytically solvable system in a continuous interpretation of the model, linking the equilibrium states and transitions to bifurcations of stable states of a corresponding self-consistency equation. Finally, we include a detailed description of the data used for the world transport network simulations in the main manuscript and a brief comparison of the exact solution for the world transport network to the approximate solution with our local percolation algorithm, which is shown in the main manuscript.

THE NETWORK SUPPLY PROBLEM

To begin, we reiterate the basic idea of the network formation model as defined in the main manuscript: Starting with an underlying network of N nodes and M *potential* transport links, we consider each node i to be an agent with a fixed demand D_i that it satisfies by purchases S_{ki} from nodes k , possibly internally from $k = i$. These purchases incur a cost of two parts: First, the production cost $K_{ki}^P = S_{ki}p_k(S_k)$ at node k depending on the total amount of production S_k at that node. Second, a transaction cost (e.g., transport cost) in the underlying network, $K_{ki}^T = p_T S_{ki} T_{ki}$, where $T_{ki} = \sum_{e \in \Pi(k,i)} t_e$ is the sum over the transaction costs of all links e along the shortest (cheapest) path $\Pi(k, i)$ from k to i . The parameter p_T describes a global scaling of transaction cost across the network, i.e., the importance of transaction costs relative to production costs. We assume that each agent only decides its purchases and the production at a node is always given as the sum of all purchases made from that node, $S_k = \sum_i S_{ki}$. As in the main manuscript we assume that sums run over all nodes $i \in \{1, \dots, N\}$ unless noted otherwise. Similarly, each agent always satisfies its demand exactly, such that $D_i = \sum_k S_{ki}$. Each agent i then individually optimizes its purchases S_{ki} to minimize its cost K_i while satisfying its demand:

$$\begin{aligned}
 &\text{minimize} && K_i = \sum_k S_{ki} p_k(S_k) + p_T \sum_k \left[S_{ki} \sum_{e \in \Pi(k,i)} t_e \right] \\
 &\text{subject to} && \sum_k S_{ki} = D_i \\
 &&& S_{ki} \geq 0 \quad \forall k.
 \end{aligned} \tag{1}$$

We consider the evolution of the network of trades described by the purchases S_{ki} , when the transaction cost p_T decrease, starting from $p_T = \infty$ with initially only internal production $S_{ii} = D_i$ and $S_{ki} = 0$ for $i \neq k$. Large transaction costs p_T mean that no external purchases are made across the network. As p_T decreases, some agents start buying at cheaper neighbors and small, localized connected components (clusters) start to grow. Specifically, as in standard percolation theory, we are interested in the evolution of the size of the largest cluster $C_1(p_T)$.

To allow easier readability, we use different indices to denote nodes or agents based on their role in the current context wherever possible. We summarize the variables together with their meaning in the following table:

N	The number of nodes in the system, system size
e_{ij}	Link (potential transport link) between nodes i and j
t_e, t_{ij}	Transaction cost per unit for the link e from node i to j
$\Pi(k, i)$	Shortest (cheapest) path from k to i , $\Pi(k, i) = \{e_{kj_1}, e_{j_1j_2}, \dots, e_{j_ni}\}$
T_{ki}	Total transaction cost per unit between the nodes k and i , $T_{ki} = \sum_{e \in \Pi(k, i)} t_e$
p_T	Specific transaction cost, relative importance of transaction cost
i, j	Index of an agent (currently looking to make purchases)
k, l	Index of a node (currently considered as a supplier)
i', j'	Index of the current supplier of agent i, j
i^*, j^*	Index of the/an optimal supplier of agent i, j
D_i	The (fixed) demand of agent i
S_k	The current, total production of node k
S_k^i	The production of node k disregarding possible purchases by agent i , $S_k^i = S_k - S_{ki}$
S_{ki}	The amount agent i purchases from node k
$p_k(S_k)$	The specific production cost at node k with a total production S_k
K_i	The total cost, including transaction costs and production costs, of all purchases of agent i
$K_i(k)$	The total cost of agent i when <i>only</i> purchasing from node k

I. MAPPING TO A LOCAL PERCOLATION MODEL

In general the optimization problem described above is very complex and quickly grows intractable for larger systems. Here we derive a general updating scheme in terms of a local percolation model, showing that it exactly solves all individual optimization problems under certain conditions. We first consider the cost function of a single agent i

$$K_i = \sum_k [S_{ki} p_k (S_k^i + S_{ki}) + p_T S_{ki} T_{ki}], \quad (2)$$

where $S_k^i = \sum_{j \neq i} S_{kj}$ denotes the production of node k ignoring purchases by agent i , such that $S_k = S_k^i + S_{ki}$, and we write the total transaction cost per unit over all links from k to i as T_{ki} . The problem becomes considerably easier for the family of cost functions describing economies of scale, this means decreasing specific production costs with increasing production. In this case the problem of finding the optimal S_{ki} reduces to finding a single optimal supplier.

Lemma 1. *Given an individual minimization problem of agent i defined by Eq. (1) with non-increasing specific production costs at each node, $\frac{\partial p_k}{\partial S_k}(S_k) \leq 0$ for all $k \in \{1, \dots, N\}$, there is a node i^* such that K_i is minimal with $S_{i^*i} = D_i$ and $S_{ki} = 0$ for $k \neq i^*$.*

Proof. Choose the node l for which

$$p_l^{\text{eff}}(S_l^i + D_i) = p_l(S_l^i + D_i) + p_T T_{li} \quad (3)$$

is smallest. Together with $\partial p_k / \partial S_k \leq 0$ for all nodes k we then obtain

$$p_l^{\text{eff}}(S_l^i + D_i) \leq p_k^{\text{eff}}(S_k^i + D_i) \leq p_k^{\text{eff}}(S_k^i + S_{ki}) \quad (4)$$

for all nodes $k \in \{1, \dots, N\}$ and arbitrary purchases $0 \leq S_{ki} \leq D_i$. Using the constraint that $D_i = \sum_k S_{ki}$ this implies

$$\begin{aligned} K_i(S_{1i} = 0, S_{2i} = 0, \dots, S_{li} = D_i, S_{l+1,i} = 0, \dots) \\ &= D_i p_l^{\text{eff}}(S_l^i + D_i) \\ &\leq \sum_k S_{ki} p_k^{\text{eff}}(S_k^i + S_{ki}) \\ &= K_i(S_{1i}, S_{2i}, S_{3i}, \dots) \end{aligned} \quad (5)$$

for all possible purchases $(S_{1i}, S_{2i}, S_{3i}, \dots)$. Thus the costs K_i assume a global minimum if agent i satisfies its entire demand by purchases from a single node $i^* = l$, this means for

$$S_{i^*i} = D_i \quad \text{and} \quad S_{ki} = 0 \text{ for } k \neq i^*.$$

□

In particular this class of non-decreasing functions includes the affine-linear specific production cost $p_k(S_k) = b_k - a \cdot S_k$ with $a \geq 0$ used in the main manuscript. However, we are not restricted to identical slopes or even cost functions of the same form for different nodes. We therefore cover a broad range of cases where the cost functions of all nodes are non-increasing in the range of possible production $S_k \in [0, \sum_i D_i]$.

With the above simplification, we now consider the optimization problem of agent i of finding the best i^* to minimize

$$K_i(i^*) = D_i p_{i^*}(S_{i^*}^i + D_i) + p_T D_i \sum_{e \in \Pi(i^*, i)} t_e. \quad (6)$$

To find an equilibrium, we can in principle simply try all possible alternatives, as formulated in the following algorithm. However, such a brute-force approach can become quickly infeasible for large networks such that we will consider further simplifications in the following sections.

Definition 1 (Equilibrium). *Consider the optimization problem defined above [Eq. (1)] with non-increasing specific production costs $p_k(S_k)$ for all agents $k \in \{1, 2, \dots, N\}$. An agent i is in equilibrium at its current supplier i' , if and only if there is no node k such that $K_i(k) < K_i(i')$. This means the agent cannot change its supplier to reduce its costs, given that all other agents keep their current supplier.*

We say that the network is in equilibrium if all agents in the network are in equilibrium.

Note that this corresponds to the notion of a Nash-equilibrium, where no agent can decrease its cost by changing only its own supplier. Note also, while the network might be in equilibrium, this does not necessarily mean that the total cost for all agents, $\sum_i K_i$, is minimal.

General algorithm

Consider the optimization problem defined above with *non-increasing specific production costs* $p_k(S_k)$ for all agents $k = \{1, 2, \dots, N\}$. Given a fixed value p_T , for each agent i , individually,

$$\begin{aligned} & \text{find} && i^* \in \{1, \dots, N\} \\ & \text{minimizing} && K_i(i^*) = D_i p_{i^*} (S_{i^*}^i + D_i) + p_T D_i \sum_{e \in \Pi(i^*, i)} t_e. \end{aligned} \quad (7)$$

In general, the following algorithm yields an equilibrium (though not necessarily a global optimum) when it terminates:

Algorithm 1 General optimization

```

1: repeat
2:    $\mathcal{U} \leftarrow \emptyset$ 
3:   for all  $i \in \{1, 2, \dots, N\}$  do
4:     for all  $k \in \{1, 2, \dots, N\}$  do
5:       Calculate  $K_i(k)$  [where  $S_{ki} = D_i$  and  $S_{li} = 0$  for all  $l \neq k$ ].
6:       if  $K_i(k) < K_i(i')$  where currently  $S_{i'i} = D_i$  then
7:          $\mathcal{U} \leftarrow \mathcal{U} \cup \{K_i(k)\}$ 
8:       end if
9:     end for
10:  end for
11:  if  $\mathcal{U} \neq \emptyset$  then
12:    Calculate  $\{i, i^*\}$  such that  $K_i(i^*) = \min_{\mathcal{U}} K_i(k)$ 
13:     $S_{i^*i} \leftarrow D_i$ 
14:     $S_{i'i} \leftarrow 0$ 
15:  end if
16: until  $\mathcal{U} = \emptyset$ 

```

In words, for every step we calculate the cost for all agents and possible suppliers. If an agent i would prefer a new supplier k to its current supplier i' we add the corresponding cost to the list of possible updates \mathcal{U} . Here, we use $K_i(k)$ to denote both the final cost for agent i as well as the corresponding update itself. In the end we execute the update with the smallest final cost (clearly any update will find the currently optimal choice i^* for agent i , since we check all possible suppliers in the network). We repeat this process until no further update is possible.

Theorem 2. *Algorithm 1 will always terminate in finite time. When it terminates the network will be in an equilibrium state, where no agent can reduce its costs by changing its supplier.*

Proof. Firstly, the algorithm only terminates when there are no possible updates, i.e., when no single agent can further reduce its costs by changing its supplier. As we assumed non-increasing specific production costs, a single supplier is always an optimal solution. Thus no single agent i can reduce its costs at all and we always end in an equilibrium.

Conversely, if the network is in equilibrium, no agent can reduce its cost. The update list will be empty and the algorithm will terminate.

Secondly, consider the possibility of the algorithm not terminating at all. This can only happen if the algorithm runs into a loop, executing the same updates repeating a finite set of states of the network. We now show that such a loop is impossible:

Assume an initial state \mathcal{I} of the network, defined by the pairs of agents and current suppliers $\mathcal{I} = \{(i, i'), \dots\}$. A loop can only occur if the network now reaches a state $\mathcal{I} - 1$, where a single agent i has a different supplier k , but would like to return to the supplier i' it has in state \mathcal{I} ($K_i(i') < K_i(k)$).

To reach this state we need a chain of updates including, at least, one update u_a where agent i switches supplier to k . However, if the reverse update u_a^{-1} is strictly beneficial in state $\mathcal{I} - 1$, clearly u_a is not beneficial in state \mathcal{I} . Thus the chain needs to consist of more than a single update.

Given that one update alone is not possible, we need another update to enable it. We thus have to consider a chain of updates $u_b \circ u_a \circ u_b^{-1}$, where u_b enables u_a and is reversed in the end to arrive at state $\mathcal{I} - 1$. The main idea is illustrated in Fig. S1. There are only two possible ways how a beneficial update u_b can enable u_a :

(1) either an agent j switches from its supplier j' to the new supplier k . But then the update u_a decreases the cost $K_j(k)$ further. The reverse update u_b^{-1} would then never be executed since $K_j(k) < K_j(j')$.

(2) u_b could have an agent j switch away from the current supplier $j' = i'$ to a different node l . But then u_a further increases the cost $K_k(i')$ and we again find $K_j(l) < K_j(i')$, making the reverse update impossible.

Following this line of argument, we need another update u_c to enable the update u_b^{-1} , e.g., a chain of updates $u_b \circ u_a \circ u_c \circ u_b^{-1} \circ u_c^{-1}$. This clearly requires a partial ordering of the updates, such that u_b , u_c , u_b^{-1} and u_c^{-1} are executed in this order (otherwise the update u_c will have no effect on the update u_b^{-1}). However, by the same logic as above u_b^{-1} would change the cost unfavorably for u_c^{-1} . We then find the same problem that u_c^{-1} would never be executed at the end of the chain. Repeating the argument, we find the same problem for every finite chain of updates.

Thus, we cannot construct a finite chain of updates to reach $\mathcal{I} - 1$. Consequently, we cannot repeat a state and thus would visit every possible state, but then the algorithm would terminate in (one of) the equilibrium states. Together, we find that the algorithm must terminate in finite time. \square

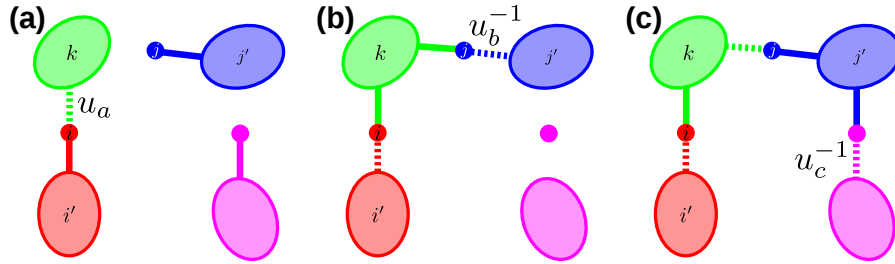


FIG. S1. Illustration of the main argument. Starting from an initial state \mathcal{I} , we cannot find a chain of updates that leaves us in a state $\mathcal{I} - 1$ where agent i buys from k but wants to switch back to i' . (a) Clearly, a single update is not enough, as i would not want to switch to k . (b) Using another agent j to enable the switch of i to k leaves j buying at k , the reverse update required to reach state $\mathcal{I} - 1$ is not possible. (c) Following the same logic, a chain with three (or more) updates will also not be able to reach state $\mathcal{I} - 1$ (see text).

A. When local information is sufficient

With the above simplification of non-increasing specific production costs, we still consider the individual optimization problem of node i of finding the best i^* to minimize

$$K_i(i^*) = D_i p_{i^*}(S_{i^*}^i + D_i) + p_T D_i \sum_{e \in \Pi(i^*, i)} t_e, \quad (8)$$

in principal considering every node in the network as a possible supplier. We now show that in certain cases we can further simplify this optimization by only considering the local neighborhood of agent i to find its best supplier i^* .

Lemma 3. *Given the optimization problem defined above with non-increasing specific production costs $p_k(S_k)$ and homogeneous demand $D_i = D$ for all agents $i \in \{1, \dots, N\}$. If all other agents are in equilibrium, then either i is its own optimal supplier or there exists a neighbor j of i such that the optimal supplier j^* of j is also an optimal supplier of i .*

Proof. Obviously, if $i^* = i$, we are done.

Otherwise, consider the external optimal supplier i^* of i . The path from i^* to i will pass through a neighbor j of i , such that

$$\Pi(i^*, i) = (e_{i^* k_1}, e_{k_1 k_2}, \dots, e_{j, i}) \quad (9)$$

$$T_{i^*, i} = t_{i^* k_1} + t_{k_1 k_2} + \dots + t_{j, i}. \quad (10)$$

We give a proof by contradiction, assuming that j^* is not an optimal supplier of i , i.e., $j^* \neq i^*$ and $K_i(j^*) > K_i(i^*)$. Using the fact that j^* is the optimal supplier of j we obtain the inequality

$$\frac{K_j(j^*)}{D} = p_{j^*}(S_{j^*}^j + D) + p_T \sum_{e \in \Pi(j^*, j)} t_e \quad (11)$$

$$\leq p_{i^*}(S_{i^*}^i + D) + p_T \sum_{e \in \Pi(i^*, j)} t_e = \frac{K_j(i^*)}{D}. \quad (12)$$

Since j is in equilibrium we have

$$\begin{aligned} S_{j^*}^j + D &= S_{j^*} \\ S_{i^*}^j + D &= S_{i^*} + D, \end{aligned} \quad (13)$$

and for agent i we have also

$$\begin{aligned} S_{j^*} &\leq S_{j^*}^i + D \leq S_{j^*} + D \\ S_{i^*} &\leq S_{i^*}^i + D \leq S_{i^*} + D, \end{aligned} \quad (14)$$

where $S_{j^*}^j$ again denotes the production of j^* ignoring purchases of agent j and the second set of inequalities simply states that i may currently be buying from any node, possibly even i^* or j^* . Using these observations together with the fact that all p_k are non-increasing we find

$$\begin{aligned} &-p_{i^*}(S_{i^*}^i + D) + p_{i^*}(S_{i^*}^j + D) - p_{j^*}(S_{j^*}^j + D) + p_{j^*}(S_{j^*}^i + D) \\ &= \underbrace{-p_{i^*}(S_{i^*}^i + D) + p_{i^*}(S_{i^*}^j + D)}_{\leq 0} - \underbrace{p_{j^*}(S_{j^*}^j + D) + p_{j^*}(S_{j^*}^i + D)}_{\leq 0} \\ &\leq 0. \end{aligned} \quad (15)$$

We can combine the above inequalities to obtain

$$\begin{aligned}
\frac{K_i(j^*)}{D} &= p_{j^*} (S_{j^*}^i + D) + p_T \sum_{e \in \Pi(j^*, i)} t_e \\
&\leq p_{j^*} (S_{j^*}^i + D) + p_T \sum_{e \in \Pi(j^*, j)} t_e + p_T t_{ji} \\
&= p_{j^*} (S_{j^*}^j + D) + p_T \sum_{e \in \Pi(j^*, j)} t_e - p_{j^*} (S_{j^*}^j + D) + p_{j^*} (S_{j^*}^i + D) + p_T t_{ji} \\
&\leq p_{i^*} (S_{i^*}^j + D) + p_T \sum_{e \in \Pi(i^*, j)} t_e - p_{j^*} (S_{j^*}^j + D) + p_{j^*} (S_{j^*}^i + D) + p_T t_{ji} \\
&= p_{i^*} (S_{i^*}^i + D) + p_T \sum_{e \in \Pi(i^*, j)} t_e + p_T t_{ji} \\
&\quad - p_{i^*} (S_{i^*}^i + D) + p_{i^*} (S_{i^*}^j + D) - p_{j^*} (S_{j^*}^j + D) + p_{j^*} (S_{j^*}^i + D) \\
&\leq p_{i^*} (S_{i^*}^i + D) + p_T \sum_{e \in \Pi(i^*, i)} t_e \\
&= \frac{K_i(i^*)}{D},
\end{aligned} \tag{16}$$

where the first inequality is due to j^* being an optimal supplier for j and the second inequality follows from Eq. (15).

Hence, we know that

$$K_i(j^*) \leq K_i(i^*), \tag{17}$$

contradicting our original assumption. Therefore j^* has to be an optimal supplier of i which concludes the proof. \square

We note that this optimal supplier $i^* = j^*$ does not necessarily have to be unique, it is possible that there are multiple optimal suppliers with identical (minimal) costs.

We briefly summarize the results so far: if we have non-increasing specific production costs and sufficiently homogeneous demand [Eq. (15)], we can in principle solve the optimization problem of a single agent i by considering only its local, direct neighborhood. In particular, this condition is fulfilled under the stronger assumption of identical demand $D_i = D_j = D$ for all agents. However, we assumed a network in (almost) equilibrium, specifically we assumed that j is already buying at j^* . If this is not the case, it remains to be shown that we can order simultaneous updates in such a way that we perform only local updates and still always find the optimal supplier for each agent.

Corollary 3.1. *We call the set $\{i_1, i_2, \dots\}$ of agents with i^* as their optimal (and current) supplier the cluster $C(i^*)$ of i^* . In a network in equilibrium the individual clusters are connected: for every node $i \in C(i^*)$ there is a shortest (cheapest) path $\Pi(i^*, i) = (i^*, j_1, j_2, \dots, j_n, i)$ from i^* to i such that all nodes $j_1, j_2, \dots, j_n \in C(i^*)$.*

B. Local percolation algorithm

Above we showed that we can always find an optimal supplier for an agent i , given the other agents are in equilibrium, by considering its neighbors when we have non-increasing specific production costs and homogeneous demand $D_i = D$ for all agents $i \in \{1, \dots, N\}$. We use this result to derive a simplified, local percolation algorithm to solve this optimization problem in an efficient way. Since individual purchase decisions can lead to larger changes in the network, we still need to consider updates in a network out of equilibrium. Here, one can easily construct examples where there exists an agent that cannot find its (currently) optimal supplier locally. In these cases, however, other agents will also want to update their purchases and at least one agent can find its optimal supplier locally. Below, we introduce the algorithm and the ordering of the local updates such that we always update agents who can find their optimal supplier locally. We show that this defines a *local* percolation rule taking the network from *any* state into an equilibrium with only local updates to find optimal suppliers.

Algorithm 2 Local percolation algorithm

```

1: repeat
2:    $\mathcal{U} \leftarrow \emptyset$ .
3:   for all  $i \in \{1, \dots, N\}$  do
4:     for all  $j \in \text{Neighborhood}(i)$  do
5:       Calculate  $K_i(j')$  and  $K_i(i')$ , where currently,  $S_{i'i} = D_i$  and  $S_{j'j} = D_j$ 
6:       if  $K_i(j') < K_i(i')$  where currently  $S_{i'i} = D_i$  then
7:          $\mathcal{U} \leftarrow \mathcal{U} \cup \{K_i(j')\}$ 
8:       end if
9:     end for
10:    Calculate  $K_i(i)$ 
11:    if  $K_i(i) < K_i(i')$  then
12:       $\mathcal{U} \leftarrow \mathcal{U} \cup \{K_i(i)\}$ 
13:    end if
14:  end for
15:  if  $\mathcal{U} \neq \emptyset$  then
16:    Calculate  $\{i, i^*\}$  such that  $K_i(i^*) = \min_{\mathcal{U}} K_i(k)$ 
17:     $S_{i^*i} \leftarrow D_i$ 
18:     $S_{i'i} \leftarrow 0$ 
19:  end if
20: until  $\mathcal{U} = \emptyset$ 

```

This algorithm is a local version of the general algorithm, where we check for every agent i only the suppliers j' of its neighbors j and itself as possible new suppliers. As before we then execute the update with the smallest final cost and repeat the process until no further update is possible.

Theorem 4. *Consider the optimization problem defined above with non-increasing specific production cost $p_k(S_k)$ and homogeneous demand $D_i = D$ for all agents $i \in \{1, \dots, N\}$. Algorithm 2 only executes updates finding the optimal supplier for an agent and terminates in finite time in an equilibrium state, where no agent can reduce its costs further.*

We now proof this theorem with two supporting Lemmas:

Lemma 5. *Given algorithm 2, for every $K_i(k) \in \mathcal{U}$, either $k = i^*$ or there is $K_j(l) \in \mathcal{U}$ with $K_j(l) < K_i(k)$.*

Proof. If $k = i^*$ we are done.

Otherwise, assuming that $k \neq i^*$, then also $i' \neq i^*$ since $K_i(i^*) < K_i(k) < K_i(i')$, thus i is not already buying at i^* . Then, following the path $\Pi(i^*, i)$ backwards from i to i^* , we will find an agent j (possibly $j = i^*$ or $j = i$), where a local update $K_j(i^*)$ is possible and the current supplier $j' \neq i^*$.

We now consider this update $K_j(i^*)$. First we show that this update is in the update list by contradiction. Assume

$K_j(j') \leq K_j(i^*)$, then similar to the above proof we find

$$\begin{aligned}
\frac{K_i(j')}{D} &= p_{j'} (S_{j'}^i + D) + p_T T_{j'i} \\
&\leq p_{j'} (S_{j'}^i + D) + p_T T_{j'j} + p_T T_{ji} \\
&\leq p_{j'} (S_{j'}^i) + p_T T_{j'j} + p_T T_{ji} \\
&= K_j(j') + p_T T_{ji} \\
&\leq K_j(i^*) + p_T T_{ji} \\
&\leq p_{i^*} (S_{i^*}^i + D) + p_T T_{i^*j} + p_T T_{ji} \\
&= \frac{K_i(i^*)}{D},
\end{aligned} \tag{18}$$

leading to the contradiction $K_i(j') < K_i(i^*)$ (i^* is not the optimal supplier). Otherwise we have $K_i(j') = K_i(i^*)$ and we can simple repeat the above argument with $i^* \leftarrow j'$ (at most until we reach $j = i$). This shows that there is an update $K_j(i^*)$ on the list \mathcal{U} for some optimal supplier i^* of i and an agent j on the path $\Pi(i^*, i)$. Finally, we find

$$\begin{aligned}
\frac{K_j(i^*)}{D} &= p_{i^*} (S_{i^*}^j + D) + p_T T_{i^*j} \\
&\leq p_{i^*} (S_{i^*}^i + D) + p_T T_{i^*j} + p_T T_{ji} \\
&= p_{i^*} (S_{i^*}^i + D) + p_T T_{i^*i} \\
&= p_{i^*} (S_{i^*}^i) + p_T T_{i^*i} \\
&= \frac{K_i(i^*)}{D},
\end{aligned} \tag{19}$$

since $S_{i^*}^j = S_{i^*}^i = S_{i^*}$ since none of the two agents are buying at i^* . This means that the update $K_j(i^*)$ with $K_j(i^*) \leq K_i(i^*) < K_i(k)$ is in the update list before $K_i(k)$, which concludes the proof. \square

It then follows directly that:

Corollary 5.1. *The update on the list $K_i(k) = \min_{\mathcal{U}} K_j(l)$ is optimal with $k = i^*$.*

Lemma 6. *The network is in an equilibrium, if and only if $\mathcal{U} = \emptyset$.*

Proof. We proof this by showing that the network is not in equilibrium, if and only if $\mathcal{U} \neq \emptyset$.

Clearly, if there is a possible update on the list, the network is not in equilibrium.

Otherwise, if the network is not in equilibrium there is an agent i currently buying from $i' \neq i^*$ which is not its optimal supplier. We then have $K_i(i^*) < K_i(i')$. If there was no update on the list, then also the update $K_i(i^*)$ via i 's neighbor j is not on the list. However, this can only be because j is buying from $j' \neq i^*$. But then either

- (1) the node j' would also be an optimal supplier for i (following the same chain of inequalities as above) and the update $K_i(j')$ would be on the list
- (2) or j would also want to update to buy at i^* .

In the second case we simply repeat the argument above with $i \leftarrow j$ (at most until $i \leftarrow i^*$ itself) until we find a possible update which must be on the list. This shows that there is always an update on the list when the network is not in equilibrium and thus concludes the proof. \square

Proof of Theorem 3. By means of Corollary 4.1 we know that the update executed is always an optimum one. Hence every iteration of the algorithm will find the optimal supplier for an agent. Lemma 5 shows that, when the algorithm terminates as soon as \mathcal{U} is empty, the network is in equilibrium. The proof that the algorithm always terminates in finite time is identical to the one given above for Algorithm 1. Together these two results proof theorem 3. \square

In contrast to $\mathcal{O}(N^2)$ updates that are checked in the general algorithm, we further reduce the number of possibilities to $\mathcal{O}(M)$ local updates, where M is the number of links in the network. In sparse networks, where $\mathcal{O}(M) = \mathcal{O}(N)$, networks this significantly reduces the required operations of the algorithm.

Together, we have now shown that for non-increasing specific production cost $p_k(S_k)$ and identical demand of all agents $D_i = D_j = D$, we can *exactly* map the interacting individual optimization problems to a local percolation model when ordering simultaneous updates by their new final cost. These updates will always lead to an equilibrium state where no agent wants to change suppliers and agents never make sub-optimal updates. Note, however, that due to the individual optimization the resulting equilibrium is, in general, not the globally optimal state with minimal total cost $\sum_i K_i$ over all agents.

Here, we always use the algorithm in the following way: We start in an equilibrium state, usually the trivial equilibrium $i^* = i$ for all i at $p_T = \infty$. Then we consider a decrease (or increase) of p_T and calculate the next value p_T^{ji} for which a agent would change its supplier (i.e., when $\mathcal{U} \neq \emptyset$). We then execute the update and iterate until no further updates occur at the current value of p_T . As we have shown that the algorithm always makes correct updates, it is also possible to study other initial conditions, reactions to changes in the network structure or sudden large changes of p_T .

Finally, in the last part of this section we briefly give small examples illustrating some assumptions in the proofs above such as homogeneous demand, before describing the actual implementation in the next section.

C. Examples

In order to illustrate the model, the proofs given above and especially the assumptions going into the proofs, we give two examples of simple systems. For all examples we will use the same cost per unit functions as in the main manuscript $p_k(S_k) = b_k - a \cdot S_k$ with $a = 1$.

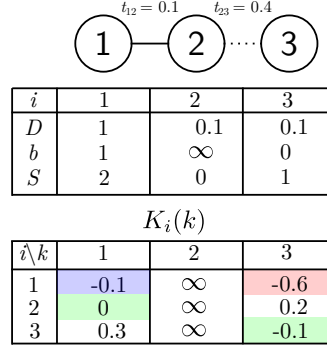


FIG. S2. Example for non-homogeneous demand at $p_T = 1$. The tables describe the model parameters and the costs if agent i would buy (is buying) at node k . Current suppliers are marked in green, if they are optimal, otherwise they are marked in blue. Optimal decisions are marked in yellow or red, depending on whether they can/cannot be found by a local update, respectively. Here, agent 1 could buy optimally from node 3 but cannot find out about it via a local update (marked in red), since agent 2 is buying from 1. Conversely, agent 2 and 3 are already buying optimally at node 1 and 3, respectively (marked in green). Therefore the only possible update is 1 buying from 3 but it cannot be found with a local update. This illustrates the requirement of homogeneous demand for the local percolation algorithm.

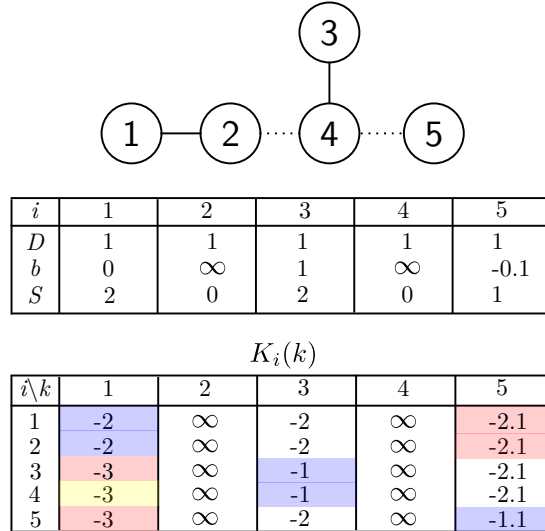


FIG. S3. Example for homogeneous demand at $p_T = 0$ and the importance of update order. The tables describe the model parameters and the costs if agent i would buy (is buying) at node k . Current suppliers are marked in green, if they are optimal, otherwise they are marked in blue. Optimal decisions are marked in yellow or red, depending on whether they can/cannot be found by a local update, respectively. Here, all agents want to update their current purchases (marked in blue), but only agent 4 can find its optimal supplier locally through its neighbor (marked in yellow). Of all possible local updates, this update has the smallest $K_i(k)$ and would be executed first, allowing other agents to follow and find their optimal supplier locally. As shown above, with homogeneous demand if the network is not in equilibrium we will always find an optimal local update by sorting for the smallest $K_i(k)$.

II. SIMULATION

With the algorithm introduced above we can solve the the individual optimization problems with only local updates for a given value of p_T when the specific production costs are non-increasing and the demand of all agents is identical. We now discuss how we exactly track the evolution of the system for the whole range of decreasing transaction costs p_T (increasing transaction costs are handled analogously). First, we note that it is not necessary to calculate the transaction costs per unit T_{ki} for all shortest (cheapest) paths, since these paths will be found automatically during the optimization.

We track for each agent i its current supplier i' and the current transaction costs per unit $T_{i'i}$, starting from the trivial initial condition $i' = i$ for all agents at $p_T = \infty$ ($T_{ii} = 0$). For all (directed) links we then explicitly calculate the values p_T^{ji} where the link would be active, i.e., when i starts buying via j from j' , where j' is the supplier of i 's neighbor j . As in the algorithm above, we also explicitly consider returning to internal production as a special case. Comparing the cost of the current supplier i' and the new supplier j' :

$$K_i(i') = D \cdot p_{i'} [S_{i'}^i + D] + D \cdot p_T \cdot T_{i'i}, \quad (20)$$

$$K_i(j' \text{ via } j) = D \cdot p_{j'} [S_{j'}^i + D] + D \cdot p_T \cdot (T_{j'j} + t_{ji}), \quad (21)$$

we find that i would prefer to buy from j' when

$$K_i(j' \text{ via } j) < K_i(i') \\ p_T \cdot \underbrace{(t_{ji} + T_{j'j} - T_{i'i})}_{\Delta T} < \underbrace{p_{i'} [S_{i'}^i + D] - p_{j'} [S_{j'}^i + D]}_{\Delta p}. \quad (22)$$

With this inequality we can easily calculate p_T^{ji} . We create the sorted list $\mathcal{U} = \left\{ \left(p_T^{ji}, K_i(j'), e_{ji} \right), \dots \right\}$, ordering all links in descending order in p_T^{ji} . If multiple links have the same p_T^{ji} , e.g., for simultaneous updates during cascades, we use the new final costs $K_i(j')$ as a secondary criterion for ordering the updates (increasing in $K_i(j')$), as discussed above. We then have to distinguish three different cases to determine p_T^{ji} :

- $\Delta T > 0$
in this case we simply have the condition $p_T < \frac{\Delta p}{\Delta T}$. We thus set $p_T^{ji} = \frac{\Delta p}{\Delta T}$ if the condition is not yet fulfilled and $p_T^{ji} = p_T$ otherwise. To calculate $K_i(j')$ for the secondary ordering we use the transaction costs p_T^{ji} , as this is the value when the ordering will be relevant. If no other changes occur and this link is not updated again, this secondary ordering is then already accurate when p_T decreases far enough (and will be updated in between otherwise).
- $\Delta T = 0$
in this case the switch does not depend on the transaction costs at all but the condition is $\Delta p > 0$. If this condition is fulfilled, we set $p_T^{ji} = p_T$, such that the switch will be applied immediately. As above, the secondary ordering is given by the final cost. Otherwise we set $p_T^{ji} = -1$ and the switch is never executed.
- $\Delta T < 0$
in this case the condition becomes $p_T > -\frac{\Delta p}{|\Delta T|}$ (note the change in the sign of the inequality). In case the condition is true, the change is applied instantly and we set $p_T^{ji} = p_T$, otherwise we set $p_T^{ji} = -1$. This case will mostly be relevant when considering increasing transaction cost.

We then use this list \mathcal{U} to iteratively execute single updates. The link with the largest p_T^{ji} (and smallest final total cost $K_i(j')$, if applicable) will become active first, meaning i will start buying via j from j 's supplier j' . The algorithm then follows an event-based methodology:

Starting from $p_T = \infty$ and only internal production $i^* = i$ for all nodes $i \in \{1, \dots, N\}$, we first decrease the value of p_T to the value of the next update $p_T \leftarrow \max_{\mathcal{U}} p_T^{ji}$. Then we update $i' \leftarrow j'$ and $T_{i'i} \leftarrow T_{j'j} + t_{ji}$. We also adjust the production of the involved suppliers i' and j' accordingly. Due to the scale effects in production these changes will affect all p_T^{kl} of all links adjacent to agents currently buying from i' or j' , for which we recalculate p_T^{kl} .

We then repeat the updating process working through all links with $p_T^{ji} = p_T$ and keep updating the purchases until no further changes occur and the next entry in the update list occurs for $p_T^{ji} < p_T$. Finally, we repeat this whole

process by further decreasing p_T . In contrast to the algorithm for a single update step given above, we will never have $\mathcal{U} = \emptyset$. Instead, we terminate once all $p_T^{ji} \leq 0$ and the transaction cost cannot be reduced further.

The final state of the network will then be given at $p_T = 0$, where generally all agents have the same supplier and form one large cluster. Note that this updating scheme models a very slow decrease of the transaction costs. In principle it is also possible to include sudden jumps in p_T . Due to the individual optimization, this generally leads to a different evolution of the clusters (see Sec. VI for examples).

We then reverse the process and consider increasing transaction cost in the same way by adjusting the conditions for p_T^{ji} and reversing the ordering (executing updates with the smallest p_T^{ji} first). In this case we terminate when $p_T^{ji} = \infty$ for all possible updates. Overall, this allows us to exactly track the evolution of all clusters in the network for both increasing as well as decreasing transaction costs.

III. NETWORK MODELS

A. Spatially embedded random networks

We create a spatially embedded random network with N nodes by distributing N points uniformly at random in the unit square. The links of the network are then created using the Delaunay-triangulation of these points. Each link e is assigned a transaction cost t_e equal to the euclidean distance between its two endpoints. We further assume an identical demand $D_i = 1/N$ for all nodes.

In order to reasonably study different system sizes (see Supplementary Information Sec. III for a detailed discussion) we define the parameters b_i of the cost functions as values of a random function on the unit square. The function is generated by spectral synthesis of 1024^2 Fourier-modes with Gaussian distributed amplitudes with mean 0 and expected power spectrum $(\omega_x^2 + \omega_y^2)^{-2}$ [1]. This function is then interpolated at the positions of the nodes and finally scaled such that $b_i \in [0, 1]$ with $\min_i(b_i) = 0$ and $\max_i(b_i) = 1$.

While this technically allows negative specific production costs when including the scale effects, a constant shift $b_k \rightarrow b_k + \text{const.}$ for all nodes $k \in \{1, \dots, N\}$ does not change the solution of the minimization problem. Therefore, for simplicity we only consider normalized $b_k \in [0, 1]$ in all examples.

B. Small world networks

To illustrate the effects of the underlying network topology we consider Watts-Strogatz small-world networks [2]. This model allows generating networks that interpolate between regular and random topologies with a large or small network diameter, respectively: Start from a ring of N nodes where each node is connected to its $k/2$ nearest neighbors on each side. Each link in the network is then rewired with probability q_{rew} , i.e., disconnected on one end and connected to another node in the network, chosen uniformly at random. This procedure interpolates between regular networks with a large diameter for $q_{\text{rew}} \rightarrow 0$ and completely random networks with a small diameter for $q_{\text{rew}} \rightarrow 1$.

For the parameters of our model we assume identical transaction costs for all links $t_e = 1/N$ as well as identical demand $D_i = 1/N$ for all nodes. The cost function parameters are taken as $a = 1$ and uniformly randomly chosen $b_i \in [0, 1]$ (see also Supplementary Information Sec. IV).

IV. EXPANSION VS. FINE-GRAINING

Usually, in percolation models the thermodynamic limit of large systems $N \rightarrow \infty$ is relevant for studying the transition to global connectedness. We will now motivate our choice to consider an increased system size N as a fine-graining of the system.

Apart from the obvious inconsistencies with the models motivation arising for infinitely expanding systems, e.g., an infinite system size would also mean infinite demand, other problems arise due to the interplay between scale effects and transaction costs. To illustrate this, we assume finite, bounded intervals for all parameters. Specifically we take the specific production cost to be $p_k(S) = b_k - aS(k)$ with $b_k \in [b_{\min}, b_{\max}]$ and similarly, $D_i = D$ and $t_{ij} \in [t_{\min}, t_{\max}]$ with $t_{\min} > 0$. For simplicity, we consider the model without scale effects in the limiting case $a = 0$.

For a given value of the transaction cost p_T we will now derive a (very loose) upper bound for the size of a cluster. Consider the longest shortest path $\Pi_{ki, \max}$ in one cluster between agent i and its supplier i^* . The transaction cost along those paths cannot be higher than the difference between the specific production costs of the two nodes:

$$\begin{aligned} b_{i^*} + p_T T_{i^*i} &\leq b_i \\ \Rightarrow p_T T_{i^*i} &\leq b_i - b_{i^*} \leq b_{\max} - b_{\min}, \end{aligned}$$

which means that there are at most d_{\max} links involved in the path, where

$$d_{\max} \leq \frac{T_{i^*i}}{t_{\min}} \leq \frac{b_{\max} - b_{\min}}{p_T t_{\min}}. \quad (23)$$

If the underlying network has a maximum degree k_{\max} , we can then find an upper bound for number of nodes in the largest cluster as

$$C_1(p_T) \leq 1 + k_{\max} + k_{\max}^2 + \dots + k_{\max}^{d_{\max}} = \sum_{d=0}^{d_{\max}} k_{\max}^d < \infty \quad (24)$$

for any $p_T > 0$. Thus for a given range of parameters as assumed above in any network with bounded degrees we find that in the limit of $N \rightarrow \infty$ the size of the largest cluster will be finite for any $p_T > 0$. Thus the fraction of nodes in the largest cluster will follow

$$C_1/N = \begin{cases} 0 & \text{for } p_T > 0 \\ 1 & \text{for } p_T = 0, \end{cases} \quad (25)$$

describing a trivial transition at $p_T = 0$. More importantly, this even holds in cases where, for example, $k_{\max} \sim \ln(N)$, such as simple random graphs. Similarly, for $a > 0$ the infinite demand will (in most cases) cause an infinite hysteresis loop for the system in the limit $N \rightarrow \infty$.

While these are interesting results, the more reasonable assumption of fine-graining the system maintains the balance between scale effects and transaction cost such that the transitions and all related phenomena can be observed in a finite, non-trivial parameter interval and results for different system sizes can be compared more easily. Of course this requires the correct scaling of transaction costs, demand and network structure when increasing the system size. For two-dimensional, spatially embedded networks, for example, this is simple. However, for networks without an underlying geometry this can be more difficult and often arbitrary, as we discuss below for small world networks.

V. SMALL WORLD NETWORKS

In the main manuscript we studied the model on an underlying small world network [3]: Start from a ring of N nodes where each node is connected to its $k/2$ nearest neighbors in each direction. Each link in the network is then rewired with probability q_{rew} , i.e., disconnected on one end and connected to another node in the network chosen uniformly at random.

In the main manuscript we showed for systems of size $N = 10^4$ how the topology changes the abruptness of the transition and the number of large clusters emerging. For completeness, Fig. S4 shows the (inverse) size of the largest gap for different scale effects a for the same system parameters ($N = 10^4$, $t_e = 1/N$, $D_i = 1/N$, b_i uniformly random in $[0, 1]$). A network with large diameter ($q_{\text{rew}} \rightarrow 0$) allows different clusters to emerge in remote parts of the network since the distances are large. Hence the jumps of the size of the largest cluster, denoted by ΔC_1 , are typically smaller ($1 - \langle \max \Delta C_1 / N \rangle$ is large). However, very strong scale effects ($a > 1$) can compensate for the large distances. Small fluctuations are amplified and one cluster will grow in a single cascade even in the network with large diameter. In contrast, for small diameters ($q_{\text{rew}} = 0.01$) the jump ΔC_1 grows already for moderate values of a .

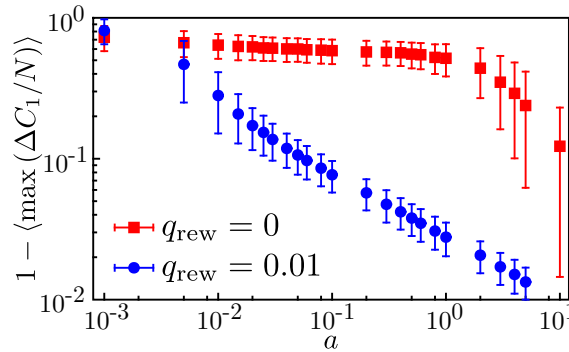


FIG. S4. Large scale effects lead to large cascades, independent of network topology. The inverse size of the largest gap in the size of the largest cluster ($1 - \text{largest gap}$) depending on the size of the scale effects a . For very small scale effects the cascades are small (the inverse is large), independent of the network topology a single agent does not have a large enough effect on the production costs. As the scale effects become larger, the average size of the largest cascade increases quickly in small world networks; due to the shortcuts a large cluster can affect all parts of the network. For regular networks with a large diameter ($q_{\text{rew}} \rightarrow 0$) cascades stay small as transaction costs are still comparatively large. Unsurprisingly, for very large scale effects a single agent will change the production cost enough that the largest cluster grows discontinuously in a single cascade, even if the distances in the network are large.

Additionally, we want to highlight some difficulties with this network structure in our model. As discussed above, to avoid trivial behavior in the limit $N \rightarrow \infty$, we need to scale all parameters, such that an increased system size corresponds to a fine graining of the system. However, in this case choosing a scaling of all variables to effectively fine-grain the system with increasing system size is not possible across different rewiring probabilities: the required scaling for a regular network will not be appropriate for a random network.

Similarly, the scaling of k and the choice of the b_i over different system sizes is not uniquely defined. The behavior of the model in the limit of large systems $N \rightarrow \infty$ thus depends on the choice of the other parameters of the system. For example, if we choose b_i uniformly at random without correlations between neighboring nodes and keep the cost per link constant, the limiting behavior will be a single cascade for very small p_T where a microscopic cluster grows discontinuously to span the whole network (similar to the discussion above).

VI. ANALYTICALLY SOLVABLE EXAMPLE

In addition to the complete graph discussed in the main manuscript we present another analytically solvable example. We consider a square lattice with a well defined single source and the same affine linear production costs per unit as in the main manuscript. The qualitative results are similar to those obtained for the complete graph, however, the actual evolution is more similar to the results obtained for the random spatial network.

We place $N = L^2$ nodes as a square grid in the unit square. We index the nodes as node (x, y) at position $(x/L, y/L)$, $x, y \in \{1, 2, \dots, L\}$. As in all other cases we have a homogeneous demand distribution $D_{(x,y)} = 1/N$. The production cost parameter $b_{(x,y)}$ at each node is given as $b_{(x,y)} = 1$ for all nodes, except for the node in the middle of the grid, where $b_{(\lfloor L/2 \rfloor, \lfloor L/2 \rfloor)} = 0$ (see Fig. S5). Increasing the number of nodes in this system is equivalent to a fine-graining of the unit square.

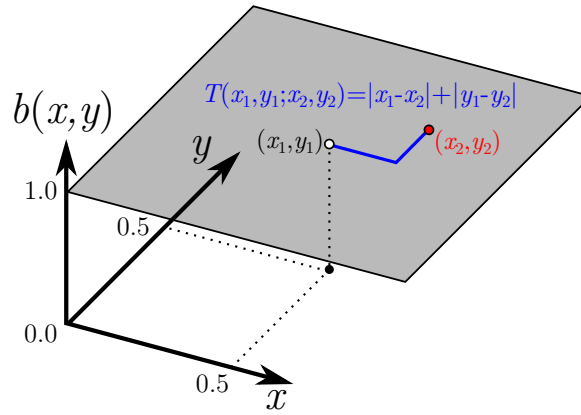


FIG. S5. Continuum illustration of the single source square lattice model. We have $b_{(x,y)} = 1$ for all points except for the center point in the lattice at $(1/2, 1/2)$. The distance between two points $T_{x_1, y_1; x_2, y_2}$ is given by the 1-norm, corresponding to the shortest path in the square lattice. This continuum form of the model can easily be solved analytically while the corresponding lattice model can be simulated efficiently.

Due to the symmetry and the single source only one cluster will emerge and we can easily calculate its size. For simplicity we consider the continuous version of the problem illustrated in Fig. S5, i.e., the limit $N \rightarrow \infty$. Due to the symmetry of the problem, all agents at the same distance $\text{dist}[(x, y); (1/2, 1/2)] = |x - 1/2| + |y - 1/2|$ to the central supplier will behave identically. Note that distances are given as the 1-norm due to the underlying square grid network structure. Thus, we can derive a self-consistency equation, describing the size of the cluster in terms of the maximum distance d of an agent still willing to buy from the central supplier. For $d < 1/2$ the shape of a cluster is a diamond (due to the 1-norm), but for $d > 1/2$ the corners are cut off, as we only consider points in the unit square. For given scale effects a and transaction costs p_T we determine d as the value where the cost of internal production is equal to the cost of buying from the central supplier:

$$-a(2d^2) + p_T d = 1 \quad d \leq 1/2 \quad (26)$$

$$-a[2d^2 - 4(d - 1/2)^2] + p_T d = 1 \quad d > 1/2, \quad (27)$$

where the first term describes the scale effects due to the size of the cluster and the second term is the transaction cost.

Solving these equations for d and keeping in mind that $0 \leq d(p_T) \leq 1$ yields

$$d(p_T) = \begin{cases} \frac{p_T \pm \sqrt{p_T^2 - 8a}}{4a} & \text{for } d \leq 1/2 \\ \frac{4a - p_T \pm \sqrt{8a^2 + 8a - 8ap_T + p_T^2}}{4a} & \text{for } d > 1/2. \end{cases} \quad (28)$$

This equation has no, one or two possible solutions for d for each p_T , however, each expression is valid only for some values of d . Therefore, as long as the bifurcation in the $d \leq 1/2$ (lower branch) solution happens for $d > 1/2$, only

one valid solution exists. We calculate the position of the bifurcation as $p_T^{\text{crit},1} = \sqrt{8a}$ and the corresponding distance $d_{\text{crit},1} = 1/\sqrt{2a}$. Thus, as long as $a < 2$ there is only one continuous solution for $d < 1/2$. We find a similar result for the solution of the $d > 1/2$ branch. Together, for $a < 2$ the two solutions join continuously and describe a single, continuous solution for $d(p_T)$. As the size of the largest cluster is a continuous function of d , these results apply equally to C_1 . Together, we find that the evolution of the largest cluster is continuous for $a < 2$, the slope of C_1 becomes infinite for $a = 2$ [see Fig. S6(a)-(c)].

For $a > 2$, however, the first solution for $d < 1/2$ disappears at $p_T^{\text{crit},1} = \sqrt{8a}$ with $d_{\text{crit},1} = 1/\sqrt{2a} < 1/2$. There is no solution for smaller values of $p_T < p_T^{\text{crit},1}$ and two valid solutions for larger values. Similarly, the solution for the $d > 1/2$ branch does not exist for $p_T > p_T^{\text{crit},2} = 4a - \sqrt{8a(a-1)}$ with $d_{\text{crit},2} > 1/2$, but has two valid solutions for smaller values of p_T . The two solutions are still joined continuously, but now by an unstable branch: a slightly smaller cluster will rapidly shrink to a small stable size, while a slightly larger cluster will grow to the larger stable size. The transition for decreasing/increasing transaction costs becomes discontinuous and a hysteresis loop emerges in between [see Fig. S6(d)].

We illustrate these analytical results in Fig. S6 together with simulations for $N = 25600$ for different values of a , showing good agreement with our analytical calculations. Contrary to the complete graph example in the main manuscript, the cluster grows by a finite amount already before the discontinuous transition takes place (similarly for increasing transaction costs). This is explained by the fact that close to the source in the middle of the grid, the difference in production cost per unit is constant while the transaction costs disappear for agents that are arbitrarily close to the center. Since the growth of the cluster is driven by these differences, the cluster grows as soon as $p_T < \infty$ whereas transaction costs (per link) are constant in the complete graph model and thus the cluster can only grow once p_T decreases below a fixed, finite value.

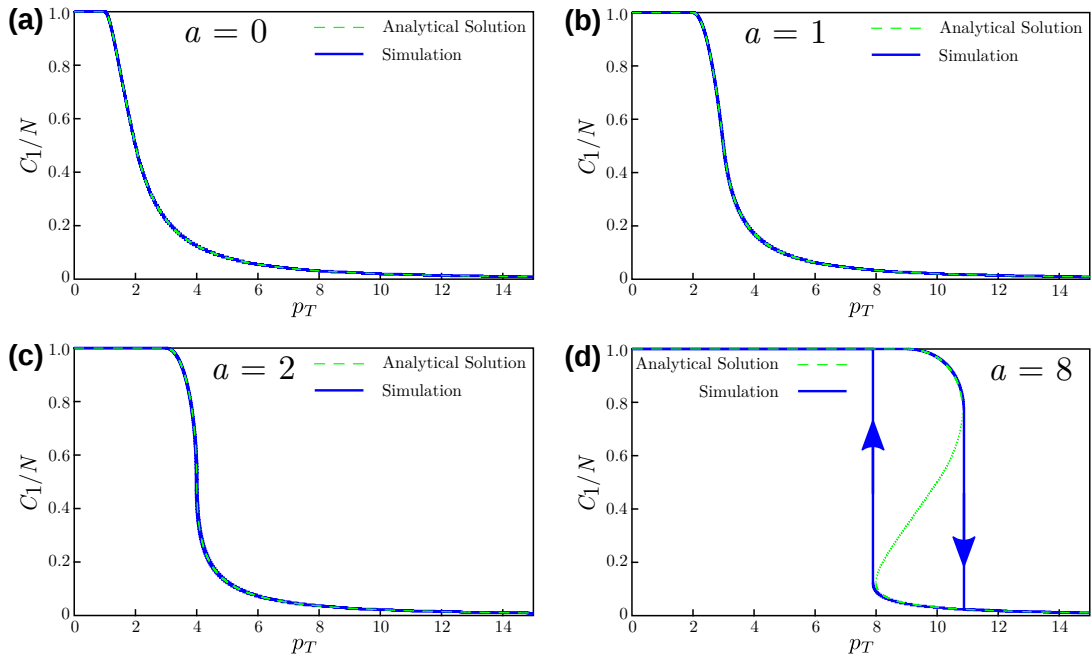


FIG. S6. Simulation of the single source square lattice model for $L = 160$ ($N = 25600$) and different a . (a),(b) The transition is continuous for small a , becoming steeper as a increases. (c) At $a = 2$ the slope becomes infinite. (d) For larger a the transition becomes discontinuous with a hysteresis loop. In the continuum self-consistency equation this structure emerges from a bifurcation of the possible solutions (dashed green lines) with an unstable branch in the middle (dotted line).

VII. WORLD TRANSPORT NETWORK

In the following we describe the data used to create the world transport network. The nodes of the network describe individual countries or smaller regions of large countries (China, Russia, USA, Canada). The location of each node is given as the centroid of the respective country or the largest city in the corresponding region. For each of these countries we collected the total gross domestic product (GDP) and the total population as additional parameters. Further, we assign a harbor to each country with access to the sea, using the largest harbor in the country/region. The full world map with the location of all nodes is shown in Fig. S7. See the Supplementary Data for all data and sources.

Transport links exist via land between countries sharing a border. The transaction cost of these links is simply given as the geodesic distance between the connected nodes. Transport links via sea exist between all countries/regions with a harbor. Here, the transaction cost is given as the sum of the distance of the corresponding node to the respective harbor for both countries and the length of the actual sea route, $t_{e_{ij}} = t_i^{\text{harbor}} + t_j^{\text{harbor}} + T_S t_{\text{sea}}$, where t_i^{harbor} is the distance of the node corresponding to country i to its largest harbor. These sea routes were determined as shortest paths across a triangulation of the world's oceans, explicitly including important routes and channels (for example the Panama and Suez channel). In order to change the preferred mode of transport we introduce the factor T_S describing the relative cost of sea travel compared to land travel. To determine the transaction cost for sea links we scale the length of the sea route with this factor, such that $T_S < 1$ effectively means that sea travel is preferred, while for $T_S > 1$ land travel is comparatively cheaper.

To determine the parameters for our model we used the population and GDP of each country: $b_i = b = \text{const.}$, $a = 1$ and $D_i \sim 1.15P_i + G_i$, where P_i and G_i are the population and GDP of the corresponding region relative to the world total. The demand is then scaled such that, as in the other examples, the total demand is $\sum_i D_i = 1$. Even though the demand is not identical for all nodes, we use the local percolation algorithm to solve the optimization problem. Differences to the exact solution are negligible.

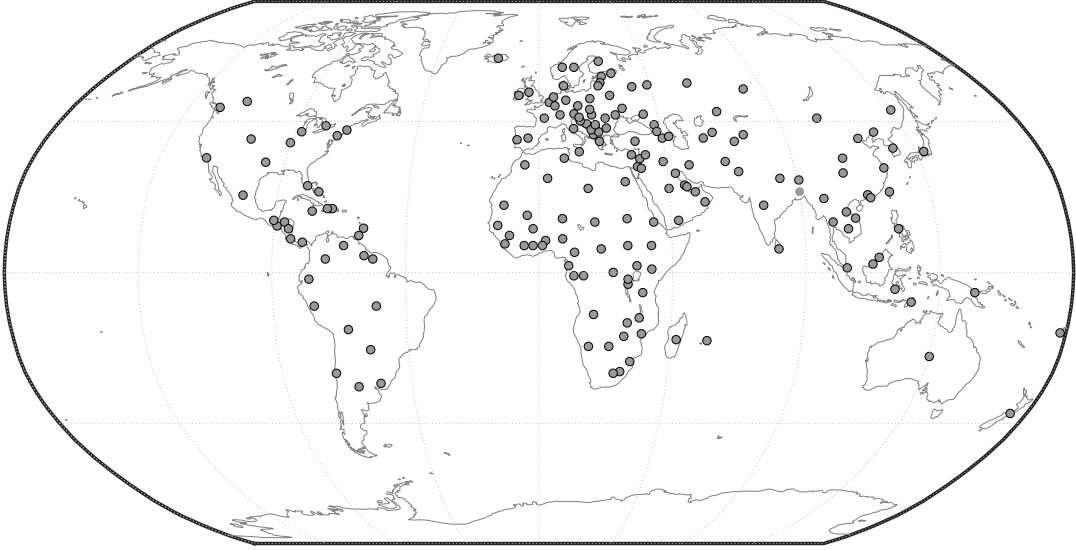


FIG. S7. Location of all nodes for the world transport network. Possible transport links (not shown) exist between countries sharing a border (via land) and between all countries with access to a harbor (via sea).

In the simulations shown in the main manuscript we defined the demand of each agent as $D_i \sim 1.15P(i) + G(i)$, where $P(i)$ and $G(i)$ are the population and GDP of the region relative to the world total. We then applied the local percolation algorithm (algorithm 2) to solve the optimization problem, even though the demand is not homogeneous (as required in the proof given above). As discussed above, this means the algorithm does not necessarily make an optimal update each step and the solution using the exact algorithm may be different. For comparison, we show the exact solution (algorithm 1) in Fig. S8. The qualitative and quantitative behavior is almost identical to the results shown in the main manuscript.

Note, that this similarity is not a-priori obvious. Especially, differences will be larger in networks with a more heterogeneous demand distribution. Additionally, the structure of the network will affect the results: in this case, the differences are smaller when transport via sea is cheap. The shortest paths are the direct connections between nodes via sea such that the local percolation algorithm checks almost all nodes regardless. Results are more likely to differ for very sparse networks with a large diameter.

We also illustrate here the impact of large changes in the transaction costs. In the main manuscript and all previous discussions we assumed a slow, gradual decrease/increase of the transaction cost, but the algorithms also allow to simulate arbitrary changes. In Fig. S9 we compare an instant decrease from very large transaction costs to a given value with the results obtained for a gradual decrease. Different suppliers emerge in both cases. More importantly, in the case of instantaneous reduction of transaction cost, the center of the emerging cluster changes, depending on the final value of the transaction costs, in a different way compared to the results for the gradual decrease. This further illustrates the impact of history on the evolution of the model.

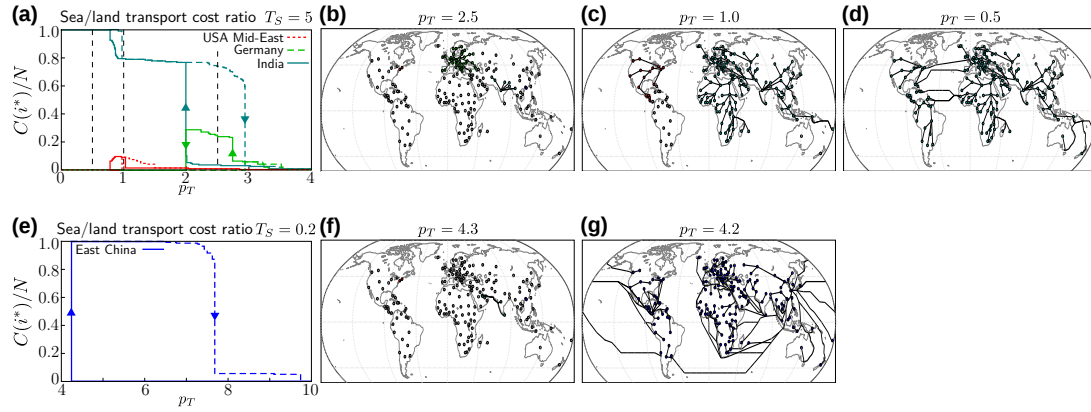


FIG. S8. Exact solution of the optimization model in a global transportation network. (a) Evolution of the connected components of individual suppliers for large sea travel costs ($T_S = 5$). (b)-(d) Evolution of the individual clusters for different values of the transaction cost parameter p_T , land routes are preferred to transaction via sea. (e) Evolution of the connected components of individual suppliers for small sea travel costs ($T_S = 0.2$). (f),(g) State of the network immediately before and after the cascade. The results are almost identical to those obtained with the local percolation algorithm (see main manuscript). The only difference is a small shift of p_T for some of the transitions.

-
- [1] D. Saupe, in *The science of fractal images* (Springer, New York, 1988) pp. 71–136.
 - [2] D. J. Watts and S. H. Strogatz, *Nature*, **393**, 440 (1998).
 - [3] D. Watts and S. Strogatz, *Nature*, **393**, 440 (1998).

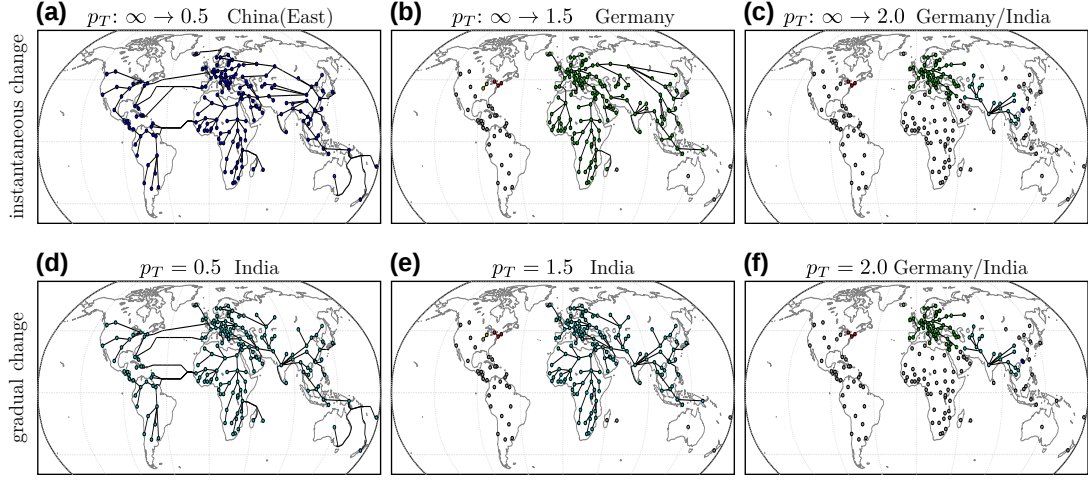


FIG. S9. Comparison of instantaneous and gradual decrease of the transaction cost. (a)-(c) Resulting network, when the transaction cost p_T are reduced instantaneously to a given value (parameters are identical to other simulations, large sea travel costs $T_S = 5$). (d)-(f) Resulting network, when the transaction cost p_T are reduced gradually. Depending on the final value of the transaction cost, different nodes emerge as suppliers of large clusters. Interestingly, if transaction costs are decreased far enough, the resulting network differs from the one obtained by gradual reduction of p_T [panels (d-f)]. If only a few changes occur in the network [large transaction costs, panels (c) and (f)], there is no difference in the history and the resulting networks are identical. Otherwise, a gradual reduction leads to different updates before the final value of p_T is obtained and consequently to generally different states. This further illustrates the multitude of stable states for any given value of the transaction costs.

Chapter 8

Discussion

In this thesis we investigate collective dynamical and topological phenomena in networked systems in different contexts. We analyze the impact of economic considerations when modelling these systems and designing control schemes for these systems. In particular, we focus on efficient control schemes for dynamical processes, and the effect of costs and optimization on network formation. We derive and analyze efficient control schemes to synchronize coupled chaotic systems and to control random network percolation and study their impact on the dynamics. Additionally, we consider costs as a driving force of network formation, studying how local optimization changes the network formation process compared to stochastic percolation models.

Control of networked systems

One major challenge, beyond understanding network dynamical processes and the interplay between topology and dynamics, is the control of these processes, either to prevent failures or to guide the system to a desired state [114, 115, 125, 144]. Such interventions are difficult and costly in many systems, for example in biological or ecological contexts [199, 201], and there is great interest in understanding how to efficiently control network dynamical systems [55, 74, 118, 190, 198]. One major application of control in networked systems is the synchronizability of coupled oscillators, depending on both the topology as well as the actual dynamics and the coupling between the units. While a number of different mechanisms have been suggested to ensure synchronization, such as adaptively or periodically changing the dynamics of the coupling [39, 40, 97, 108, 202] or optimizing the topology of the interaction network itself [62, 131, 132, 171], most of the methods require strong access to the individual units as well as ongoing, potentially strong interactions.

In chapters 3 and 4 we introduce and analyze *transient uncoupling*, a new and efficient control mechanism of the interactions in networked systems. Transient uncoupling modifies the coupling of the systems with a state-dependent binary factor, controlling whether the coupling is active or inactive. We show that this not only enhances synchronizability of coupled chaotic oscillators but may even enable and guarantee it, independent of the topology of the interaction network. By limiting the interactions to states where coupling is strongly beneficial, this state-dependent uncoupling guarantees stable synchronization while reducing the overall interaction between the units to as little as coupling only 5% of the time. This is especially relevant when interaction and communication between the individual units is costly.

Additionally, the proposed coupling scheme requires minimal accessibility to the system, not changing the coupling dynamics or permanently modifying the structure of the interaction network, in contrast to many previously proposed control schemes. Instead, we need to be able to only turn interactions on or off, depending on the systems state. Furthermore, we illustrate the robustness of the method, showing that it can be effective even when controlling only a fraction of the units or observing the units' state only at discrete times. The simplicity of the method also allows its application in combination with other control schemes to efficiently control the dynamics of large systems with arbitrary interaction topology.

In contrast to structural control [199, 201], transient uncoupling requires no knowledge of the interaction network, providing a complementary approach. Since we take advantage of the internal dynamics, however, we have to adapt the state dependence of the uncoupling to the dynamics of the individual units. Even with a known model, it is not trivial to predict the effect of uncoupling in a given state. While it is possible to make an educated guess about an appropriate uncoupling region, as demonstrated in [182], in the manuscripts in chapters 3 and 4, we have chosen the state-dependence of the coupling by (extensive) trial-and-error. A more efficient way to apply transient uncoupling would combine the (un)coupling scheme with an optimization approach in order to automatically learn the state-dependence of the coupling. A successful optimization approach might effectively eliminate the need for knowledge of the dynamics. This could potentially lead to a coupling control scheme that guarantees synchronizability of arbitrary coupled systems, without knowledge of the internal dynamics nor of the interaction network structure [199, 201].

Similarly, transient uncoupling would become even more effective if it becomes possible to analytically predict efficient or optimal (un)coupling states depending on the units' dynamics. Such analysis would relate the local divergence of trajectories and state-dependent observability to the controllability of networked systems [11, 135, 136]. However, this is not a local problem. The effectiveness of (un)coupling the system in one state may be strongly affected by whether the system is coupled in other states. Therefore, this analysis would require more than instantaneous measures of stability like local Lyapunov exponents [182] to also include these non-local effects. Finally, one may consider transient uncoupling to optimize other measures relevant to synchronization or to the stability of networked systems. For example, a similar approach might be able to efficiently contain the spreading of perturbations in networked systems.

Synchronization

As one of the most common collective dynamic phenomena, synchronization is studied in a broad range of systems [12, 142]. Especially the prototypical Kuramoto-model and its variants are still widely studied and find new and relevant applications, for example in modeling the dynamics of power grids [152, 157, 194, 195]. Such applications require a way to accurately measure the phase coherence in these systems, for example

to quantify and compare the effects of control strategies. Previous works have already introduced extensions of the original Kuramoto order parameter to networks in order to analyze the formation of synchronized clusters in specific network topologies or to allow more detailed mean field calculations [77, 91, 148, 149]. However, none of these order parameters covers the whole dynamical range from incoherent oscillations to complete synchrony.

In chapter 5 we introduce a new, universal order parameter for such phase oscillators that accurately quantifies the coherence of the oscillators with respect to the interaction network through all stages of phase locking and synchrony. In particular, the new order parameter is directly related to the stability of the phase locked state, enabling analytical calculations as exemplified in the manuscript. We show that the order parameter can also be interpreted for other related systems, such as the second order Kuramoto model describing the dynamics of power grids. The direct correspondence of the stable operating state of power grids to the fully phase locked state of the standard Kuramoto model even allows us to transfer some of the analytic results. Importantly, the order parameter also quantifies the phase coherence accurately in small networks. This is especially relevant when considering possible applications, for example to power grids, where real networks are necessarily finite or only a small part of the whole network may be of interest.

Beyond measuring the global phase coherence in networks, the order parameter may also be used to study local phase coherence, for example to identify network motifs or specific regions of a network that are susceptible to desynchronization. Such a local measure may help to better identify links in power grids that are critical for the stability of the phase locked state [194, 195]. In general, a universal order parameter for Kuramoto-like models allows efficient and unified prediction of transitions in different models [12, 142], a necessary preliminary step to design control schemes for these systems and to quantify and compare their efficiency.

Control of percolation

Efficient control of networks is also relevant from a topological point of view, for example to prevent the spreading of diseases, to facilitate marketing in social networks, or to prevent or delay the failure of networked financial systems. Based on this motivation, a number of recent works have studied the identification of optimal spreaders and, related to that, the optimal immunization of networks [8, 46, 51, 101, 123, 137, 188]. There is great interest in understanding how to delay or, if possible, prevent failure of complex networks due to this general relevance for many social and economic systems [75, 83, 86, 175]. However, even before, research in percolation theory considered the impact of control and interventions to delay and modify the percolation transition, leading to the well-known “explosive” percolation model by Achlioptas et al. [3, 47, 147]. This triggered a broad investigation into the possibility of and the requirements for a discontinuous percolation transition [42, 43, 63, 156, 163]. These studies showed, for example, that local information may not be sufficient to control a

stochastic link addition process [150, 151]. Many models were studied that extended the idea of control and indeed result in discontinuous transitions [10, 24, 160, 161]. However, none of these models explicitly include the cost of interventions and (implicitly) assume unlimited resources to enact control decisions, an unrealistic condition in any real world system.

In chapter 6 we derive and study an efficient control scheme to delay random network percolation under the constraint of limited resources. By targeting interventions to prevent only those links that would create clusters that can be delayed significantly, we are able to delay the percolation transition with very few interventions. We extend the proposed control scheme to include limited observability, where control is only applied after some time has already passed, for example because the impending transition might not be apparent at first. We show with a detailed finite size analysis that the percolation transition remains smooth and in the same universality class as random percolation, characterized by a supercritical scaling of the relative size of the largest cluster $C_1/N \sim (p - p_c)^\beta$ with $\beta = 1$. This is in stark contrast to the known abrupt transition with $0 < \beta \ll 1$ of explosive percolation. Furthermore, we establish an analytical approximation to estimate optimal control parameters and reveal that optimal control, leading to maximal delay of the transition, also causes a discontinuous percolation transition. This transition becomes effectively uncontrollable since a single link can cause a macroscopic change in the size of the largest cluster. In summary, we reveal how optimal control to delay the transition has the unintended consequence to make the system uncontrollable once the transition happens [167].

While the presented framework is efficient, it is not strictly optimal. In particular, comparison to the Bohman-Fireze-Wormald model suggests that a linearly increasing intervention rate instead of a constant one might yield even better results [25, 44]. Similarly, while we discussed different cost functions in the framework of our control scheme, optimal control strategies for different cost functions likely differ qualitatively as well. Finding a way to directly apply optimal control theory to understand the impact of the cost function would therefore be an interesting and relevant next step [85, 86]. Another relevant extension of the model would focus on the information and observability of the system. As presented, the model assumes perfect information of the current cluster size distribution and observation of all links added to the network. While we considered limited observability affecting the start of the control, other constraints will likely have a bigger impact on the control. Ideally, it would be possible to estimate or predict the transition point during the percolation process [45, 162] to adjust the control and budget usage adaptively, independent of the underlying link addition process. Finally, the presented framework might be adapted or extended to also allow control of other network formation or percolation models, for example percolation on an underlying scale-free or small world network. In particular, in such networks one might consider control optimizing different objectives, such as path lengths or searchability, instead of simply the size of the largest cluster and the possible (unintended) consequences resulting from optimal control in these settings [51, 123].

Network formation

Many network formation models are based on stochastic link addition, such as the percolation models discussed above [3, 42, 43, 63, 156, 163]. However, many real world networks are not formed by (entirely) random processes but are designed or optimized for a certain function. Examples include technical and transport networks, like road or airline connections [76, 189], as well as distribution networks in biological systems, like leaf venation networks [99, 153]. This optimization is often done on the scale of the entire network governed by a single, global objective function. In contrast, many networks are formed by individual decisions of the nodes of the network, such as people in social networks. Some game theoretic models exist to describe network formation in this way based on individual optimization [94, 95]. Yet, due to the complexity of the optimization problems, most models that describe network formation in such a way are difficult to analyze and often provide only limited general insight. At the same time, while simpler stochastic models may be capable of recreating observed structural properties [14, 191] and perhaps hint at underlying mechanisms, they may be missing important interactions and effects. Establishing a connection between these two types of network formation models would enable an efficient analysis of the effects of individual decisions and economic considerations on network formation. This would be one important step to developing a general theory describing the underlying mechanisms. However, such a connection has not yet been established.

In chapter 7 we introduce and analyze a new model of network formation driven by global optimization by the individual nodes in the network. Based on a simple, economically motivated supply problem, the network structure is determined by these interacting individual optimization problems. We derive an exact mapping of our model to a local percolation rule. This mapping makes it possible to exactly solve the optimization problem, even for large networks, and study the formation of a large component as interaction costs between the nodes decrease. We reveal the existence of a parameter dependent continuous or discontinuous transition as well as hysteresis, in contrast to standard percolation transitions. Additionally, we analyze the impact of model parameters as well as of the structure of the underlying network of potential transport links. We analytically show how the transition is related to the collective interactions in the model, similar to dynamics observed in other models such as co-infection models and cascading failures [30, 41, 112]. This may explain the existence of discontinuous transitions and rare events beyond power law fluctuations expected from continuous transitions [112, 113, 176].

The proposed model may provide a way to understand aspects of network formation processes introduced by individual optimization and economic considerations as driving forces. Future work will show in detail how the formation of global connectivity depends on the underlying structure of potential transport links, for example analyzing the correlation of the cluster growth with the centrality or other connectivity measures. Some extensions of the presented model are immediately possible, such as multiple goods resulting in multiple parallel (uncoupled) supply problems. Many other extensions, however, would not retain the mapping to a local percolation model.

Such extensions include, for example, variable interaction costs or more general cost functions. Here, more detailed investigations are required to determine if simplifications in other aspects are possible to maintain the relation to local percolation models and keep the system solvable for large networks.

A different view on the model is possible from the perspective of a stochastic system. Instead of strict optimization, we can interpret the cost minimization as energy minimization, potentially resulting in a model similar to exponential random network models [128]. In general, the model opens a way to study the impact of economic forces and individual optimization on the formation of large scale networks as well as providing an efficient basic model to study control of network formation in such a setting.

Summary

Networked systems are steadily increasing in complexity with new and often unpredictable phenomena and interactions emerging between the different parts. This requires the development of new, more detailed theories to describe, understand, and predict these interactions. This is particularly relevant for ubiquitous constraints such as limited resources, information, or observability that affect almost any system. It is important to develop a general understanding of how these constraints affect network dynamical systems and of the underlying mechanisms. This will not only help to better predict the dynamics of such systems but also to design more efficient control methods.

Overall, we hope that this thesis will help to develop such a fundamental understanding of the effects of economic considerations on the formation and dynamics of networked systems.

Bibliography

- [1] CCBY license at <https://creativecommons.org/licenses/by/4.0/legalcode>, retrieved on 09/11/2017.
- [2] J. A. Acebrón, L. L. Bonilla, C. J. P. Vicente, F. Ritort, and R. Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.*, 77(1):137, 2005.
- [3] D. Achlioptas, R. M. D’Souza, and J. Spencer. Explosive percolation in random networks. *Science*, 323(5920):1453–1455, 2009.
- [4] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman. Search in power-law networks. *Phys. Rev. E*, 64:046135, 2001.
- [5] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47, 2002.
- [6] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.
- [7] U. Alon. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.*, 8(6):450–461, 2007.
- [8] F. Altarelli, A. Braunstein, L. Dall’Asta, J. R. Wakeling, and R. Zecchina. Containing epidemic outbreaks by message-passing techniques. *Phys. Rev. X*, 4:021024, 2014.
- [9] T. Antal, P. L. Krapivsky, and S. Redner. Dynamics of social balance on networks. *Phys. Rev. E*, 72(3):036121, 2005.
- [10] N. A. M. Araújo and H. J. Herrmann. Explosive percolation via control of the largest cluster. *Phys. Rev. Lett.*, 105:035701, 2010.
- [11] A. Arenas, A. Diaz-Guilera, and C. J. Pérez-Vicente. Synchronization reveals topological scales in complex networks. *Phys. Rev. Lett.*, 96(11):114102, 2006.
- [12] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Phys. Rep.*, 469(3):93–153, 2008.
- [13] V. Bala and S. Goyal. A noncooperative model of network formation. *Econometrica*, 68(5):1181–1229, 2000.
- [14] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

- [15] A.-L. Barabási and R. Albert-László. Scale-free networks: a decade and beyond. *Science*, 325(5939):412–413, 2009.
- [16] M. Barahona and L. M. Pecora. Synchronization in small-world systems. *Phys. Rev. Lett.*, 89:054101, 2002.
- [17] A. Barrat, M. Barthélemy, and A. Vespignani. *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge, 2008.
- [18] E. Barter and T. Gross. Meta-food-chains as a many-layer epidemic process on networks. *Phys. Rev. E*, 93:022303, 2016.
- [19] C. Bick. *Chaos and chaos control in network dynamical systems*. PhD thesis, University of Göttingen, 2013.
- [20] S. Boccaletti, C. Grebogi, Y.-C. Lai, H. Mancini, and D. Maza. The control of chaos: theory and applications. *Phys. Rep.*, 329(3):103–197, 2000.
- [21] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Phys. Rep.*, 424(4):175–308, 2006.
- [22] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin. The structure and dynamics of multilayer networks. *Phys. Rep.*, 544(1):1–122, 2014.
- [23] S. Boettcher, B. Gonçalves, and H. Guclu. Hierarchical regular small-world networks. *J. Phys. A*, 41(25):252001, 2008.
- [24] S. Boettcher, V. Singh, and R. M. Ziff. Ordinary percolation with discontinuous transitions. *Nat. Commun.*, 3:1774, 2012.
- [25] T. Bohman, A. Frieze, and N. C. Wormald. Avoidance of a giant component in half the edge set of a random graph. *Random Struct. Algor.*, 25(4):432–449, 2004.
- [26] S. Bohn and M. O. Magnasco. Structure, scaling, and phase transition in the optimal transport network. *Phys. Rev. Lett.*, 98:088702, 2007.
- [27] B. Bollobás. Random graphs. In *Modern Graph Theory*, pages 215–252. Springer-Verlag, Berlin, 1998.
- [28] B. Bollobás and O. M. Riordan. Mathematical results on scale-free random graphs. In *Handbook of graphs and networks: from the genome to the internet*, pages 1–34. Weinheim: Wiley-VCH, 2003.
- [29] S. P. Borgatti and M. G. Everett. Models of core/periphery structures. *Soc. Networks*, 21(4):375–395, 2000.
- [30] L. Böttcher, J. Nagler, and H. J. Herrmann. Critical behaviors in contagion dynamics. *Phys. Rev. Lett.*, 118(8):088301, 2017.
- [31] S. Brin and L. Page. Reprint of: The anatomy of a large-scale hypertextual web search engine. *Comput. Netw.*, 56(18):3825–3833, 2012.

- [32] D. Brockmann and D. Helbing. The hidden geometry of complex, network-driven contagion phenomena. *Science*, 342(6164):1337–1342, 2013.
- [33] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. *Nature*, 439(7075):462–465, 2006.
- [34] C. D. Brummitt, P. D. Hines, I. Dobson, C. Moore, and R. M. D’Souza. Transdisciplinary electric power grid science. *Proc. Natl. Acad. Sci. USA*, 110(30):12159–12159, 2013.
- [35] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin. Catastrophic cascade of failures in interdependent networks. *Nature*, 464(7291):1025–1028, 2010.
- [36] R. S. Burt. *Structural holes: The social structure of competition*. Harvard university press, Cambridge, MA, 2009.
- [37] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85(25):5468, 2000.
- [38] C. Castellano and R. Pastor-Satorras. Thresholds for epidemic spreading in networks. *Phys. Rev. Lett.*, 105:218701, 2010.
- [39] L. Chen, C. Qiu, and H. Huang. Synchronization with on-off coupling: Role of time scales in network dynamics. *Phys. Rev. E*, 79(4):045101, 2009.
- [40] L. Chen, C. Qiu, H. Huang, G. Qi, and H. Wang. Facilitated synchronization of complex networks through a discontinuous coupling strategy. *Eur. Phys. J. B*, 76(4):625–635, 2010.
- [41] L. Chen, F. Ghanbarnejad, and D. Brockmann. Fundamental properties of cooperative contagion processes. *New J. Phys.*, 19(10):103041, 2017.
- [42] W. Chen and R. M. D’Souza. Explosive percolation with multiple giant components. *Phys. Rev. Lett.*, 106:115701, 2011.
- [43] W. Chen, X. Cheng, Z. Zheng, N. N. Chung, R. M. D’Souza, and J. Nagler. Unstable supercritical discontinuous percolation transitions. *Phys. Rev. E*, 88(4):042152, 2013.
- [44] W. Chen, J. Nagler, X. Cheng, X. Jin, H. Shen, Z. Zheng, and R. M. D’Souza. Phase transitions in supercritical explosive percolation. *Phys. Rev. E*, 87(5):052130, 2013.
- [45] W. Chen, M. Schröder, R. M. D’Souza, D. Sornette, and J. Nagler. Microtransition cascades to percolation. *Phys. Rev. Lett.*, 112(15):155701, 2014.
- [46] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley. Finding a better immunization strategy. *Phys. Rev. Lett.*, 101(5):058701, 2008.

- [47] Y. S. Cho, J. S. Kim, J. Park, B. Kahng, and D. Kim. Percolation transitions in scale-free networks under the achlioptas process. *Phys. Rev. Lett.*, 103:135702, 2009.
- [48] Y. S. Cho, B. Kahng, and D. Kim. Cluster aggregation model for discontinuous percolation transitions. *Phys. Rev. E*, 81(3):030103, 2010.
- [49] Y. S. Cho, S. Hwang, H. J. Herrmann, and B. Kahng. Avoiding a spanning cluster in percolation models. *Science*, 339(6124):1185–1187, 2013.
- [50] A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM Rev.*, 51(4):661–703, 2009.
- [51] P. Clusella, P. Grassberger, F. J. Pérez-Reche, and A. Politi. Immunization and targeted destruction of networks using explosive percolation. *Phys. Rev. Lett.*, 117(20):208301, 2016.
- [52] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Phys. Rev. Lett.*, 85:4626, 2000.
- [53] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Breakdown of the internet under intentional attack. *Phys. Rev. Lett.*, 86:3682, 2001.
- [54] R. Cohen, S. Havlin, and D. ben Avraham. Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.*, 91(24):247901, 2003.
- [55] S. P. Cornelius, W. L. Kath, and A. E. Motter. Realistic control of network dynamics. *Nat. Commun.*, 4:1942, 2013.
- [56] N. J. Corron, J. N. Blakely, and S. D. Pethel. Lag and anticipating synchronization without time-delay coupling. *Chaos*, 15(2):023110, 2005.
- [57] M. Cross and H. Greenside. *Pattern Formation and Dynamics in Nonequilibrium Systems*. Cambridge University Press, Cambridge, 2009.
- [58] R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. Explosive percolation transition is actually continuous. *Phys. Rev. Lett.*, 105:255701, 2010.
- [59] P. Dayan and L. F. Abbott. *Theoretical Neuroscience*. MIT Press, Cambridge, MA, 2001.
- [60] M. De Domenico and A. Arenas. Modeling structure and resilience of the dark network. *Phys. Rev. E*, 95:022313, 2017.
- [61] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas. Mathematical formulation of multilayer networks. *Phys. Rev. X*, 3:041022, 2013.
- [62] L. Donetti, P. I. Hurtado, and M. A. Muñoz. Entangled networks, synchronization, and optimal network topology. *Phys. Rev. Lett.*, 95:188701, 2005.

- [63] R. M. D’Souza and J. Nagler. Anomalous critical and supercritical phenomena in explosive percolation. *Nat. Phys.*, 11(7):531–538, 2015.
- [64] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge University Press, Cambridge, 2010.
- [65] M. Elliott, B. Golub, and M. O. Jackson. Financial networks and contagion. *Am. Econ. Rev.*, 104(10):3115–3153, 2014.
- [66] P. Erdős and A. Rényi. On random graphs i. *Publ. Math. Debrecen*, 6:290–297, 1959.
- [67] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 5(1):17–60, 1960.
- [68] G. Filatrella, A. H. Nielsen, and N. F. Pedersen. Analysis of a power grid using a kuramoto-like model. *Eur. Phys. J. B*, 61(4):485–491, 2008.
- [69] P. Fleurquin, J. J. Ramasco, and V. M. Eguiluz. Systemic delay propagation in the us airport network. *Sci. Rep.*, 3:1159, 2013.
- [70] P. J. Flory. Molecular size distribution in three dimensional polymers. I. Gelation. *J. Am. Chem. Soc.*, 63(11):3083–3090, 1941.
- [71] V. Flunkert, S. Yanchuk, T. Dahms, and E. Schöll. Synchronizing distant nodes: A universal classification of networks. *Phys. Rev. Lett.*, 105:254101, 2010.
- [72] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.
- [73] P. Gai and S. Kapadia. Contagion in financial networks. *Proc. R. Soc. A*, 466:2401, 2010.
- [74] J. Gao, Y.-Y. Liu, R. M. D’souza, and A.-L. Barabási. Target control of complex networks. *Nat. Commun.*, 5:5415, 2014.
- [75] J. Gao, B. Barzel, and A.-L. Barabási. Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312, 2016.
- [76] M. T. Gastner and M. E. J. Newman. Optimal design of spatial distribution networks. *Phys. Rev. E*, 74:016117, 2006.
- [77] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas. Paths to synchronization on complex networks. *Phys. Rev. Lett.*, 98:034101, 2007.
- [78] P. Grassberger, C. Christensen, G. Bizhani, S.-W. Son, and M. Paczuski. Explosive percolation is continuous, but with unusual finite size behavior. *Phys. Rev. Lett.*, 106:225701, 2011.
- [79] J. Grimm, E. M. Elçi, Z. Zhou, T. M. Garoni, and Y. Deng. Geometric explanation of anomalous finite-size scaling in high dimensions. *Phys. Rev. Lett.*, 118:115701, 2017.

- [80] G. R. Grimmett. *Percolation*. Springer-Verlag, Berlin, 1999. 2nd edition.
- [81] J. W. Grossman and P. D. F. Ion. On a portion of the well-known collaboration graph. *Congressus Numerantium*, 108:129–132, 1995.
- [82] J. Guare. *Six degrees of separation: A play*. Vintage, 1990.
- [83] A. G. Haldane and R. M. May. Systemic risk in banking ecosystems. *Nature*, 469(7330):351–355, 2011.
- [84] T. Hara and G. Slade. Mean-field critical behaviour for percolation in high dimensions. *Commun. Math. Phys.*, 128(2):333–391, 1990.
- [85] S. Havlin, D. Y. Kenett, E. Ben-Jacob, A. Bunde, R. Cohen, H. Hermann, J. W. Kantelhardt, J. Kertész, S. Kirkpatrick, J. Kurths, J. Portugali, and S. Solomon. Challenges in network science: Applications to infrastructures, climate, social systems and economics. *Eur. Phys. J. ST*, 214(1):273–293, 2012.
- [86] D. Helbing. Globally networked risks and how to respond. *Nature*, 497(7447):51–59, 2013.
- [87] M. W. Hirsch, S. Smale, and R. L. Devaney. *Differential equations, dynamical systems, and an introduction to chaos*. Academic Press, Cambridge, MA, 2012.
- [88] P. Holme. Core-periphery organization of complex networks. *Phys. Rev. E*, 72:046111, 2005.
- [89] P. Holme and J. Saramäki. Temporal networks. *Phys. Rep.*, 519(3):97–125, 2012.
- [90] L. Hufnagel, D. Brockmann, and T. Geisel. Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci. USA.*, 101(42):15124–15129, 2004.
- [91] T. Ichinomiya. Frequency synchronization in a random oscillator network. *Phys. Rev. E*, 70:026116, 2004.
- [92] M. O. Jackson. A survey of network formation models: stability and efficiency. In *Group Formation in Economics: Networks, Clubs, and Coalitions*, pages 11–49. Cambridge University Press, Cambridge, 2005.
- [93] M. O. Jackson. *Social and Economic Networks*. Princeton University Press, Princeton, 2008.
- [94] M. O. Jackson and A. Watts. The evolution of social and economic networks. *J. Econ. Theory*, 106(2):265–295, 2002.
- [95] M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *J. Econ. Theory*, 71(1):44–74, 1996.
- [96] S. Jahnke, R.-M. Memmesheimer, and M. Timme. Stable irregular dynamics in complex neural networks. *Phys. Rev. Lett.*, 100:048102, 2008.

- [97] L. Junge and U. Parlitz. Synchronization using dynamic coupling. *Phys. Rev. E*, 64:055204, 2001.
- [98] D. Jungnickel. *Graphs, Networks and Algorithms*. Springer-Verlag, Berlin, 2008.
- [99] E. Katifori, G. J. Szöllősi, and M. O. Magnasco. Damage and fluctuations induce loops in optimal transport networks. *Phys. Rev. Lett.*, 104:048704, 2010.
- [100] R. Kenna and B. Berche. Universal finite-size scaling for percolation theory in high dimensions. *J. Phys. A*, 50(23):235001, 2017.
- [101] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse. Identification of influential spreaders in complex networks. *Nat. Phys.*, 6(11):888–893, 2010.
- [102] M. Kivelä, A. Arenas, M. Barthélemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *J. Complex Netw.*, 2(3):203–271, 2014.
- [103] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pages 163–170. ACM, 2000.
- [104] J. Kleinberg and S. Lawrence. The structure of the web. *Science*, 294(5548):1849–1850, 2001.
- [105] J. Kleinberg, R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. The web as a graph: measurements, models, and methods. In *Proceedings of the International Conference on Combinatorics and Computing*, number 1627 in Lecture Notes in Computer Science, pages 1–17. Springer-Verlag, Berlin, 1999.
- [106] J. M. Kleinberg. Authoritative sources in a hyperlinked environment. *J. ACM*, 46(5):604–632, 1999.
- [107] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845–845, 2000.
- [108] J. Klinglmayr, C. Kirst, C. Bettstetter, and M. Timme. Guaranteeing global synchronization in networks with stochastic interactions. *New J. Phys.*, 14(7):073031, 2012.
- [109] G. Kossinets and D. J. Watts. Empirical analysis of an evolving social network. *Science*, 311(5757):88–90, 2006.
- [110] Y. Kuramoto. Self-entrainment of a population of coupled non-linear oscillators. In *International Symposium on Mathematical Problems in Theoretical Physics*, Lecture Notes in Physics Vol. 39, page 420, New York, 1975. Springer-Verlag.
- [111] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence*. Springer-Verlag, Berlin, 1984.
- [112] D. Lee, W. Choi, J. Kertész, and B. Kahng. Universal mechanism for hybrid percolation transitions. *Sci. Rep.*, 7(1):5723, 2017.

- [113] Y. Lin, K. Burghardt, M. Rohden, P.-A. Noël, and R. M. D'Souza. The self-organization of dragon kings. *arXiv preprint arXiv:1705.10831*, 2017.
- [114] Y.-Y. Liu and A.-L. Barabási. Control principles of complex systems. *Rev. Mod. Phys.*, 88(3):035006, 2016.
- [115] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási. Controllability of complex networks. *Nature*, 473(7346):167–173, 2011.
- [116] E. N. Lorenz. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20(2):130–141, 1963.
- [117] S. Mangan and U. Alon. Structure and function of the feed-forward loop network motif. *Proc. Natl. Acad. Sci. USA*, 100(21):11980–11985, 2003.
- [118] R. M. May, S. A. Levin, and G. Sugihara. Complex systems: Ecology for bankers. *Nature*, 451(7181):893–895, 2008.
- [119] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.
- [120] P. Misra and P. Enge. *Global Positioning System: signals, measurements and performance*. Ganga-Jamuna Press, Massachusetts, 2006. 2nd edition.
- [121] C. Moore and M. E. J. Newman. Epidemics and percolation in small-world networks. *Phys. Rev. E*, 61(5):5678, 2000.
- [122] J. L. Moreno. *Who shall survive?* Beacon House, Beacon, NY, 1934.
- [123] F. Morone and H. A. Makse. Influence maximization in complex networks through optimal percolation. *Nature*, 524(7563):65–68, 2015.
- [124] A. E. Motter. Bounding network spectra for network design. *New J. Phys.*, 9(6):182, 2007.
- [125] A. E. Motter. Networkcontrology. *Chaos*, 25(9):097621, 2015.
- [126] J. Nagler, A. Levina, and M. Timme. Impact of single links in competitive percolation. *Nat. Phys.*, 7(3):265–270, 2011.
- [127] M. E. J. Newman. The structure and function of complex networks. *SIAM Rev.*, 45(2):167–256, 2003.
- [128] M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, Oxford, 2010.
- [129] M. E. J. Newman and R. M. Ziff. Fast monte carlo algorithm for site or bond percolation. *Phys. Rev. E*, 64:016706, 2001.
- [130] I. Newton. *Philosophiæ naturalis principia mathematica (Mathematical principles of natural philosophy)*. London, 1687.

- [131] T. Nishikawa and A. E. Motter. Synchronization is optimal in nondiagonalizable networks. *Phys. Rev. E*, 73(6):065106, 2006.
- [132] T. Nishikawa and A. E. Motter. Network synchronization landscape reveals compensatory structures, quantization, and the positive effect of negative interactions. *Proc. Natl. Acad. Sci. USA*, 107(23):10342–10347, 2010.
- [133] B. H. Obama. Presidential policy directive 21: Critical infrastructure security and resilience., Washington, DC, 2013.
- [134] J. F. Padgett and C. K. Ansell. Robust action and the rise of the medici, 1400-1434. *Am. J. Sociol.*, 98(6):1259–1319, 1993.
- [135] U. Parlitz. Estimating model parameters from time series by autosynchronization. *Phys. Rev. Lett.*, 76(8):1232, 1996.
- [136] U. Parlitz. Estimating lyapunov exponents from time series. In C. H. Skokos, G. A. Gottwald, and J. Laskar, editors, *Chaos Detection and Predictability*, pages 1–34. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [137] R. Pastor-Satorras and A. Vespignani. Immunization of complex networks. *Phys. Rev. E*, 65(3):036104, 2002.
- [138] L. M. Pecora and T. L. Carroll. Synchronization in chaotic systems. *Phys. Rev. Lett.*, 64(8):821, 1990.
- [139] L. M. Pecora and T. L. Carroll. Master stability functions for synchronized coupled systems. *Phys. Rev. Lett.*, 80:2109–2112, 1998.
- [140] L. M. Pecora and T. L. Carroll. Synchronization of chaotic systems. *Chaos*, 25(9):097611, 2015.
- [141] L. M. Pecora, T. L. Carroll, G. A. Johnson, D. J. Mar, and J. F. Heagy. Fundamentals of synchronization in chaotic systems, concepts, and applications. *Chaos*, 7(4):520–543, 1997.
- [142] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, Cambridge, 2003.
- [143] P. Pillai, A. Gonzalez, and M. Loreau. Metacommunity theory explains the emergence of food web complexity. *Proc. Natl. Acad. Sci. USA*, 108(48):19293–19298, 2011.
- [144] M. Pósfai, J. Gao, S. P. Cornelius, A.-L. Barabási, and R. M. D’Souza. Controllability of multiplex, multi-time-scale networks. *Phys. Rev. E*, 94(3):032316, 2016.
- [145] W. H. Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge University Press, Cambridge, 2007.
- [146] D. d. S. Price. A general theory of bibliometric and other cumulative advantage processes. *J. Assoc. Inf. Sci. Technol.*, 27(5):292–306, 1976.

- [147] F. Radicchi and S. Fortunato. Explosive percolation in scale-free networks. *Phys. Rev. Lett.*, 103:168701, 2009.
- [148] J. G. Restrepo, E. Ott, and B. R. Hunt. Onset of synchronization in large networks of coupled oscillators. *Phys. Rev. E*, 71:036151, 2005.
- [149] J. G. Restrepo, E. Ott, and B. R. Hunt. Synchronization in large directed networks of coupled phase oscillators. *Chaos*, 16(1):015107, 2006.
- [150] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science*, 333(6040):322–324, 2011.
- [151] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *Ann. Appl. Probab.*, 22(4):1450–1464, 2012.
- [152] M. Rohden, A. Sorge, M. Timme, and D. Witthaut. Self-organized synchronization in decentralized power grids. *Phys. Rev. Lett.*, 109(6):064101, 2012.
- [153] H. Ronellenfitsch and E. Katifori. Global optimization, local adaptation, and the role of growth in distribution networks. *Phys. Rev. Lett.*, 117:138301, 2016.
- [154] M. Rosenblum, A. Pikovsky, and J. Kurths. Phase synchronization of chaotic oscillators. *Phys. Rev. Lett.*, 76(11):1804, 1996.
- [155] O. E. Rössler. An equation for continuous chaos. *Phys. Lett. A*, 57(5):397–398, 1976.
- [156] A. A. Saberi. Recent advances in percolation theory and its applications. *Phys. Rep.*, 578:1–32, 2015.
- [157] B. Schäfer, M. Matthiae, M. Timme, and D. Witthaut. Decentral smart grid control. *New J. Phys.*, 17(1):015002, 2015.
- [158] B. Schäfer, C. Beck, K. Aihara, D. Witthaut, and M. Timme. Non-gaussian power grid frequency fluctuations characterized by lévy-stable laws and super-statistics. *Nat. Energy*, 2017.
- [159] E. Schöll and H. Schuster. *Handbook of chaos control*. Wiley-VCH, Weinheim, 2nd edition, 2008.
- [160] K. J. Schrenk, N. A. M. Araújo, and H. J. Herrmann. Gaussian model of explosive percolation in three and higher dimensions. *Phys. Rev. E*, 84(4):041136, 2011.
- [161] K. J. Schrenk, A. Felder, S. Deflorin, N. A. M. Araújo, R. M. D’Souza, and H. J. Herrmann. Bohman-frieze-wormald model on the lattice, yielding a discontinuous percolation transition. *Phys. Rev. E*, 85:031103, 2012.
- [162] M. Schröder. Discontinuous transitions in growing networks. Master’s thesis, University Göttingen, 2014.
- [163] M. Schröder, S. H. E. Rahbari, and J. Nagler. Crackling noise in fractional percolation. *Nat. Commun.*, 4:2222, 2013.

- [164] M. Schröder, M. Mannattil, D. Dutta, S. Chakraborty, and M. Timme. Transient uncoupling induces synchronization. *Phys. Rev. Lett.*, 115:054101, 2015.
- [165] M. Schröder, S. Chakraborty, D. Witthaut, J. Nagler, and M. Timme. Interaction control to synchronize non-synchronizable networks. *Sci. Rep.*, 6:37142, 2016.
- [166] M. Schröder, W. Chen, and J. Nagler. Discrete scale invariance in supercritical percolation. *New J. Phys.*, 18(1):013042, 2016.
- [167] M. Schröder, N. A. M. Araújo, D. Sornette, and J. Nagler. Controlling percolation with limited resources. *Phys. Rev. E*, 96:062302, 2017.
- [168] M. Schröder, M. Timme, and D. Witthaut. A universal order parameter for synchrony in networks of limit cycle oscillators. *Chaos*, 27(7):073119, 2017.
- [169] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White. Economic networks: The new challenges. *Science*, 325(5939):422–425, 2009.
- [170] J. Sethna. *Statistical mechanics: entropy, order parameters, and complexity*. Oxford University Press, Oxford, 2006.
- [171] R. Sevilla-Escoboza, J. Buldú, S. Boccaletti, D. Papo, D.-U. Hwang, G. Huerta-Cuellar, and R. Gutiérrez. Experimental implementation of maximally synchronizable networks. *Physica A*, 448:113–121, 2016.
- [172] S. Singh. Erdos-bacon numbers, 2002. <https://simonsingh.net/media/articles/maths-and-science/erdos-bacon-numbers/> retrieved 01.11.2017.
- [173] R. V. Solé, R. Pastor-Satorras, E. Smith, and T. B. Kepler. A model of large-scale proteome evolution. *Adv. Complex Syst.*, 5(01):43–54, 2002.
- [174] D. Sornette. *Why stock markets crash: critical events in complex financial systems*. Princeton University Press, Princeton, 2017.
- [175] D. Sornette and P. Cauwels. 1980–2008: The illusion of the perpetual money machine and what it bodes for the future. *Risks*, 2(2):103–131, 2014.
- [176] D. Sornette and G. Ouillon. Dragon-kings: mechanisms, statistical methods and empirical evidence. *Eur. Phys. J. ST*, 205(1):1–26, 2012.
- [177] D. Stauffer and A. Aharony. *Introduction to percolation theory*. Taylor & Francis, London, 1992. 2nd edition.
- [178] S. H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press, Boulder, CO, 1994.
- [179] S. H. Strogatz. From kuramoto to crawford: Exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143:1, 2000.

- [180] S. H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.
- [181] S. H. Strogatz, D. M. Abrams, A. McRobie, B. Eckhardt, and E. Ott. Theoretical mechanics: Crowd synchrony on the millennium bridge. *Nature*, 438(7064):43–44, 2005.
- [182] A. Tandon, M. Schröder, M. Mannattil, M. Timme, and S. Chakraborty. Synchronizing noisy nonidentical oscillators by transient uncoupling. *Chaos*, 26(9):094817, 2016.
- [183] G. Timár, A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes. Mapping the structure of directed networks: Beyond the bow-tie diagram. *Phys. Rev. Lett.*, 118:078301, 2017.
- [184] M. Timme. Does dynamics reflect topology in directed networks? *Europhys. Lett.*, 76(3):367, 2006.
- [185] M. Timme, L. Kocarev, and D. Witthaut. Focus on networks, energy and the economy. *New J. Phys.*, 17(11):110201, 2015.
- [186] J. Travers and S. Milgram. The small world problem. *Psychology Today*, 1:61–67, 1967.
- [187] J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4):425–443, 1969.
- [188] T. W. Valente. Network interventions. *Science*, 337(6090):49–53, 2012.
- [189] T. Verma, F. Russmann, N. A. M. Araújo, J. Nagler, and H. J. Herrmann. Emergence of core-peripheries in networks. *Nat. Commun.*, 7:10441, 2016.
- [190] A. Vespignani. Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.*, 8(1):32–39, 2012.
- [191] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, 1998.
- [192] D. J. Watts, P. S. Dodds, and M. E. J. Newman. Identity and search in social networks. *Science*, 296(5571):1302–1305, 2002.
- [193] R. J. Wilson. *An Introduction to Graph Theory*. Pearson Education Limited, Essex, 2010. 5th edition.
- [194] D. Witthaut and M. Timme. Braess’s paradox in oscillator networks, desynchronization and power outage. *New J. Phys.*, 14(8):083036, 2012.
- [195] D. Witthaut, M. Rohden, X. Zhang, S. Hallerberg, and M. Timme. Critical links and nonlocal rerouting in complex supply networks. *Phys. Rev. Lett.*, 116:138701, 2016.
- [196] M. Wittmann and A. P. Young. Finite-size scaling above the upper critical dimension. *Phys. Rev. E*, 90:062137, 2014.

- [197] Z. Wu, C. Lagorio, L. A. Braunstein, R. Cohen, S. Havlin, and H. E. Stanley. Numerical evaluation of the upper critical dimension of percolation in scale-free networks. *Phys. Rev. E*, 75:066110, 2007.
- [198] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási. Spectrum of controlling and observing complex networks. *Nat. Phys.*, 11(9):779–786, 2015.
- [199] G. Yan, P. E. Vértes, E. K. Towilson, Y. L. Chew, D. S. Walker, W. R. Schafer, and A.-L. Barabási. Network control principles predict neuron function in the *caenorhabditis elegans* connectome. *Nature*, 2017.
- [200] G. U. Yule. A mathematical theory of evolution, based on the conclusions of Dr. J.C Willis, FRS. *Phil. Trans. R. Soc. B.*, 213:21–87, 1925.
- [201] J. G. T. Zañudo, G. Yang, and R. Albert. Structure-based control of complex networks with nonlinear dynamics. *Proc. Natl. Acad. Sci. USA*, 114(28):7234–7239, 2017.
- [202] C. Zhou and J. Kurths. Dynamical weights and enhanced synchronization in adaptive complex networks. *Phys. Rev. Lett.*, 96(16):164102, 2006.

Acknowledgements

My gratitude goes to my parents, friends, and colleagues for all their support during my PhD, it would have certainly been a much duller affair without you.

First, I would like to thank my thesis committee for their continued support during my PhD. Especially Marc and Dirk for always supporting my work and for sharing their experience and time writing papers, abstracts, letters to editors and whatever else there was to do. Thank you also for letting me work on so many different things, always trusting me that something useful will come of it and for sometimes pushing me to try more new things, go to conferences, and meet new people. I would also like to thank Prof. Kree for his interest in my work and the always insightful discussions during our meetings.

I would like to thank all of my co-authors, both of the manuscripts included in this thesis as well as of those still in preparation, for their time, effort, and helpful comments, not only concerning our manuscripts. I want to especially thank Jan Nagler for his support, his never-ending memes, and of course for getting me interested in network dynamics in the first place.

I want to thank the staff of the MPI DS for all their support with travel arrangements and all other administrative things, especially Barbara, Ayse and Monika. I also thank the members of the GGNB and PBCS office for their support with all the administration and red tape during my PhD and I would like to acknowledge GGNB for financial support during most of my PhD in form of the GGNB excellence stipend (DFG Grant GSC 226/2).

Thank you to all my colleagues and fellow network dynamicists for the many funny stories shared over lunch, the wonderful atmosphere in our group, and for always supplying interesting problems to think about. Thank you also for bringing cake or preparing food on every possible (and impossible) occasion.

Finally, I want to thank my friends, especially Benjamin, Meike, Jana, Diemut, and Alexander, for all their time spent with me, playing boardgames, getting as invested in the stories and the rules, and always supporting me in all possible ways. Thanks to you there were always more enjoyable things to look forward to than the next referee report.

Last, but not least, thank you to my parents, who always supported me.