Robustness and Entropy for Dynamics on Networks



Alice Chun-Hua Ulrike Schwarze

Pembroke College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Contents

1 Introduction **5**

1.1 Motivation 5

1.2 Structural and functional redundancy in biological networks 8

1.2.1 Distinguishing structural and functional redundancy 8

1.2.2 Quantifying structural and functional redundancy 9

1.3 Structure of this dissertation 12

2 Robustness of networks to structural changes **14**

2.1 Introduction 16

2.2 Simple networks, subnetworks, and matrix representations 18

2.2.1 Networks and simple networks 18

2.2.2 Subnetworks 19

2.2.3 Walks and closed walks 20

2.2.4 Matrix representations of networks 20

2.3 System robustness and related concepts 21

2.3.1 Defining robustness 21

2.3.2 Different name, same concept? 22

2.3.3 Same name, different concept? 23

2.3.4 Reviews of robustness and resilience 24

2.4 Specification of a network robustness problem 26

2.4.1 A system and a model 26

2.4.2 A measure of robustness 30

2.4.3 A set of perturbations 31

2.4.4 A measure of performance 40

2.5 Conclusions and outlook 59

2.5.1 Conclusions 59

2.5.2 Outlook 63

3 Plausiblity of performance measures for protein-interaction networks 64
3.1 Introduction . . . 66
3.1.1 Glossary of biological concepts . . . 69
3.2 Materials and methods . . . 71
3.2.1 The HIPPIE interactome . . . 71
3.2.2 Random-graph models . . . 72
3.2.3 Sampling function-specific protein-interaction networks . . . 73
3.2.4 Impacts of structural perturbations . . . 75
3.2.5 Performance measures . . . 76
3.2.6 Targeting strategies . . . 77
3.2.7 A measure of plausibility . . . 77
3.3 Results . . . 80
3.3.1 Structure of function-specific protein-interaction networks . . . 81
3.3.2 Robustness of structural properties under node removal . . . 91
3.3.3 Plausibility scores of performance measures . . . 96
3.4 Conclusions and discussion . . . 101
3.4.1 Summary . . . 101
3.4.2 Low clustering coefficients . . . 102
3.4.3 Similar observations for different sampling methods . . . 103
3.4.4 Similar robustness profiles for different targeting strategies . . . 103
3.4.5 Plausible measures of performance and viability . . . 104
3.4.6 Outlook . . . 104

4 Entropy, redundant paths, and cycles in a network with Ornstein–Uhlenbeck dynamics 107
4.1 Introduction . . . 109
4.2 Background: Entropy and networks . . . 111
4.2.1 Shannon entropy and Kolmogorov–Sinai entropy . . . 112
4.2.2 Network entropy . . . 114
4.3 The Ornstein–Uhlenbeck process and its differential entropy . . . 115
4.3.1 The Ornstein–Uhlenbeck process . . . 115
4.3.2 Variance and differential entropy of the Ornstein–Uhlenbeck process . . . 116
4.3.3 Short-range signal decay . . . 117
4.3.4 Network structure matters . . . 118
4.3.5 Finite-walk approximation . . . 119

4.4 Walk graphs as network motifs . . . 121
4.4.1 Some graph-theoretic concepts . . . 121
4.4.2 Characterizing and counting walk graphs . . . 122
4.4.3 Walk graphs and subgraphs . . . 123
4.5 Covariance-inducing motifs of the Ornstein–Uhlenbeck process . . . 125
4.5.1 Structure of covariance-inducing motifs . . . 125
4.5.2 Weighting factors for covariance-inducing motifs . . . 127
4.5.3 Origins of variance in the Ornstein–Uhlenbeck process . . . 128
4.6 Entropy-inducing motifs of the Ornstein–Uhlenbeck process . . . 128
4.6.1 Structure of entropy-inducing motifs . . . 129
4.6.2 Weighting factors for entropy-inducing motifs . . . 129
4.6.3 Entropy contribution of a 3-node feedforward loop and 3-node feedback loops . . . 131
4.7 Conclusions and discussion . . . 133
4.7.1 Summary . . . 133
4.7.2 An extended information-theoretic framework . . . 134
4.7.3 Disorder versus coherence . . . 135
4.7.4 Outlook . . . 136

5 Log-minor distributions and an application to estimating mean subsystem entropy 138
5.1 Introduction . . . 140
5.2 Notation . . . 141
5.3 Bounds on the distribution of log-minors . . . 142
5.4 Proofs of bounds on the distribution of log-minors . . . 144
5.4.1 Proof of Theorem 1 . . . 144
5.4.2 Proof of Theorem 2 . . . 147
5.4.3 Proof of Theorem 3 . . . 148
5.5 Examples . . . 149
5.6 Estimating mean subsystem entropy . . . 152
5.7 Conclusions . . . 158

6 Conclusions and outlook 161
6.1 Robustness of networks to structural changes . . . 161
6.2 Plausibility of performance measures for protein-interaction networks 162
6.3 Walk graphs in a multivariate Ornstein–Uhlenbeck process on a network 164
6.4 Sampling guarantees for the mean differential entropy of subsystems . 165

6.5 Outlook 166

A Comparison of structural properties for different sampling methods 168

A.1 Comparison of degree distributions and local clustering coefficients . . 168

A.2 Comparison of structural properties for different network sizes 169

B Covariance matrix of an Ornstein–Uhlenbeck process with signal decay 170

B.1 Derivation of the covariance matrix 170

B.2 Conditions for short-range signal decay 171

Bibliography 172

List of Figures

1.1 Examples of structural and functional redundancy. 10

2.1 Combinations of specifications of robustness problems in studies of robustness of networks to node and edge removal. 32
2.2 A map from the video game Mini Metro with a connected transportation network. 56
2.3 A map from the video game Mini Metro with a fragmented transportation network. 57

3.1 Sampling methods for function-specific protein-interaction networks. . 74
3.2 Sizes of network samples. 82
3.3 Degree distributions and mean local clustering coefficients in the HIPPIE interactome and fPINs. 84
3.4 Mean degree distributions and mean distributions for sets of fPINs. . 86
3.5 Comparison of structural properties of fPINs for different sampling methods. 88
3.6 Structural properties of fPINs. 90
3.7 Robustness profiles for performance measures on SP-sampled fPINs . 92
3.8 Robustness profiles for performance measures on SP-sampled fPINs. . 93
3.9 Mean robustness profiles for fPINs from several sampling methods. . . 95
3.10 Criterion scores and plausibility scores for performance measures. . . 97
3.11 Criterion scores and plausibility scores for performance measures for AP sampling. 98
3.12 Criterion scores and plausibility scores for performance measures for SP-sampled fPINs and associated sets of random networks. 100

4.1 Examples of walk graphs. 111
4.2 The differential entropy $h_{[\infty]}(\epsilon, \mathbf{A})$ for a multivariate OU process. . . 118
4.3 Accuracy of $h_{[L]}(\epsilon, \mathbf{A})$ as an approximation of $h_{[\infty]}(\epsilon, \mathbf{A})$. 120
4.4 Walk graphs that can occur on a 3-node feedforward loop. 124

4.5 Walk graphs that can occur on a 3-node feedback loop. 125
4.6 Structures and configurations of covariance-inducing motifs, variance-inducing motifs , and entropy-inducing motifs. 126
4.7 Weighting factors for CMs. 127
4.8 Weighting factors for short EMs. 130
4.9 Weighting factors for long EMs. 130
4.10 Largest weighting factors for EMs. 132

5.1 Approximate probability density functions for log-minors. 150
5.2 Empirical tails of the distribution of log-minors. 153
5.3 Bounds on the standard error and coefficient of variation for a sample mean of log-minors and a sample mean of subsystem entropy. 157

A.1 Degree distribution and mean local clustering coefficient for fPINs from different sampling methods. 168
A.2 Structural properties for fPINs with different sizes. 169

List of Tables

2.1 List of symbols in Chapter 2. 15
2.2 List of symbols in Chapter 2 (continued). 16
2.3 Examples of studies of robustness of network models for different areas of application. 29
2.4 Examples of centrality measures for node removal. 34
2.5 Examples of centrality measures for edge removal. 36
2.6 Performance measures in studies of network robustness. 43

3.1 List of symbols for Chapter 3. 65
3.2 Acronyms of sampling methods for fPINs. 75
3.3 Numbers and structural properties of sampled fPINs. 83
3.4 Slopes of a linear least-squares regression of the mean local clustering coefficient C and the global clustering coefficient T. 90

4.1 List of symbols for Chapter 4. 108

5.1 List of symbols for Chapter 5. 139
5.2 Expectation and variance for log-minors for Examples E1, E2, E3, and E4. 151

Abstract

Robustness and entropy are widely used concepts in the study of networks in various disciplines. There exist various notions of network robustness and network entropy, because there are many network properties that can be robust to perturbations and there are many network-associated distributions for which one can calculate a notion of entropy. In this dissertation, we explore several notions of robustness and entropy — and connections between them — on networks. We investigate the robustness of structural properties of protein-interaction networks to the targeted removal of nodes and propose a test for identifying structural properties that are plausible proxies for measures of performance or viability of a protein-interaction network. Several researchers have suggested that structural redundancy and functional redundancy are important mechanisms by which a biological system can achieve robustness. A connection between redundancy and the mean subsystem entropy of a network with Ornstein–Uhlenbeck dynamics motivates our study of the relationship between the entropy of a multivariate Ornstein–Uhlenbeck process and the structure of an underlying network. We introduce *walk graphs* as a type of network motif that can help one understand the relationship between network structure and the covariance and entropy of a multivariate Ornstein–Uhlenbeck process. Because of the large number of subsystems in systems with as few as 20 nodes, studying the mean entropy of subsystems computationally is a difficult task. To address this issue, we derive sampling guarantees that indicate that one can calculate a very accurate sample mean of subsystem entropy by sampling very few subsystems.

The results in this dissertation offer new approaches to studying performance, viability, and robustness of systems on the basis of network structure and dynamics on networks. They demonstrate the importance of the choice of an appropriate performance measure for the study of network robustness and provide a framework for identifying plausible performance measures or plausible proxies of performance measures. The analysis of walk graphs and their contribution to covariance and entropy demonstrates that it can be useful to consider dynamics on a network (instead of just a network's structure) as a composite entity that one can decompose into many small parts. This decomposition of dynamics on networks provides a framework for studying the mechanisms by which dynamics and network structure contribute to covariance, entropy, redundancy, and other properties of dynamics on networks.

Acknowledgements

This work would not have been possible without the support of many people in many ways. I thank my advisor Mason Porter for exceeding all expectations as an advisor and providing invaluable advice, support, criticism, encouragement, mentorship, and friendship throughout my DPhil. I am grateful to my co-advisor Jonny Wray for his advice, support, mentorship, and for inspiring me to pursue unusual research questions in unusual ways. I also thank my co-advisor Philip Maini for his advice, support, mentorship, and his help in the various challenges that come with conducting research on one continent and pursuing a degree on another. My gratitude goes to my college mentor Alfons Weber for the warm welcome that I received when I arrived in Oxford and the cordial support when I decided to leave for Los Angeles. I also thank my advisors on previous research projects — James Anderson, Tobias Brandes, Thomas Dahms, Heather Harrington, Judith Lehnert, Antonis Papachristodoulou, Christina Pöltl, Eckehard Schöll — for their advice and support in the years that lead up to my DPhil studies. I am very grateful to Udo Mannel whose influence was crucial in igniting my passion for science and mathematics.

I thank Charlotte Deane, Victoria Flores, Heather Harrington, Brendan Jackson, Adam Sardar, and Marie Weston for helpful discussions on my research on protein-interaction networks. I am also grateful to Alex Arenas, Lionel Barnett, Heather Brooks, Bingni Brunton, Michelle Feng, Kameron Decker Harris, Neave O'Cleary, Gesine Reinert, Denis Serre, and David Spencer for helpful discussions and comments on my research of entropy-inducing motifs. Clément Canonne, Kameron Decker Harris, Michael Neely, and participants of the *IPAM Quantitative Linear Algebra Tutorials* provided helpful comments on my research on log-minor distributions.

During my time as a DPhil student, I was fortunate to take part in several programs that helped me find and claim my identity as a mathematician and as a scientist. I thank the Park City Mathematics Institute (PCMI) and the Institute for Pure and Applied Mathematics (IPAM) for hosting exceptional mathematical programs. I thank Tim Austin, Persi Diaconis, Tom Garrity, Henry Segerman, and Terry Tao, whose lectures at programs at the PCMI and the IPAM sparked and furthered my passion for mathematics pf various degrees of "purity". My thanks goes to David Feldman, Paul Hooper, and Carla Shedivy for organizing the Complex Systems Summer School (CSSS) 2018, which helped me mature greatly as a scientist. I thank the participants of the CSSS 2018, especially "Team Normal Networks" —

i.e., George Cantwell, Yanchen Liu, Benjamin Maier, Carlos Serván, Jordan Snyder, and Guillaume St-Onge — and "Team Neutral Networks" — i.e., Sarah Berkemer, Kofi Khamit-Kush, and Jacob Rutten — for being partners in research and crime during this program. I also thank Mark Newman and the Society of Young Network Scientists (SYNS) — especially Phil Chodrow, Ryan Gallagher, Brennan Klein, Tim LaRock, Sarah Shugars, Leo Torres, Cindy Wang, and Sara Williams — for giving me a warm welcome to the network science community. Special thanks goes to Phil Chodrow for being an awesome collaborator from whom I learned a lot.

I am also grateful to the many people who contributed to the research environments in which I loved to study, research, and write. My sincere thanks goes to Keith Gillow at the Oxford Mathematical Institute, and Mihai Duta, Dugan Witherick, and Steven Young of the support team at Oxford Advanced Research Computing (ARC) for guiding my first steps and enduring my first mistakes as a student with access to cluster computing facilities. Alison Franklin, Annette Miller, Sandhya Patel, John-Paul Roche, Karen Walker, and Francesca Wright helped me navigate the multitude of institutions that make up the University of Oxford. Judy Spiegelmann and Ana Josefa helped me keep balance and rhythm in my life after moving to Los Angeles. I thank Roxana Pamfil, Fabian Ying, and "Team Sparrowhawk" — i.e., Liam Brown, Joshua Bull, Chris Hutchinson, Laurence Hutton-Smith, Florian Klimm, Tom Lewin, Hannah Patel, Tammo Rukat, Nafisa Sharif, Bernadette Stolz-Pretzer, and Naushad Velgy — for their friendship and support during my time in Oxford and afterwards. I am grateful to "Team Mason" — i.e., Michelle Feng, Yacoub Kureh, and Will Oakley — for their friendship and support during my time in Los Angeles. I thank Sewook Oh for being on all teams and no teams and having boba with me on the lawn.

I am very grateful to the Clarendon Fund, the Engineering and Physical Sciences Research Council, and e-Therapeutics plc for funding my DPhil studies. I thank e-Therapeutics, the Society for Applied and Industrial Mathematics, PCMI, and SYNS for their financial support with visits to conferences, workshops, and summer schools that I undertook during my DPhil.

Finally, I thank my husband and my family who have been with me on this journey from the very beginning. I am grateful to my parents and my brother for their love, encouragement, support, and patience. I thank my husband Fachreddin for accompanying me on this journey and supporting me in every imaginable way through all of it.

Publications related to this DPhil

I adopted Chapters 3 and 4 from paper drafts that are currently in preparation. Chapter 5 is based on a paper that is available on arXiv [474]:

> A. C. Schwarze, P. S. Chodrow, and M. A. Porter. Log-minor distributions and an application to estimating mean subsystem entropy. *arXiv preprint arXiv:1901.09456*, 2019.

I co-authored a paper on the structure of networks that one obtains from thresholding correlated normally distributed data [79]:

> G. Cantwell, Y. Liu, B. F. Maier, A. C. Schwarze, C. A. Serván, J. Snyder, and G. St-Onge. Thresholding normally distributed data creates complex networks. *arXiv preprint arXiv:1902.08278*, 2019.

The latter work is not part of this dissertation.

Chapter 1

Introduction

Robustness is an important concept for many applications in various fields of research. Many researchers have reported that network robustness is linked to dynamics on networks [13, 139, 144, 198, 203, 225, 288, 302, 379, 490, 522] and to various notions of entropy [89, 139, 207, 294, 352, 353, 528, 563, 578, 590].

In this dissertation, we review the study of network robustness and examine several notions of robustness in protein-interaction networks (PINs). We consider structural redundancy and functional redundancy as mechanisms by which systems can achieve robustness. We also extend an information-theoretic framework that one can use to quantify structural redundancy and functional redundancy in networks by examing the links between entropy of a dynamical process on a network and network motifs. We also derive sampling guarantees for mean differential entropy of subsystems.

In this chapter, we give a brief motivation for our approach to the study of robustness, and we motivate our interest in structural redundancy and functional redundancy in biological systems. We end this chapter by outlining the structure of this dissertation.

1.1 Motivation

The study of robustness is relevant to many scientific disciplines. Our motivation for the study of robustness arises from applications of robustness in biomedical research. In biomedicine, there are systems for which researchers investigate ways to increase their ability to fulfill a desired function "robustly" [144, 286]. Examples include many systems that contribute to human health and survival, such as the human gut microbiome [263, 282] or biological pathways that are necessary for the healthy growth of cells or tissues [189, 575]. For other biomedical systems,

robustness can be an undesired property. Researchers have investigated biological systems with the aim of determining interventions that can cause large disruptions [23, 25, 128, 241, 242, 286, 287, 523]. Examples include parasites [300], tumors [25, 242, 287, 523], and biological pathways of diseases [23, 128, 241, 242]. The search for effective therapeutic interventions that can inhibit tumor growth is a particularly active field of research [26, 283, 432, 457]. Researchers have attempted to use drugs to block pathways that facilitate tumor growth [432]. Some of these studies have demonstrated that some tumors are able to adapt to use new biological pathways when therapeutic interventions block the biological pathways that previously facilitated the tumor's growth [26]. By using alternative biological pathways, a tumor thus becomes resistant to the therapy.

A major challenge facing attempts to discover effective tumor therapies is the prediction of a tumor's resistance to a drug [155, 340]. Many candidate drugs for tumor therapy turn out to be ineffective because the targeted tumor can become resistant to the candidate drug [361, 493]. Frequently, a candidate drug goes through several long and costly stages of drug discovery and drug design, tests in animals, and clinical trials before researchers ultimately find that the targeted tumor can become resistant to the drug. An improved understanding of the mechanisms through which tumors can become resistant to therapeutic interventions can greatly improve the efficiency of drug discovery and decrease the numbers of animals and persons that participate in tests of ultimately unsuccessful drugs [493].

Many experimental studies of tumor resistance rely on animal tests [493], which tend to be slower and more expensive than mathematical and computational studies [183]. By contrast, the study of molecular interactions between proteins and other biomolecules, as well as the collection of gene-interaction data, has become fast and affordable through high-throughput screening methods [183]. These experiments can yield relational data between hundreds or thousands of genes, proteins, and other biologically relevant molecules [179, 232, 546]. These relational data facilitate the construction of gene-interaction networks and PINs [541]. For many biological systems, data for constructing associated genetic and biomolecular networks are publicly available [86, 465]. Motivated by the availability of these data, one can ask whether one can further understanding of the mechanisms of tumor resistance from studying associated genetic and biomolecular networks.

In reviewing the study of network robustness (see Chapter 2), we find that the outcome of a study of robustness depends crucially on the choice of a relevant notion of robustness and that the consideration of dynamical processes on networks and

other features of network models can guide the search for relevant notions of robustness. However, when one lacks information about what dynamics are occurring on a network and other features of network models, it is typically difficult to find relevant notions of robustness. This is the case for the study of robustness of PINs. An interaction in a PIN can reflect one or several of many different types of interactions that can occur between two proteins[1] [38, 151, 229] and, to our knowledge, there do not exist mechanistic models for dynamics on protein-interaction networks.

In this dissertation, we consider two approaches to this problem. First, we test the plausibility of several notions of robustness by examining the change of structural properties of protein-interaction networks under several types of structural perturbations (see Chapter 3). Second, we consider a multivariate Ornstein–Uhlenbeck process (see Chapter 4) as a proxy dynamical process for the study of robustness in biological networks. Various researchers have argued that viewing biological systems as systems that transmit and process information can yield insights into their functions [3, 497, 528, 578]. Researchers have identified many proteins that contribute to cellular functions by participating in *signaling pathways* [281, 382, 525], which are pathways that enable communication between cells [218]. For an Ornstein–Uhlenbeck process on a network, one can calculate both its differential entropy and the differential entropy of its subsystems. Tononi et al. (1999) proposed a framework for calculating structural redundancy and functional redundancy from the mean differential entropy of subsystems [535]. Many scholars have proposed that *structural redundancy* and *functional redundancy* are important mechanisms by which biological systems achieve robustness [99, 154, 309, 437, 455, 535, 559]. Motivated by this link between robustness of biological systems and differential entropy of dynamical processes on networks, we study an Ornstein–Uhlenbeck process on a network and link its differential entropy to motifs in networks (see Chapter 4). We also derive sampling guarantees for the mean differential entropy of subsystems (see Chapter

[1] Various biochemical reactions require physical bonds between a pair of proteins or the formation of a protein complex [38, 151, 229]. For example, *molecular chaperones* are proteins that assist in the assembly, disassembly, folding, or unfolding of proteins and other biomolecules [229]. A molecular chaperone A that assists in the folding of a protein B binds to an unfolded protein B, induces a change in the conformation of B, and unbinds again. This process corresponds to a single edge $e_{A,B}$ in a PIN. As another example, one can consider *protein redox reactions*, which are important reactions for cellular signal transduction. In a protein redox reaction, an electron is transferred from a protein A to a protein B [38]. A direct transfer of an electron from A to B can correspond to a single edge $e_{A,B}$ in a protein-interaction networks. Redox reactions can also involve a protein C that acts as an *electron transfer protein* [38, 219]. In that case, the protein C binds to A and accepts an electron from it. Simultaneously or subsequently, it binds to B and transfers the electron to B. In a protein-interaction network, such a redox reaction corresponds to two or three edges in a protein-interaction network: $e_{A,C}$, $e_{B,C}$, and (if A, B, and C form a complex) $e_{A,B}$.

5). We anticipate that these sampling guarantees can facilitate the study of links between network structure, the mean differential entropy of subsystems, structural redundancy, functional redundancy, and robustness in biological systems.

1.2 Structural and functional redundancy in biological networks

In this section, we briefly explain the concepts of structural redundancy and functional redundancy. We also summarize a framework that Tononi et al. (1999) proposed for quantifying structural and functional redundancy [535]. This framework links redundancy in networks to the mean differential entropy of subsystems of a dynamical process on a network. This link motivates our study of differential entropy and mean differential entropy of subsystems in Chapters 4 and 5.

1.2.1 Distinguishing structural and functional redundancy

Redundancy indicates the existence of two or more components in a system that can perform the same function [309]. *Structural redundancy* indicates the existence two or more components that are structurally similar and that can perform the same function [154, 437, 535]. *Functional redundancy* (which has also been called "degeneracy" [154, 323, 359, 535, 579] and "distributed robustness" [344, 444, 559, 582]) indicates the existence of two or more components that are structurally different and that can perform the same function [309, 336, 455]. In many studies of biological robustness, researchers do not distinguish between structural and functional redundancy [349]. However, several researchers have pointed out that the mechanisms through which structural redundancy and functional redundancy contribute to a system's robustness are different and have consequently argued that it is important to distinguish between these two concepts [154, 309, 437, 503, 530, 535, 579–582].

From an engineering perspective, it is easy to see how structural redundancy contributes to the robustness of a system. If the performance of a system relies on the functionality of several essential subsystems, having duplicates of these subsystems makes the system robust to failure of one of its subsystems. Let us illustrate this with a couple of brief examples. In some respects, the existence of multiple alleles of one gene or the existence of multiple identical organs (such as kidneys and lungs in the human body) suggest that biological redundancy is relevant for robustness in real biological systems. However, carefully conducted studies of biological robustness (e.g., studies of gene knock-outs in yeast or bacteria [252, 558]) found that the

robustness of these systems cannot rely solely on redundancy in their genomes. Instead, multifunctionality of genes and overlaps in gene function (i.e., functional degeneracy) are more important than mere gene redundancy [154, 360, 558, 559].

1.2.2 Quantifying structural and functional redundancy

Tononi et al. (1999) proposed a measure of redundancy based on information propagation in a set of nodes [535]. Consider a set V of nodes $v_1, v_2, \ldots, v_N$ with state variables $x_1, x_2, \ldots, x_N \in \mathcal{D}$ that are determined by a stochastic process. At a stationary state, every variable x_i has a stationary probability distribution $p_i(x)$. When the domain $\mathcal{D}$ is a set of finite size, one can use these probability distributions to compute the Shannon entropy

$$H_S(v_i) := -\sum_{x \in \mathcal{D}} p_i(x) \log(p_i(x)) \tag{1.1}$$

for each node. Using joint distributions $p_{i,j,k,\ldots}(x, y, z, \ldots)$, one can also assign a value of Shannon entropy or differential entropy to a set $\{v_i, v_j, v_k, \ldots\}$ of nodes. That is,

$$H_S(\{v_i, v_j, v_k, \ldots\}) = -\sum_{x,y,z \ldots \in \mathcal{D}} p_{i,j,k\ldots}(x, y, z \ldots) \log(p_{i,j,k\ldots}(x, y, z, \ldots)) \,.$$

Tononi et al. (1999) used mutual information

$$\mathrm{MI}(V_1, V_2) := H_S(V_1) + H_S(V_2) - H_S(V_1 \cup V_2) \tag{1.2}$$

as a measure of influence[2] of a node set $V_1 \subseteq V$ on a node set $V_2 \subseteq V$ [535].

Components of a system can be redundant with respect to one function of a system while not being redundant with respect to a different function of a system. To measure redundancy with respect to a system function, one needs a formal description of a "function" of a system. Tononi et al. (1999) considered the function of a system to be to maintain a given state for a subset of nodes [535]. They partitioned

[2] A shortcoming of mutual information as a measure of influence is that it is symmetric in its input (i.e., $\mathrm{MI}(V_1, V_2) = \mathrm{MI}(V_2, V_1)$). Tononi et al. (1999) proposed a definition of *asymmetric mutual information* AMI to ensure that one measures only the influence of V_1 on V_2 and not the influence of V_2 on V_1. However, the proposed definition of AMI was ambiguous. In subsequent work on information-theoretic definitions of functional redundancy, Randles et al. (2011) and Li et al. (2012) used symmetric mutual information in Eq. (1.2) to measure the influence of V_1 on V_2 [323, 444]. In this dissertation, we follow their approach, but point out that the development of an unambiguous definition of AMI is an important avenue of future research. Possible solutions to this problem include the use of transfer entropy [472, 507, 553], the use of Granger causality [30, 73, 205], and the study of impulse-responses in multivariate linear input-output systems [466].

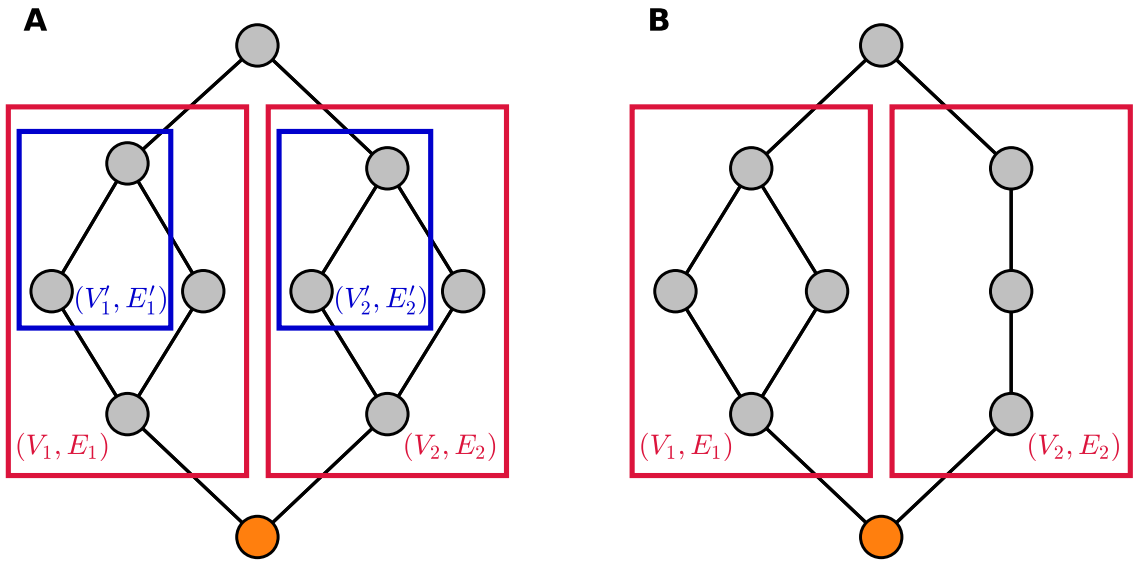


Figure 1.1: Examples of structural and functional redundancy. In (**A**), we show a network (V, E) with a partition of V into a kernel set V_K (grey nodes) and an output set V_O (orange node). For two structurally redundant subnetworks (V_1, E_1) and (V_2, E_2), a subnetwork (V_1', E_1') of (V_1, E_1) has a correspondng subnetwork (V_2', E_2') of (V_2, E_2) such that (V_1', E_1') and (V_2', E_2') are structurally redundant. In (**B**), we show a partition of (V, E) into a kernel set and an output set with functionally redundant subnetworks (V_1, E_1) and (V_2, E_2).

V into a *kernel set* V_K and an *output set* V_O. They then associated the influence of the kernel set on the output set with $\mathrm{MI}(V_K, V_O)$. They also associated the influence of each node $v_i \in V_K$ on the output with $\mathrm{MI}(\{v_i\}, V_O)$.

Tononi et al. (1999) argued that if the sum of influences $\mathrm{MI}(\{v_i\}, V_O)$ of individual nodes on the output set is equal to the influence $\mathrm{MI}(V_K, V_O)$ of the kernel set on the output set, then every node $v_i \in V_K$ contributes differently to the output set. Removing any node from V_K changes the influence of V_K on V_O. In that case, there is no redundancy in V_K with respect to V_O. However, if the sum of influences of individual nodes on the output set is larger than the influence of the kernel set on the output set, then subsets of V_K must overlap in their contribution to the kernel set's influence on the output set. In that case, there is redundancy in V_K with respect to V_O. Tononi et al. (1999) defined the following measure of redundancy:

$$R(V_K, V_O) := \sum_{v_i \in V_K} \mathrm{MI}(\{v_i\}, V_O) - \mathrm{MI}(V_K, V_O)\,. \tag{1.3}$$

For distinguishing structural and functional redundancy, it can be helpful to

consider V to be the set of nodes of a network[3] (V, E) and subsets of V to be the node sets of subnetworks[4] of (V, E). In Fig.1.1, we show examples of a system with structurally redundant subsets and a system with functionally redundant subsets. If two sets $V_1, V_2 \subset V_K$ are structurally redundant with respect to V_O, the networks (V_1, E_1) and (V_2, E_2) have similar structures and for each subnetwork (V_1', E_1') of (V_1, E_1), there exists a subnetwork (V_2', E_2') of (V_2, E_2) such that (V_1', E_1') and (V_2', E_2') are also structurally redundant (see Fig. 1.1(**A**)). By contrast, if $V_1, V_2 \subset V_K$ are not structurally redundant with respect to V_O, the networks (V_1, E_1) and (V_2, E_2) have different structures and there exists at least one subnetwork of (V_1, E_1) such that one cannot find a structurally redundant subnetwork in (V_2, E_2) (see Fig. 1.1(**B**)).

Based on the above considerations, Tononi et al. (1999) used the redundancy of subsets of nodes to distinguish between structural and functional redundancy [535]. They defined

$$R_F(V_K, V_O) := \sum_{n=1}^{N} \left[\frac{n}{N} R(V_K, V_O) - \langle R(V_K', V_O) \rangle_{|V_K'|=n} \right] \tag{1.4}$$

as a measure of functional redundancy. In Eq. (1.4), we use V_K' to denote a subset of V_K and $\langle \cdot \rangle_{|V_K'|=n}$ to denote the mean over all subsets V_K' with n nodes. One can then define $R_S(V_K, V_O) := R(V_K, V_O) - R_F(V_K, V_O)$ as a measure of structural redundancy.

Tononi et al. (1999) also showed that one can express R_F as a function of mutual information of subsets of nodes. That is [535]:

$$R_F(V_K, V_O) = \sum_{k=n}^{N} \left[\langle \mathrm{MI}\,(V_K', V_O) \rangle_{|V_K'|=n} - \frac{n}{N} \mathrm{MI}(V_K, V_O) \right] . \tag{1.5}$$

Using Eq. (1.2), one can thus express R, R_F, and R_S as functions of the mean entropy of subsystems. Equations (1.2), (1.3), and (1.5) establish links between the mean entropy of subsystems of a stochastic process on a network and the network's structural and functional redundancy.

When the probability distributions $p_i(x)$ are continuous (i.e., the size of the domain $\mathcal{D}$ is infinite), one cannot compute the Shannon entropy in Eq. (1.1). Tononi et al. (1999) considered continuous probability distributions and replaced Shannon entropy with the differential entropy $\int_{x \in \mathcal{D}} p_i(x) \log(p_i(x))$ (see Section 4.2.1). A

[3] We consider a *network* (V, E) to be an object that is defined by a set V of nodes and a set E of edges (see Section 2.2).

[4] A *subnetwork* of a network (V, E) is a network (V_S, E_S) with node set $V_S \subseteq V$ and edge set $E_S \subseteq E$ (see Section 2.2).

caveat of differential entropy that it is not the limit of the Shannon entropy of a distribution on $\mathcal{D}$ as the size of $\mathcal{D}$ tends to infinity. However, continuous mutual information, which one can calculate via Eq. (1.2) from differential entropy, is the limit of discrete mutual information of two distributions on $\mathcal{D}$ as the size of $\mathcal{D}$ tends to infinity [112, p. 251]. One can thus calculate R, R_F, and R_S for continuous distributions $p_i(x)$ of node states from continuous mutual information and, in turn, from differential entropy.

This link between redundancy and differential entropy motivates our study of the differential entropy of an Ornstein–Uhlenbeck process on a network (see Chapter 4) and our study of sampling guarantees for mean differential entropy of subsystems (see Chapter 5).

1.3 Structure of this dissertation

In Chapter 2, we review the study of robustness of networks to structural changes. A major challenge for research on this topic is that studies have reported different and seemingly contradictory network properties to be correlated with a network's robustness. Using the framework by Alderson and Doyle (2010), one can resolve many seemingly conflicting results of studies of network robustness by distinguishing different notions of robustness [14]. We categorize various studies by their notions of robustness and discuss the conditions under which one can expect different studies of robustness to yield similar results. We also discuss how knowledge about a given application can guide researchers in identifying relevant notions of robustness.

In Chapter 3, we compare several notions of robustness for protein-interaction networks. Using deterministic seed-based sampling methods, we construct function-specific protein-interaction networks for biological pathways. We examine several notions of robustness to node removal on these networks. To identify relevant measures of performance, we construct a plausibility test for measures of performance of function-specific protein-interaction networks. Using this test, we find that the relative size of a network's largest connected component (see Section 2.4.4.3) and its algebraic connectivity (see Section 2.4.4.6) tend to be plausible measures of performance.

In Chapter 4, we examine links between motifs in a network's structure and the differential entropy of a multivariate Ornstein–Uhlenbeck process on the network. We introduce *walk graphs* as a type of motif for processes on networks. Walk graphs in a network can contribute to covariance and differential entropy of an Ornstein–Uhlenbeck process on that network. We calculate the contributions of walk graphs

to covariance and differential entropy and identify the walk graphs that contribute the most. With this study of walk graphs and their contributions to covariance and differential entropy, we demonstrate that it can be useful to consider dynamics on a network (instead of just a network's structure) as a composite entity that one can decompose into many small parts. This decomposition of dynamics on networks into walk graphs opens new avenues for studying the mechanisms by which dynamics and network structure contribute to covariance, differential entropy, redundancy, and other properties of dynamics on networks.

In Chapter 5, we derive sampling guarantees for the mean differential entropy of subsystems. Motivated by the link between structural redundancy, functional redundancy, and mean differential entropy of subsystems (see Eqs. (1.2), (1.3), and (1.5)), we consider the problem of calculating the mean differential entropy of subsystems for a multivariate normal distribution. A system with N variables has $\binom{N}{n}$ subsystems of size n. The large number of subsystems for a system with as few as $N = 20$ nodes makes the exact computation of mean values over the set of n-size subsets computationally very expensive. We derive upper bounds for the standard error of the sample mean of the differential entropy of subsystems. Our results demonstrate that, despite the rapid growth of the set of subsystems with N, the number of samples that are needed to bound the sampling error is asymptotically independent of N. Instead, it is sufficient to increase the number of samples in linear proportion to n to achieve a desired sampling accuracy.

In Chapter 6, we conclude this dissertation. We summarize salient results from Chapters 2, 3, 4, and 5 and indicate their relevance to possible avenues of future research.

Chapter 2

Robustness of networks to structural changes

The study of robustness of networks is important in many scientific disciplines. An improved understanding of network robustness can aid the development of robust infrastructure, transportation, and communication networks. It can also help anticipate major disruptions in ecological systems and financial systems, and it can facilitate the development of cancer therapies.

For these and other applications, scholars have studied the robustness of networks to the removal of nodes or edges (or other network components) to identify and characterize robust network structures. A major challenge in this endeavor is that studies have reported different and seemingly contradictory network properties to be correlated with a network's robustness. Alderson and Doyle (2010) proposed a framework for distinguishing various notions of robustness by four specifications: a network model, a measure of robustness, a set of permissible perturbations, and a measure of performance [14]. Using this framework, one can resolve many seemingly conflicting results of studies of network robustness. We categorize various notions of robustness that researchers have considered using these four specifications. We survey studies of network robustness with a focus on (1) identifying robustness specifications that researchers have used often, (2) understanding when these specifications are appropriate for applications, and (3) understanding the conditions under which one can expect different studies of robustness to yield similar results.

Symbol	Meaning
(V, E)	a network with node set $V = \{v_i\}_{i=1,\dots,N}$ and edge set $E = \{e_{i,j}\}$
(V_S, E_S)	a subnetwork with node set $V_S \subseteq V$ and edge set $E_S \subseteq E$
(V_P, E_P)	a path in a network
$\mathcal{S}$	a set of networks
$E_i^{(\ell)}$	the edge set of the ℓ-hop neighborhood of node v_i
∂V_S	the border of a node subset V_S
w	a walk in a network
ℓ	the length of a path or walk
$\tilde{n}_{i,j}$	a number of shortest paths from node v_i to node v_j
$n_\circ(\ell)$	a the number closed length-ℓ walks
n	number of removed nodes
f	a fraction of removed nodes
I	the impact of a perturbation
k	a node's degree
C_i	a node's local clustering coefficient of node v_i
d_{ij}	the geodesic distance from node v_i to node v_j
Ξ_{ij}	the indicator for a path from v_i to v_j
N	a network's size (i.e., number of nodes)
m	a network's number of edges
n_c	the number of components in a network
$\overline{N}_c$	the mean component size in a network
N_{LCC}	the size of a network's largest connected component
F	the relative size of a network's largest connected component
r	a network's reachability
$\mathbf{A} = (a_{ij})$	a network's adjacency matrix
$\mathbf{L} = (l_{ij})$	a network's combinatorial Laplacian
$p(k)$	a network's degree distribution
H	a network's graph entropy
C	a network's mean local clustering coefficient
T	a network's global clustering coefficient
L	a network's mean shortest-path length
$\mathcal{E}$	a network's efficiency
κ_v	a network's edge connectivity
κ_e	a network's node connectivity
a	a network's algebraic connectivity
h_e	a network's edge expansion
Ω	a network's resistance distance
$\kappa(i, j)$	a network's communicability function
K	a network's natural connectivity

Table 2.1: List of symbols in Chapter 2.

Symbol	Meaning
λ_i	eigenvalue of a matrix
$\boldsymbol{\eta}$	eigenvector of a matrix
ρ	the spectral radius of a matrix
τ_c	the epidemic threshold
$\mathbf{X}$	the state of a dynamical system
$\mathbf{S}$	the stoichiometric matrix of a chemical reaction network
$\tilde{\mathbf{v}}$	a flux vector
μ	a biomass production rate

Table 2.2: List of symbols in Chapter 2 (continued).

2.1 Introduction

Across many scientific disciplines, researchers have modeled various real-world systems as networks and have studied the robustness of these network models to perturbations [76, 85, 118, 192, 306, 584]. Robustness is a crucial concept for many applications of networks. For example, the study of the robustness of infrastructure networks (e.g., power grids or road networks) can help anticipate the performance of these systems in the event of failures of components or various forms of adversarial attacks [118, 192]. Examining robustness of social networks can help understand how communities of people, animals, or other individuals can react to various perturbations (e.g., as a natural disaster or the loss of one or several community members) [45, 380]. The study of the robustness of ecological systems can inform predictions on the effects of climate change or species extinction on an ecological system [152, 469]. Understanding the robustness of biological systems is an important part of understanding the existence of life on earth [403, 579]. Organisms are subjected frequently to many environmental changes, and genetic changes from one generation to the next are abundant [383]. The study of genetic robustness can help understand how a genome that is subjected to many changes can robustly code for a phenotype [357, 579]. Robustness of biological systems is also relevant to applications in biomedicine [286]. Depending on the application, one can view robustness as a desirable property and search for ways to increase it. For example, one can view the human body as a system whose function or task is to survive. One can also study the robustness of subsystems of an organism (e.g., organs, cells, or microbiomes) [286]. For these applications, an understanding of system robustness can help identify mechanisms by which one can therapeutically enhance robustness.

By contrast, one can view a cancer, a viral infection, and other diseases as biological systems that one wants to disrupt by a therapeutic intervention. Understanding the robustness of these systems can help in the design of therapeutic interventions that can circumvent fail-safe mechanisms in these systems and thus cause their disruption [287].

These and other applications have generated interest in identifying "robust network structures" [231, 423, 549, 604]. A major problem in this endeavor is that many researchers have reported various (and sometimes seemingly contradictory) network properties to be correlated with a network's robustness. For example, Dunne et al. (2002) investigated the robustness of a food web to species loss and concluded that robustness increases with increasing network density [152]. By contrast, Vieira and Almeida-Neto (2015) examined the robustness of a mutualistic network to species loss and concluded that robustness increases with decreasing network density [555]. Newman (2002) reported that a network's robustness to node removal increases with degree–degree assortativity [392]. However, Zhou et al. (2012) reported that the robustness of a multilayer network to node removal decreases with degree–degree assortativity [609]. These examples demonstrate that it is difficult, and perhaps impossible, to identify network properties that are positively correlated with robustness in different systems. Van Mieghem (2010) stated that "the rather simple but highly relevant question 'What is a robust network?' seems beyond the realm of present understanding" [550].

Alderson and Doyle (2010) argued that one can resolve seeming contradictions between studies of robustness by distinguishing different types of robustness by several specifications [14]. They proposed to define robustness as a measure of invariance of a property of a system to a set of perturbations. In general, different choices for the measure of invariance, the system, the system property, and the set of perturbations lead to different notions of robustness that do not need to be correlated with each other. This framework helps recognize that many studies of robustness consider different notions of robustness and thus do not need to lead to similar conclusions regarding the characteristics of a "robust network structure". Recognizing that studies of robustness consider various different notions of robustness, one can ask if there are general lessons to be learned from the research of network robustness. Are there conditions under which one can expect different notions of robustness to lead to similar results? Which notion of robustness is most appropriate for the study of a given application? In this chapter, we review research on network robustness. The focus of this review is on (1) identifying robustness specifications that researchers use often, (2) understanding when these specifications are appropriate

for applications, and (3) understanding the conditions under which one can expect different studies of robustness to yield similar results.

To discuss these questions, it is important to distinguish between real-world systems and the models that researchers use to study those systems. Throughout this chapter, we use the word "system" to refer to a real-world object that a researcher wishes to study. A "model" is an abstraction of a system. For this review, we focus on network models, which are abstract models that capture the interactions between components of a system. These models include a *network* (see Section 2.2.1) and can also include additional features (see Section 2.4.1).

This chapter proceeds as follows. In Section 2.2, we give definitions of graph-theoretic concepts that are useful for the developments in this chapter and the following chapters. In Section 2.3, we adapt Alderson and Doyle's definition of robustness and compare this notion of robustness to related concepts. In Section 2.4, we survey robustness specifications that researchers have considered in the study of network robustness and discuss the relevance of these specifications to real-world systems. We conclude our review in Section 2.5.

2.2 Simple networks, subnetworks, and matrix representations

In this section, we give definitions for graph-theoretic concepts that are useful for the study networks and their robustness to structural perturbations.

2.2.1 Networks and simple networks

We consider a *network* (V, E) to be an object that is defined by a set $V = \{v_i\}$ of nodes (i.e., vertices) and a set $E = \{e_{i,j}\}$ of edges. In a *directed* network, edges are directed (i.e., the order of indices i, j matters). In an *undirected* network, the order of indices i, j does not matter and the notations $e_{i,j}$ and $e_{j,i}$ denote the same edge. We call an edge $e_{i,i}$ a *self-edge*. One can construct network models $(V, E, \dots)$ that include additional information about a system by, for example, including a set of node labels or a set of edge weights. We briefly discuss such models in Section 2.4.1. We call an undirected network (V, E) with distinct edges $e_{i,j}$ and no self-edges a *simple* network.

We denote the *size* (i.e., the number of elements) in a set $\mathcal{S}$ by $|S|$. For a network (V, E), we write $N := |V|$ and $m := |E|$. We call N the *size of a network*.

2.2.2 Subnetworks

A *subnetwork* (or *subgraph*) of a network (V, E) is a network (V_S, E_S) with node set $V_S \subseteq V$ and edge set $E_S \subseteq E$ [60].

Paths. A *path* is a subnetwork (V_P, E_P) with node set $V_P = \{v_i\}_{i=1,\dots,\ell+1} \subseteq V$ and edge set $E_P = \{e_{i,i+1}\}_{i=1,\dots,\ell}$, where an edge $e_{i,j}$ starts at node v_i and ends at node v_j and $v_1, \dots, v_{\ell+1}$ are distinct nodes[1] [60, p. 4]. We say that the nodes $v_i, v_j \in V$ are *connected by a length-ℓ path* if there exists a path with $v_1 = v_i$ and $v_{\ell+1} = v_j$. We say that two paths (V_{P_1}, E_{P_1}) and (V_{P_2}, E_{P_2}) are non-overlapping if $E_{P_1} \cap E_{P_2} = \emptyset$.

Node-induced subgraphs. A *(node-)induced subgraph* of (V, E) on a node set V_S is a subnetwork (V_S, E_S) of (V, E), where E_S includes all edges in E that connect pairs of nodes in V_S [529]. The *ℓ-hop neighborhood* of a node $v_i \in V$ is the node set that consists of all nodes $v_j \in V$ that are connected to v_i by a length-ℓ path [60]. The *degree* k_i of a node v_i is the number of nodes in V that are connected to v_i by an edge in E. In a simple network, the degree of a node is equivalent to the number of nodes in its 1-hop neighborhood.

Connected subnetworks and connected components. In an undirected network (V, E), a *connected* subnetwork (V_S, E_S) is a subnetwork such that for every pair of nodes $v_i, v_j \in V_S$, there exists a path from v_i to v_j in (V_S, E_S). One can partition a network into one or several connected subnetworks. The coarsest partition in connected subnetworks gives the *connected components* of a network [156, p. 54]. In directed networks, one distinguishes between strongly and weakly connected components[2] [156, p. 53]. Unless we specify otherwise, we use the term "connected component" to refer a connected component in an undirected network or a strongly connected component in a directed network. A *connected network* has only one connected component. A network is *fragmented* if it is not connected. The *size* of a connected component is its number of nodes. Many researchers have used measures

[1] In directed networks, one can distinguish *directed paths* and *undirected paths*. The definition for directed paths in directed networks is equivalent to the definition of paths in undirected networks. An undirected path in a directed network is a subnetwork (V_P, E_P) with $V_P = \{v_i\}_{i=1,\dots,\ell+1}$ such that for every pair of nodes $v_i, v_{i+1} \in V_P$, the edge set E_P includes either $e_{i,i+1}$ or $e_{i+1,i}$.

[2] A *strongly connected* subnetwork of a network is a subnetwork (V_S, E_S) such that for every pair of nodes $v_i, v_j \in V_S$, there exists a directed path from v_i to v_j and a directed path from v_j to v_i in (V_S, E_S) [536, p. 43]. A *weakly connected* subnetwork of a network is a subnetwork (V_S, E_S) such that for every pair of nodes of $v_i, v_j \in V_S$, there exists an undirected path from v_i to v_j in (V_S, E_S). Every strongly connected subnetwork of a network is also a weakly connected subnetwork, but, in general, not every weakly connected subnetwork is a strongly connected subnetwork.

of performance that are related to the number and size of a network's connected components.

2.2.3 Walks and closed walks

A *walk* in (V, E) is a sequence

$$w = (v_1, e_{1,2}, v_2, e_{2,3}, \ldots, e_{\ell-1,\ell}, v_\ell, e_{\ell,\ell+1}, v_{\ell+1})$$

of nodes $v_1, v_2, \ldots, v_\ell, v_{\ell+1} \in V$ and edges $e_{1,2}, e_{2,3}, \ldots, e_{\ell-1,\ell}, e_{\ell,\ell+1} \in E$ such that each edge $e_{i,j}$ starts at the node v_i and ends at the node v_j [60, p. 4]. A *length-ℓ closed walk* is a sequence

$$(v_1, e_{1,2}, v_2, e_{2,3}, \ldots, e_{\ell-1,\ell}, v_\ell, e_{\ell,1}, v_1)$$

of nodes $v_i \in V$ and of edges $e_{i,j} \in E$ [60]. In walks and closed walks, neither the nodes v_i nor the edges $e_{i,j}$ need to be distinct. Because walks and closed walks can be sequences of non-distinct nodes and edges, walks and closed walks are not subnetworks.

2.2.4 Matrix representations of networks

Two popular matrix representations of networks are the adjacency matrix and the (combinatorial) Laplacian matrix [398, p. 142]. For a network (V, E) with N nodes, the *adjacency matrix* is the $N \times N$ matrix $\mathbf{A} = (a_{ij})$, where [169]

$$a_{ij} := \begin{cases} 1\,, & \text{if } e_{i,j} \in E\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

The (combinatorial) *Laplacian matrix* of (V, E) is the $N \times N$ matrix $\mathbf{L} = (l_{ij})$, where [398]

$$l_{ij} := \begin{cases} k_i\,, & \text{if } i = j\,, \\ -1\,, & \text{if } (i,j) \in E \text{ and } i \neq j\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

One can represent networks with edge weights by a *weighted* adjacency matrix[3] or a weighted (combinatorial) Laplacian[4]. There are several approaches to representing the structure of multilayer networks in matrix form or tensor form [289]; we do not consider these in this review.

[3] In a weighted adjacency matrix, elements a_{ij} are equal on the weight of the edge $e_{i,j}$ if $e_{i,j} \in E$ and 0 otherwise.

[4] In a weighted (combinatorial) Laplacian, non-diagonal elements l_{ij} are equal to the negative of the weight of an edge if $e_{i,j} \in E$ and 0 otherwise. Diagonal elements are equal to a node's *weighted degree* (i.e., *node strength*), which is equal to the sum of the weights of incident edges.

2.3 System robustness and related concepts

In this section, we give a definition of robustness and discuss many related concepts. We also give an overview of reviews of robustness in various fields of research.

2.3.1 Defining robustness

In a sequence of several publications, Carlson, Doyle, Alderson, and collaborators developed a frequently referenced approach of system robustness [14, 81–83, 149, 150]. Their approach characterizes most systems as "robust yet fragile" [150]. They proposed that, for most systems, one can find some aspects that are "robust" and others that "fragile" and that an assessment of "robustness" and "fragility" in a system requires several specifications [14]. In particular, Alderson and Doyle (2010) proposed that a "[property] of a [system] is *robust* if it is [invariant] with respect to a [set of perturbations]", where they used brackets to indicate the concepts that require specification before one can attempt to study a system's robustness [14]. For example, it is impossible to answer the question "Are bacteria robust?" without specifying the system (i.e., the bacterial species), the system property (i.e., some property of bacteria, such as movement, growth, reproduction rate, and so on), the set of permissible perturbations (e.g., changes in temperature, pressure, nutrient availability, and so on), and how we measure "invariance" of a system property under perturbation. Here, we use the term "invariant" to be faithful to the definition in Ref. [14]. However, in most practical settings, system properties almost always change under perturbations, and they are very rarely invariant (i.e., unchanged). We thus adopt the following slightly modified definition by Alderson and Doyle (2010):

> *Robustness* is the [insensitivity] of a [property] of a [model of a system] to a [set of perturbations].

We call the problem of determining a system's robustness a *robustness problem* and require the following specifications for a robustness problem:

1. the specification of a system and (if the study is not an observational study) the specification of an experimental, computational, or theoretical model for the specified system;
2. the specification of a measure of insensitivity (i.e., robustness);
3. the specification of a set of perturbations to which one subjects the specified performance measure;

4. and the specification of a property of the specified model as a measure of system performance.

Considering these specifications can help resolve many apparent contradictions between the conclusions of different studies of robustness [14]. In Section 2.4, we discuss the specifications of a robustness problem for computational studies of network models of systems. In the remainder of the current section, we review concepts that are related to robustness. Some of these concepts have different names but a similar meaning to the definition of robustness from Alderson and Doyle (2010). Other concepts are different from the notion of robustness that we use in this review, but researchers have considered them to be synonyms of robustness in previous studies.

2.3.2 Different name, same concept?

Researchers in different fields have used a variety of different words to describe responses of systems to perturbations. Examples include stability [147, 317, 509], reliability [118, 330, 445], and resilience [47]. In many texts, researchers have used these terms as synonyms for robustness and to each other [47, 118, 147, 562]. Therefore, a clear disambiguation of these terms is not possible without disregarding a vast amount of literature on this topic. In this subsection, we consider words that researchers have used to refer to robustness, special cases of robustness, and concepts that are closely linked to robustness.

2.3.2.1 Stability

In the study of dynamical systems, it is very common to investigate the *asymptotic stability* of a system to small perturbations[514, p. 18]. A system is *asymptotically stable* to small perturbations if a small perturbation to a system's state dies out over time so that the system soon returns to its state before the perturbation [211, p. 3]. We can understand asymptotic stability to perturbations as a special case of robustness; asymptotic stability is the robustness of a dynamical system's state to a small perturbation [317, 509].

2.3.2.2 Reliability

In the engineering sciences, the reliability of a human-made system is the ability of a system "to perform a required function, under given environmental and operational conditions and for a stated period of time" [445]. Researchers have proposed several reliability definitions that are closely related to a system's robustness [118] or

resilience [103]. Recently, Cuadra et al. (2015) reviewed research on the robustness of power grids. They characterized reliability as "a beneficial property for a power grid that refers to its ability to supply electric loads with a high level of probability, during a given time interval" [118]. They considered robustness to measures the extent to which a power grid has high reliability.

2.3.2.3 Resilience

Robustness is the insensitivity of some property of a system to some perturbation. One measures robustness as some function of the difference between a system property before and immediately after a perturbation. Many scholars consider "resilience" to be a measure of the ability of a system to recover its previous state after a perturbation has caused a change in a system property [118, 244]. A system's resilience can therefore include processes in a system that can have several phases and can happen over some period of time. For example, Cuadra et al. (2015) considered a power-grid to be "robust" if black-outs are rare events [118]. They considered a power-grid to be "resilient" if blackouts are short-lived and the power-grid is able to recover its functioning state quickly. Bhamra et al. (2011) presented a comprehensive literature survey of the resilience of systems [47].

2.3.3 Same name, different concept?

The word "robustness" is widely-used across many fields of research. In this subsection, we list concepts that researchers have called "robustness" but that which differ from the notion of robustness in Section 2.3.1.

2.3.3.1 Performance

Consider a system and a measure X of performance. A possible measure of the impact of a perturbation is the absolute value $|X_0 - X'|$ of the difference of the performance X_0 of the unperturbed system and the performance X' of the system after perturbation. A measure of robustness should be large if perturbations lead to small impacts on a system's performance. If $X_0 \geq X'$, large values of X' lead to small impacts and, in turn, to large values of a robustness measure. Because of this positive correlation between performance after perturbation and robustness, researchers have used performance after perturbation as a proxy for robustness [597, 606].

2.3.3.2 Mechanisms of robustness

In many disciplines, the mechanisms by which the performance of a system can become robust to some type of perturbation are of great interest [285, 509]. Researchers have proposed many mechanisms by which systems can achieve robustness [154, 285, 356, 509]. In some studies, indicators and requirements for such mechanisms — such as measures of network redundancy [285, 298, 330, 509, 559] or modular structure of networks [285, 343, 489, 489, 509, 567] — are used as synonyms to robustness. For example, in several disciplines, researchers have proposed that a system can become robust by having several structurally different components that fulfill the same function within a system [309, 344, 535, 559, 579, 582]. Different researchers have referred to this property of a system's or network's structure by "functional redundancy" [309, 336, 455], "degeneracy" [535, 579], or "distributed robustness" [344, 559, 582].

2.3.3.3 Insensitivity of a system to parameter changes

Stelling et al. (2004) defined robustness as the "persistence of a system's characteristic behavior under perturbations or conditions of uncertainty" [509]. This definition includes the robustness definition in Section 2.3.1. However, it also includes the insensitivity of system properties with respect to uncertainty in system parameters. Robustness (i.e., insensitivity to perturbations) and persistence (i.e., insensitivity to parameter uncertainty) are different concepts. In studies of robustness, one compares the state of a system before and after a perturbation. This requires that one chooses an initial (i.e., unperturbed) system state or an initial system trajectory. The study of persistence does not require the choice of an initial state. Instead, it is common to consider some distribution of parameter values. Measures of persistence are functions of this distribution [509, 562].

2.3.4 Reviews of robustness and resilience

The study of robustness and resilience is a very active field of research in many disciplines. We aim to give an overview and a basic framework for comparing studies of robustness of networks. Many scholars have provided reviews of robustness and resilience of network models [76, 85, 192, 584], as well as other models of systems in different fields of research [278, 285, 288, 509, 544, 562].

Many researchers have studied robustness of systems in biology and biomedicine [285–288, 306, 509, 560]. Kitano (2004) examined robustness in biological systems and discussed several mechanisms that can lead to robustness in biological systems

[285]. He later presented a formal definition of robustness as the expectation of the impact of a perturbation [288]. He also considered cancers to be robust biological systems [286] and argued that an improved understanding of the origins of the robustness of cancers can help in the design of cancer drugs [287]. Stelling et al. (2004) reviewed robustness and persistence of cellular processes [509]. Wagner (2007) wrote a monograph about the robustness of biological systems to genetic perturbations [560]. Larhlimi et al. (2011) reviewed several definitions of robustness in the study of network models of metabolic systems [306].

Robustness and resilience are also relevant concepts in the study of social systems [4, 45, 85, 120, 347, 431, 532]. In sociology, many researchers have investigated the resilience of human communities [45, 120, 347, 532]. Plough et al. (2013) reviewed approaches for improving the resilience of human communities to, for example, natural disasters [265, 431]. One can also consider mechanisms for reducing community resilience. Catanese et al. (2016) reviewed studies of the resilience of crime networks [85]. In the networks that they studied, nodes represent criminals and edges correspond to some interaction between two criminals (e.g., a phone call or a crime committed together). Many of these studies have focused on efficient strategies for disrupting crime networks as much as possible. Agreste et al. (2016) studied the robustness of social networks of Sicilian Mafia organizations [4]. Their article includes a brief review of the role of social ties in criminal organizations.

One can study the robustness of systems of economic ties between people, businesses, and banks, and other entities [47, 76, 244, 270, 373, 540]. For example, supply systems are systems of economic ties between producers, distributors, and consumers of goods [524]. Many researchers have studied the robustness and resilience of supply systems to various perturbations [47, 244, 270, 373, 540]. Another example of a system of economic ties is financial markets. Caccioli et al. (2018) reviewed studies of robustness of network models of financial markets to cascading failures caused by debtor default and other perturbations [76]. For the study of the effect of perturbations on financial system, it is more common to consider the concepts of "systemic risk" and "contagion risk", which relate to the likelihood of failure propagation in a system [228]. Several scholars have reviewed studies of systemic risk [18, 35, 43, 52, 132, 157, 215] and contagion risk [222, 421]. Neveu (2018) presented a survey of network-based approaches to measuring systemic risk [390].

Robustness and resilience are also important for infrastructure systems [118, 195, 248, 544, 584]. Gay and Sinha (2013) reviewed system resilience with a focus on civil infrastructure systems [195]. Cuadra et al. (2015) reviewed models of network robustness, resilience, and reliability in power grids. Their article includes a critical

discussion of the potential of network analysis to provide insights into the robustness of real power grids [118]. Williams and Musolesi (2016) examined the robustness of spatio-temporal network models of urban transportation systems [584]. Urruty et al. (2016) proposed a disambiguation of the concepts of stability, robustness, vulnerability, and resilience in the study of agricultural systems and reviewed various studies of these concepts in agronomy [544]. Huang et al. (2015) reviewed studies of network robustness and resilience for applications to wireless sensor networks [248].

Other areas of research for which scholars have reviewed notions of robustness and resilience include psychological sciences [562], ergonomic design [234], software design [479], and perception of music [278].

Some scholars have reviewed robustness from a theoretical perspective and focussed on a specific type of model. Cohen and Havlin wrote a monograph in which they explored, among other aspects, the robustness of simple networks [102]. Gao et al. (2015) reviewed studies of robustness of "networks of networks" [192]. Their review focused on the work of various researchers who have studied the robustness of random-graph models, but it also includes studies of networks of networks (see Section 2.4.1) that model infrastructure systems such as traffic in cities. Williams and Musolesi (2016) focused on spatio-temporal network models (see Section 2.4.1) of urban traffic systems [584].

2.4 Specification of a network robustness problem

The specification of a robustness problem requires the specification of a system, a model, a measure of robustness, a set of perturbations, and a measure of performance. In this section, we review how researchers have specified robustness problems that use networks as models (or as parts of models) of real-world systems.

2.4.1 A system and a model

Across different scientific disciplines, the word "model" can refer to many different things. For example, in experimental biology, it is common to use mice as a *model organism* to study various aspects of mammalian life [95, 111, 245, 312]. In theoretical sciences, models are simplified mathematical or computational descriptions of real-world systems. Such models can take the form of a set of differential equations [255], agent-based models [386], decision trees [562], and so on. In this review, we focus on theoretical models that include networks. For example, researchers have used road networks as a model of a city's infrastructure [319] or protein-interaction

networks as a model of cellular processes [214, 261]. It is also common to use models that consist of a network and additional features. For example, popular models for the spread of disease consist of a dynamical system on a network [417].

For many real-world systems, there exist more than one way to construct a network model. Consider, for example, a system of chemical reactions. As a model for this system, one can construct, for example, a bipartite network in which each node corresponds either to a metabolite or a reaction and edges between metabolites and reaction indicate that a species takes part in a reaction [585]. However, one can also consider a metabolite-based network in which nodes correspond to metabolites and edges indicate chemical reactions that can transform one species into another [585]. One can also construct a reaction-based network in which nodes correspond to reactions and edges indicate that reactions share products or educts [585]. Each of these three network constructions can lead to a suitable model of a system of chemical reactions. However, the study of robustness of these three different types of networks for the same system of chemical reactions can lead to different results [306, 585]. In particular, the same structural network property can reflect different system properties when using different network constructions. As an example, consider the size N_{LCC} of the largest connected component (LCC; see Section 2.4.4.3). When using a reaction-based network to model a system of chemical reactions, the number N_{LCC} indicates the largest number of connected chemical reactions. By contrast, in a metabolite-based network, the number N_{LCC} reflects that the number of metabolites that take part in the largest component of connected reactions . In a bipartite network model of a system of chemical reactions, the number of nodes in the LCC is the sum of the number of metabolites and the number of reactions in the LCC of reactions. Using the same structural network property as a measure of system performance on different network models of a system can reflect very different notions of system performance. It is thus possible that studies of a system's robustness that use different network models lead to very different results — even when all other specifications of the robustness problem are identical.

This discussion of network models for metabolic networks demonstrates that network models for the same system can be very different, because nodes and edges in a network can represent different types of entities and different types of interactions in a system. Network models can also differ in the type of detail that one includes. For many real-world systems, an appropriate characterization requires more information than binary data on pairwise interactions between components. Researchers have proposed many network models that can include additional knowledge and data about a system. For example, one can include annotations for nodes and edges.

Nodes and edges can have annotations that reflect different types of entities or interactions [9, 519]. One can use node annotations in a network model of a power grid, for example, to distinguish power producers and power consumers [9]. One can use edge annotations to distinguish different types of social interactions (e.g., friendship, communication, trade, and so on) when constructing a social network model [519]. In a network with node weights, nodes have weights that can reflect measurable node properties [123], such as the capacity of a power plant [428], an individual's susceptibility to an infection [204], and so on. In a network with edge weights, edges have weights that can reflect various aspects of interactions, such as velocity [319], duration [8], or intensity [188] of an interaction. In a *spatial network*, nodes and edges can have spatial coordinates [34]. In a temporal network, one can annotate nodes and edges with a set of time points or time intervals during which they are active components of a network [236]. When a system depends on several different types of interactions, interactions at several times, or interactions in several different places, a popular model is a *multilayer network* [289], such as *networks of networks* [192]. If a model should include interactions between more than two entities, simplicial complexes [412, 599] or a *hypergraph* (sometimes also called *hypernetwork*) can be an appropriate model [290]. If one wants to study a system's dynamics, one can combine a network with a mathematical or computational model for dynamical processes. One can consider models of dynamical processes on networks (e.g., a set of coupled differential equations on a network [561]) or dynamics of a network's structure (i.e., *evolving networks* [547]).

There are many more possibilities for constructing advanced network models for real-world systems than we can enumerate in this review. The list above is far from complete. In addition, many combinations of the approaches listed above yield further network models. Examples include "spatio-temporal" networks [584], dynamics on spatial networks [124], or combinations of dynamics on network and dynamics of networks. Models of the latter type are called "co-evolving networks" [48] or "adaptive networks" [208, 464].

In Table 2.3, we sort various publications on robustness of networks by area of application and network model. References appear in multiple cells if they include several different network models or a network model that combines several of the extensions that we listed above. The references that we considered for this table are a small sample of the thousands of publications on network robustness in various fields of research. We consider five application areas: biological systems, ecological systems, social systems, and infrastructure systems. Network models of biological systems are models of organisms or cellular systems. In these networks,

Network model	General systems	Biological systems	Ecological systems	Social systems	Infrastructure systems
simple network	[11, 46, 77, 238, 253, 470, 552]	[125, 261, 320, 569]	[387]	[4, 145]	[50, 80, 101, 133, 193, 454, 564]
directed network	[58, 438]	[306, 520, 585]	[152, 153]	[27]	[386, 518, 524]
network with node weights	[565, 566]			[380]	[93, 191, 331, 386]
network with edge weights	[549]	[578]	[28, 305, 469, 555]	[27, 380]	[243, 518]
multilayer network	[192, 376]	[440]	[126]	[27]	[447, 610]
dynamics on networks	[565, 566]	[255, 306, 585]	[28, 126, 152, 153, 305, 461, 469, 555]	[601]	[93, 191, 331, 386]

Table 2.3: Examples of studies of robustness of network models for different areas of application.

nodes correspond to cells, genes, proteins, protein complexes, or chemical species that are relevant for metabolism. Ecological networks are models of ecosystems. In these networks, nodes correspond to species of plants, animals, fungi, or unicellular organisms. Edges can indicate various relationships between species (e.g., predator–prey relationships, mutualistic relationships, or competitive relationships). In social networks, edges indicate some form of social interaction (e.g., friendships, acquaintanceships, collaborations, economic transactions, and so on) and nodes correspond to people, animals, and other individuals that can form social interactions. Infrastructure networks are models of human-built systems, such as the Internet, power grids, and traffic systems. In the category "general systems", we list theoretical studies and studies that have considered several networks for areas of applications.

2.4.2 A measure of robustness

In this subsection, we explain the general structure of many computational and some theoretical studies of network robustness. We discuss common approaches for calculating a measure of robustness or, alternatively, a measure of vulnerability from computed changes of performance measures under perturbation.

The impact of a perturbation. When examining the effect of a perturbation on a performance measure X, it is common to calculate the value X_0 of the performance measure in the unperturbed network and the value X' of a performance measure after the perturbation. One can consider the *impact* I of a perturbation to be the absolute change $X_0 - X'$ of a performance measure or the relative change $(X_0 - X')/X_0$ of a performance measure [9, 42, 126, 305, 308, 331, 385]. When one wants to ensure that impacts are positive, one can consider the absolute value $|X_0 - X'|$ and $|X_0 - X'|/X_0$ as measures of impact.

Sets and sequences of impacts. In studies of network robustness, one generally examines the impacts of many perturbations. For example, one can study the robustness of networks to single-node removal or single-edge removal. In a study of single-node removal, one calculates the set $\{I_v\}_{v \in V}$ of impacts for the removal of nodes $v \in V$ [524, 566, 583]. Similarly, in a study of single-edge removal, one calculates the set $\{I_e\}_{e \in E}$ of impacts I_e for the removal of edges $e \in E$ [308].

One can also study the robustness of networks to the removal of multiple nodes or multiple edges. We distinguish studies of *simultaneous* node/edge removal and studies of *sequential* node/edge removal. In a study of simultaneous node removal, one examines the impacts of removing sets of nodes from a network. A set of N nodes has $\binom{N}{n}$ subsets of size n, which makes an exhaustive computation of the set $\{I_{V_S}\}_{V_S \subset V, |V_S|=n}$ of impacts of removals of size-n node sets computationally challenging. Researchers have circumvented this computational difficulty either by sampling $\{I_{V_S}\}_{V_S \subset V, |V_S|=n}$ [524] or by using theoretical results from probability theory and statistical mechanics to calculate expectations for I_{V_S} [58, 101, 102].

In a study of sequential node removal, one constructs one sequence or several sequences $(v_1, v_2, v_3, \ldots)$ of nodes. For each sequence of nodes, one then calculates a sequence $(I_{v_1}, I_{v_2}, I_{v_3}, \ldots)$ of impacts, where I_{v_i} is the impact of removing the set $\{v_1, \ldots, v_i\}$ of nodes from a network [11, 253, 552]. In studies of sequential edge removal, one calculates a similar sequence of impacts for the removal of a sequence of edges [243, 604].

Scalar measures of robustness and vulnerability. Sets or sequences of impacts for structural perturbations give detailed information about the robustness (i.e., insensitivity) and vulnerability (i.e., sensitivity) of a performance measure to a set of perturbations. However, to facilitate a simple comparison of the robustness of different networks, researchers have developed several approaches for summarizing this information in a single number as a scalar measure of robustness [308, 470, 566]. For sets of impacts of single-node or single-edge removal, it is common to use the set's mean [566] or maximum [308] as a measure of vulnerability.

For studies of simultaneous node removal, a widely-used measure of robustness is the fraction f of nodes for which the expectation of I_{V_S} with $|V_S| = \lceil fN \rceil$ is greater than or equal to a specified value [101, 152, 484, 568].

For studies of sequential node or edge removal, Schneider et al. (2011) proposed to use the normalized mean of relative performances X'/X_0 as a measure of robustness and the mean of relative changes $|X_0 - X'|/X_0$ as a measure of vulnerability [470]. They named their measure of robustness the "r index". The normalization guarantees that the r index is in $[1/N^2, 0.5]$. It lies in $[1/N, 0.5]$, if the initial network is connected [470]. Many researchers have subsequently used the r index in other studies of network robustness [5, 231, 238, 253, 337, 364, 365, 470, 552, 604, 611].

In Fig. 2.1, we visualize our categorization of studies of network robustness. For some studies of network robustness, it can be difficult to see how they fit into this categorization because the authors of these studies did not explicitly define impact measures. Instead, they calculated measures of robustness directly from performance measures (see, for example, Ref. [470]).

2.4.3 A set of perturbations

A robustness problem requires the specification of a set of perturbations that a system experiences. There are various possibilities for modeling internal and environmental perturbations that a system can experience. For parameter-dependent models (e.g., networks with node weights or edge weights or dynamical systems on networks) one can consider changes in model parameters as perturbations [509, 562]. However, when using networks to model real-world systems, one is often interested in effects that changes in network structure can have on a system's ability to perform a specified function. It is possible to consider a rewiring of edges [274] or the addition of nodes or edges [586] as structural perturbations. However, the most widely-studied structural perturbations are node and edge removals [102, 192]. In

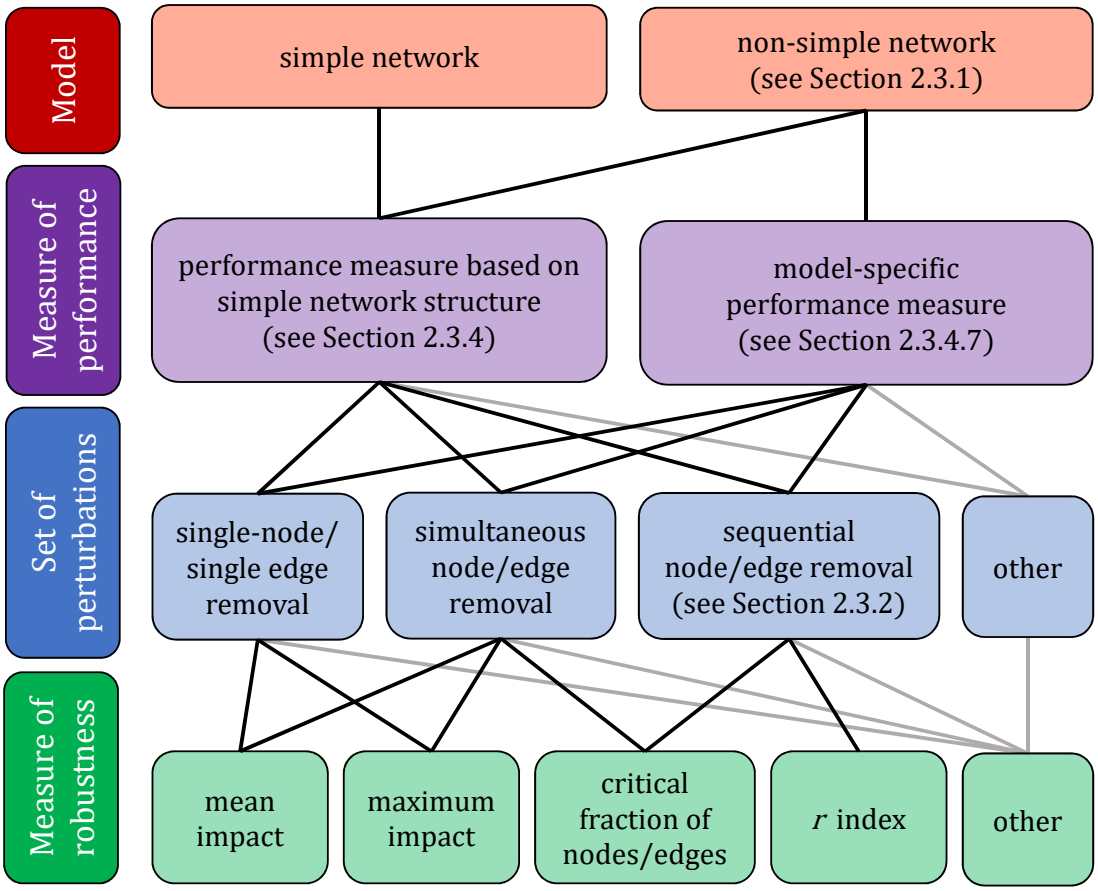


Figure 2.1: Combinations of specifications of robustness problems in studies of robustness of networks to node and edge removal. In each row, we categorize widely-used choices for specifications of a network robustness problem. Black and gray lines connect two categories for different specifications if, to our knowledge, scholars have chosen specifications from this pair of categories in a study of network robustness. For improved visibility of all connections, lines that connect to "other" sets of perturbations and "other" measures of robustness are gray.

this subsection, we consider several strategies for constructing sequences of nodes and sequences of edges for sequential node or edge removal.

2.4.3.1 Structural perturbations and centrality measures

Albert et al. (2000) presented perhaps the most influential paper on the topic of network robustness [11]. They examined the robustness of the mean size of components, the size of the LCC, and the mean shortest-path length of several networks under sequential node removal. They observed that the choice of the sequence $\{v_1, v_2, v_3, \ldots\}$ crucially affects the robustness of performance measures. When

one chooses $v_1, v_2, v_3, \ldots$ uniformly at random from the node set V, the change of performance measures tends to be smaller than when choosing nodes with the largest degree. In various subsequent studies, researchers have reported similar observations for various types of networks and various strategies for node selection [231, 335, 379, 552] or edge selection [243, 335]. We call a strategy for selecting nodes or edges for removal a *targeting strategy*. The most common targeting strategies use *centrality measures*. *Node centrality measures* assign a value of "centrality" (i.e., importance) to each node in a network [398]. One can calculate these values from the network's structure. Similarly, *edge centrality measures* assign a centrality value to each edge in a network. One can rank nodes or edges by their centrality values[5] and use this ranking to choose a sequence of nodes or edges for removal. Holme et al. (2002) were among others to point out that the removal of a node or edge can change the centrality values of the remaining nodes or edges [235]. Holme et al. thus distinguished targeting strategies that select nodes or edges based on their initial centrality values and targeting strategies that select nodes or edges based on centrality values that one recalculates after each removal of a node or an edge. They proposed to call targeting strategies that use recalculated centrality values as *dynamical* targeting strategies. However, many subsequent studies of network robustness instead use the term *adaptive* targeting strategies [231, 238, 342, 596, 611]. Holme et al. (2002) demonstrated that adaptive targeting strategies tend to yield larger changes in performance measures than non-adaptive targeting strategies. Subsequent studies have reported similar results [217, 237, 339, 399, 400].

2.4.3.2 Node centrality measures

In Table 2.4, we list node centrality measures and studies of network robustness that have used these centrality measures for adaptive or non-adaptive targeting strategies for node removal. In this subsection, we give definitions of the centrality measures in Table 2.4.

Degree centrality. A very popular centrality measure is *degree centrality* (or just degree) [10]. One ranks nodes by their degree and considers nodes with large degree to be more important than nodes with small degree.

[5] It is common to rank nodes or edges in decreasing order of centrality, but one can also consider rankings of nodes or edges from lowest to highest centrality [126, 152, 153].

Node centrality measure	References
Degree	[4, 5, 11, 16, 46, 67, 125, 126, 133, 140, 152, 153, 161, 193, 220, 231, 235, 238, 243, 253, 261, 269, 280, 305, 335, 380, 386, 387, 399, 419, 430, 438, 447, 470, 478, 484, 498, 518, 526, 552, 564, 605, 611]
Betweenness centrality	[4, 5, 67, 80, 133, 140, 180, 193, 235, 238, 243, 253, 280, 335, 380, 399, 430, 518, 552, 605]
Closeness centrality	[4, 5, 180, 253, 280, 430, 518, 552, 605]
Eigenvector centrality	[5, 180, 193, 253, 280, 430, 518, 605]
Subgraph centrality	[140, 280]
Damage centrality	[280, 308, 335, 518, 564]

Table 2.4: Examples of centrality measures for node removal.

Betweenness centrality. The *(geodesic node) betweenness centrality* of a node is a measure of how many shortest paths in a network traverse the node. Formally, the betweenness centrality BC_i of a node v_i in a network (V, E) is given by [169, p. 152]

$$\mathrm{BC}_i := \sum_{\substack{v_j, v_k \in V \setminus \{i\}, \\ j \neq k}} \frac{\tilde{n}_{jk,i}}{\tilde{n}_{jk}},$$

where $\tilde{n}_{jk}$ the number of shortest paths from v_j to v_k, and $\tilde{n}_{jk,i}$ is the number of shortest paths from v_j to v_k that traverse v_i [187]. One can define other types of betweenness centrality, such as one based on random walks rather than on strictly shortest paths [264].

Closeness centrality. There exists a family of node centrality measures that researchers call *closeness centrality*. These centrality measures use the geodesic distances d_{ij} between a node v_i and nodes $v_j \in V$ to assign a centrality value CC_i to v_i. It is very common to define close centrality as the inverse of the mean of geodesic distance between a node and all nodes in V [398, p. 171]. That is,

$$\mathrm{CC}_i := \left(\frac{1}{N} \sum_{j=1}^{N} d_{ij} \right)^{-1}. \tag{2.1}$$

This definition of CC_i has several issues. For example, if a network has more than one connected component, no node in the network is connected to all other nodes in

the network. For every node v_i, there exists at least one node v_j with $j \neq i$ such that there is no path from v_i to v_j. If one sets $d_{ij} = \infty$ for such node pairs, one obtains $\mathrm{CC}_i = 0$ for all $v_i \in V$. Newman (2018) pointed out an alternative definition,

$$\mathrm{CC}_i := \frac{1}{N-1} \sum_{j \neq i}^{N} d_{ij}^{-1} ,$$

that avoids this and other issues of Eq. (2.1) [398, p. 172]. However, many studies of network robustness consider targeting strategies based on the closeness centrality in Eq. (2.1) [4, 5, 430, 518, 605] or other notions of closeness [180, 253, 552].

Eigenvector centrality. One can consider a node's importance to depend on the importance of its neighbors. Eigenvector centrality is a centrality measure for capturing this notion of importance. One can consider eigenvalue centrality to be an extension of degree centrality [398, p. 330]. Degree centrality uses the number of neighbors of a node as a measure of its importance. Eigenvector centrality uses a normalized sum of the importance of neighboring nodes as a measure of a node's importance [63]. To obtain the eigenvector centralities of the nodes in a network, one sets the initial centrality of each node to 1 and subsequently updates the centrality of each node by the sum of centrality values of its neighbors. After normalizing the sum of the updated centrality values to 1, one again updates the centrality of each node by the sum of centralities of its neighbors, and so on. This iterative calculation converges to the eigenvector centralities of a network's nodes. The eigenvector centrality of a node v_i is equal to the ith entry of the principal eigenvector of the network's adjacency matrix [398, p. 169].

Subgraph centrality. Estrada and Rodriguez-Velazquez (2005) proposed a centrality measure based on the number of closed walks in which a node participates [170]. The *subgraph centrality* of a node v_i is the weighted sum of the numbers $\tilde{n}_{\circ,i}(\ell)$ of closed walks of length ℓ that traverse v_i. Using a common choice of weighting, subgraph centrality is [166]

$$\mathrm{SC}_i := \sum_{\ell=0}^{\infty} \frac{\tilde{n}_{\circ,\ell}(\ell)}{\ell!} .$$

One can calculate the subgraph centrality of a node using the eigenvalues λ_i and eigenvectors $\boldsymbol{\eta}_i$ of a network's adjacency matrix $\mathbf{A}$. Specifically,

$$\mathrm{SC}_i = \sum_{j=0}^{N} \left(\eta_j^{(i)}\right)^2 e^{\lambda_j} ,$$

where $\eta_j^{(i)}$ is the ith component of the eigenvector $\boldsymbol{\eta}_j$.

Damage centrality. For any performance measure X, one can define a node centrality measure as the change of X under the removal of a node. That is, [308],

$$\mathrm{DAC}_i := X_{(V,E)} - X_{(V\setminus\{v_i\},E')}\,,$$

where $X_{(V,E)}$ denotes the performance measure for the initial network and $X_{(V\setminus\{v_i\},E')}$ the performance measure forthe network after removal of node v_i. Latora and Marchiori (2005) introduced this family of centrality measures and called them *damage (centrality)*. Many researchers have considered targeting strategies based on the damage to the size of the LCC (see Section 2.4.4.3) of a network [308, 335, 564]. Sun et al. (2017) examined the impact of targeting nodes by damage centrality on efficiency (see Section 2.4.4.5) [518]. It is possible to view other centrality measures as a damage centrality measure. For example, one can view the degree of a node as a node's damage centrality to the number of edges in a simple network.

2.4.3.3 Edge centrality measures

In Table 2.5, we list edge centrality measures and studies of network robustness that have used these centrality measures for adaptive or non-adaptive targeting strategies for edge removal. In this subsection, we give definitions of the centrality measures in Table 2.5.

Edge entrality measure	References
Edge betweenness	[223, 235, 243, 604]
Edge clustering coefficient	[268, 604]
Degree product	[235, 268, 604]
Degree difference	[268]

Table 2.5: Examples of centrality measures for edge removal.

Edge betweenness. The *(geodesic edge) betweenness centrality* of an edge $e_{i,j}$ is a measure of how often a shortest path in a network traverses the edge. That is [243],

$$\mathrm{BC}_{ij} := \sum_{\substack{k,l=1,\\ j\neq i}}^{N} \frac{\tilde{n}_{kl,ij}}{\tilde{n}_{kl}}\,,$$

where $\tilde{n}_{kl}$ the number of shortest paths from v_k to v_l, and $\tilde{n}_{kl,ij}$ is the number of shortest paths from v_k to v_l that traverse $e_{i,j}$. Edge betweenness centrality is a popular edge centrality measure for studies of network robustness [223, 235, 243, 604].

Edge clustering coefficient. In an undirected network, one can define the edge clustering coefficient of an edge $e_{i,j}$ between two nodes v_i, v_j with degrees $k_i, k_j \geq 2$ as [442]

$$C_{ij} := \frac{|V_{(i,j)}|}{\min\{k_i, k_j\} - 1},$$

where $V_{(i,j)} \subset V \backslash \{v_i, v_j\}$ is the set of nodes that are adjacent to v_i and v_j. Kaiser and Hilgetag (2004) and Zeng and Liu (2012) used C_{ij} in studies of network robustness [268, 604].

Degree product and degree difference. Several researchers have assigned centrality values to edges based on the degrees k_i and k_j of the nodes that an edge $e_{i,j}$ connects [235, 268, 604]. For example, one can use the degree product $k_i k_j$ as a measure of edge centrality [235]. The degree product is large for edges that connect nodes with large degrees. Another example is (the absolute value of) the degree difference, $|k_i - k_j|$, which is an edge centrality measure that assigns large centrality values to edges that connect large-degree nodes to small-degree nodes [268].

The distribution of degree products and the distribution degree differences in a network are related to the network's degree–degree assortativity. In a *degree–degree assortative* network, edges tend to connect nodes with similar degrees [392]. In a degree–degree disassortative network, edges tend to connect nodes with different degrees. Several researchers have explored links between degree–degree assortativity and robustness in networks [231, 392, 393, 487, 537, 609].

2.4.3.4 Target strategies and applications

We review similarities and differences in the impact of different targeting strategies on performance measures. We subsequently discuss motivations for considering different targeting strategies in studies of network robustness.

Correlations between centrality measures. Many researchers have compared the impacts of uniformly random node/edge removal to the impacts of different targeting strategies on performance measures of networks [231, 243, 253, 335, 335,

379, 552]. Their results indicate that many strategies for targeted node or edge removal affect performance measures more strongly than removing nodes or edges uniformly at random, and that the difference between the impacts of targeted and uniform random removal is much greater than the difference between the impacts of different targeting strategies. Some researchers have argued that similar impacts of different targeting strategies are due to strong positive correlations between different centrality measures [235, 253, 335].

There are links between many popular centrality measures, suggesting that one should expect positive correlations between them [170, 398]. For example, one can derive eigenvector centrality as the result of an infinite iterative process that updates the centrality of a node with the sum of the centralities of its neighboring nodes in every iteration [398, p. 169]. In the first iteration, this centrality is identical to degree centrality. A node with many edges tends to be connected to many other nodes by short paths, which suggests that (for many networks) one can expect a positive correlation between degree and closeness centrality [398, p. 171]. Subgraph centrality assigns more importance to closed walks of a short length ℓ than to closed walks of lengths greater than ℓ. If one considers only closed walks of length $k = 2$, subgraph centrality and degree centrality result in identical node rankings in an undirected, unweighted network. If one considers only closed walks of length $k = 3$, subgraph centrality indicates the number of triangles associated with a node and is thus correlated with the numerator of the local clustering coefficient (see Eq. (2.3)).

In addition to the above considerations, many researchers have reported positive correlations of centrality measures when comparing targeting strategies for node removal [235, 253, 408]. However, in comparing the numerical results of different studies, one observes that the strength of correlations of centrality measures varies strongly across networks. The strength of correlations and the functional form of the relationship between two centrality measures depends on network structure [235, 548]. Uncovering the effect of network structure on correlations of centrality measures is an active field of research [55, 97, 253, 254, 402, 408, 605].

Why consider more than one targeting strategy? If one tends to observe similar impacts on performance measures for different targeting strategies, does one need to consider more than one targeting strategy in a study of network robustness? The answer to this question depends on a study's aim. For example, if one wants to demonstrate that the impact of sequential node removal can be very different for different sets of removed nodes, it is likely sufficient to compare the impact of removing nodes uniformly at random to the impact of targeting nodes by largest

degree (see, for example, Ref. [11]). However, a study's aim can be to find a good heuristic approach to identifying a fixed-size set of nodes that, if removed, leads to the largest possible change of a performance measure. For such a study, a comparison of different targeting strategies can be useful. For many performance measures, neither degree nor any other known node centrality measure can reliably identify fixed-size sets of nodes or edges whose removal leads to the largest change in the performance measure [72]. Several studies have demonstrated that, on some networks, targeting nodes by largest betweenness centrality and other by other centrality measures can lead to larger impacts on performance measures than targeting nodes by largest degree [235, 253, 335, 552]. In our study of the robustness of various performance measures on function-specific protein-interaction networks (see Chapter 3), we obtain results that support this assertion.

Different centrality measures can also be relevant when modeling adversarial attacks on networked systems. For example, if one is interested in the impact of a hacker's attack on the Internet, one can study centrality measures that one assumes to reflect a hacker's preferences for selecting target nodes. As another example, consider extinctions in an ecological network. If one assumes that a centrality measure reflects the likelihood of extinction for plant and animal species, one can use that centrality measure to construct realistic node-removal perturbations. In some studies, researchers have used observations on real systems to rank nodes or edges [321, 380]. Mourier et al. (2017) used data from a catch-and-release program to assign blacktip reef sharks a value of "catchability" [380]. Li et al. (2015) used measurements of traffic flow to assign "velocities" to edges in a road network [319]. They then examined the size of the LCC when removing edges from highest to lowest velocity.

Which targeting strategy leads to the greatest impacts? If one is interested in reliably identifying fixed-size sets of nodes that, if removed, lead to the largest possible change in a performance measure, it is possible that no node centrality measure (and no combination of node centrality measures) can help solve this problem. In general, one cannot calculate the impact of the removal of a set of nodes from the impacts of the single-node removals. Researchers refer to the impact of the removal of a set of nodes as the set's *collective influence* [338, 377, 378, 420]. The study of collective influence is an active area of research [62, 72, 338, 381, 613]. If one is interested in the impact of the removal of a set of nodes on the size of a network's LCC, one can consider strategies for *network dismantling* [72, 448, 603]. In a network-dismantling problem, one aims to find a smallest set of nodes that,

if removed, reduces the size of the LCC to some prescribed value [72]. Such sets of nodes are called *optimal dismantling sets* [72, 122, 129]. Many researchers have proposed heuristic algorithms for finding (near) optimal dismantling sets on various networks [72, 122, 129, 377].

2.4.4 A measure of performance

A crucial part of the specification of a robustness problem is the specification of a performance measure. In this subsection, we discuss challenges in identifying an appropriate performance measure for a given system. We also review several structural properties of networks that researchers have used as performance measures.

2.4.4.1 Performance measure and system function

When studying the robustness of a system, one tends to be interested in whether a system can continue to perform some function after it has experienced some perturbations. In a study of system robustness, one thus typically examines model properties that one assumes to be measures of the extent to which a system can perform some function. We discuss such model properties in this subsection.

The specification of a performance measure is a crucial part of specifying a robustness problem. The insensitivity of one model property does not, in general, indicate the insensitivity of another model property. Studies of robustness that use different performance measures can differ considerably in their conclusions [552].

Whether a model property is an appropriate measure of performance depends on the system that one studies and the model that one uses to study the system. However, the choice of a performance measure also reflects the system function (or functions) in which a researcher is most interested. Consider a system of chemical reactions again. One can be interested in the system's ability to produce a specific metabolite (e.g., some metabolites that are important for a cell's survival, such as adenosine triphosphate (ATP)) or its ability to degrade a specific metabolite (e.g., toxins that can hinder the survival of a cell or organism). Alternatively, one can be interested in the system's ability to turn one specific metabolite into another specific metabolite (e.g., the production of ATP from glucose). It is likely that different model properties are appropriate performance measures for these different system functions.

In models with many features, one can use the information from features as guidance for finding appropriate measures of performance [307]. In simplified models, in which there are no additional features available to guide a scholar's search

for relevant performance measures, the identification of an appropriate performance measure can be hard. For example, protein–interaction networks (PINs) are popular models of cellular processes. However, these models include only information about protein-protein interactions and lack, for example, information about interactions of proteins with other molecules. It is mostly unclear which properties of a PIN's structure can measure the viability (i.e., the ability to survive) of a cell or an organism [214]. When it is difficult to identify appropriate measures of performance, it is common to examine the robustness of several structural properties of a network model [50, 552, 585, 595] that one suspects to be positively or negatively correlated with system performance.

As in most previous work, we focus on scalar measures of performances. However, one can consider constructing multidimensional performance measures. Such measures can be appropriate when, for example, a researcher is interested in a system's ability to perform several functions. We are currently not aware of studies of the robustness of non-scalar performance measures. However, many researchers have studied the robustness of the output of algorithms for community detection [274, 321, 487, 517] and node ranking [65, 110, 186, 402, 476] to changes in network structure. The motivations for studies of robustness of algorithms tend to be different to those of studies of system robustness, but studies of robustness of node rankings and community detection can have a similar structure (see Fig. 2.1) to studies of robustness of non-scalar performance measures on network.

For some applications, it can be reasonable to assume that a network (V, E) has the uniquely optimal structure for performing a required function (or a set of required functions) and that network structures that are very different to the structure of (V, E) are very unlikely to perform the required function(s) well. These applications can motivate the use of graph distances for measuring impacts of structural perturbations. Researchers have used graph distances instead of performance measures to compare network structures before under structural perturbations [59, 467, 468]. In these studies, researchers chose a graph distance $\Delta(\cdot, \cdot)$ and measured the impact of a perturbation by the distance $\Delta((V', E'), (V, E))$ between the perturbed network (V', E') and the original network (V, E). This approach eliminates the need to specify a performance measure. Instead, it requires a researcher to specify a graph distance that is sensitive to the structural changes that are important to a given application. For a review of the use of graph distance for different applications, see Ref. [146].

In Table 2.6, we list performance measures and studies of them. Many studies of robustness of networks to structural perturbations do not distinguish between

measures of performance and measures of robustness. It is thus difficult to divide the set of structural network properties that researchers have studied in the context of network robustness into these two types of measures. Therefore, the list of performance measures in Table 2.6 includes some structural properties that researchers have used as measures of performance or as measures of robustness. In the remainder of this subsection, we give definitions and discuss applications of the structural properties in Table 2.6.

2.4.4.2 Performance measures based on the number of nodes or the number of edges

A basic way for measuring the impact of a perturbation on a network is to measure "how much is left of it". Two straightforward measures for quantifying how much is left of a network are the number of nodes (i.e., size) and the number of edges in a network.

Number of nodes. When considering the robustness of a simple network's number of nodes, the study of the impact of node removal perturbations is a trivial problem. The removal of n nodes from a network with N_0 nodes yields a network with $N' = N_0 - n$ nodes. For node-removal perturbations on a simple network, the number of nodes is a performance measure that is independent of a network's structure and thus is not very interesting. However, one can consider dynamics on a network that model successive failures of nodes or edges following the removal of an initial set of nodes or edges. Models that include such dynamics are often called models of *cascading failures* or *cascading behavior* [66, 116, 227, 273, 434]. For example, it is common to study *trophic cascades* in food webs [28, 126, 152, 153, 305, 555]. The extinction of one species in an ecosystem can cause the extinction of further species that, for example, rely on the first species as prey [152]. In the study of power grids, it is common to investigate perturbations that can cause power blackouts for entire cities or countries [93, 118, 191, 331]. Many researchers have modeled such blackouts as *cascading failures* in a power grid [93, 191, 331]. In a network model of a power grid, nodes correspond to power generators or consumers and edges correspond to power transmission lines. Each generator node contributes to the overall power generation in the power grid, and each edge contributes to the distribution of power to consumers. If a node or edge fails, the contributions (i.e., "loads") of other nodes and edges tend to increase to meet the consumers' demand. The redistribution can cause the failure of further nodes or edges and thus trigger a cascading failure. The number of nodes that remain after a cascade depends, in

Performance measure	References
Number of nodes	[16, 93, 133, 152, 153, 269, 305, 364, 365, 459, 469, 520, 565]
Number of edges	[27, 518]
Mean degree	[50, 335]
Number of components	[50, 71, 88, 243]
Mean component size	[11, 88, 484]
Size of the largest connected component	[4, 11, 61, 88, 101, 102, 140, 192, 231, 235, 335, 337, 386, 399, 419, 454, 478, 484, 498, 518, 526, 563–566, 595, 604]
Relative LCC size	[5, 46, 50, 67, 133, 145, 161, 243, 253, 380, 447, 470, 552, 611]
Reachability	[50, 243, 430, 552, 605]
Node connectivity	[160, 549]
Edge connectivity	[160, 481, 549]
Mean local clustering coefficient	[180, 335, 585]
Global clustering coefficient	[50, 386, 585]
Graph entropy	[28, 552]
Mean shortest-path length	[4, 11, 46, 50, 71, 80, 88, 125, 139, 235, 335, 386, 447, 478]
Efficiency	[46, 180, 193, 235, 243, 308, 335, 399, 424, 447, 552]
Spectral radius of $\mathbf{A}$	[50, 258, 569, 595]
Spectral gap of $\mathbf{A}$	[569, 595, 597]
Natural connectivity	[71, 140, 193, 387, 424, 569, 588, 589, 595]
Algebraic connectivity	[50, 549, 569, 595, 597]
Resistance Distance	[158, 569, 570, 595]
Model-specific performance measures	[9, 27, 42, 160, 306, 385, 386, 475, 583, 585, 601]

Table 2.6: Performance measures in studies of network robustness. Horizontal lines separate the performance measures from Section 2.4.4.2, Section 2.4.4.3, Section 2.4.4.4, Section 2.4.4.5, Section 2.4.4.6, and Section 2.4.4.7.

general, on a network's structure [118]. The number of nodes remaining after a cascade (or alternatively, the number of nodes that "become extinct" or "fail" during a cascade) is thus a non-trivial measure of performance for network models with cascades.

Number of edges and mean degree. One can also consider the number of edges as a measure of performance. Baggio et al. (2016) proposed "interconnectedness" as a measure of performance in a multilayer network of ecological and social relationships between households [27]. In their study, interconnectedness is proportional to the number of interlayer edges in a multilayer network. Sun et al. (2017) used the number of edges in an air transportation network as a measure of performance [518]. Lordan et al. (2014) also studied air transportation networks [335]. They used the mean degree, which is proportional to the number of edges in a network, as a measure of performance.

2.4.4.3 Performance measures based on connected components

Number of connected components and mean component size. For many real-world systems, it is common to assume that its ability to perform a function requires some form of interactions between its parts. Under this assumption, a necessary (but not sufficient) condition for a good system performance is that most nodes of an associated network are part of the network's LCC. The number n_c of components is an example of a performance measure that is likely to be negatively correlated with a system's ability to perform a function, because a large value of n_c indicates that a network consists of many components that do not interact with each other. The mean component size $\overline{N}_c$ is another network property that researchers have studied frequently in the context of network robustness. It is related to n_c by

$$\overline{N}_c = \frac{N}{n_c} \,.$$

Albert et al. (2000) and Shargel et al. (2003) considered $\overline{N}_c$ as a measure of system performance [11, 484]. Hossain et al. (2013) examined n_c under node and edge removal [243]. Chaverri (2010) considered n_c and $\overline{N}_c$ as a performance measure in a social network of bats [88].

Size of the largest connected component and the relative size of the largest connected component. A very popular measure of system performance is the size N_{LCC} of the LCC. (See Table 2.6 for references.) The relative size of the largest

connected component is the fraction $F := N_{\mathrm{LCC}}/N$ of the number N_{LCC} of nodes in a network's LCC and the number N of nodes in the network. The size of the LCC takes values in $\{0, 1, \ldots, N\}$. The relative LCC size has values in $[0, 1]$, which facilitates a comparison of relative LCC size for networks of different sizes. One can use results from percolation theory [279, 458, 508] to calculate expectations of the size of the LCC and the relative LCC size under several types of node-removal or edge-removal perturbations [58, 101, 102].

Reachability. The *reachability* r of a network is proportional to the number of node pairs that are connected by at least one path [50, 389, 443]:

$$r := \frac{1}{2}\binom{N}{2}^{-1} \sum_{\substack{i,j=1 \\ j \neq i}}^{N} \Xi_{i,j} \,,$$

where $\Xi_{i,j} = 1$ if there exists a path from v_i to v_j and $\Xi_{i,j} = 0$ else. In an undirected network, one can calculate a network's reachability from the sizes of its connected components [552]:

$$r = \frac{1}{2} \sum_{i} N_{c_i}(N_{c_i} - 1) \,,$$

where N_{c_i} is the size of the ith component in a network. Many researchers have studied reachability but have given it different names (e.g., "pairwise connectivity" [552], "flow robustness" [430, 605], and "average two-terminal reliability" [50, 389]).

Node connectivity and edge connectivity. For a connected network, one can consider how many nodes or edges one needs to remove to turn the connected network into a fragmented one. If the removal of a set of nodes or edges causes a network to become fragmented, we call that set of nodes or edges a *node cut set* or *edge cut set*, respectively. A network's *node connectivity* κ_v is the size of a smallest node cut set in a network [481]. A network's *edge connecitivity* κ_e is the size of a smallest edge cut set [481]. A node cut set can include nodes in the LCC, so the size of the LCC can change when removing a node cut set. The node connectivity indicates how many nodes one can remove from a network without increasing the number of nodes outside the LCC. The edge connectivity indicates how many edges one can remove without decreasing the size of the LCC. These relationships to a network's LCC can motivate the study of node connectivity [160, 549] and of edge connectivity [160, 481, 549] in the context of network robustness.

2.4.4.4 Performance measures based on node degrees and triangles

In this subsection, we discuss performance measures that are based on properties of nodes and small subgraphs: An entropy measures that is a function of a network's degree distribution, and the mean local clustering coefficient and the global clustering coefficient, which are measures of the frequency of triangles in a network.

Graph entropy For a network with degree distribution $p(k)$ and maximum degree $k_{\max}$, one can define a notion of graph entropy[6] as [563]

$$H = -\sum_{k=1}^{k_{\max}} p(k) \log(p(k)) \,. \tag{2.2}$$

Wang et al. (2006) used H as a measure of degree heterogeneity in a network and demonstrated that it is related to the robustness of a network's LCC to removing nodes uniformly at random [563]. This relationship motivated Ventresca and Aleman (2014) to consider H as a performance measure [552]. Bähner et al. (2017) use "interaction evenness", which is proportional to a network's graph entropy, as a measure of performance [28]. We discuss other notions of entropy on networks in Section 4.2.

Global clustering coefficient. The *global clustering coefficient* (i.e., *transitivity*) of a network (V, E) is [397]

$$T := 3 \times \frac{\text{number of triangles in } (V,E)}{\text{number of connected triples in } (V,E)} \,,$$

where the "number of connected triples" is the number of length-2 paths in (V, E) and the "number of triangles" is the number of subgraphs that consist of a length-2 path and an edge that connects the starting node and ending node of that path.

Mean local clustering coefficient. Watts and Strogatz (1998) defined the *local clustering coefficient* C_i of a node v_i as [574]

$$C_i := \begin{cases} 2\frac{|E_i^{(1)}|}{k_i(k_i-1)} \,, & \text{if } k_i \geq 2 \,, \\ 0 \,, & \text{if } k_i < 2 \,, \end{cases} \tag{2.3}$$

[6] The base b of the logarithm in Eq. (2.2) determines the units of entropy. If one chooses $b = 2$, one measures entropy values in bits. If one chooses $b = e$, one measures entropy values in nats. Throughout this dissertation, we use $b = e$ when computing numerical values of entropy.

where k_i is the degree of node v_i and $|E_i^{(1)}|$ is the number of edges in the induced subgraph on the 1-hop neighborhood of v_i (i.e., the number of edges between neighbors of v_i). For a network with N nodes, the *mean local clustering coefficient* is $C = \sum_{i=1}^{N} C_i/N$.

The mean local clustering coefficient and the global clustering coefficient are measures of the frequency of triangles in a network, relative to the number of connected triples. Several researchers have linked the number of triangles in a network to system performance [333, 406]. Triangles in a network are related to a notion of redundancy [398, p. 186], because a triangle in an undirected network indicates that there are at least two paths by which each pair of nodes in the triangle are connected. Redundancy is a concept that is closely related to a system's ability to perform a function robustly [356, 363, 444, 471, 559, 579, 597, 612]. Another motivation for linking the global clustering coefficient and the mean local clustering coefficient to a system's function is that many researchers have proposed high values of C and T to be characteristic of many real-world networks [174, 397, 441, 574] and thus potential proxies for a system's ability to fulfill a function [17, 333]. For example, several scholars have proposed that triangular network motifs are components of important functional units in biological systems [17, 285, 302]. The link between triangular motifs and system function motivates our study of the plausibility of C and T as measures of performance for function-specific protein-interaction networks in Chapter 3.

2.4.4.5 Performance measures based on geodesic distance

The geodesic distance from a node v_i to a node v_j is the length of the shortest path from v_i to v_j. The geodesic distance thus indicates the minimum number of edges that a message, people, good, or other object needs to traverse to go from v_i to v_j.

Mean shortest-path length. The *mean shortest-path length* L is a measure of a mean geodesic distance between pairs of nodes. Fragmented networks have node pairs v_i, v_j for which there do not exist connecting paths. To avoid the difficulty of defining a distance between such node pairs, one can define L as the mean shortest-path distance in the LCC of a network:

$$L := \binom{N_{\mathrm{LCC}}}{2}^{-1} \sum_{i=1}^{N_{\mathrm{LCC}}} \sum_{j>i}^{N_{\mathrm{LCC}}} d_{ij} \,,$$

where d_{ij} is the geodesic distance from v_i to v_j.

Efficiency. One can define the *(communication) efficiency* $\mathcal{E}$ as the mean of inverse geodesic distances of pairs of nodes in a network [46] or its LCC [200]:

$$\mathcal{E} := \frac{1}{2}\binom{N'}{2}^{-1} \sum_{\substack{i,j=1 \\ j\neq i}}^{N'} d_{ij}^{-1},$$

where N' is either N [46] or N_{LCC} [200]. If one chooses $N' = N$, the inverse geodesic distance between nodes in different connected components is 0 [46].

Mean shortest-path length and efficiency are especially popular measures of performance for infrastructure and communication networks, which have the function to transport and distribute people, goods, messages, or other things between various sites [80, 243, 335, 386, 447]. In such systems, it is possible that people or algorithms with complete information about a network's structure guide agents or objects to take a shortest path to get from one site to another. For example, when traveling from one location in a city to another, it is likely that people plan their route using GoogleMaps or other route planning services. For many other systems, the assumption of shortest-path-based routing is unlikely [313]. For example, in metabolic systems, metabolites do not choose their reactions and thus do not choose shortest paths in a metabolic network. Even in systems in which humans or algorithms aim to select shortest paths, it can be impossible to reliably find shortest paths if only incomplete information about a network's structure is available [293]. For example, in Milgram's experiments on the "six degrees of separation", Milgram asked people to deliver a parcel to a person via handing it to one of their acquaintances [367]. Although the participants of this experiment may have sought a shortest path for the delivery of their letters, it is likely that some letters traversed the network of acquaintanceships on paths other than the shortest ones, because the participants did not have complete information on the network of acquaintanceships. Erlebach et al. (2006) argued that even for the Internet, the assumption of shortest-path-based routing is unrealistic due to the restrictions that peer-to-peer protocols impose on the paths that data packages on the Internet can traverse [160].

2.4.4.6 Performance measures based on spectral network properties

For some directed networks and for all undirected networks, it is possible to diagonalize $\mathbf{A}$ and $\mathbf{L}$ and thus represent these networks by a set of eigenvalues and a set of eigenvectors. Spectral graph theory (i.e., the study of eigenvalues and eigenvectors of matrix representations of networks) has uncovered various relationships between a network's structural properties and its eigenvalues and eigenvectors [550]. Using

these relationships, one can construct functions of eigenvalues and eigenvectors of $\mathbf{A}$ and $\mathbf{L}$ that can be measures of system performance. We review some of these functions in this subsection.

Algebraic connectivity. The algebraic connectivity a of a network is the second-smallest eigenvalue of the network's combinatorial Laplacian matrix $\mathbf{L}$ [169]. The algebraic connectivity of a network is positive if and only if the network is connected [259]. In that case, one can view a as a measure of how well a graph is connected [130]. It is linked to the node connectivity κ_v and the edge connectivity κ_e by the relation [121, 549]

$$2\kappa_e \left(1 - \cos\left(\pi/N\right)\right) \leq a \leq \kappa_v \,.$$

This relationship has motivated the study of algebraic connectivity in various contexts [130, 181, 259, 260, 372, 502]. Specifically, several researchers have considered a as a measure of performance [50, 595, 597] or as a measure of robustness [49, 259, 260, 355, 549, 550, 569].

Spectral gap of A. The edge expansion of a network is the smallest number of edges that connect the two parts of a bipartition of the network divided by the size of the smaller part of the bipartition. Formally, one defines the edge expansion (which is also called the isoperimetric constant or Cheeger constant [550]) of a network (V, E) with N nodes as [550, p. 95]

$$h_e = \min_{0<|V_S|\leq \frac{N}{2}} \frac{|\partial V_S|}{|V_S|} \,,$$

where $\partial V_S := \{e_{i,j} \in E : \; v_i \in V_S \wedge v_j \in V \setminus V_S\}$, $S \subset V$. The edge expansion of a network is closely related to a network's edge connectivity. Researchers have investigated the relationship between several notions of network robustness and its expansion properties [161, 163]. For undirected k-regular[7] network, the Cheeger inequality [550, p. 96] links h_e to a network's algebraic connectivity:

$$\frac{a}{2} \leq h_e \leq \sqrt{2ka} \,.$$

This relationship between a and h_e gives an additional motivation for the study of a in the context of network robustness. It also motivates the study of the spectral gap of $\mathbf{A}$ (i.e., the difference between the largest and second-largest eigenvalue of $\mathbf{A}$

[7] An undirected network is *k-regular* if every node in the network has a degree of value k [7, p. 2].

[550, p. 55]), which is equal to a for undirected regular networks [240, 506]. Yazdani and Jeffrey (2011) and Yamashita et al. (2019) considered the spectral gap of $\mathbf{A}$ as a measure of performance in studies of network robustness [595, 597]. Wu and Holme (2011) considered the spectral gap of $\mathbf{A}$ as a measure of network robustness [591]. The spectral gap of an undirected network (whether it is regular or not) is also a measure of the time of convergence for a random walk on the network [550, p. 63].

Spectral radius of A. The spectral radius $\rho(\cdot)$ of a matrix is the largest absolute value of its eigenvalues [310, p. 195]. For many models of dynamics on networks, the spectral radius $\rho(\mathbf{A})$ of a network's adjacency matrix is an important quantity [31, 418] (see also Chapter 4). For example, a popular model for the spread of infectious diseases is the susceptible–infected–susceptible (SIS) model [417]. For several random-graph models (e.g., Erdős–Rényi random networks[8] [551], a degree–degree assortative generalization of a configuration model [57], and a generalized configuration model with prescribed triangle distribution [368]), one can find a critical threshold τ_c for the effective spreading rate[9] [57, 368, 551]. When the effective spreading rate is smaller than τ_c, infections die out. When the effective spreading rate is equal to or greater than τ_c, infections can persist and a positive fraction of a network's nodes remain infected [551]. Using a mean-field approach [434, p. 37], van Mieghem et al. (2009) showed that $\rho(\mathbf{A})^{-1}$ is a lower bound on τ_c [318, 551]. The spectral radius of the adjacency matrix is thus related to a system's ability to resist an epidemic. Several researchers have used $\rho(\mathbf{A})$ as a measure of performance in studies of network robustness [50, 258, 595].

Resistance distance. Interpreting every edge in a network as a resistor in an electrical circuit, one can calculate the *effective resistance* $\tilde{\Omega}_{i,j}$ between any two nodes in the network using Kirchhoff's laws. Klein and Randić (1993) showed that one can use $\tilde{\Omega}_{i,j}$ as a distance on a network and thus called $\tilde{\Omega}_{i,j}$ the *resistance distance* [292]. Several researchers have considered the *(total) resistance distance*

$$\Omega := \frac{1}{N(N-1)} \sum_{\substack{i,j=0 \\ i \neq j}} \tilde{\Omega}_{i,j}$$

[8] See Section 3.2.2 for explanation of Erdős–Rényi random networks and other random-graph models.

[9] The *effective spreading rate* is the ratio of the rate recovery rate (i.e., the rate with which an infected node recovers) and the infection rate (i.e., the rate with which a node that is adjacent to an infected node becomes infected) [418].

as a measure of system performance [104, 158, 209, 346, 531, 570]. The link between resistance distance and electric circuits has contributed to the popularity of Ω as a measure of performance for power grids [104, 209, 346].

One can compute Ω using

$$\Omega = N\mathrm{tr}(\mathbf{L}^{+}) ,$$

where $\mathbf{L}^{+}$ is the Moore–Penrose pseudoinverse [426] of $\mathbf{L}$ [346]. Several researchers have used Ω as a measure of performance or robustness [104, 158, 209, 346, 531, 570].

Natural connectivity. In several studies, researchers have proposed to use network properties based on the exponential of the adjacency matrix as measures of network functionality [164, 166, 170]. Exponentials of $\mathbf{A}$ and $\mathbf{L}$ are closely linked to linear dynamics on networks, because the system

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}\mathbf{x} \tag{2.4}$$

of linear differential equations has the solution $\mathbf{x}(t) = \exp(\mathbf{M}t)\mathbf{x}_0$ for some diagonalizable matrix $\mathbf{M}$ and initial conditions $\mathbf{x}_0$. If $\mathbf{M} = -\mathbf{L}$, one can consider Eq. (2.4) as a model of a linear diffusion process on a network [371]. When $\mathbf{M} = \mathbf{A}$, one can consider Eq. (2.4) as a model of a spreading process on a network [332]. In particular, the (i, j)th element of this exponential are equivalent to a weighted sum over ℓ of the counts of length-ℓ walks from v_i to v_j. It is thus related to the number of walks along which a person, object, or signal can travel from v_i to v_j. Estrada and Hatano (2008) interpreted the (i, j)th element of the matrix exponential of the adjacency matrix as the number of walks along which v_i can "communicate" with v_j. They defined the *communicability function* [166]

$$\kappa(i, j) := \left(e^{\mathbf{A}}\right)_{i,j} . \tag{2.5}$$

Jun et al. (2010) examined the natural logarithm of the mean of "self-communicability" $\kappa(i, i)$ in a network[588]. They considered the *natural connectivity*

$$K := \log\left(\sum_{i=1}^{N} \kappa(i, i)/N\right) = \log\left(\mathrm{tr}(e^{\mathbf{A}})\right) - \log N$$

as a measure of system performance [588]. In the last decade, natural connectivity has become a popular measure for the study of network robustness [424, 569, 588, 589, 595]. Researchers have considered natural connectivity for applications

in communication systems [193], transportation systems [594], and road networks [140].

In the discussion above of performance measures based on spectral network properties, we have pointed out several links between these performance measures and dynamical systems that model the flow of people, goods, or electricity or the spread of diseases or information on networks. Spectral network properties are structural properties of a network in the sense that one can compute them from a network's structure. However, much of the motivation for considering these structural properties as quantities of interest in the study of robustness is based on models of dynamics on networks. For example, the relevance of the spectral radius of the adjacency matrix to spreading processes on networks has motivated researchers to study $\rho(\mathbf{A})$ in the context of network robustness. One can compute $\rho(\mathbf{A})$ from a network's structure, but the interpretation of $\rho(\mathbf{A})$ as a characteristic quantity of a system is tied to the spectral radius' relevance to spreading processes.

Wang et al. (2018) considered several performance measures based on spectral network properties to rank the robustness of a set of networks and found that different spectral properties lead to different rankings of networks [569]. One likely reason for these differences is that the performance measures that they considered are related to different aspects of different dynamical systems on networks.

2.4.4.7 Model-specific performance measures

In Section 2.4.4.6, we discussed that many spectral measures of performance are closely linked to dynamical processes on networks and that, in using these measures of performance, one tends to make implicit assumptions on the movement of people, goods, electricity, information, and other things on a network. In this subsection, we discuss examples of measures of performance that depend explicitly on dynamical processes or other aspects of a model. The computation of these performance measures requires one to explicitly calculate solutions for a dynamical process on a network or obtain solutions via numerical simulation. We also consider examples of performance measures that depend on node labels and multilayer structure.

Performance measures for dynamics on metabolic networks. It is common to model metabolic networks as chemical reaction networks with kinetic reaction fluxes [306]. Such a network consists of a set of chemical species with concentrations $x_1, x_2, x_3, \ldots$ and a set of reactions with fluxes $\tilde{v}_1, \tilde{v}_2, \tilde{v}_3, \ldots$. The stochiometric matrix $\mathbf{S}$ represents the stoichiometric information for reactions (i.e., the number of

molecules of a species that a reaction consumes or produces) [306]. For a closed and well-mixed system, the concentrations of species are related to reaction fluxes by

$$\frac{d\mathbf{x}}{dt} = \mathbf{S}\tilde{\mathbf{v}}\,,$$

where $\mathbf{x} := (x_1, x_2, x_3, \ldots)$ is a *concentration vector* and $\tilde{\mathbf{v}} := \tilde{v}_1, \tilde{v}_2, \tilde{v}_3, \ldots$ is a *flux vector*. Many researchers have studied steady-state flux vectors (i.e., the vectors $\tilde{\mathbf{v}}$ that satisfy $0 = \mathbf{S}\tilde{\mathbf{v}}$) and proposed links between steady-state flux vectors and notions of robustness of metabolic networks [42, 583, 585]. For example, a popular concept in the study of robustness of metabolic networks are *elementary flux modes*, which are minimal[10] subsets of reactions that have a steady-state flux vector[11] [42, 583]. Using a different appoach, Winterbach et al. (2011) measured the performance of a metabolic system by its maximal[12] "rate of biomass production", which is a linear combination of the steady-state fluxes of "key reactions" for biomass production [585].

Performance measures for dynamics on transportation networks. Scott et al. (2006) considered a measure of congestion for the study of robustness of highway networks to edge removal [475]. They used a game-theoretic model [106, 572] for the decision process of drivers as they decide which route to take. This dynamical model has a so-called "Wardrop equilibrium" [106], at which no driver can improve their travel time by unilaterally changing routes [475]. Scott et al. (2006) used the Wardrop equilibrium to assign values of road congestion to compute their measure of performance.

Performance measures for networks with node labels or multilayer structure. Network models of systems can include node labels [143] or affiliations of nodes to layers [289]. For many applications, performance measures that use information of node labels and layer affiliations are more clearly linked to the performance of a system than performance measures that ignore node labels and layer affiliations. For example, if one considers the delivery of electricity from generators to consumers to be the function of a power grid, paths from generators to consumers are more

[10] These sets are minimal in the sense that the removal of any reaction from an elementary flux mode yields a set of reactions that is not an elementary flux mode.

[11] This definition of elementary flux modes applies only to closed systems. Many researchers have used a definition of elementary flux modes that facilitates the study of elementary flux modes on open systems as well [473, 602].

[12] Formally, the maximal rate μ of biomass production is $\max_{\tilde{\mathbf{v}}} \mathbf{c}^{\mathbf{T}}\tilde{\mathbf{v}}$ over all steady-state flux vectors $\tilde{\mathbf{v}}$. The elements $\mathbf{c}$ represent the contribution of a "key reaction" to biomass production [585].

relevant for performance than paths between generators. Albert et al. (2004) used a version of reachability as a measure of performance for the North American power grid that only considers paths from generators to distribution substations [9]. Erlebach et al. (2006) modeled the Internet as a network in which nodes are customers or providers [160]. They suggested to view providers to form an "upper" layer and consumers as a "lower" layer of a 2-layer network. They argued that common routing policies prevent data packages from taking "valley paths" (i.e., producer–consumer–producer paths) and subsequently proposed performance measures based on the path lengths of "valley-free paths".

The examples that we considered in this subsection demonstrate that one can use node labels, layer affiliations, and aspects of dynamics on networks to construct performance measures that are more closely linked to the performance of a system than performance measures depend only on a simple network's structure. We also see that there can be more than one way to include aspects of dynamics on networks in the construction of a performance measure. For example, for reaction kinetics on metabolic networks, we discussed two different approaches to constructing performance measures; One approach is based on elementary flux modes, and the other is based on the production rates of a set of "key metabolites". The different approaches reflect different perspectives on a system's function.

2.4.4.8 Popularity and applicability of performance measures

The large number of performance measures that researchers have considered in the study of network robustness problems leads to several challenges in the further study of such problems: Should one expect different performance measures to lead to similar results of a study of robustness? If not, why are the results different? Which performance measures should one consider? We discuss these and further questions in this subsection.

Can one expect different performance measures to lead to similar results on robustness? In general, the answer is 'no' [14, 118, 160, 552, 569, 585, 595]. In their definition of robustness, Alderson and Doyle (2010) included the performance measure as a crucial specification [14]. Wang et al. (2018) considered different spectral network properties to assess the robustness of a set of networks [569]. They then ranked networks by their robustness. Their results demonstrated that different spectral network properties yield different robustness rankings of networks. Winterbach et al. (2011) and Yamashita et al. (2019) considered several performance measures

and examined correlations between performance measures [585, 595]. Yamashita et al. (2019) reported that many, but not all, considered performance measures were positively correlated [595]. Winterbach et al. (2011) computed a measure of performance from reaction kinetics in a chemical reaction network. They compared this measure of performance to several measures that one can derive from a network's structure. They reported weak positive correlations between the performance measure that they calculated using reaction kinetics and the performances measures that they calculated from network structure. They concluded that these correlations were too weak to justify using the structure-based performance measures as proxies for the performance of reaction kinetics on a metabolic network.

Why do some performance measures lead to similar results on robustness? In general, one cannot expect performance measures to be positively correlated or to lead to similar results of studies of network robustness. However, some performance measures have positive correlations on some sets of networks. For example, Yamashita et al. (2019) found that several widely-used performance measures are positively correlated on a set of networks of various sizes N and various numbers m of edges [595]. A possible reason for this observation is that many structural network properties that researchers have considered as performance measures are measures of similar things. For example, many structural properties in Table 2.6 are measures of some notion of "connectedness". Reachability, mean component size, and size and relative size of an LCC are positively correlated with the probability that a pair of nodes chosen uniformly at random is connected by some path. The number of components is negatively correlated with this notion of connectedness. The spectral gap, algebraic connectivity, node connectivity, and edge connectivity are positively correlated with the number of nodes and edges that one can remove without changing the size of an LCC. Efficiency is related to a notion of "well-connectedness": It is positively correlated to the likelihood that a pair of nodes chosen uniformly at random is connected by a short path. The mean shortest-path length is negatively correlated with this notion of well-connectedness. The mean local clustering coefficient and global clustering coefficient are positively correlated with the probability that a pair of nodes chosen uniformly at random is connected by two short paths (specifically, one path of length 1 and one path of length 2). Natural connectivity is positively correlated with the number of closed walks that connects a node in a network to itself.

Connectivity measured via any of these structural properties tends to increase as the number m of edges in a network increases. Consider a set $\mathcal{S}$ of networks



Figure 2.2: A map from the video game Mini Metro with a connected transportation network. White symbols with black borders indicate train stations. Small black symbols indicate passengers who are waiting to embark on a train. The shape of a station indicates its type. The shape of a passenger indicates their desired destination type.

that differ in their values of m and another set $\mathcal{S}'$ of networks that have the same values of m. In $\mathcal{S}$, the relationship between various notions of connectivity and m can lead to positive correlations between performance measures that are related to some notion of connectivity. By contrast, in $\mathcal{S}'$, correlations between performance measures can be smaller than on $\mathcal{S}$ or non-existent. In studies of sequential node or edge removal, it is common to compare performance measures on networks with different values of m. In such studies, it is thus possible to observe positive correlations between performance measures because of the link between many performance measures and connectivity. For a set of networks with identical values of m, Wang et al. (2018) found that different performance measures lead to different robustness rankings [569].

Is connectedness a necessary condition for good performance? The answer to this question will be 'yes' in some cases and 'no' in others. Performance measures that are related to some notion of connectedness are very popular in the study of robustness problems (see Table 2.6 for references). Their popularity reflects the widely-applied assumption that only a (well-)connected network can be a well-performing network. Whether this is an appropriate assumption in the study of a given network robustness problem depends, among other things, on the focus



Figure 2.3: A map from the video game Mini Metro with a fragmented transportation network.

of a study. For example, if a relevant function of an urban transportation system is to transport people from any point of departure to any destination, the associated network needs to be connected to fulfill this function [176]. However, if one is interested in a different function of a transportation system, connectedness does not have to be a necessary condition. As a toy example, we show a map from the video game *Mini Metro* [119] in Figs. 2.2 and 2.3. In this game, a player's task is to connect stations (white symbols with black borders) on a map by train lines so that all passengers (small black symbols) can travel to their desired type of destination, which is given by a passenger's symbol shape. One can construct a connected network of train lines that fulfills this task (see Fig. 2.2). Because every type of symbol occurs at least twice in this example, one can construct a fragmented network that achieves the same function (see Fig. 2.3). Another example of a system function for which the connectedness of a network model is not a requirement for performance is the rate of biomass production of a metabolic system, because several connected components of a metabolic network can represent functional units that produce one or several biomolecular species without interacting with other connected components [341]. Winterbach et al. (2011) studied the rate of biomass production of a metabolic system and demonstrated that this rate is not strongly correlated to several performance measures that relate to some notion of connectedness [585].

Is connectedness a sufficient condition for good performance? The answer to this question can also vary between studies, but for many cases, the answer is 'no'. Connectedness measured by the size of an LCC, the relative LCC size, mean component size, and reachability measures the probability that a pair of nodes that one chooses uniformly at random is connected by a path. There are many reasons why such a path can be irrelevant in practice. For example, a system can have restrictions on interactions that depend on node labels and edge labels, regardless of whether one includes node labels or edge labels in a model of the system. Erlebach et al. (2006) explained that peer-to-peer protocols impose restrictions on the paths that data packages can traverse on the Internet [160]. Whether or not a path is relevant for data transmission on the Internet depends on the labeling of a path's nodes as consumers and producers. Paths that do not satisfy a protocol's criteria are irrelevant for the transmission of data packages on the Internet [160] because they are the "wrong type" of path.

How is connectedness related to performance? In several of the examples that we discussed to demonstrate that connectedness is neither a necessary nor a sufficient condition for good performance of a system, we needed node labels to explain why connectedness was neither necessary nor sufficient for good performance. For the Mini Metro network, passengers select a node type instead of an individual node as their destination. Because a player understands this aspect of the game's dynamics and because every node type occurs at least twice on the map, the player can construct a fragmented transportation network that fulfills the required function. For the example of biomass production in a metabolic system, one assumes that some nodes that correspond to key metabolites are more relevant to biomass production than others. Node labels and edge labels can indicate that some nodes and edges contribute to performance differently than other nodes and edges, and that some paths on a network are more relevant than other paths to system functions. Many studies of network robustness consider networks that do not have node labels, either because there exists only one type of nodes and one type of edges or because information about node labels and edge labels is not available. In such studies, it is common to assume that all nodes contribute in a similar way to system performance and that all paths are equally relevant to a system function. Under these assumptions, the size of an LCC or other performance measures that relate to some notion of connectedness tend to be relevant measures of performance.

2.5 Conclusions and outlook

We conclude this review with a discussion of the relevance of different specifications of a robustness problem. We also give a brief outlook on research avenues that can further understanding of network robustness.

2.5.1 Conclusions

A major challenge in the identification of "robust network structures" is that many studies of network robustness lead to different (and sometimes seemingly contradictory) conclusions on the characteristics of robust network structures. Alderson and Doyle (2010) proposed to distinguish different notions of robustness by several specifications [14]. Their framework helps resolve apparent conflicts between results of different studies of robustness, because it helps recognize that many studies consider different notions of robustness that do not need to be correlated with one another. However, it leads to the challenge of choosing specifications for a robustness problem. In this chapter, we reviewed various choices for specifying a model of a system, a measure of insensitivity (i.e., robustness), a set of perturbations, and a performance measure.

2.5.1.1 How should one choose a model?

It is very important to note that there are usually several reasonable approaches to constructing a network model of a system. Different network models tend to capture different aspects of a system [585]. Notions of robustness on different network models are different and do not need to correlate with each other. The study of robustness of different network models can lead to different conclusions. While we cannot give general guidance on the construction of network models, we emphasize the choice of a model is an important specification robustness problem.

2.5.1.2 How should one choose a measure of robustness?

Scalar versus non-scalar measures of robustness. In a computational study of a robustness problem, it is common to choose a set of perturbations and compute the impact of each perturbation in that set. The resulting set of impacts reveals a lot of information on the robustness of the system under investigation. The study of non-scalar robustness measures (e.g., a distribution of impacts of single-node removal or a sequence of impacts of sequential node removal [11, 125, 133, 191, 243, 269, 335, 391, 399, 438, 478]) can yield detailed insights into a robustness

problem. One inevitably loses information when computing a scalar measure of robustness from a set of impacts. However, scalar measures of robustness facilitate simple comparisons of the robustness values for different systems [552], different sets of perturbations [253], and so on. For some scalar measures of robustness, it is possible to calculate their exact values of bounds on their values from theoretical models using mathematical tools (e.g., tools from percolation theory [279, 508] or spectral graph theory [550]).

Impacts versus perturbation sizes. Many scalar measures of robustness are either a maximum value or a mean value over a set of impacts [5, 231, 238, 308, 337, 364, 365, 470, 604, 611] or a minimum or a mean value over a set of sizes[13] of perturbations that yield an impact that exceeds a given threshold [101, 102, 152, 153, 419, 454, 481, 484, 498, 526, 549, 563, 568]. For example, Cohen et al. (2000) examined the fraction f of nodes such that I_{V_S} on N_{LCC} with $|V_S| = \lceil fN \rceil$ is greater than or equal to a specified value [101]. Other examples of robustness measures that are based on perturbation size include the edge connectivity. The node connectivity measures the size of a smallest node set V_S that one needs to remove to achieve an impact $I_{V_S} > 0$ on the size of an LCC.

How should one choose a scalar measure of robustness? Robustness measures based on the mean value of a set of impacts are related to the expected impact of a perturbation on a system's performance. By contrast, robustness measures based on the maximum value of a set of impacts are related to a "worst-case scenario" of a system's performance under perturbation. The choice of mean impact or maximum impact for measuring robustness thus reflects a researcher's focus on either the expected-case scenario or a worst-case scenario for a system's performance. Both scenarios are interesting aspects of a system's robustness, and a system's robustness measures for the expected-case scenario and for a worst-case scenario can be very different.

2.5.1.3 How should one choose a set of perturbations?

For the study of robustness of networks to structural perturbations, it is common to consider a set of node removals or edge removals as the set of perturbations. We distinguished the removal of individual nodes or edges, simultaneous node or edge removal, and sequential node or edge removal.

[13] We consider the *size* of a perturbation to be number of nodes or edges that one removes.

Studies of trophic cascades and cascading failures. In the study of trophic cascades or cascading failures, researchers have investigated the impact of single-node removal and single-edge removal [115, 152, 153, 379]. One can use such an approach to study questions like: How likely is the failure of a single node to cause a major disruption in a system? In systems in which the failure of a single node or edge is unlikely to cause a major disruption, researchers have considered the fraction of nodes, if simultaneously removed, one can expect lead to a major disruption [101, 419, 454, 484, 498, 526, 563].

Simultaneous and sequential removal of nodes and edges. In models that do not have cascades, many researchers have studied the impact of simultaneous node or edge removal [105, 191, 207, 454, 526, 596] and sequential node removal [4, 231, 235, 238, 470, 604, 611]. If one is interested in modeling failures in a system, it is common to examine the impact of removing nodes uniformly at random [11, 101, 116]. If a network model neither includes dynamics on a network nor includes dynamics of a network's structure, the impact of simultaneous removal or sequential removal of nodes that one chooses uniformly at random leads to the same impact. When studying the impact of attacks on a network, it is common to consider the removal of nodes or edges according to a centrality-based ranking of nodes or edges [11, 253, 335, 552]. The distinction between simultaneous and sequential removal of nodes or edges matters in such studies, because their centrality-based rankings can change after the removal of a node or edge [4, 235].

Targeting strategies. For demonstrating that targeted node or edge removal can lead to larger impacts than removing nodes or edges uniformly at random, it usually suffices to compare targeting nodes by largest degree to removing nodes uniformly at random. For modeling the impact of an adversary's attack on a network, it is sensible to consider centrality measures that best approximate the adversary's attack strategy. If possible, the choice of a centrality measure should reflect one's knowledge of an adversary's behavior. One can also be interested in algorithms that can identify the sets of nodes or edges that, if removed, lead to the largest impact. It is possible that no node or edge centrality measure can identify such sets. Tools that researchers have developed for the study of collective influence [377, 378, 420, 613] and network dismantling [72, 448, 603] are likely to find node sets and edge sets that, if removed, lead to larger impacts than sets that one can construct from centrality-based rankings [72, 448]. Comparative studies of the impact of targeted attacks using different node and edge centrality measures can be useful for identifying such

measures that are good heuristics for finding node or edge sets that, if removed, have a large impact on a system's performance.

2.5.1.4 How should one choose a performance measure?

Guidance from additional features. Many studies of robustness use the size of an LCC as a measure of performance. However, for several models that include additional features (e.g., node labels or dynamics on networks), researchers compared the size of an LCC to one or more model-specific measures of performance, and they concluded that the size of an LCC is not a good proxy for performance measures that take model-specific features into account [307, 585]. When it is possible to include additional features in a network model, these additional features can guide a researcher in the specification of a notion of robustness that is more relevant for the study of a system than a notion of robustness that is based only on a simple network's structure [307].

Guidance from relevant types of walks and paths. When it is not possible to include additional features in a network model, it can be helpful to consider the types of walks and paths that one expects to contribute to functions of a system. For example, if one expects system function to depend on the existence of a path between as many node pairs as possible and if one assumes that all paths are permissible, then the size of an LCC, the relative size of an LCC, and reachability can be relevant measures of performance. If one assumes that walks or paths need to be short to contribute to a system function, the mean shortest-path length or efficiency can be relevant measures of performance. If one expects that one can approximate dynamics on a network that contribute to a system function reasonably well by a linear diffusion process or a spreading process, several spectral network properties (e.g., resistance distance, the spectral radius of $\mathbf{A}$, and natural connectivity) can be relevant measures of performance.

Robustness of simple networks. Choosing a relevant performance measure is especially difficult if no additional features are available, and it is unclear which types of walks or paths contribute to a system function. In such cases, many researchers have either used the size of an LCC as a measure of performance [61, 101, 231, 337, 419, 454, 498, 526, 563, 604] or examined the robustness of several measures of performance [11, 243, 335, 552]. In Chapter 3, we propose a framework for identifying relevant performance measures by testing their plausibility as performance measures for a given system.

2.5.2 Outlook

The search for "robust network structures" has been a very active field of research for at least two decades, but progress has been slow. A very challenging aspect of the study of network robustness is that a network's ability to perform a function depends in many cases on a dynamical process on the network [13, 139, 144, 198, 203, 225, 302, 379, 490, 522]. Understanding the links between network robustness and network structure thus depends on understanding the links between dynamics on networks and network structure, which is also an active field of research, with many open problems [37, 56, 277, 396, 401, 416].

We anticipate that the study of robustness of network models with additional features (e.g., dynamics on networks, node labels, or multilayer structure) can help accurately characterize network structures that are beneficial for a model-specific notion of robustness. A comparison of robust network structures for different notions of robustness has the potential to identify groups of robustness notions that yield similar "robust network structures".

Chapter 3

Plausiblity of performance measures for protein-interaction networks

The study of protein-interaction networks can facilitate the discovery of essential genes and proteins, drug targets, and biological pathways[1]. When representing a cellular function by a protein-interaction network, one typically constructs a network from a subset of all known protein interactions. The choice of such a subset of interactions matters greatly, because it can significantly affect structural properties of the resulting network and, in turn, the outcome of the study of such a network. We use deterministic seed-based sampling methods to sample *function-specific* protein-interaction networks (fPINs) from protein interactomes using sets of seed proteins that are associated with biological pathways, and we examine the robustness of these fPINs under node removal. We find that the structure of fPINs tends to be similar for several seed-based sampling methods. A test of the plausibility of structural network properties as measures of performance (i.e., viability) of fPINs indicates that the relative size of a network's largest connected component and its algebraic connectivity tend to be plausible measures of performance.

Our results demonstrate how the choice of sampling method, performance measure, and targeting strategy affects the outcome of a study of fPIN robustness. We anticipate that our plausibility test can be the basis for the study of plausibility of performance measures for many applications of network models and can thus help in the identification of suitable performance measures for a given network system.

[1] A pathway is a sequence of reactions between molecules in a cell [213, 341].

Symbol	Meaning
(V, E)	a network with node set $V = \{v_i\}_{i=1,\dots,N}$ and edge set $E = \{e_{ij}\}$
(V_S, E_S)	a network sample with node set $V_S \subseteq V$ and edge set $E_S \subseteq E$
V_{S_0}	a set of seed nodes
$\mathcal{S}$	a set of networks
N	a network's size (i.e., number of nodes)
m	a network's number of edges
ℓ	the length of a walk or a path
$G(N, m)$	an Erdős–Rényi random-graph model with fixed number of edges
k	a node's degree
$\mathbf{k} = (k_i)$	a network's degree sequence
$\mathbf{P} = (p_{ij})$	a network's assortativity matrix
$p(k)$	a network's degree distribution
C_i	local clustering coefficient of a node v_i
$c(k)$	mean local clustering coefficient of nodes with degree k
C	mean local clustering coefficient in a network
T	global clustering coefficient in a network
F	relative size of a network's largest connected component
a	algebraic connectivity of a network
L	mean shortest-path length in a network
$\mathcal{E}$	efficiency of a network
Ω	resistance distance of a network
K	natural connectivity of a network
$\kappa(i, j)$	a network's communicability function
$\tilde{\theta}(i, j)$	communicability angle between nodes v_i and v_j
$\tilde{\xi}(i, j)$	communicability distance between nodes v_i and v_j
θ	mean communicability angle in a network
ξ	mean communicability distance in a network
$I_{X,t,n,s,w}$	impact on a performance measure X of the sequential removal of n nodes via targeting strategy t in a network sample that we obtain via sampling strategy s for pathway w
σ	standard deviation
c_v	coefficient of variation
spr	Spearman rank correlation
$\mathrm{SR}_{X,t,n,s}$	score for selective robustness
$\mathrm{SS}_{X,t,n,s}$	score for structural sensitivity
$\mathrm{SM}_{X,t,n,s}$	score for perturbation–size monotonicity
$\mathrm{PS}^{\mathrm{min}}_{X,t,n,s}$	plausibility score based on minimum of criterion scores
$\mathrm{PS}^{\mathrm{prod}}_{X,t,n,s}$	plausibility score based on product of criterion scores

Table 3.1: List of symbols for Chapter 3.

3.1 Introduction

Proteins are crucial components of almost all cellular processes for the growth and survival of cells and organisms [201, 299]. Many databases of protein–protein interactions (PPIs) offer data for constructing protein interaction networks (PINs) [201, 465]. Researchers have studied structural properties of PINs to further understanding of pathways in biological systems [40, 91, 190, 262, 488], predict the importance of proteins to a cellular function [74, 109, 483, 510, 607], and assess its potential as a drug target [96, 242, 556, 614].

Many studies of PINs are studies of protein interactomes; that is, such PINs include all known PPIs for an organism [554]. However, several researchers have called into question whether protein interactomes are suitable network models of biological systems [214, 328]. One of the criticisms of the study of protein interactomes is that most interactions in a protein interactome are irrelevant for the study of a single pathway or cellular function, because many PPIs are not present when or where the cellular function of interest takes place [214]. For example, some PPIs are specific to a tissue or cell type and thus are not present in other tissues or cell types of an organism [69, 251]. Other PPIs are specific to environmental conditions [251, 557], a disease [19, 593], or an organism's age [251, 577]. Even when one has evidence that a PPI is present in a relevant tissue and under relevant environmental conditions for a cellular function of interest, the co-occurrence of the PPI and the cellular function may be coincidental [233, 249, 375, 527]. After all, cells perform many functions in parallel to maintain their growth and ensure their survival.

To circumvent these issues, many reseachers have studied subnetworks of protein interactomes [19, 69, 251, 256, 557, 577, 593]. Examples of such subnetworks are *tissue-specific* PINs, which include only the PPIs that are present in a certain tissue [69, 256]. Researchers have studied *differential* gene-interaction networks and PINs [251, 557]. These networks include only interactions that are stronger under some specified environmental conditions [557], disease [19, 593], or at a specified age of a cell or organism [577] than in a reference state. The study of differential gene-interaction networks and PINs has led to the discovery of biological pathways [251], disease biomarkers [24], and potential drug targets [341].

When a cell or organism experiences perturbations (such as a change in environmental conditions, the suppression of a gene, or the chemical inhibition of a protein), it can change its behavior. As part of such a change of behavior, alternative pathways that are inactive in the original state can become active in the perturbed state [286, 315]. In turn, such alternative pathways can use gene interactions and PPIs

that one does not observe in the unperturbed state [251]. Because one does not observe these interactions in the unperturbed state, they typically are not included in associated differential gene-interaction networks or PINs. However, they can be crucial for the study of robustness and resilience in biological systems [246]. In this chapter, we use *function-specific* PINs to represent cellular functions by a PIN that includes a known biological pathway and PPIs that can be part of alternative pathways. Function-specific PINs (fPINs) include PPIs that — with respect to the structure of a protein interactome, tissue-specific PIN, or differential PIN — are "close to" (specifically, in the neighborhood of) or "between" (specifically, on short paths between) proteins that are associated with a given cellular pathway. We construct fPINs by sampling neighborhoods of and paths between pathway-associated proteins.

Many studies of protein interactomes and samples of them aim to identify proteins that are important for the viability of a cell or organism [68, 127, 224, 415] or for a cellular function [74, 109, 483, 510, 607]. When the suppression of a gene leads to the death of a cell or organism, the gene and proteins that it encodes are called *essential* [68, 127, 162, 224, 415, 615]. When the inhibition of a protein leads to a desired change of a cellular function, one can interpret the protein as a potential drug target [96, 242, 287]. Researchers have proposed various computational methods for inferring essential proteins or the potential of a protein as a drug target from the structure of a PIN [162, 170, 352, 353, 415, 615]. Many of these computational methods use tools from network analysis, such as centrality measures [170, 415, 615] and several notions of robustness of various network properties to node removal [224, 288, 352, 353].

When modeling the impact of protein inhibition in a cell by node removal in a PIN, it is not clear which structural properties of a PIN are suitable proxies for the viability of an organism or the performance of a cellular function. In Chapter 2, we argued that one should use dynamics on networks, node labels, or other model features to guide one's search for relevant notions of robustness. However, to our knowledge, there do not exist mechanistic models of dynamics on PINs. One can use knowledge about proteins and their functions in biological systems to construct node labels for nodes in a PIN but, for many proteins, one knows none or only a subset of their functions in biological systems [141, 291, 369].

Many studies have used the size of the largest connected component (LCC) [552], the mean shortest-path length [592], and other structural properties of PINs

to indicate the performance[2] (i.e., viability) of a biological system [592]. Different studies of robustness of cellular functions have considered different network models of biological systems [53, 242, 256, 358], different network properties as measures of performance [552, 592], and different strategies for selecting nodes for removal [94, 138, 322, 552]. The variability of these specifications for computational studies of PIN robustness makes it difficult to build an intuition for the robustness of PINs to structural perturbations.

We construct fPINs that are much smaller than the corresponding protein interactome[3]. The relatively small size of fPINs makes it possible to consider many more networks, structural properties, and strategies for node removal when studying the robustness of fPINs than when studying the robustness of protein interactomes. The focus on fPINs thus enables us to conduct a large comparative study of the robustness of fPINs under node removal. In our investigation, we consider different methods for sampling fPINs, different measures of performance, and different strategies for selecting nodes for removal. We present the results of this study in this chapter.

Comparing the change of a performance measure on many networks and under many perturbations yields an impression of "typical" behavior of the performance measure under node removal. We develop *plausibility scores*, which measure the extent to which typical behavior of a performance measure is consistent with observations of the robustness of real biological systems. We use plausibility scores to identify the most plausible and the least plausible measures of performance among ten network properties that we selected based on proposals in previous studies [22, 49, 158, 259, 260, 355, 470, 552, 570, 588, 589]. For most choices of sampling methods for fPINs and targeting strategies, we find that the most plausible performance measures are the relative size of a network's LCC and algebraic connectivity (i.e., the second-smallest eigenvalue of a network's combinatorial Laplacian matrix). The least plausible among the examined measures are three measures related to a communicability function (see Section 3.2.4) and the mean local clustering coefficient.

As in previous studies of the robustness of PINs, the conclusions of our study are limited to our choices of data sets, methods for PIN constructions, sampling methods, performance measures, and targeting strategies. The results that we present

[2]Measures of performance quantify the success of a system at performing a specified function. For biological systems, one is almost always interested in a system's ability to survive. In the context of biological systems, we thus use the word "performance" and "viability" interchangeably.

[3]Sizes of fPINs range from 1000 to 2000 nodes. The corresponding protein interactome has 16431 nodes.

in this chapter are not "universal" properties of PINs or of networks in general. However, the comparison of several sampling methods, performance measures, and targeting strategies can contribute to building an intuition for the robustness of simple PINs to node removal. In particular, we anticipate that the comparison of results for different choices of sampling methods, performance measures, and targeting strategies can inform expectations of how a change of one of these choices may affect the results of a study of PIN robustness.

This chapter proceeds as follows. In Section 3.1.1, we give a brief overview of biological concepts that are relevant for this chapter. In Section 3.2, we describe the data sources and the methods that we use for this study. In Section 3.2.1, we give a brief introduction to the "Human Integrated Protein-Protein Interaction rEference" (HIPPIE) and describe how we use this data bank to construct a human protein interactome. In Section 3.2.2, we describe three random-graph models that we use in this chapter. In Section 3.2.3, we explain the sampling methods that we use to construct fPINs from a protein interactome. When studying the robustness of an fPIN, one has to choose a network property as a measure of performance or viability and a type of perturbation. We discuss the necessary specifications for a study of network robustness in Section 3.2.4. In Section 3.2.5, we list the structural properties that we consider for measures of performance of fPINs and give definitions for structural properties that we have not defined in Chapter 2. In Section 3.2.6, we list the centrality measures that we use to specify structural perturbations for fPINs. In Section 3.2.7, we introduce three criteria for the plausibility of performance measures on PINs and propose a method for combining measures of these criteria to a plausibility score of a performance measure. In Section 3.3, we present the results of our computational study of fPINs on the HIPPIE interactome. We investigate several structural properties of fPINs in Section 3.3.1. In Section 3.3.2, we study the robustness of several performance measures on fPINs under targeted node removal. In Section 3.3.3, we compare the plausibility scores of several measures of performance for several targeting strategies for node removal and for several fPIN-sampling methods. We summarize and discuss our results in Section 3.4.

3.1.1 Glossary of biological concepts

In this section, we give a brief overview of biological concepts that are relevant to this chapter.

Genes and proteins. A **gene** is a sequence of nucleotides in DNA that encodes a single hereditary characteristic [12, p. G:13]. Such a hereditary characteristic can be a single protein or a set of related proteins [12, p. 7]. The process by which a cell uses a gene's information to construct macromolecules is called **gene expression** [12, p. 7]. The resulting macromolecules are **gene products**. Many gene products are proteins [12, p. 7]. A **protein** is a large molecule that consists of amino acid residues [12, p. 4]. Proteins are important for many functions in cells and organisms. See Chapter 1 for examples.

Gene networks. When a gene is expressed, one says that the gene is *active* [571]. Otherwise, the gene is *inactive*. The activity of a gene can affect the activity of other genes. If the activity of gene i increases the activity of a gene j, one says that gene i *activates* gene j. If the activity of gene i decreases the activity of gene j, one says that gene i *inhibits* gene j. One can use these relationships between genes to construct **gene (interaction) networks** [15, 89, 252, 581].

Protein–protein interaction networks. There are various definitions of **protein–protein interactions** (PPIs) [465, 527]. In this dissertation, we focus on physical PPIs. A physical PPI indicates that two proteins can form a physical bond. These physical bonds are important for many protein functions, such as intra-cell and inter-cell signaling [12, p. 814] and enzymatic catalysis of metabolic reactions [12, p. 51]. One can use PPIs to construct a **protein–protein interaction network** (PIN) [221, 510].

Biological pathways. A **(biological) pathway** is a sequence of biochemical reactions between molecules in a cell [213, 341].

Interactomes, tissue-specific PINs, and function-specific PINs. A **protein interactome** is a PIN that includes all known PPIs for an organism [554]. A **tissue-specific PIN** includes all known PPIs for a tissue [256]. In this chapter, we introduce **function-specific PINs** as PINs that include PPIs that we anticipate to be relevant to a biological function or pathway.

Viability and essentiality. The **viability** of a system is its ability to survive. It is very common to associate the performance of a biological system with its viability. An **essential gene** is a gene that is necessary for an organism's viability [12, p. 9].

An **essential protein** is a protein that is necessary for an organism's viability [12, p. 15]. Essential proteins are gene products of essential genes.

Genotype, fitness, and genetic interactions. An organism's **genotype** is the set of genes that contribute to its physical characteristics [12, p. 485]. The **fitness** of a genotype is a measure of its ability to reproduce. The suppression of single gene can affect a genotype's fitness [257]. A **genetic interaction** between a pair i, j of genes indicates that the simultaneous suppression of genes i and j leads to a much larger change of fitness than one would expect from the change of fitness under suppression of a single gene i or j [257].

Disease biomarkers and drug targets. A **disease biomarker** is an experimentally measurable feature (e.g., the concentration of a protein) that indicates the existence or severity of a disease in a cell or organism [513]. A drug can achieve a therapeutic effect by binding to its **drug target** in an organism [598]. Many drug targets are proteins that are important to a disease pathway [573].

3.2 Materials and methods

3.2.1 The HIPPIE interactome

The group for Computational Biology and Data Mining (CBDM) at the Institute for Molecular Biology in Mainz has published an integrated PPI-data bank, which they called the "Human Integrated Protein-Protein Interaction Reference" (HIPPIE) [465]. This data bank includes data on human PPIs from several PPI databases, including HPRD [436], BioGRID [86], IntAct [230], and MINT [327], as well as data from studies that the CBDM group selected manually. Interactions in HIPPIE have a confidence value between 0 and 1 that depends on the number of studies that detected the interaction, the reliability of the experimental methods that were used in these studies, and the number of non-human organisms in which the interactions have been reproduced. For the results in this chapter, we use the variant of HIPPIE that Schaefer et al. released on 24 June 2016 [465]. By using all interactions irrespective of their confidence score, we prioritize network completeness over the reliability of interactions. The protein interactome that we construct from the HIPPIE data is a simple network. We refer to it as the "HIPPIE interactome".

3.2.2 Random-graph models

When investigating structural properties and robustness of networks, it can be helpful to compare the networks of interest with random networks. To generate random networks, we use three random-graph models: an Erdős–Rényi random-graph model, a simple configuration model, and a configuration model that preserves degree–degree assortativity. We describe these random-graph models in this subsection.

Erdős–Rényi random-graph model. The $G(N, m)$ Erdős–Rényi (ER) random-graph model samples graphs uniformly at random from a space of graphs with N nodes and m edges. For convenience, we refer to a network that we generate using the ER model as an "ER network". The expected degree distribution of an ER network is binomial. That is, the probability $p(k)$ for a node in an ER network to have degree k is

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

where $p = m/\binom{N}{2}$ [398, p. 346]. We use the implementation of the Erdős–Rényi random graph model in Python's graph-tool library [422].

Configuration model. A configuration model (CM) generates samples uniformly at random from a space of graphs with a fixed degree sequence or a sequence drawn from a specified degree distribution [184]. We consider a configuration model that samples a space of simple graphs (i.e., unweighted, undirected graphs without self-edges or multiedges) with a fixed degree sequence. To sample this CM, we use the implementation of CM in Python's graph-tool library [422]. The algorithm starts with a graph with the given degree sequence and then randomizes the structure via a Markov chain of edge swaps. We refer to a network that we generate using the CM as a "CM network".

Degree–degree assortativity-preserving configuration model. Newman (2002) proposed a model for random graphs with prescribed degree–degree assortativity [392]. Consider a graph (V, E) with a degree sequence $\mathbf{k} = (k_i)$ and assortativity matrix $\mathbf{P} = (p_{ij})$. An element p_{ij} denotes the probability that an edge in (V, E), chosen uniformly at random, connects a node with degree i to a node with degree j. For simple graphs, the matrix $\mathbf{P}$ is symmetric.

Newman (2002) proposed that one can uniformly sample the space of graphs with a given degree sequence $\mathbf{k}$ and assortativity matrix $\mathbf{P}$ using the following algorithm:

1. One generates a graph with the desired degree sequence $\mathbf{k}$ (e.g., using CM).

2. One randomizes the structure via a Markov chain of edge swaps. For each edge swap, one selects a pair of edges, e_{i_1,i_2}, $e_{j_1,j_2} \in E$, uniformly at random. One accepts an edge swap with probability

$$\min\left\{1, \frac{e_{k_{i_1}k_{j_1}} p_{k_{i_2}k_{j_2}}}{p_{k_{i_1}k_{i_2}} p_{k_{j_1}k_{j_2}}}\right\},$$

where k_i is the degree of node i. In our implementation of this model, we ensure that graphs are simple by rejecting edge swaps that lead to self-edges.

Newman (2002) did not indicate the time that the Markov chain requires to converge to a graph that has a similar assortativity matrix to $\mathbf{P}$. Using Newman's algorithm for the HIPPIE interactome, we find that the assortativity matrix of a sample of DA tends to be identical or very similar to the assortativity matrix of the HIPPIE interactome after each edge has undergone ten swaps, on average. For constructing a random network with a prescribed degree sequence and degree–degree assortativity, we thus use Newman's algorithm and terminate the Markov chain after $10|E|$ edge swaps[4]. We refer to this random-graph model as the *degree–degree-assortativity-preserving configuration model* (DA). We call a network that we generate using the DA a "DA network".

3.2.3 Sampling function-specific protein-interaction networks

An fPIN is a subnetwork of a protein interactome. Sampling methods are sets of rules for constructing a subnetwork. For a network (V, E), we call a subnetwork (V_S, E_S) obtained via sampling a *network sample* of the *parent network* (V, E). We use deterministic seed-based sampling methods: Given a set $V_{S_0} \subset V$ of seed nodes, we use sets of deterministic rules to construct (V_S, E_S). For examples of stochastic seed-based sampling methods, see Refs. [113, 297, 314, 316].

We consider three basic deterministic, seed-based sampling methods: 1-hop snowball (SB) sampling [314], shortest-path (SP) sampling [435], and 3-hop all-path (AP) sampling. When using ℓ-hop snowball sampling, a network sample for a seed-node set V_{S_0} is the induced subgraph on the node set V_S that is the union V_{S_0} and the q-hop neighbourhoods of all nodes in V_{S_0}, where $q \leq \ell$ (and q is positive) [314]. A shortest-path sample for a seed-node set V_{S_0} is the induced subgraph on

[4]Bassler et al. (2015) presented an algorithm for sampling networks with a prescribed degree sequence and assortativity matrix [39]. Their algorithm does not rely on the convergence of a Markov chain of edge swaps. We became aware of this algorithm after completing our simulations.

the node set V_S that is the union of all node sets $V_{P_{ij}}$ of shortest paths between pairs of nodes $v_i, v_j \in V_{S_0}$ [435]. An ℓ-hop all-path sample for a seed-node set V_{S_0} is the induced subgraph on the node set V_S that is the union of all node sets $V_{P_{ij}}$ of length-q paths between pairs of nodes $v_i, v_j \in V_{S_0}$, where $q \leq \ell$ (and q is positive).

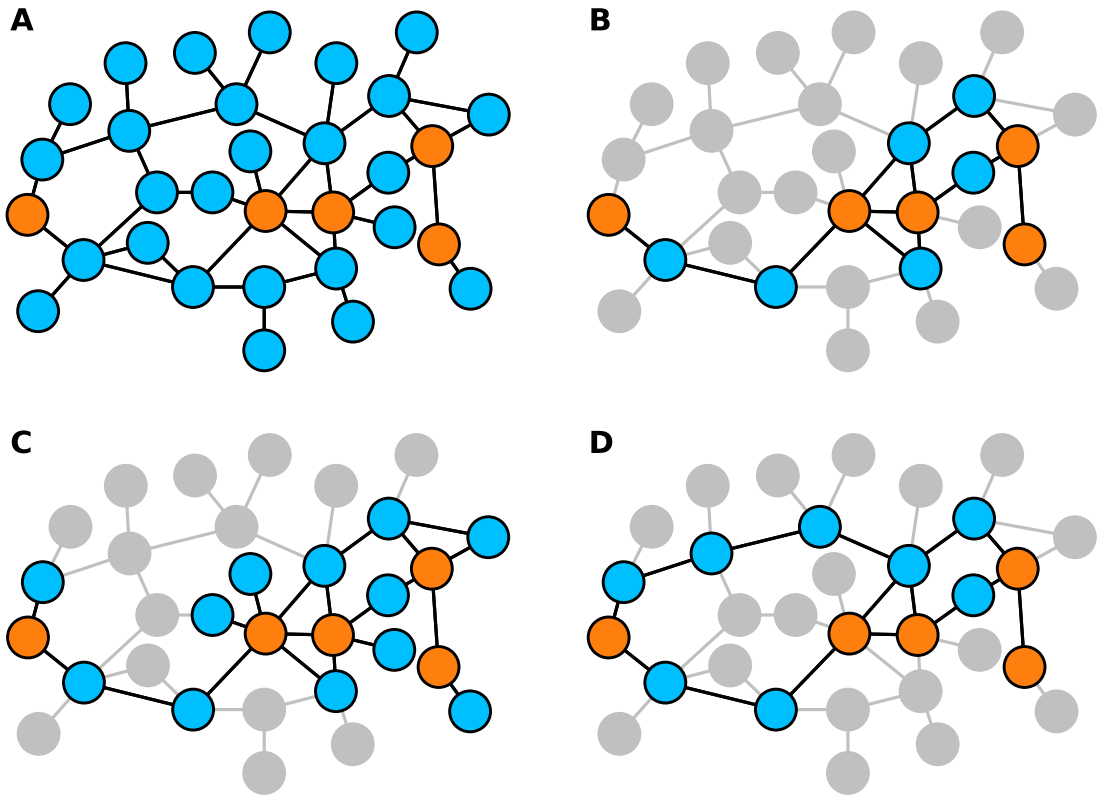


Figure 3.1: Sampling methods for fPINs. In (**A**), we show an example network with an example set of seed nodes. We show seed nodes in orange and the remaining nodes in blue. In (**B**)–(**D**), we show the network's nodes and edges in grey. We indicate network samples by black edges and blue and orange nodes. The network samples in (**B**), (**C**), and (**D**) are network samples that we construct via AP sampling, SB sampling, and SP sampling, respectively.

In Fig. 3.1, we show examples of AP samples, SB samples, and SP samples of a small example network. We see that the AP sample is a subnetwork of the SB sample. A node that is part of a length-3 path between seed nodes is a neighbor of a seed node. Therefore, it is generally the case that AP samples are subnetworks of corresponding SB samples.

We also consider combinations of AP sampling, SB sampling, and SP sampling. When combining two sampling methods, the resulting network sample is the induced subgraph of the union of node sets that one obtains for each of the sampling methods. Because the union of AP sample and the corresponding SB sample identical to the

SB sample, we consider only the combination of AP sampling and SP sampling and the combination of SB and SP sampling. We list the sampling methods and their respective acronyms in Table 3.2.

Acronym	Sampling method
AP sampling	3-hop all-path sampling
SB sampling	1-hop snowball sampling
SP sampling	shortest-path sampling
APSP sampling	combination of 3-hop all-path sampling and shortest-path sampling
SBSP sampling	combination of 1-hop snowball sampling and shortest-path sampling

Table 3.2: Acronyms of sampling methods for fPINs. Underlined characters in the name of a sampling method contribute to its acronym.

To sample fPINs, we construct seed-node sets from biological pathway data from the Reactome Pathway Knowledgebase [173]. Each set of seed nodes is a set of proteins that are known to be relevant to a disease, a metabolic pathway, a signal transduction pathway, or some other cellular function in humans.

3.2.4 Impacts of structural perturbations

To examine the robustness of an fPIN, we study the impact of structural perturbations on structural properties of fPINs. We define the *impact* I of a structural perturbation to be the absolute value $|(X_0 - X')/X_0|$ of the relative change of a structural property X, where X_0 is the value of X in the unperturbed network and X' is the value of X in the perturbed network. The impact of a perturbation depends on the choice of a performance measure X, the perturbation, and the network (V, E). As structural perturbations, we consider the removal of a prescribed number n of nodes, which we select according to a targeting strategy t. We can thus specify a perturbation by t and n. We measure impacts of perturbations on fPINs that we obtain via a deterministic seed-based sampling method s with a set of seed nodes that corresponds to a pathway w. We can thus specify an fPIN by s and w and, consequently, we can specify an impact I by X, t, n, s, and w: The impact $I_{X,t,n,s,w}$ is the relative change of the performance measure X when one removes n nodes according to the targeting strategy t from a network that we obtain via the sampling strategy s for a biological pathway w. We call a sequence $(I_{X,t,n=0,s,w}, I_{X,t,n=1,s,w}, I_{X,t,n=2,s,w}, \dots)$ of

impacts a *robustness profile*. To avoid the inconvenience of keeping track of many subscripts, we adopt the following conventions:

1. We use $I_{X,t,n,s}$ to denote the mean impact over a set of fPINs that we obtain using all w in the Reaction Pathway Knowledgebase.

2. We drop further subscripts when statements about impacts apply for all considered choices of the respective specifier. For example, we write $I_{n=0} = 0$ to convey that $I_{X,t,n,s} = 0$ for $n = 0$ and all considered choices of X, t, and s.

3. In the subscript that indicates the performance measure X, we drop the explicit assignment "$X=$". For example, we write $I_F \in [0, 1]$ to denote $I_{X=F} \in [0, 1]$.

4. We use $\langle\cdot\rangle$ to denote mean values and use subscripts on brackets to denote the specifiers over which we average. For example, we write $\langle I_{X,t,n}\rangle_s$ to denote the mean of $I_{X,t,n,s}$ over all considered sampling methods. We write $\langle I_{X,n}\rangle_{t,s}$ to denote the mean of $I_{X,t,n,s}$ over all considered sampling methods and all considered targeting strategies.

When averaging over sampling methods, we take the mean over AP, SB, SP, APSP, and SBSP sampling unless we specify otherwise. When averaging over targeting strategies, we take the mean over the four targeting strategies in Section 3.2.6. A mean $\langle\cdot\rangle_n$ denotes the mean over $n \in \{1, \ldots, 100\}$.

In Section 3.2.7, we use the came conventions for subscripts on *criterion scores* and *plausibility scores*.

3.2.5 Performance measures

We consider ten different network properties as measures X of an fPIN's performance: relative LCC size F, algebraic connectivity a, mean-shortest path length L, communication efficiency $\mathcal{E}$, global clustering coefficient T, mean local clustering coefficient C, resistance distance Ω, natural connectivity K, mean communicability angle θ, and mean communicability distance ξ. In addition to structural properties that we have discussed in Section 2.4.4, this list includes the mean communicability angle, and mean communicability distance. We give definitions for these network properties in this subsection.

Using the *communicability function* (see Eq. (2.5)), one can define the *communicability distance* [164]

$$\tilde{\xi}(i, j) := \kappa(i, i) + \kappa(j, j) - 2\kappa(i, j)$$

and the *communicability angle* [167, 172]

$$\tilde{\theta}(i,j) := \arccos\left(\frac{\kappa(i,j)}{\sqrt{\kappa(i,i)\kappa(j,j)}}\right).$$

Estrada and collaborators have proposed to use mean communicability distance [164]

$$\xi := \binom{N}{2}^{-1} \sum_{\substack{i,j=0,\\ i\neq j}}^{N} \tilde{\xi}(i,j)$$

and the mean communicability angle [167, 172]

$$\theta := \binom{N}{2}^{-1} \sum_{\substack{i,j=0,\\ i\neq j}}^{N} \tilde{\theta}(i,j)$$

as measures of network functionality.

3.2.6 Targeting strategies

The impact of removing n nodes from a network depends crucially on the choice of the n nodes [117, 600]. In this chapter, we consider five strategies for choosing nodes for sequential removal: choosing nodes uniformly at random or successively removing the most "important" node, where we measure node "importance" by largest degree, largest betweenness centrality, largest eigenvector centrality, or largest subgraph centrality. (See Section 2.4.3.2 for definitions of centrality measures.) Generally, centrality measures are sensitive to changes in a network's structure, and several researchers have stressed the importance of using *adaptive targeting strategies* for which one continuously updates centrality measures when sequentially removing several nodes [231, 235]. We therefore recalculate centrality measures after each node removal before choosing the next node to be removed.

3.2.7 A measure of plausibility

We design a test to quantify the "plausibility" of a structural property for measuring the performance of fPINs. To achieve a high *plausibility score* (PS), we require a performance measure to satisfy three criteria: *selective robustness*, *structural sensitivity*, and *perturbation-size monotonicity*. In Secs. 3.2.7.1–3.2.7.3, we discuss these criteria and introduce *criterion scores* that measure the extent to which a performance measure satisfies one of these criteria.

3.2.7.1 Selective robustness

Many scholars have proposed that biological systems are generally robust to random failures, but vulnerable to targeted attacks (with various types of targeting) [98, 177, 221, 284]. We call this criterion for plausibility of performance measures *selective robustness*. We introduce a criterion score $\mathrm{SR}_{X,t,n,s}$ as a measure of the extent to which impacts are larger under targeted attacks than under uniformly random node removal. We define

$$\mathrm{SR}_{X,t,n,s} := \left\langle \frac{I_{X,t,n,s,w} - I_{X,t=R,n,s,w}}{I_{X,t,n,s,w}} \right\rangle_w ,$$

where $I_{X,t=R,n,s}$ denotes the impact of removing n nodes uniformly at random[5]. If $I_X > 0$, then $\mathrm{SR}_X \in [0, 1]$.

3.2.7.2 Structural sensitivity

The impact of node removal should depend on the structure of the network from which one removes nodes. Therefore, for a plausible performance measure, the impact of removing n nodes should, in general, be different for different networks. We use the coefficient of variation c_v of impacts on different networks to measure the sensitivity of impacts to network structure. We define the criterion score

$$\mathrm{SS}_{X,t,n,s} := c_v(\{I_{X,t,n,s,w}\}_w) ,$$

where $\{I_{X,t,n,s,w}\}_w$ is the set of impacts $I_{X,t,n,s,w}$ for all considered pathways w.

3.2.7.3 Perturbation-size monotonicity

Many researchers have studied the change of fitness of model organisms (e.g., the yeast species *S. cerevisiae*) when one suppresses one or several genes in the organism's genome [107–109, 329, 351, 545]. In these studies, researchers assign a *fitness* $f \in [0, 1]$ to a genotype. A widely used model for the fitness $f(\{i, j\})$ of combined gene suppression is the "product rule": Based on experimental observation, one expects the fitness after suppression of a gene i and a gene j to be approximately $f(\{i\})f(\{j\})$, where $f(\{i\})$ is the fitness after the suppression of gene i and $f(\{j\})$ is the fitness after the suppression of gene j [108, 257, 429, 505]. For some gene pairs i, j, the product rule does not hold. *Positive* (or *negative*) *genetic interactions* indicate gene pairs for which $f(\{i, j\})$ is much greater (or much smaller) than one

[5]For the results that we present in this dissertation, we use the mean over 20 realisations of removing nodes uniformly at random.

would expect from the product rule[6] [108]. Extreme cases of genetic interactions in which the suppression of two genes instead of one gene changes whether an organism is viable are known as *syntethic rescue* and *synthetic lethality* [351]. Various studies of gene suppression in model organisms have indicated that genetic interactions are rare [107–109, 329, 545]. For example, Costanzo et al. (2016) tested 36 million gene pairs in the yeast *S. cerevisiae* and reported that genetic interactions account for roughly 2.5% of the tested gene pairs [109]. The product rule thus appears to hold for the vast majority of gene pairs.

When using an impact measure I_f to model the impact of gene suppression on fitness, the product rule implies that

$$I_f(\{v_i, v_j\}) \gtrapprox I_f(\{v_i\}),$$

where $I_f(\{v_i, v_j\})$ denotes the impact of removing the nodes v_i and v_j from a gene interaction network or PPI network. If one removes more than two nodes,

$$I_f(V') \gtrapprox I_f(V''), \tag{3.1}$$

where $V'' \subset V' \subset V$. If impacts I_X satisfy Eq. (3.1), we say that X satisfies the criterion *perturbation-size monotonicity*[7].

We introduce a criterion score $\mathrm{SM}_{X,t,s}$ as a measure of the extent to which typical behavior of a performance measure X leads to impacts $I_{X,n}$ that increase monotonically with perturbation size n. We define $\mathrm{SM}_{X,t,s}$ to be the mean Spearman rank correlation $\mathrm{spr}(n, I_{X,t,n,s,w})$ of the number n of removed nodes and the impact $I_{X,t,n,s,w}$. That is,

$$\mathrm{SM}_{X,t,s} := \langle \mathrm{spr}(n, I_{X,t,n,s,w}) \rangle_w \in [0, 1].$$

3.2.7.4 Plausiblity scores

We intend the plausibility score $\mathrm{PS}_{X,t,s}$ to be a number that signifies the extent to which the typical behavior of X satisfies selective robustness, structural selectivity, and perturbation-size monotonicity. There are many ways to define a function that maps a set of criterion scores to a single number $\mathrm{PS}_{X,t,s}$. We consider selective robustness, structural selectivity, and perturbation-size monotonicity to be necessary criteria for the plausibility of a performance measure. We thus require $\mathrm{PS}_{X,t,s}$ to be

[6] For example, Jaffe et al. (2019) measured fitness values f for the suppression of one or two genes and computed the standard deviation σ of the deviation $f(\{i,j\}) - f(\{i\})f(\{j\})$ from the product rule [257]. They considered a gene pair i and j to have a genetic interaction if $f(\{i,j\}) - f(\{i\})f(\{j\}) > 2\sigma$.

[7] We use *perturbation size* to mean the number n of removed nodes.

(1) approximately 0 if at least one of the three criterion scores is approximately 0 and (2) monotonically non-decreasing in all three criterion scores. Some of the most obvious ways of combining a list of numbers into a single number — e.g., averages and the Euclidean vector norm — do not satisfy the first requirement. Therefore, we do not use averages or vector norms to construct plausibility scores. Instead, we focus on two approaches. In the first approach, we take the minimum

$$\mathrm{PS}^{\mathrm{min}}_{X,t,s} = \langle \min\{\mathrm{SR}_{X,t,s,n}, \mathrm{SS}_{X,t,s,n}, \mathrm{SM}_{X,t,s}\} \rangle_n$$

of criterion scores to measure the plausibility of a performance measure. In the second approach, we take the product

$$\mathrm{PS}^{\mathrm{prod}}_{X,t,s} = \langle \mathrm{SR}_{X,t,s,n}\, \mathrm{SS}_{X,t,s,n}\, \mathrm{SM}_{X,t,s} \rangle_n \ ,$$

as a plausibility score for a performance measure. These two plausiblity scores are monotonically non-decreasing in SR, SS, and SM; and they equal 0 if SR, SS, or SM equals 0.

The plausiblity scores $\mathrm{PS}^{\mathrm{min}}$ and $\mathrm{PS}^{\mathrm{prod}}$ are just two of many functions that are monotonically non-decreasing in their input variables and approximately 0 if at least one of the input variables equals 0. We select $\mathrm{PS}^{\mathrm{min}}$ and $\mathrm{PS}^{\mathrm{prod}}$ for their simplicity, which we consider to be an advantage. However, both plausiblity scores also have disadvantages. The plausiblity score $\mathrm{PS}^{\mathrm{min}}$ is equal to the smallest criterion score. If one of the criterion scores is always smaller than the others, the plausibility score $\mathrm{PS}^{\mathrm{min}}$ of a performance measure depends only on that criterion score. This issue does not arise when using the plausibility score $\mathrm{PS}^{\mathrm{prod}}$. Its disadvantage is that if one of the criterion scores is close to 0 but positive, the product of the criterion scores does not have to be close to 0. For example, if the three criterion scores take the values 1, 1, and 0.1, respectively, one obtains $\mathrm{PS}^{\mathrm{prod}} = 0.1$. One can obtain the same plausibility score when none of the criterion scores are close to 0 (for example, when the three criterion scores have the values 0.5, 0.5, and 0.4).

3.3 Results

We present the results of our computational study of fPINs on the HIPPIE interactome. We investigate several structural properties of fPINs in Section 3.3.1. In Section 3.3.2, we study the robustness of several performance measures on fPINs under targeted node removal. In Section 3.3.3, we compare the plausibility scores of several measures of performance for several targeting strategies for node removal and for several fPIN-sampling methods.

3.3.1 Structure of function-specific protein-interaction networks

We investigate several structural properties of fPINs. In Section 3.3.1.1, we compare structural properties of fPINs to their parent network. In Section 3.3.1.2, we compare structural properties for fPINs for different sampling methods to structural properties of random networks. We investigate the relationship between several structural properties and fPIN size in Section 3.3.1.3.

3.3.1.1 Structural properties of the protein interactome and network samples

From the Reactome Pathway Knowledgebase [173], we obtain 1869 sets of proteins. After excluding duplicate sets, sets with fewer than three nodes, and sets that include proteins that we cannot map to proteins in the HIPPIE interactome, we are left with 1127 sets of proteins that we use as seed-node sets for fPINs. For these seed-node sets, we sample fPINs using one of the five sampling methods in Table 3.2.

We consider networks with fewer than 300 nodes to be too small and networks with more than 2000 nodes to be too large for the analysis of robustness of fPINs and plausibility of performance measures in subsequent sections. On one hand, network samples that include a large portion of an interactome's nodes are unlikely to be good models of a specific pathway[8]. On the other hand, for very small network samples, removing 100 nodes is either not possible (if $N < 100$) or is likely to fragment the network samples into many small components, irrespectively of the considered targeting strategy.

In Fig. 3.2, we show the number of fPINs of size N for in $N \in [300, 2000]$. We find that the number of fPINs tends to decrease as N increases. The numbers of networks with $N \geq 1000$ is particularly small in the set of SP-sampled fPINs. For the set of AP-sampled fPINs, there are more networks with $N \geq 1000$ than for the SP-sampled set. For all other sets of fPINs, the distribution of network sizes is similar to the distribution in the AP-sampled set.

In Table 3.3, we list the size of the resulting sets of fPINs and the mean values for some of their structural properties. For each sampling method, the network samples include, on average, 6% or fewer of the nodes in the HIPPIE interactome and 8% or fewer of the edges of the HIPPIE interactome. The SP samples are smaller than the SB samples and AP samples. On average, the SP samples include 5% of the

[8] In initial calculations, we also included 2-hop snowball sampling. However, almost all network samples that we obtain via 2-hop snowball sampling include more than half of the interactome's nodes, so we do not consider these network samples for further analysis.

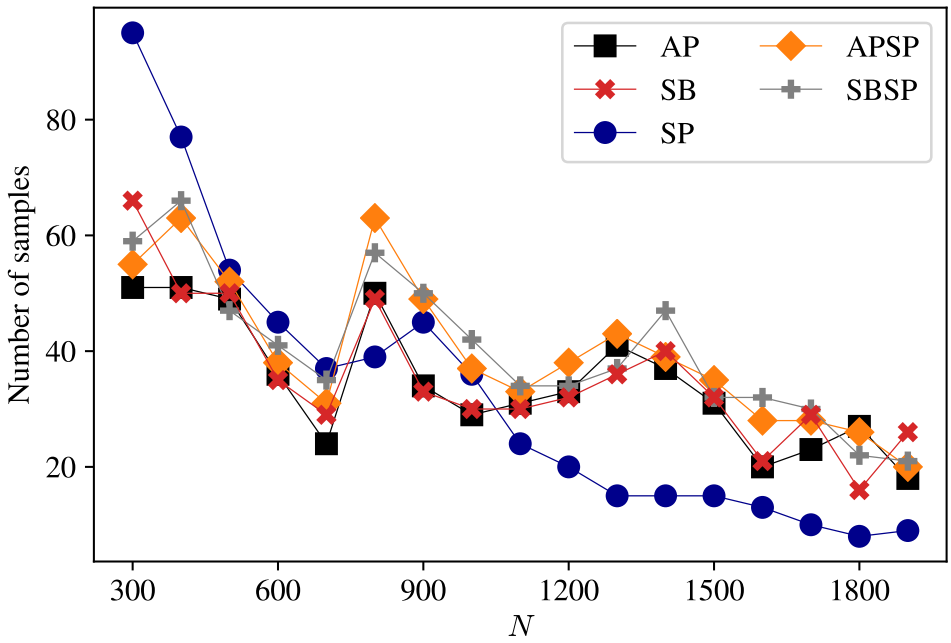


Figure 3.2: Sizes of network samples. We show the number of fPINs of size N for different sampling methods: 1-hop snowball (SB) sampling, shortest-path (SP) sampling, 3-hop all-path (AP) sampling, and combinations (APSP, and SBSP sampling). We use a bin size of $\Delta_N = 100$.

HIPPIE interactome's node set and 7% of the interactome's edge set. The mean degrees $\langle k \rangle$ in fPINs are similar to the mean degree in the HIPPIE interactome. The fPINs tend to have shorter L and much smaller values of C, T, and K than the HIPPIE interactome. The differences in mean shortest-path lengths and in natural connectivity between the HIPPIE interactome and fPINs One can attribute some of the structural differences between the HIPPIE interactome and fPINs to their different sizes. For example, a network's natural connectivity has a strong positive correlation to the network's number of edges [166]. One can thus expect a network with many edges to have a larger natural connectivity than a network with few edges. By contrast, the large differences in C and T between the HIPPIE interactome and fPINs is noteworthy. When sampling subnetworks from a variety of different networks, Smith and Moody (2013) reported that the mean local clustering coefficient and the global clustering coefficient are almost unchanged when removing up to 70% of a network's nodes uniformly at random [496]. Samples of networks do not need to have much smaller mean local and global clustering coefficients than their parent network. However, our samples of the HIPPIE interactome do have much smaller values of C and T than their parent network.

Sampling method	$\lvert\mathcal{S}\rvert$	N	$\langle k\rangle$	L	C	T	K
None (full protein interactome)	1	16431	32.8	3.08	0.1255	0.0672	161.7
AP	585	1035 (481)	36.3 (11.3)	2.29 (0.11)	0.0425 (0.0079)	0.0309 (0.0066)	73.6 (27.8)
SB	604	1022 (491)	34.5 (12.0)	2.33 (0.13)	0.0410 (0.0082)	0.0307 (0.0065)	71.2 (29.0)
SP	557	820 (428)	43.7 (12.9)	2.17 (0.11)	0.0428 (0.0078)	0.0370 (0.0081)	77.7 (27.2)
APSP	686	1038 (476)	38.5 (11.1)	2.26 (0.10)	0.0417 (0.0077)	0.0319 (0.0068)	77.4 (26.9)
SBSP	686	1027 (476)	36.9 (11.4)	2.2 (0.10)	0.0420 (0.0073)	0.0318 (0.0069)	75.4 (27.3)

Table 3.3: Numbers and structural properties of sampled function-specific protein-interaction networks (fPINs). For each sampling method, we show the number $|\mathcal{S}|$ of fPINs, the mean values of the number N of nodes, the mean degree $\langle k\rangle$, the mean shortest-path length L, the mean local clustering coefficient C, the global clustering coefficient T, and the natural connectivity K in a set $\mathcal{S}$ of fPINs. Blue numbers in parantheses indicate standard deviations. For comparison, we show structural properties of the full HIPPIE interactome in the first row.

In Fig. 3.3, we compare the degree distributions $p(k)$ for the HIPPIE interactome and a set of fPINs. As an example of a set of fPINs, we use the set of SP-sampled fPINs. We also show the mean local clustering coefficients $c(k)$ for the HIPPIE interactome and the SP-sampled fPINs, where, for each value of k, the mean local clustering coefficient $c(k)$ is the mean of C_i over all nodes with degree k in a network. We show results for other sampling methods in Appendix A.1.

The mean degree distribution of the set of SP-sampled fPINs is qualitatively different from the degree distribution of the HIPPIE interactome. For $k \in \{1, \ldots, 100\}$, the degree distribution $p(k)$ of the HIPPIE interactome tends to decrease with increasing k. Over the same range of k values, the mean degree distribution of the SP-sampled fPINs has a maximum at $k = 6$. Low-degree nodes (specifically those with $k < 6$) are less frequent in SP-sampled fPINs than in the HIPPIE interactome. Nodes with intermediate degrees, which we take to satisfy $6 \leq k \lessapprox 130$, are more frequent in the SP-sampled fPINs than in the HIPPIE interactome or have a similar frequency. Nodes with $k \geq 137$ are very rare in the set of SP-sampled fPINs. However, five large-degree nodes (with degrees $k = 944$, $k = 996$, $k = 1122$, $k = 1167$,

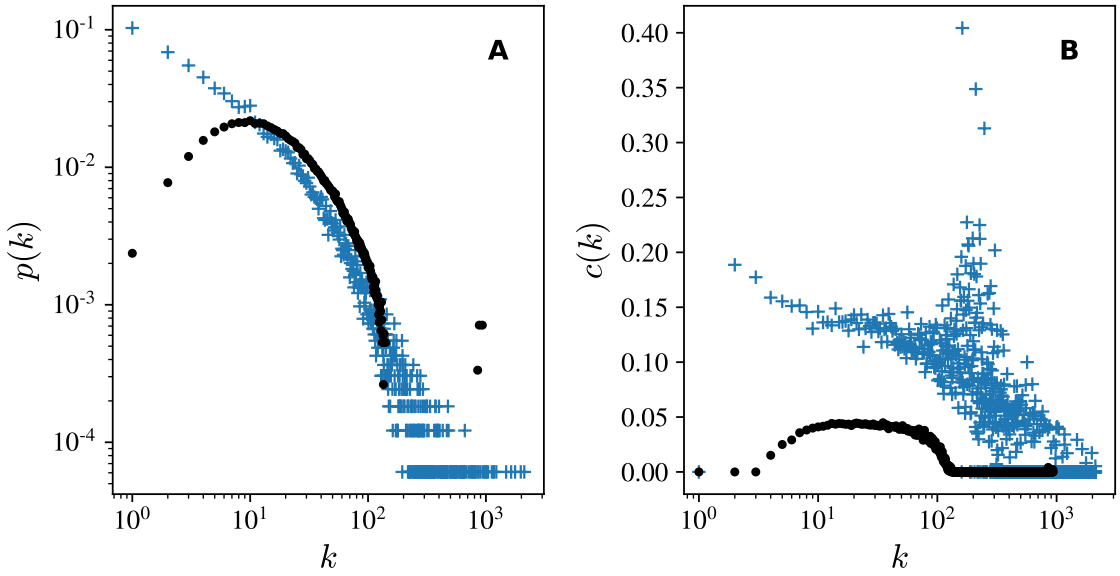


Figure 3.3: Degree distributions and mean local clustering coefficients in the HIPPIE interactome and fPINs. In (**A**), we show the degree distribution $p(k)$ of the HIPPIE interactome (blue crosses) and the mean degree distribution for the set of SP-sampled fPINs (black dots). In (**B**), we show the distributions $c(k)$ of mean local clustering coefficients.

and $k = 1637$) have a higher frequency in the SP-sampled fPINs than in the HIPPIE interactome.

The mean local clustering coefficient $c(k)$ for the SP-sampled fPINs is very different from $c(k)$ of the HIPPIE interactome. For $k \in \{2, \ldots, 100\}$, the function $c(k)$ in the HIPPIE interactome tends to decrease with increasing k. In SP-sampled fPINs, the function $c(k)$ tends to increase with k when $k \lessapprox 25$ and tends to decrease with increasing k when $k \gtrapprox 25$. For all k with $p(k) > 0$ for the SP-sampled fPINs, the mean local clustering coefficient in the fPINs is less than or equal to the value of $c(k)$ in the HIPPIE interactome. This observation is consistent with our previous observation that the mean local clustering coefficient C of the HIPPIE interactome is much larger than the mean C for the set of fPINs. A notable observation on $c(k)$ for SP-sampled fPINs is that $c(k) \approx 0$ for $k \leq 3$. A value $c(k) \approx 0$ indicates that the neighbors of a node with degree k are almost never adjacent to each other. Our observation of $c(k) \approx 0$ for $k \leq 3$ thus indicates that, in SP-sampled fPINs, the neighbors of nodes with a degree of 2 or 3 are almost never adjacent to each other. For $k = 2$, we also observe that $c(k) \approx 0$ for other sampling methods (see Appendix A.1).

The results in Fig. 3.3 are different to what one would expect when sampling

subnetworks by selecting nodes uniformly at random. Stumpf et al. (2005) demonstrated that uniform sampling of networks with power-law degree distributions leads to subnetworks that have higher probabilities $p(k)$ for nodes with degree $k = 1$ and lower probabilities for nodes with degree $k \geq 2$ [516]. Smith and Moody (2013) reported that the mean local clustering coefficient C of a uniform network sample is similar to the value of C of its parent network [496]. We therefore believe that the differences that we observe between the structural properties of the HIPPIE interactome and fPINs are related to the sampling methods that we use. However, we found that fPINs tend to have lower $p(k)$ than the HIPPIE interactome for $k \lessapprox 10$.

3.3.1.2 Sampling method and network structure

In this subsection, we compare structural properties of fPINs that we sample using different sampling methods to structural properties of several types of random networks.

In Fig. 3.4, we show the mean degree distributions $p(k)$ for sets of fPINs that we obtain via different sampling methods. On the left, we show mean degree distributions for sets of fPINs that we obtain via the path-based sampling methods SP and AP sampling, as well as a combination of both (APSP sampling). On the right, we show the mean degree distributions for sets of fPINs that we obtain via SB sampling and SBSP sampling. For all considered sampling methods, we observe that nodes with degree $k = 1$ are less frequent in fPINs than in the HIPPIE interactome. When comparing the different sets of fPINs, we find that the probability $p(k = 1)$ is lowest (specifically, $p(k = 1) < 0.002$) for AP-sampled and APSP-sampled fPINs, low (specifically, $p(k = 1) \in [0.002, 0.003)$) for SP-sampled fPINs, and largest (specifically, $p(k = 1) \geq 0.003$) for fPINs that we obtain via SB sampling or combinations of SB sampling and path-based sampling methods. Another difference between the mean degree distribution for sampling methods with SB sampling and the mean degree distributions for sampling methods without SB sampling is that nodes with very large degree (specifically nodes with $k \gtrapprox 1000$) are more frequent in fPINs that we obtain via sampling methods with SB sampling than in those that we obtain via path-based sampling methods.

For many networks, the degree distributions of network samples are qualitatively different from the degree distribution of their parent network [516]. In general, the degree distribution of a network sample depends on the degree distribution of their parent network and on the sampling method that we apply to construct the sample [314, 516]. The fPINs in our study have the same parent network (i.e.,

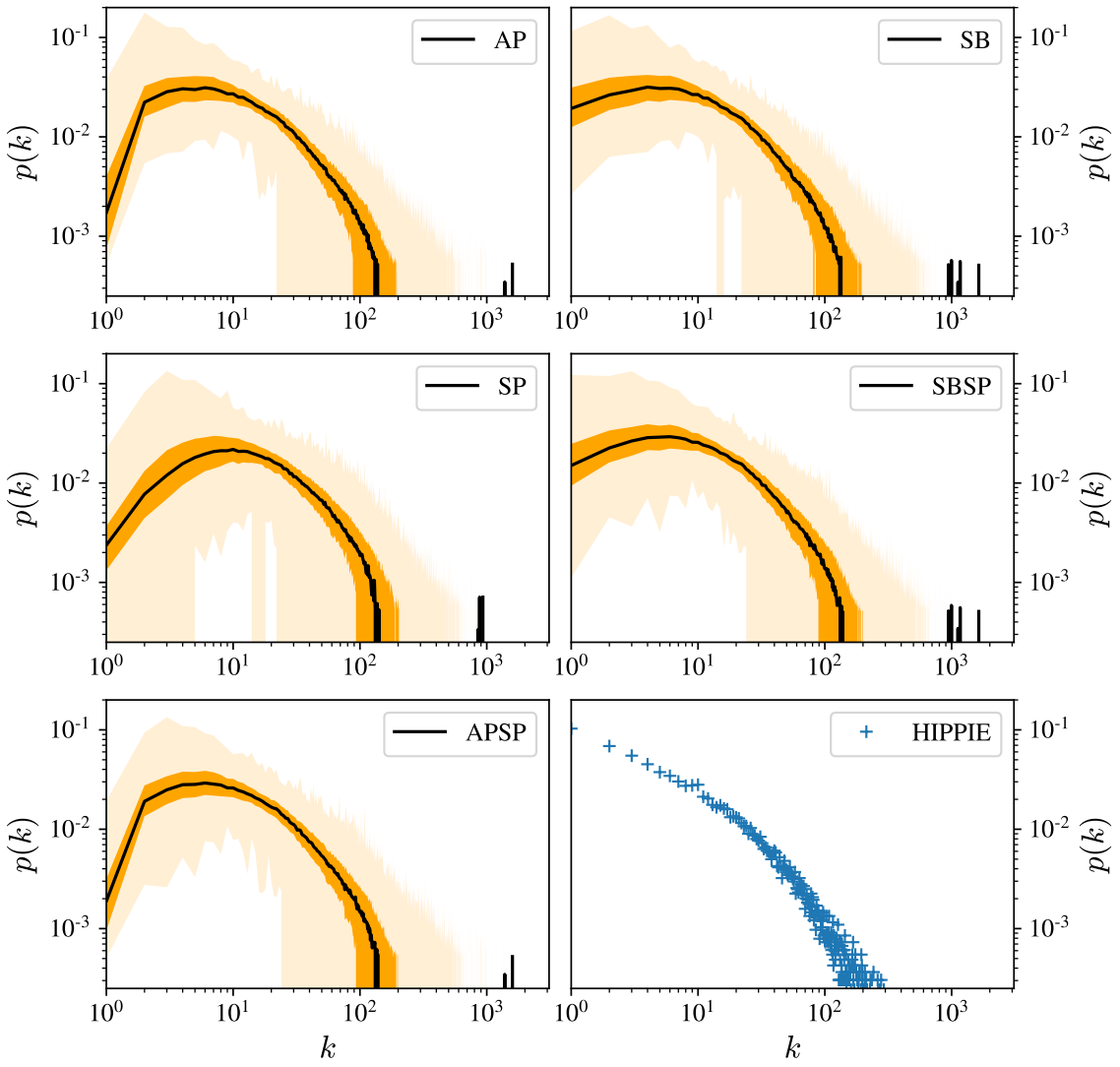


Figure 3.4: Mean degree distributions and mean distributions for sets of fPINs. Black curves indicate the mean degree distribution for a set of fPINs. The orange shaded area indicates the interval from the first to the third quartile of the values of $p(k)$ (i.e., an area in which 50% of values of $p(k)$ for fPINs lie in this interval). The light-orange area indicates the interval from smallest to largest value of $p(k)$. On the left, we show mean degree distributions for sets of fPINs that we obtain via the path-based sampling methods 3-hop all-path (AP) sampling, shortest-path (SP) sampling, and a combination of both (APSP). On the right, we show the mean degree distributions for sets of fPINs that we obtain via 1-hop snowball (SB) sampling and combinations of SB and SP sampling. For comparison, we show the degree distribution for the HIPPIE interactome in the bottom-right panel (blue crosses).

the HIPPIE interactome). By comparing how different sampling methods select nodes from a parent network, we can explain the differences between the mean degree distributions for sampling methods with SB sampling and the mean degree distributions for sampling methods without SB sampling. When using a path-based sampling method, nodes in an fPIN's node set V_S are either nodes from the seed-node set V_{S_0} or nodes that lie on a path that connects two seed nodes. A node that lies on a path between two other nodes has a degree of $k \geq 2$. Therefore, every node in an fPIN that one obtains via SP, AP, or APSP sampling must have $k \geq 2$ or be a seed node. Accordingly, the probability $p(k = 1)$ for finding a node with degree 1 in these fPINs is very low. When a sampling method includes SB sampling, a node $v \in V_S \setminus V_{S_0}$ can have $k = 1$. Accordingly, $p(k = 1)$ is larger for sampling methods that include SB sampling (i.e., SB, SBSP, ASB, and ASBSP sampling) than for sampling methods that do not include SB sampling (i.e., SP, AP, and APSP sampling). Nodes with a very large degree are more frequent in fPINs that we obtain via SB sampling (and combinations of SB sampling with other sampling methods) than in fPINs that we obtain via path-based sampling methods, because seed nodes have the same degree in the HIPPIE interactome and in an SB-sampled fPIN. A node with large degree in the HIPPIE interactome thus has an equally large degree in an SB-sampled fPIN. By contrast, to have a large degree in an fPIN from path-based sampling methods, a node must have a large degree in the HIPPIE interactome and many of its edges must be part of a length-3 path or a shortest path to other seed nodes.

In Fig. 3.5, we show the mean of the mean shortest-path length L, the natural connectivity K, the mean local clustering coefficient C, and the global clustering coefficient T over sets of fPINs for different sampling methods. Whiskers show the standard deviation of L, K, C, and T over a set of fPINs. For comparison, we also show mean values of L, K, C, and T for associated[9] sets of ER networks and associated sets of CM networks. We observe that L and K vary strongly within each set of fPINs. By contrast, the variances of the local mean clustering coefficient and the global clustering coefficient within a set of fPINs are very small.

When comparing the mean properties of sets of fPINs to those of ER and CM networks, we observe that CM networks capture the mean L and K of fPINs much better than ER networks, but ER networks capture the mean C and T of fPINs much better than CM networks. This latter observation is noteworthy, because

[9]For set of fPINs with $|S|$ networks, we generate associated sets of random networks that consist of $100|S|$ networks. For each network in a set of fPINs, an associated set of random networks has 100 networks of the same network size and the same number of edges.

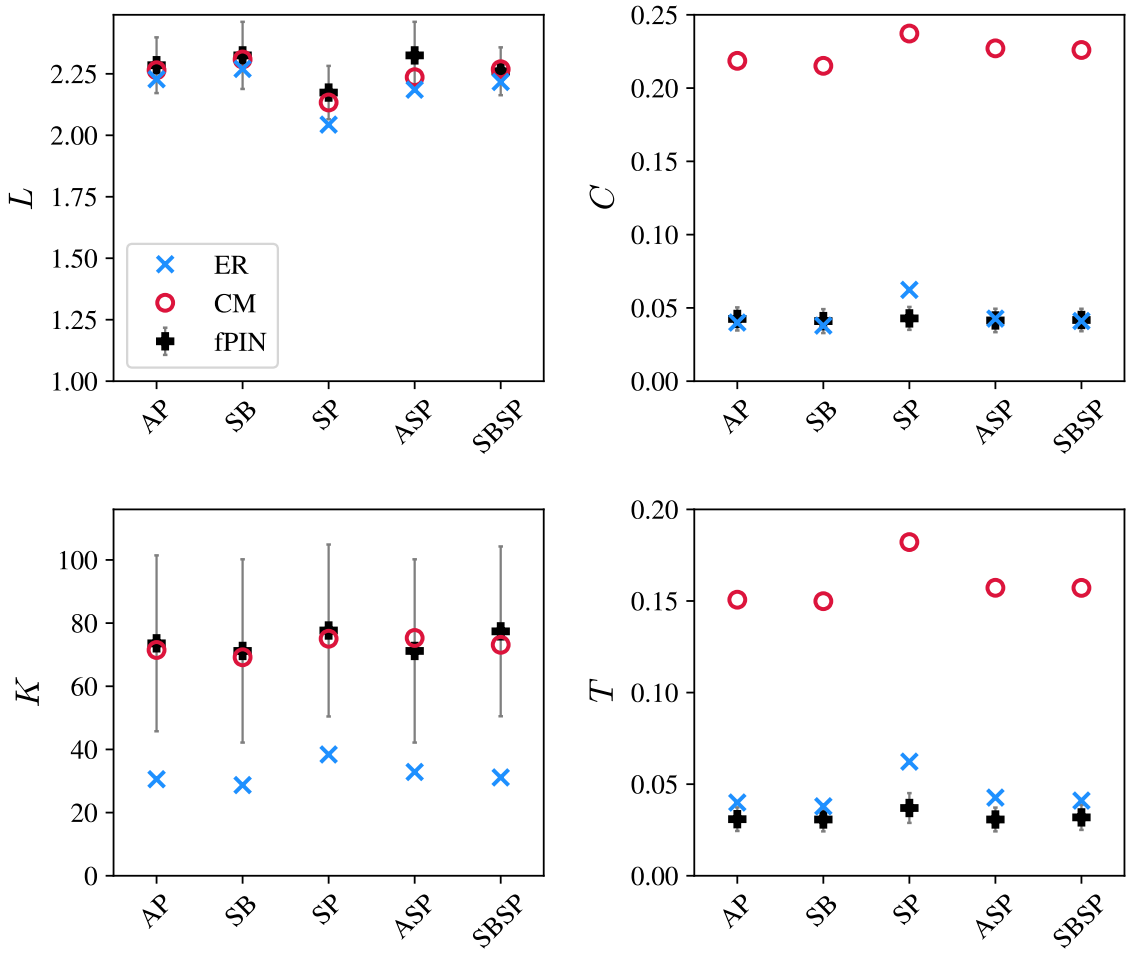


Figure 3.5: Comparison of structural properties of fPINs for different sampling methods. We show the mean shortest-path length L, the natural connectivity K, the mean local clustering coefficient C, and the global clustering coefficient T over the size N of SP-sampled fPINs. Black crosses indicate the mean of L, K, C, and T over fPINs with sizes in $\{N, \ldots, N + 99\}$. For comparison, we show means of L, K, C, and T for an associated set of ER networks (blue crosses) and an associated set of CM networks (magenta circles).

one tends to expect the opposite [398, p. 369]. For a given fPIN, an associated CM network has the same size, number of edges, and degree sequence — and thus has the same degree distribution — as the fPIN. An associated ER network has the same size and the same number of edges as the fPIN, but in general it has a different degree sequence and degree distribution. The degree distribution of a network is a very popular tool for characterizing networks and can be indicative of many other network properties [515]. In general, one thus expects CM networks to capture most structural properties of fPINs better than ER networks.

We now compare the mean properties of fPINs for different sampling methods.

Overall, we find that the sampling method has a very small effect on the mean values of K, C, and T. The largest differences in L, C, or T occur when comparing SP sampling to one of the other sampling methods. The mean values of C and T are larger, and the mean value of L is smaller, for SP-sampled fPINs than for sets of fPINs from other sampling methods. It is likely that these structural differences between the SP-sampled fPINs and the other sets of fPINs arise because of the difference in mean network size. Networks in the set of SP-sampled fPINs have, on average, 820 nodes. In all other sets of fPINs, the mean network size is larger than 1000 nodes (see Table 3.3). A comparison of the mean structural properties of the associated sets of ER networks supports this hypothesis. The set of ER networks for the SP-sampled fPINs have larger mean C and T and smaller mean L than the ER networks for the other considered sampling methods. These relations between the mean structural properties of the sets of ER networks suggests that the structural differences that we observe in Fig. 3.5 between SP-sampled fPINs and other fPINs are related to the differences in network size or number of edges of fPINs.

3.3.1.3 Network size and network structure for fPINs and random networks

In this subsection, we investigate the effect of network size on several structural properties of fPINs. As an example, we use the SP-sampled fPINs. We obtain similar results for the other sampling methods that we considered (see Appendix A.2).

In Fig. 3.6, we show the mean shortest-path length L, the natural connectivity K, the mean local clustering coefficient C, and the global clustering coefficient T for different sizes N of SP-sampled fPINs. For comparison, we show means of L, K, C, and T for associated sets of ER networks (oranges crosses) and CM networks (magenta circles). For all network sizes of SP-sampled fPINs, the CM networks capture the mean values of L and K of the fPINs better than ER networks, but ER networks capture the mean values of C and T of the fPINs better than CM networks. This observation is consistent with our results in Fig. 3.5. For fPINs and their associated ER and CM networks, natural connectivity tends to increase with network size. The clustering coefficients C and T of fPINs and the associated ER and CM networks tend to decrease with increasing N. Interestingly, the clustering coefficients for ER and CM networks tend to decrease more sharply than for fPINs. We show the slopes of linear least-squares regressions of C and T as functions of N in Table 3.4.

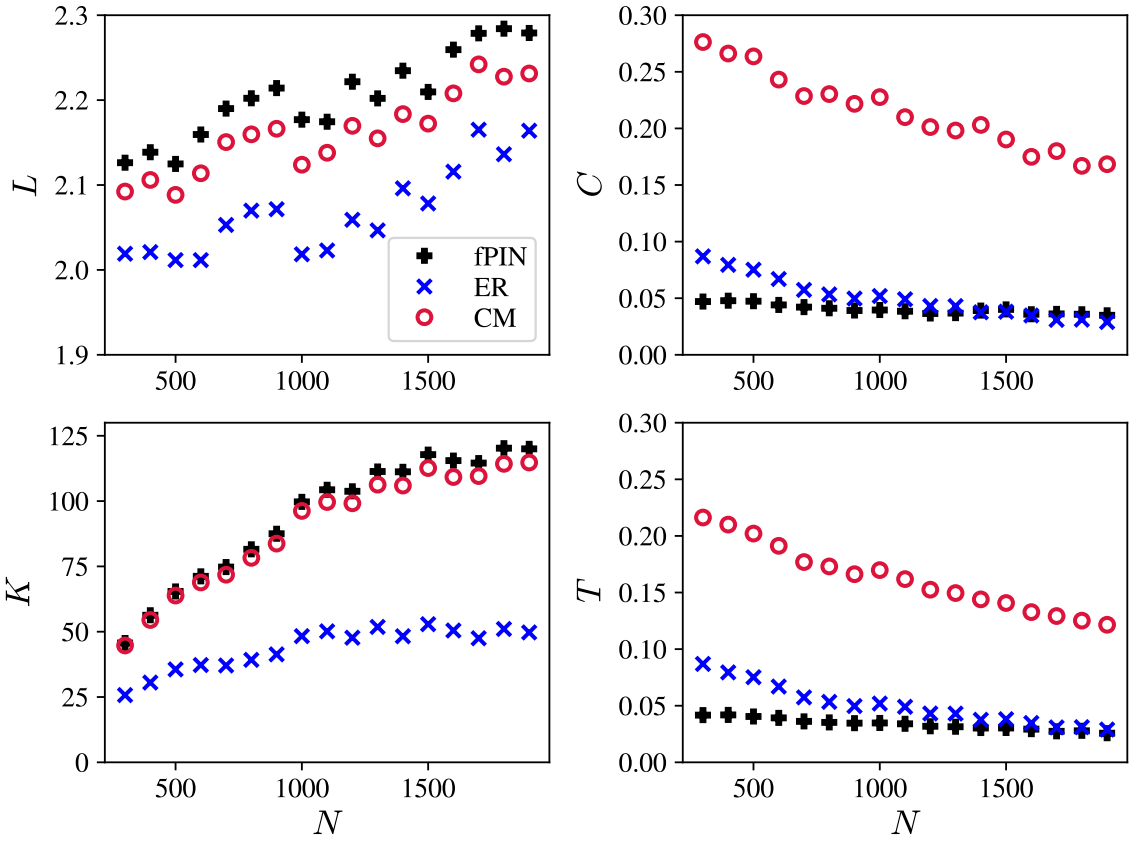


Figure 3.6: Structural properties of fPINs. We show the mean shortest-path length L, the natural connectivity K, the mean local clustering coefficient C, and the global clustering coefficient T over the size N of SP-sampled fPINs. Black crosses indicate the mean of L, K, C, and T over fPINs with sizes in $\{N, \ldots, N+99\}$. For comparison, we show the mean values of L, K, C, and T for associated sets of ER networks and CM networks.

Quantity	data set	slope $[10^{-4}]$
C	fPINs	-8 (1)
	ER	−33 (3)
	CM	−66 (4)
T	fPINs	−10 (1)
	ER	−34 (2)
	CM	−58 (3)

Table 3.4: Slopes of a linear least-squares regression of the mean local clustering coefficient C and the global clustering coefficient T for fPINs, ER networks, and CM networks. The numbers in parentheses are the associated standard deviations of the linear regression.

In Fig. 3.3, we observed that $c(k)$ tends to be much smaller in SP-sampled fPINs

than in the HIPPIE interactome for all k. That observation is consistent with our observation of a small C for fPINs in Fig. 3.6. Our comparison of fPINs with associated sets of CM networks indicates that the difference between C and T of the HIPPIE and C and T of fPINs is not just a consequence of an fPIN's degree distribution. In Section 3.4, we discuss possible explanations for the small values of C and T that we observe for all considered sampling methods and network sizes.

3.3.2 Robustness of structural properties under node removal

In this subsection, we compare robustness profiles for the relative LCC size F, mean shortest-path length L, efficiency $\mathcal{E}$, mean local clustering coefficient C, global clustering coefficient T, algebraic connectivity a, resistance distance Ω, natural connectivity K, mean communicability angle θ, and mean communicability distance ξ. First, we examine the robustness profiles of impacts on SP-sampled fPINs. Subsequently, we compare robustness profiles for different sampling methods.

In Figs. 3.7 and 3.8, we show impacts $I_{X,t,n,s=\mathrm{SP},w}$ for ten performance measures X on a small SP-sampled fPIN (with $N = 354$) and a large SP-sampled fPIN (with $N = 1216$) and the mean impacts $I_{X,t,n,s=\mathrm{SP}}$ over the set of SP-sampled fPINs. We first consider the robustness profiles for the large SP-sampled fPIN. For eight of ten structural properties, we observe that the corresponding impacts tend to increase with the number n of removed nodes. The impacts $I_{\theta,w}$ are very small and appear to be uncorrelated with n. We measure Spearman rank-correlation coefficients of 0.05, 0.11, 0.06, and 0.16 for targeting nodes by degree, BC, EC, and SC, respectively. The impacts $I_{\xi,n,w}$ increase sharply with increasing $n \leq 3$, but they are roughly constant for $n > 3$.

When comparing the robustness profiles for the large SP-sampled fPIN and the small SP-sampled fPIN, we observe that most of the robustness profiles for the small network are similar to the corresponding robustness profiles for the large network. The impacts for a given perturbation size n tend to be larger for the small SP-sampled fPIN than for the large SP-sampled fPIN. Impacts $I_{F,s=\mathrm{SP},w}$ are much larger for the small SP-sampled fPIN than for the large SP-sampled fPIN. For the small SP-sampled fPIN, we observe that $I_{F,n=100,w} \geq 0.95$. These values of $I_{F,n=100}$ indicate that after removing 100 nodes using one of the four considered targeting strategies, the LCC of the small SP-sampled fPIN includes at most 5% of its remaining nodes (i.e, at most 12 nodes). For the considered targeting strategies, we observe that the impacts $I_{L,s=\mathrm{SP},w}$, $I_{E,s=\mathrm{SP},w}$, and $I_{\Omega,s=\mathrm{SP},w}$ on the small SP-sampled PIN tend to increase with n when $I_{F,s=\mathrm{SP},w} \lessapprox 0.65$ (i.e., when the perturbed fPIN has $N \gtrapprox 100$)

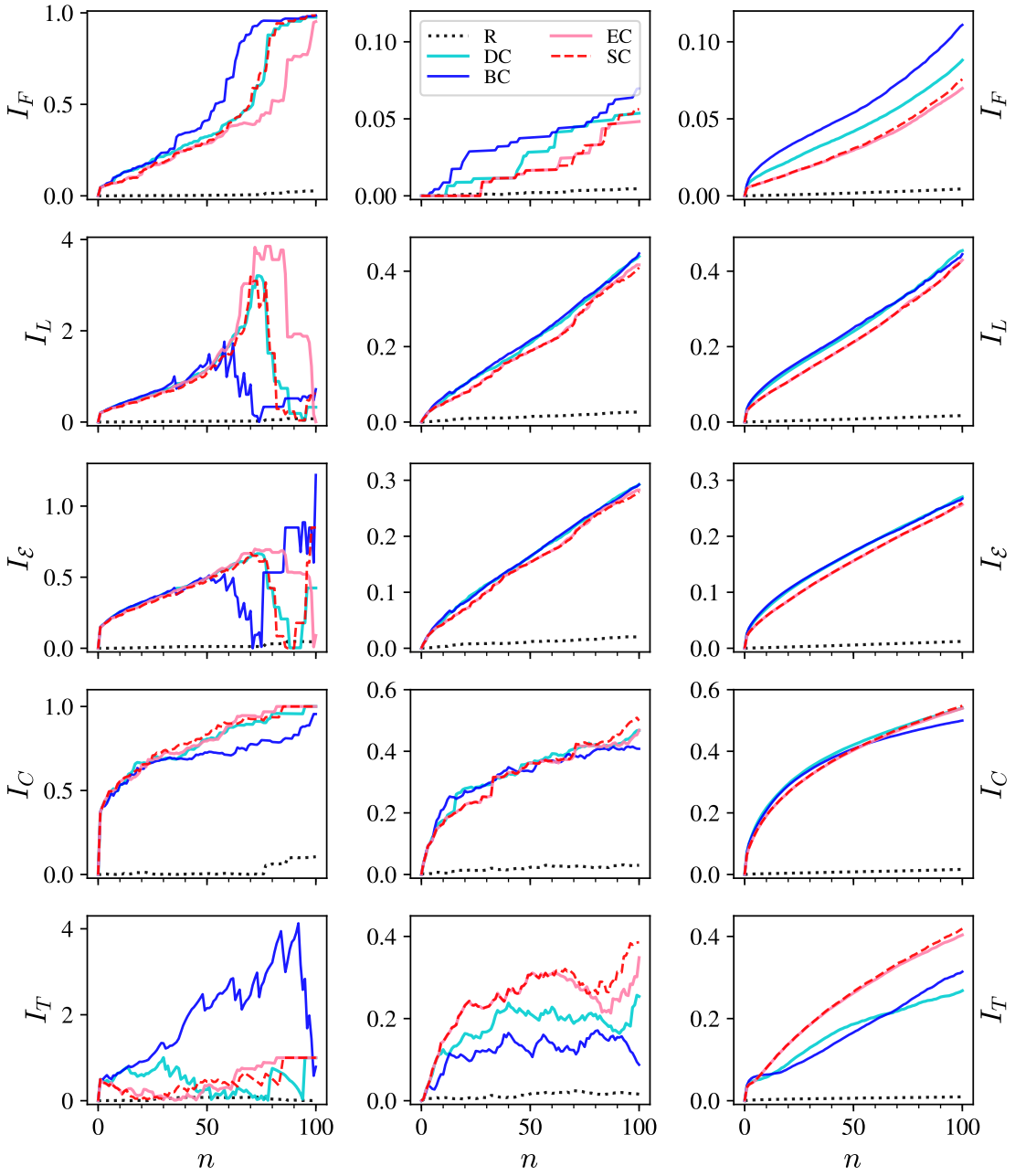


Figure 3.7: Robustness profiles for performance measures on SP-sampled fPINs. We show robustness profiles for relative LCC size F, mean shortest-path length L, efficiency $\mathcal{E}$, local clustering coefficient C, and global clustering coefficient T. In the left and center panels, we show robustness profiles for performance measures on an SP-sampled fPIN with $N = 354$ and an SP-sampled fPIN with $N = 1216$, respectively. In the right panels, we show mean robustness profiles for the set of SP-sampled fPINs. In each panel, different curves correspond to targeting nodes uniformly at random (R), by degree (DC), by betweenness centrality (BC), by eigenvector centrality (EC), and by subgraph centrality (SC).

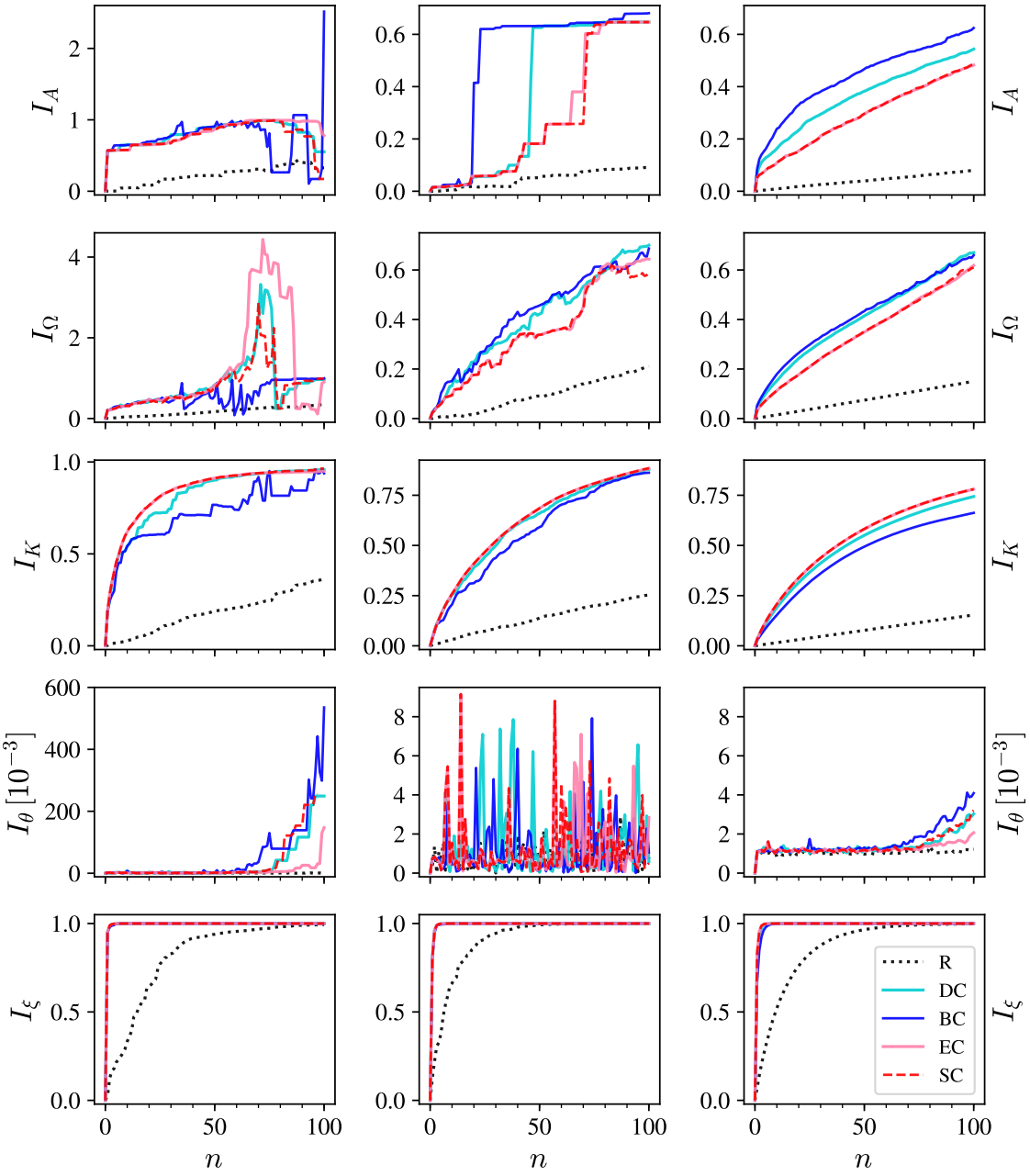


Figure 3.8: Robustness profiles for performance measures on SP-sampled fPINs. We show robustness profiles for algebraic connectivity a, resistance distance Ω, natural connectivity K, mean communicability angle θ, and mean communicability distance ξ. In the left and center panels, we show robustness profiles for performance measures on an SP-sampled fPIN with $N = 354$ and an SP-sampled fPIN with $N = 1216$, respectively. In the right panels, we show mean robustness profiles for the set of SP-sampled fPINs. In each panel, different curves correspond to targeting nodes uniformly at random (R), by degree (DC), by betweenness centrality (BC), by eigenvector centrality (EC), and by subgraph centrality (SC).

and tend to decrease with n when $I_{F,s=\mathrm{SP},w} \gtrapprox 0.65$ (i.e., when the perturbed fPIN has $N \lessapprox 100$).

The mean robustness profiles for the set of SP-sampled fPINs are similar to those for the large SP-sampled fPIN. Impacts tend to increase with perturbation size. Impacts $I_{\theta,s=\mathrm{SP}}$ tend to be very small. For ξ, we observe $I_{\xi,s=\mathrm{SP}} \approx 1$ for $n \geq 4$. Overall, the impact of targeted node removal is much larger than the impact of removing nodes uniformly at random. However, we observe only small differences between the impacts of removing nodes by degree, BC, EC, or SC. For the performance measures F, and a, the impacts tend to be larger when $t = \mathrm{BC}$ than when $t \in \{\mathrm{DC}, \mathrm{EC}, \mathrm{SC}\}$. For T and K, the impacts tend to be larger when targeting by EC or SC than when targeting by degree or BC.

Our observations on the robustness of structural properties of fPINs under targeted node removal are similar to the results of Ventresca and Aleman (2014). They examined the change of several structural properties of various networks under node removal using various targeting strategies, and reported that the change of a structural property tends to be similar for different targeting strategies [552]. For the size of the LCC, they reported that targeting nodes by BC leads to slightly larger impacts than targeting by other centrality measures. In agreement with their results, we find that the impacts on relative LCC size, which is proportional to the size of the LCC, are slightly larger when removing nodes by BC than by degree, EC, or SC. In Fig. 3.8, we observe that targeting by BC also leads to larger impacts on algebraic connectivity than targeting by degree, EC, or SC. However, for other structural properties (such as global clustering coefficient or natural connectivity), other centrality measures lead to larger impacts than BC.

We observe very similar robustness profiles for different sampling methods (see Figs. 3.7 and 3.8 and Appendix B). We find the greatest differences between robustness profiles for different sampling methods when comparing AP sampling to any of the other considered sampling methods. To investigate these differences, we compare the robustness profiles for different sampling methods for a given targeting strategy. We choose $t = \mathrm{BC}$ as an example. For each performance measure X and perturbation size n, consider the sets $\mathcal{I}_{X,t,n} := \{I_{X,t,n,s}\}_{s \neq \mathrm{AP}}$ of impacts for all considered sampling methods except AP sampling. In Fig. 3.9, we show the robustness profiles for targeting nodes by degree. Blue curves indicate $I_{X,t=\mathrm{BC},s=\mathrm{AP}}$ for different performance measures X. Black curves indicate the mean of $\mathcal{I}_{X,t=\mathrm{BC},n}$. We signify the minimum and maximum of $\mathcal{I}_{X,n}$ by the edges of an orange shaded region.

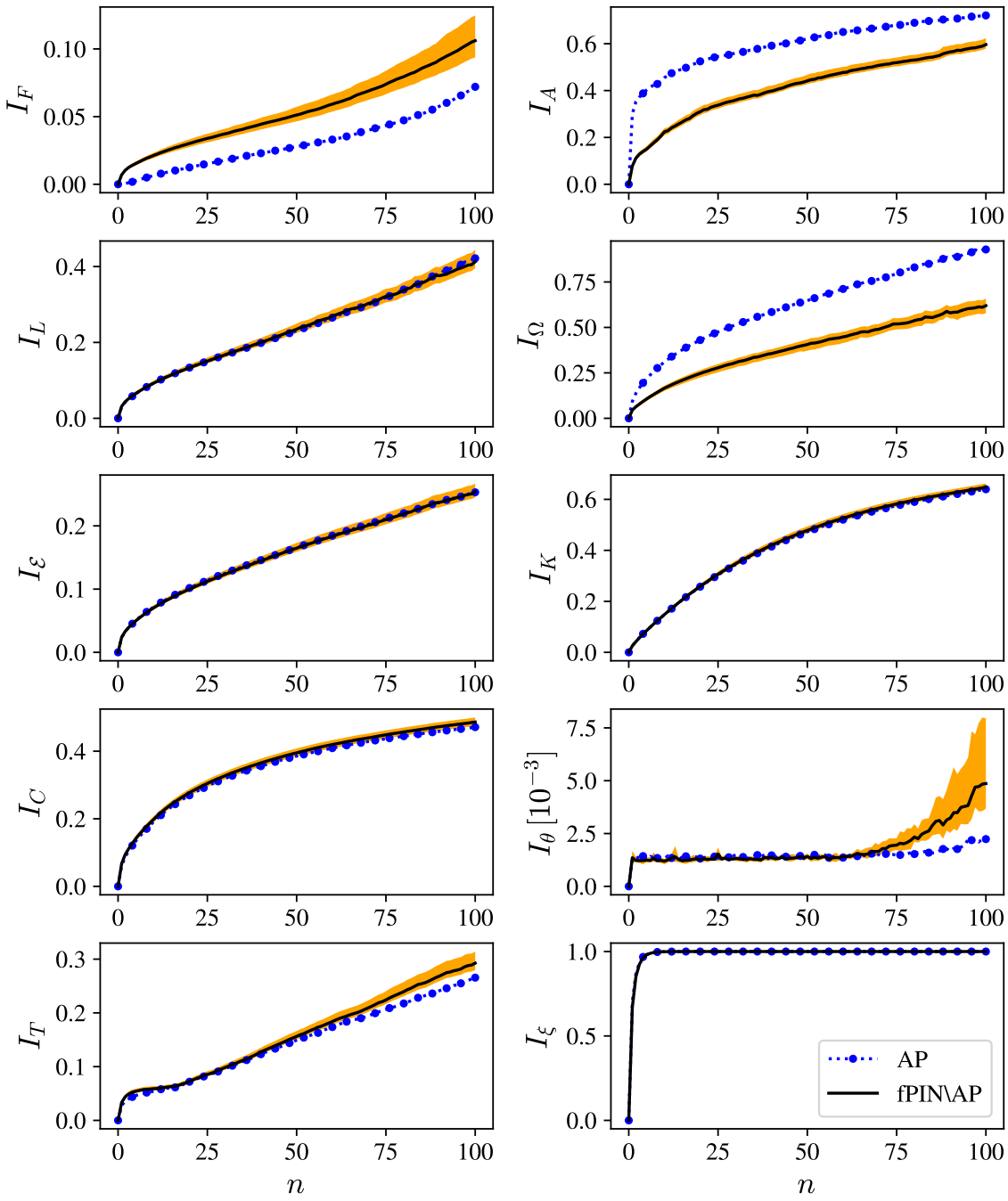


Figure 3.9: Mean robustness profiles for fPINs from several sampling methods. We show impacts $I_{X,t=\mathrm{BC},s=\mathrm{AP}}$ (blue curves; "AP") and the mean $\mathcal{I}_{X,t=\mathrm{BC},n}$ (black curves; "fPIN\AP"). The boundaries of the orange-shaded region indicate the minimum and maximum of $\mathcal{I}_{X,t=\mathrm{BC},n}$.

For all considered performance measures, the narrow orange-shaded regions indicate that the mean robustness profiles are very similar for different sampling methods, aside from AP sampling. For AP sampling, the impacts $I_{F,t=\mathrm{BC},s=\mathrm{AP}}$ are smaller than $I_{F,t=\mathrm{BC}}$ for other sampling methods. The impacts $I_{a,t=\mathrm{BC},s=\mathrm{AP}}$ and $I_{\Omega,t=\mathrm{BC},s=\mathrm{AP}}$ are larger than $I_{a,t=\mathrm{BC}}$ and $I_{\Omega,t=\mathrm{BC}}$ for other sampling methods.

3.3.3 Plausibility scores of performance measures

The observation of similar mean robustness profiles for all considered sampling methods except for AP sampling suggests that criterion scores and plausiblity scores should be similar for these sampling methods. In Fig. 3.10, we show the mean of the mean criterion scores $\langle \mathrm{SR}_X \rangle_{n,s\neq\mathrm{AP}}$, $\langle \mathrm{SS}_X \rangle_{n,s\neq\mathrm{AP}}$, and $\langle \mathrm{SM}_X \rangle_{n,s\neq\mathrm{AP}}$ and the mean of the plausibility scores $\langle \mathrm{PS}_X^{\mathrm{min}} \rangle_{n,s\neq\mathrm{AP}}$ and $\langle \mathrm{PS}_X^{\mathrm{prod}} \rangle_{n,s\neq\mathrm{AP}}$ (over all considered perturbation sizes and the fPIN sets for the four sampling methods SB) SP, APSP and SBSP. The top and bottom ends of vertical error bars indicate the maximum and the minimum of a criterion score or plausibility score over the six sampling methods.

When comparing the three criterion scores, structural sensitivity has the greatest variation for different performance measures. The scores for selective robustness and perturbation-size monotonicity tend to be high for most of performance measures. Exceptions are the mean communicability angle θ and the mean communicability distance ξ, which have low scores for selective robustness. Additionally, we obtain small values of $\langle \mathrm{SM}_\theta \rangle_{n,s\neq\mathrm{AP}}$ and very small values of $\langle \mathrm{SS}_\xi \rangle_{n,s\neq\mathrm{AP}}$. These observations are consistent with our observations on the robustness curves for θ and ξ in Section 3.3.2. We found there that, for two different SP-sampled fPINs, the impacts $I_\xi = 1$ for all $n \geq 4$ when removing nodes uniformly at random or using one of four considered targeting strategies. The similarity between these robustness profiles suggests that this performance measure is not sensitive to the examined changes in network structure. Impacts I_θ tend to have very small values and have a weak or no correlation with n. The small values of $\langle \mathrm{SM}_\theta \rangle_{n,s\neq\mathrm{AP}}$ reflects this observation.

In Fig. 3.10, we show the mean selective robustness $\langle \mathrm{SR}_X \rangle_{n,s\neq\mathrm{AP}}$, the mean structural sensitivity $\langle \mathrm{SS}_X \rangle_{n,s\neq\mathrm{AP}}$, the mean perturbation-size monotonicity $\langle \mathrm{SM}_X \rangle_{n,s\neq\mathrm{AP}}$, and the mean plausibility scores $\langle \mathrm{PS}_X^{\mathrm{min}} \rangle_{n,s\neq\mathrm{AP}}$ and $\langle \mathrm{PS}_X^{\mathrm{prod}} \rangle_{n,s\neq\mathrm{AP}}$ for several performance measures X. The plausibility scores $\mathrm{PS}_{s\neq\mathrm{AP}}^{\mathrm{prod}}$ and $\mathrm{PS}_{s\neq\mathrm{AP}}^{\mathrm{min}}$ lead to similar conclusions on the plausibility of performance measures. When using either $\mathrm{PS}_{s\neq\mathrm{AP}}^{\mathrm{prod}}$ or $\mathrm{PS}_{s\neq\mathrm{AP}}^{\mathrm{min}}$ to indicate the plausibility of a performance measure, the structural properties F and a are the most plausible, and the structural properties ξ and θ are the least plausible performance measures. Among the examined structural properties,

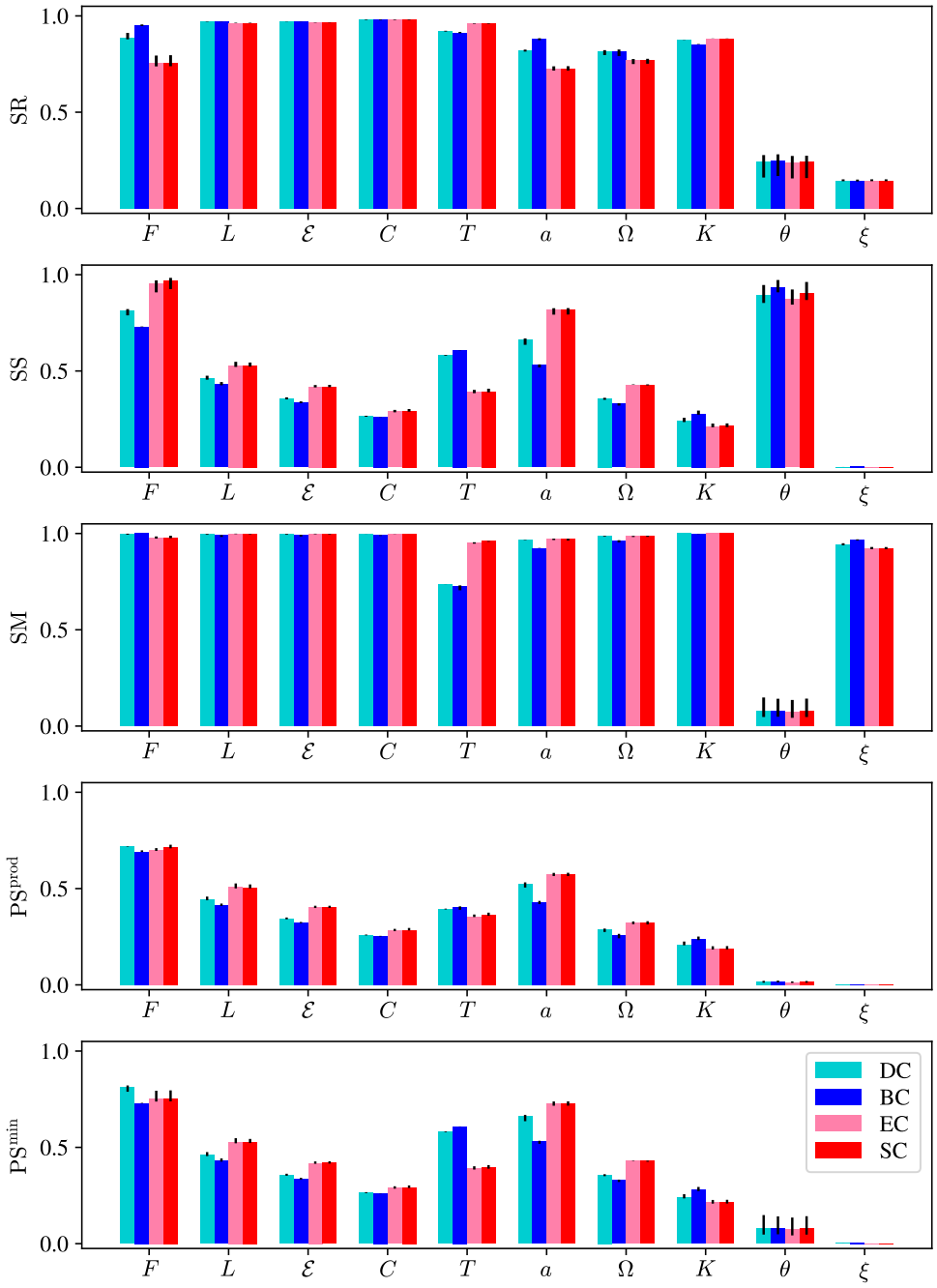


Figure 3.10: Criterion scores and plausibility scores for performance measures. We show the mean selective robustness $\langle \mathrm{SR}_X \rangle_n$, the mean structural sensitivity $\langle \mathrm{SS}_X \rangle_n$, the mean perturbation-size monotonicity $\langle \mathrm{SM}_X \rangle_n$, and the mean plausibility scores $\langle \mathrm{PS}_X^{\mathrm{min}} \rangle_n$ and $\langle \mathrm{PS}_X^{\mathrm{prod}} \rangle_n$ for several performance measures X when targeting nodes by largest degree (DC), largest betweenness centrality (BC), largest eigenvector centrality (EC), and largest subgraph centrality (SC). Colored bars indicate the mean value of a score over the sampling methods excluding AP. The upper and lower ends of the thin black bars indicate the largest and smallest values of a score over the sampling methods excluding AP.

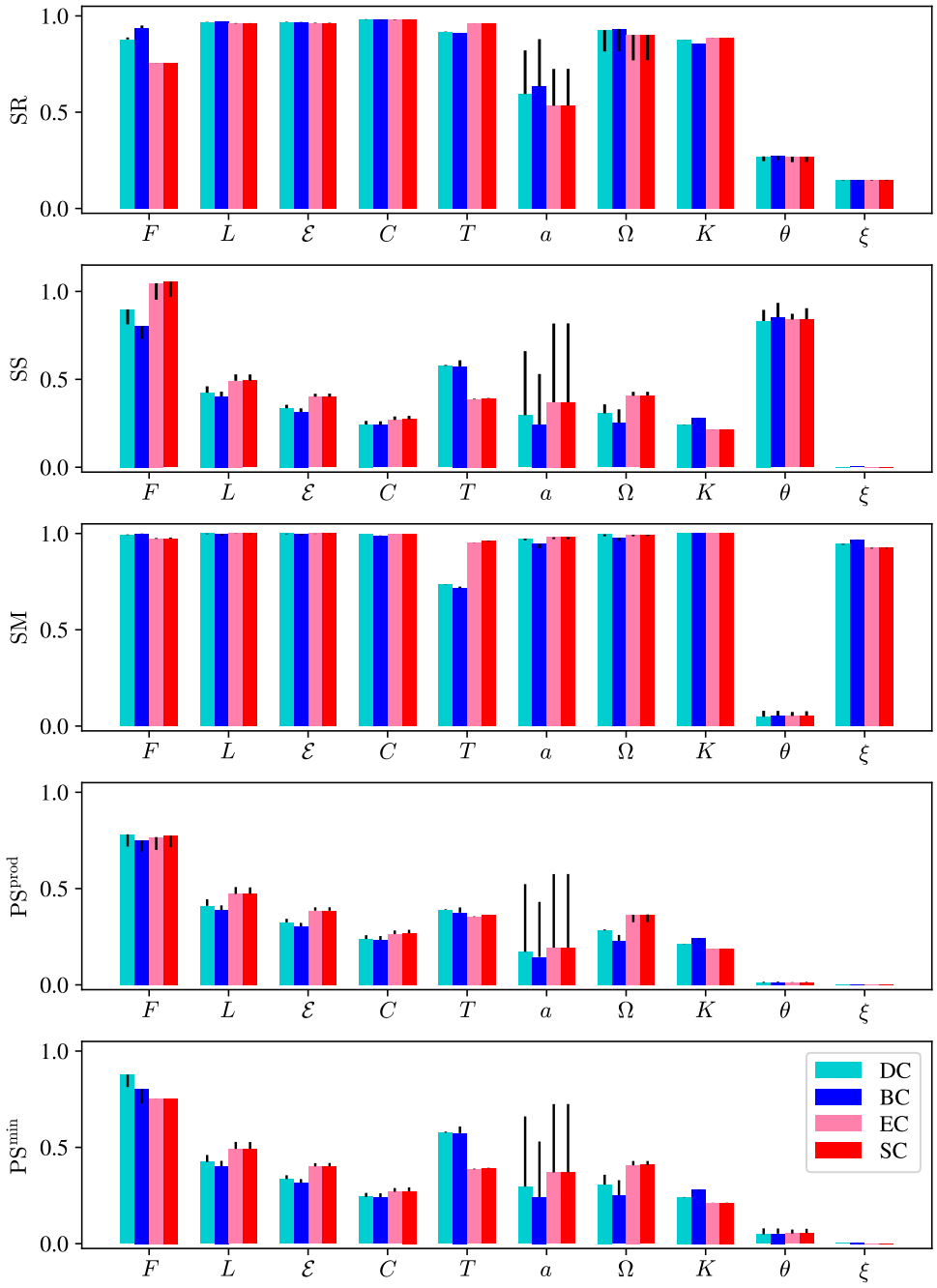


Figure 3.11: Criterion scores and plausibility scores for performance measures for AP sampling. The bars indicate mean selective robustness $\langle \mathrm{SR}_{X,s=\mathrm{AP}} \rangle_n$, mean structural sensitivity $\langle \mathrm{SS}_{X,s=\mathrm{AP}} \rangle_n$, mean perturbation-size monotonicity $\langle \mathrm{SM}_{X,s=\mathrm{AP}} \rangle_n$, and the mean plausibility scores $\langle \mathrm{PS}^{\mathrm{min}}_{X,s=\mathrm{AP}} \rangle_n$ and $\langle \mathrm{PS}^{\mathrm{prod}}_{X,s=\mathrm{AP}} \rangle_n$ for several performance measures X when targeting nodes by largest degree (DC), largest betweenness centrality (BC), largest eigenvector centrality (EC), and largest subgraph centrality (SC). The length of the thin black bars indicate the difference between the score for AP and the mean of scores for the other sampling methods.

the third and fourth least plausible performance measures are K and C. These observations are independent of the targeting strategy for removing nodes. The targeting strategy affects the plausibility of performance measures, but it does so less than than the choice of X itself. For all considered targeting strategies, the performance measures C, K, θ, and ξ have lower $\mathrm{PS}^{\mathrm{prod}}_{s\neq\mathrm{AP}}$ and $\mathrm{PS}^{\mathrm{min}}_{s\neq\mathrm{AP}}$ than the other performance measures. In particular, the plausibility scores for the local clustering coefficient are considerably lower than those for the global clustering coefficient.

In Fig. 3.11, we show mean criterion scores and the plausibility scores for the set of AP-sampled fPINs. For each score, black error bars indicate the difference between the of AP-sampled fPINs and the mean for the other six sampling methods (see Fig. 3.10).

Most criterion scores and plausibility scores are similar for all considered sampling methods. We observe the greatest differences between the criterion scores of AP and other sampling methods for the performance measure a. For the set of AP-sampled fPINs, criterion scores for selective robustness and structural sensitivity are lower than for the other sets of fPINs. Consequently, the mean plausibility scores $\langle \mathrm{PS}^{\mathrm{min}}_{a,s=\mathrm{AP}} \rangle_n$ and $\langle \mathrm{PS}^{\mathrm{prod}}_{a,s=\mathrm{AP}} \rangle_n$ are lower than $\langle \mathrm{PS}^{\mathrm{min}}_{a,s} \rangle_n$ and $\langle \mathrm{PS}^{\mathrm{prod}}_{a,s} \rangle_n$ for other sampling methods. Depending on the choice of the targeting strategy, the second most plausible performance measure is either the mean shortest-path length L or the global clustering coefficient T for AP-sampled fPINs.

From Figs. 3.10 and 3.11, we conclude that the choice of sampling method for fPINs has only a small effect on the outcome of the study of plausibility of performance measures. Across almost all considered sampling methods, targeting strategies, groups of network sizes, and plausibility scores, the structural properties F and a appear to be the most plausible performance measures, and the structural properties ξ and θ appear to be least plausible performance measures.

The results of a plausibility test depend on the sets of networks for which one computes robustness profiles. A performance measure with high plausibility on a set of networks (such as a set of fPINs) may have a low plausibility on a different set of networks. To give an illustrative example, we show in Fig. 3.12 the criterion scores and plausibility scores for performance measures on the set of SP-sampled fPINs and associated sets of ER networks and DA networks. We see that the criterion scores and plausibility scores can be very different for different sets of networks. For example, the criterion scores for selective robustness of F and Ω are high for SP-sampled fPINs and DA networks, but they are almost 0 for ER networks. The scores for selective robustness of L, $\mathcal{E}$, and a are also much lower for ER networks than for fPINs or DA networks. These and further differences between the criterion scores

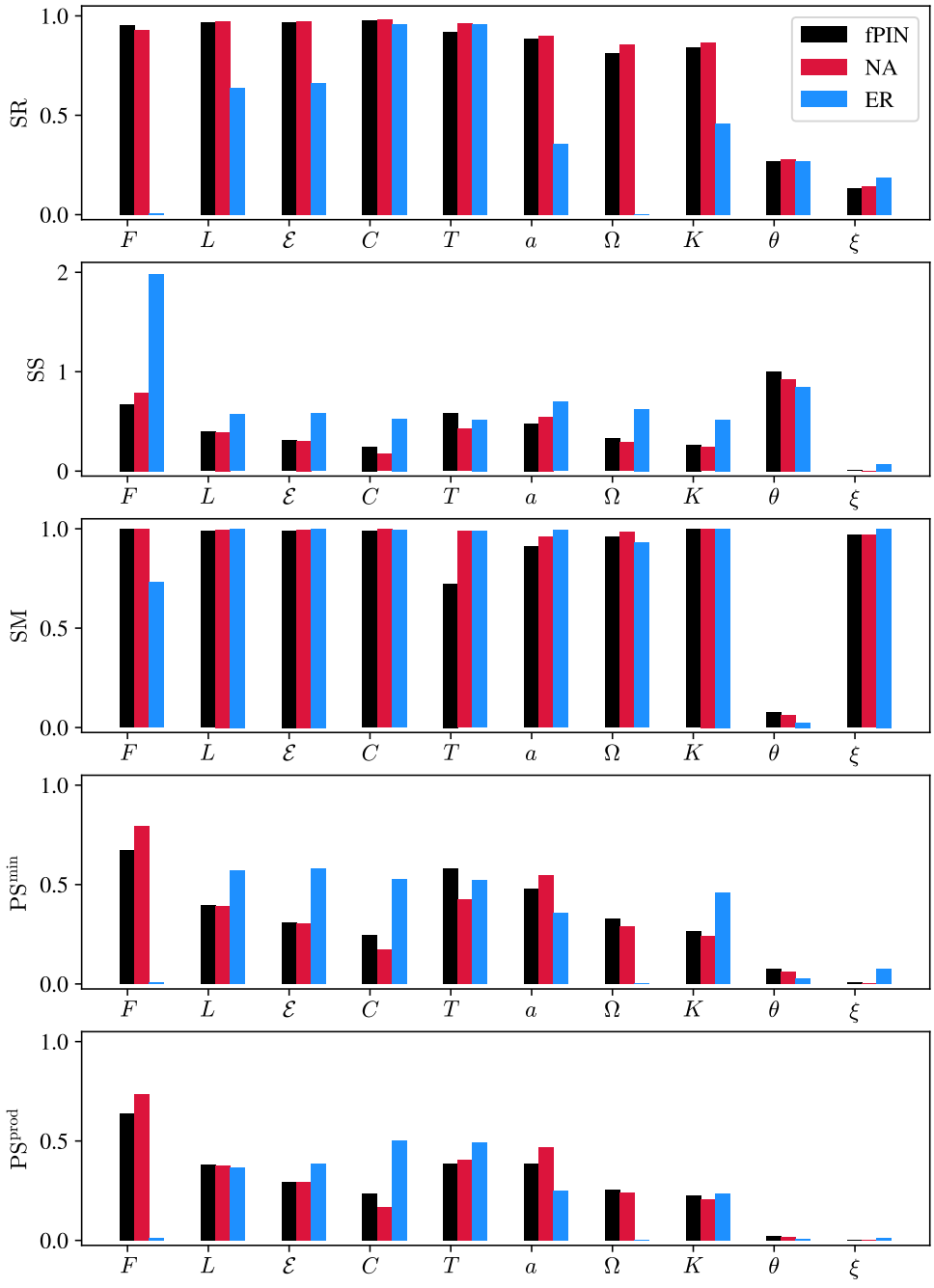


Figure 3.12: Criterion scores and plausibility scores for performance measures for SP-sampled fPINs ("fPIN") and associated sets of ER networks ("ER") and DA networks ("DA"). The bars indicate mean selective robustness $\langle \mathrm{SR}_{X,t=\mathrm{BC},s=\mathrm{SP}} \rangle_n$, mean structural sensitivity $\langle \mathrm{SS}_{X,t=\mathrm{BC},=\mathrm{SP}s} \rangle_n$, mean perturbation-size monotonicity $\langle \mathrm{SM}_{X,t=\mathrm{BC},s=\mathrm{SP}} \rangle_n$, and the mean plausibility scores $\langle \mathrm{PS}^{\mathrm{min}}_{X,t=\mathrm{BC},s=\mathrm{SP}} \rangle_n$ and $\langle \mathrm{PS}^{\mathrm{prod}}_{X,t=\mathrm{BC},s=\mathrm{SP}} \rangle_n$ for several performance measures X.

for fPINs, DA networks, and ER networks lead to differences in plausibility scores and consequently in the ranking of performance measures by plausibility. For SP-sampled fPINs, each plausibility score is highest for F, a, and T. For ER networks, we obtain low plausibility scores for F and a. The performance measures with the three largest values of $PS^{\min}_{X,t=\mathrm{BC}}$ for ER networks are $\mathcal{E}$, L, and T. The plausibility scores $PS^{\mathrm{prod}}_{X,t=\mathrm{BC}}$ for ER networks are largest for C, T, and $\mathcal{E}$.

This greater similarity between plausibility scores for fPINs and DA than for fPINs and ER indicates that, in general, the outcome of a study of plausibility of performance measures depends on the networks that one considers. For two sets of networks with very different structural properties (e.g., fPINs and associated ER networks), the plausibility scores for a given performance measure can differ considerably. For two sets of networks that have similar structural properties (e.g., SP-sampled fPINs and associated DA networks), we observe similar plausibility scores for a given performance measure.

3.4 Conclusions and discussion

We summarize our findings from this chapter and discuss interpretations of our results. We close with an outlook on possible extensions of our study of function-specific PINs.

3.4.1 Summary

When using network analysis to study the robustness of cellular functions, it is important to represent cellular functions by networks that include both active and alternative biological pathways. We represent cellular functions by function-specific protein-interaction networks (fPINs), which are network samples of a protein interactome. They include pathway-associated proteins and proteins that are "close to" pathway-associated proteins or "between" pairs of pathway-associated proteins. We examined several structural properties of fPINs that we obtained using different sampling methods. We found that fPINs from different sampling methods tend to have very similar structural properties. For all considered sampling methods, the mean local clustering coefficient C and the global clustering coefficient T have very small values.

Researchers have proposed various structural properties of networks as measures of performance for network systems [22, 49, 158, 231, 259, 260, 355, 552, 570, 588,

589]. We examined the robustness of ten such properties for several variants of targeted node removal. We found that the robustness profiles for different structural properties tend to be different. Among the performance measures, targeting strategies, and sampling methods that we considered in our study, we found that, for a given performance measure, robustness profiles for different targeting strategies or different sampling methods tend to be similar.

We formalized three criteria for the plausibility of a measure of performance (i.e., viability) for fPINs. Using these criteria to assess the plausibility of ten performance measures, we found that relative LCC size F and algebraic connectivity a are the most plausible performance measures for most sampling methods and mean local clustering coefficient C, natural connectivity K, mean communicability angle θ, and mean communicability distance ξ are the least plausible performance measures for all considered sampling methods.

3.4.2 Low clustering coefficients

The deterministic seed-based sampling methods that we used to sample fPINs lead to different network properties than one would expect when constructing subnetworks from nodes that one samples uniformly at random. Stumpf et al. (2005) demonstrated that uniform sampling of networks with power-law degree distributions leads to subnetworks that have higher probabilities $p(k)$ for nodes with degree $k = 1$ and lower probabilities for nodes with degree $k \geq 2$ [516]. By contrast, we found that fPINs tend to have lower $p(k)$ than the HIPPIE interactome for $k \lessapprox 10$. Smith and Moody (2013) reported that the values of C and T of a uniform network sample are similar to those of C and T of its parent network [496]. We found that fPINs have much lower values of C and T than the HIPPIE interactome.

We anticipate that there are several mechanisms by which the considered sampling methods lead to small values of C and T in fPINs. One potential mechanism is that SP sampling leads to network samples that tend to include many non-overlapping) paths between pairs of seed nodes (see Fig. 3.1 (**D**)). Consider a shortest path between the seed nodes v_{s_1} and v_{s_2}. Assume that this path is not overlapping with any other shortest path between two seed nodes $v_{s_i}, v_{s_j} \in V_{S_0}$. In an SP-sampled fPIN, a node v_i on this path between v_{s_1} and v_{s_2} has a degree of $k = 2$. Such a node has a clustering coefficient of $C_i = 0$, because if its neighbors were adjacent to each other, the node v_i would not be on a shortest path between v_{s_1} and v_{s_2}. Shortest-path sampling can thus lead to network samples that have many nodes with degree 2 and a clustering coefficient of 0. These nodes contribute

negatively to the mean local clustering coefficient of a network sample. Because we observe similarly small values of C and T for all considered sampling methods, we expect that other mechanisms contribute to small mean local clustering coefficients and global clustering coefficients in fPINs. Uncovering these mechanisms requires further study of these sampling methods.

3.4.3 Similar observations for different sampling methods

We found that several structural properties and robustness profiles of structural properties are similar for fPINs from different sampling methods. Consequently, the plausibility scores of performance measures are similar for the sampling methods that we considered in our study. The similarity of results for different sampling methods suggests that the considered sampling methods yield similar network models of cellular functions.

3.4.4 Similar robustness profiles for different targeting strategies

Many theoretical studies of network robustness have examined the size of the largest connected component when removing the nodes with the largest degrees from a network [10, 46, 77, 604]. However, for applications in biology and pharmacology, researchers have suggested that other structural properties of networks can be relevant to the study of performance and viability of biological systems [17, 592]. Moreover, many different centrality measures (and combinations of centrality measures) perform better as indicators of "essentiality" [162, 170, 352, 353, 415, 592] and other node properties that are relevant in the study of networks in biology and biomedicine [592].

In this chapter, we investigated the robustness of several structural properties under node removal using several targeting strategies. For four centrality measures, we found that targeted node removal tends to lead to much greater impacts on structural properties of fPINs than removing nodes uniformly at random. Comparing degree, betweenness centrality, eigenvector centrality, and subgraph centrality, we found that targeting by different centrality measures tends to lead on average to similar impacts on structural properties. This observation is consistent with results of a comparative study of network robustness to node removal [552].

Overall, when comparing impacts of node removal for various structural properties X, the targeting strategies t, and the sampling methods s, we found that the impact of the targeted removal of n nodes from an fPIN tends to depend strongly

on the choice of X and weakly on the choices of t and s. We anticipate that this result is helpful for building intuition for the effects that different specifications for studies of robustness of PINs can have on the results of such a study. However, it is important to keep in mind that there exist exceptions to these observations. For example, sampling strategy greatly affects the impact of node removal on algebraic connectivity (see Fig. 3.9), and the choice of targeting strategy can affect the mean impact on T (see Fig. 3.7).

3.4.5 Plausible measures of performance and viability

Using selective robustness, structural sensitivity, and perturbation-size monotonicity as criteria for the plausibility of structural properties as measures of performance and viability on fPINs, we compared the plausibility of ten structural properties. For almost all considered sampling methods and targeting strategies, we found that relative LCC size F and algebraic connectivity a are the most plausible performance measures of the ten structural properties. Eight of ten structural properties have large values for the scores that measure selective robustness and perturbation-size monotonicity. The plausibility scores of performance measures thus correlate strongly with the scores for structural sensitivity. Further study of structural sensitivity of performance can help inform expectations of the plausibility of performance measures for fPINs.

For this course of research, a relevant observation from our study is the following. We tested the plausibility of ten performance measures. Seven of these ten performance measures are closely related to sums or mean values of properties of single nodes or single node pairs. The three remaining structural properties are F, a, and global clustering coefficient T, which are not closely related to sums or mean values of properties of a single node or a single node pair. These three structural properties tend to yield higher plausibility scores than most of the other considered structural properties. We hypothesize that the procedure of summing or averaging structural properties over nodes or node pairs reduces the score for structural sensitivity of performance measures; we leave the test of this hypothesis for future research.

3.4.6 Outlook

In this chapter, we examined several structural properties of fPINs and the robustness of fPIN properties under node removal. The aim of many studies of PINs is to identify proteins that are important for a cellular function or the viability of a cell or organism. An interesting avenue of future research is to test the practical use of

fPINs for this aim. When one experimentally observes the activation of alternative pathways in a cell that experiences a perturbation, is that pathway included in an associated fPIN? Can centrality measures on fPINs identify essential proteins and potential drug targets better than on protein interactomes of tissue-specific PINs? Addressing these questions will require comparisons of computational results on fPINs to experimental data.

Several studies of biological networks have indicated that the predictive power of network models of biological systems increases when one integrates several types of data (e.g., PPIs, differential gene expression, and other types of data) into a network model [2, 446]. We thus anticipate that one can improve fPINs by integrating several types of interaction data. A potential approach for achieving such an improvement is to include PPIs, gene interactions, and other interactions in a multilayer fPIN [78, 289]. Alternatively, one can consider including data from differential network models of PPIs and gene interactions as node and edge labels or node and edge weights [135, 578]. One can also consider to model fPINs as hypergraph in which hyperedges can represent protein complexes or other types of interactions between three or more proteins [290].

As discussed in Chapter 2, different network models can lead to very different results of the study of network robustness. The study of weighted fPINs, multilayer fPINs, and functional PPI hypergraphs can lead to different and potentially more accurate insights than the results in this chapter. There are several ways to extend the centrality measures and performance measures that we considered in this study to weighted networks [33, 394, 409, 439], multilayer networks [134, 500], and hypergraphs [64, 171, 272, 456]. These extensions offer a potential starting point for studies of structural robustness of weighted fPINs, multilayer fPINs, and functional PPI hypergraphs.

Other possible avenues of future research concern questions of network sampling and robustness of networks and are not necessarily tied to an application to gene or protein interactions. For example, what are the mechanisms by which the sampling methods in our study lead to network samples with very small values of C and T? How can one sample a subnetwork that includes nodes "close to" and "between" a set of seed nodes and has similar values of C or T as its parent network? The study of network samples of random-graph models with prescribed values of C or T [226, 275, 395] can perhaps address this question.

The results of our study of plausibility of performance measures indicate directions for future research of network robustness. One can adapt our plausibility tests

to various network systems (e.g., biological networks, financial networks, communication networks, supply networks, and so on) by adjusting the plausibility criteria or adding new plausibility criteria. Are there performance measures that are plausible for many different types of networks? A comparison of appropriate plausibility scores across different network types can help address this question. In our plausibility tests of structural properties of fPINs, we found that the plausibility scores for performance measures correlate strongly on the scores for structural sensitivity (see Figs. 3.10 and 3.11). To understand why some performance measures have higher plausibility scores than others requires the development of further understanding of the structural sensitivity of network properties to perturbations.

Chapter 4

Entropy, redundant paths, and cycles in a network with Ornstein–Uhlenbeck dynamics

The study of motifs in networks can help researchers uncover links between structure and function of networks in biology, ecology, computational social sciences, and many other fields. Empirical studies of networks in biology and biochemistry have identified feedback loops, feedforward loops, and several other small structures as "motifs" that often occur in real-world networks and may be components of important functional units of these systems. For many of these motifs, the mechanism by which they contribute to a network's functionality are unknown. Using an Ornstein–Uhlenbeck process as a simple model of noisy information propagation on a network, we find that, when differential entropy is a measure of system functionality, motifs need to be *circular* to contribute to system functionality. We find that a non-cyclic motif contributes more to differential entropy than a cyclic motif. However, cyclic motifs can have a greater total contribution to differential entropy than non-cyclic motifs if there are more cyclic motifs than non-cyclic motifs in a network. Our results highlight the importance of cycles and redundant paths and trails to the differential entropy of network dynamics and related network characteristics.

Researchers in many fields have used measures based on differential entropy to analyze global network characteristics. Our results offer an understanding and comparison of the mechanisms by which differential entropy arises in network dynamics. We expect that our findings will enable researchers to connect measures of global network characteristics to a network's local structure.

Symbol	Meaning
$\mathbb{R}$	the set of real numbers
$\mathbb{R}_+$	the set of positive real numbers
$\mathbb{N}_0$	the set of non-negative integers
$\mathbb{N}_0^*$	the set of arbitrary-length tuples of non-negative integers
(X, p)	a random variable on a set X with probability distribution p
$(X, \mathcal{B}, p, T)$	a dynamical system with phase space X, Borel algebra $\mathcal{B}$, probability measure p, and a map T
ℓ	number of (time-)steps or length of a walk
Q	a partition of a space
$H_S(p)$	the Shannon entropy of a probability distribution p
$h_S(p)$	the differential entropy of a probability distribution p
$H_{\mathrm{KS}}(T)$	the Kolmogorov-Sinai entropy of a dynamical system with map T
$\mathbf{\Sigma} = (\sigma_{ij})$	the covariance matrix of a dynamical system
$\mathcal{N}(0, \mathbf{\Sigma})$	a 0-centered multivariate normal distribution with scale matrix $\mathbf{\Sigma}$
$h_{[\infty]}$	the differential entropy of $\mathcal{N}(0, \mathbf{\Sigma})$
$h_{[L]}$	the L-walk approximation of the differential entropy of $\mathcal{N}(0, \mathbf{\Sigma})$
$\mathcal{E}_L$	the relative error of approximation $h_{[L]}$
$\mathbf{x}_t$	the time-dependent state of a multivariate system
θ	the reversion rate of an Ornstein–Uhlenbeck process
ϵ	the couping parameter of an Ornstein–Uhlenbeck process
ς	the noise amplitude of an Ornstein–Uhlenbeck process
W_t	a Wiener process
$\mathbf{I}$	the identity matrix
$\mathbf{A}$	a network's adjacency matrix
λ_1	the principal eigenvalue of $\mathbf{A}$
$\boldsymbol{\eta}$	the principal eigenvector of $\mathbf{A}$
ρ	the spectral radius of a matrix
(V, E, W)	a weighted network with node set $V = \{v_i\}_{i=1,\dots,n}$, edge set $E = \{e_{i,j}\}$, and edge weights $W : e \to \mathbb{R}_+$
N	the number of nodes in a network
m	the number of edges in a network
d	the density of a network
(V_P, E_P)	an (unweighted) path in a network
(V_C, E_C)	an (unweighted) cycle in a network
w	a walk
w_c	a closed walk
(V_W, E_W, ℓ_W)	a walk graph with node set V_W, edge set E_W, and edge lengths ℓ_W
L	the length of a walk graph (i.e., motif)
ν	the count of a motif
ω	the weight of a motif

Table 4.1: List of symbols for Chapter 4.

4.1 Introduction

The study of motifs in networks has advanced the understanding of various systems in biology and biochemistry [17, 486, 512], ecology [450, 451], economics [407, 521], computational social science [239, 267], and other areas. When interpreting motifs as small building blocks that can contribute to a network's functionality, the identification of motifs that are crucial to a network's function can be an important step towards uncovering the relationship between network structure and network function.

Several researchers have reported that closed walks and cycles[1] in a network's structure are important to various notions of robustness in networks [301, 452] and a network's ability to fulfill various functions [165, 196]. Other scholars have identified feedforward loops and other motifs as important building blocks of functional units in networks [17, 31, 542]. What do these motifs have in common? Are some of these motifs more important than others for a network's ability to perform a function? The answers to these questions depend on how one measures a network's function and on the dynamics on a network. As an illustrative example, we consider these questions for networks with Ornstein–Uhlenbeck dynamics and study how motifs contribute to the differential entropy of this system.

For a stochastic process with a single steady state, its entropy is maximal at its steady state [366]. The maximum entropy of a stochastic process is central to the maximum-entropy principle, which is often invoked for model specification and parameter estimation for various dynamical systems and their applications [41, 197]. Steady-state differential entropy is the basis for several quantifiers — such as neural complexity [533], redundancy and degeneracy [535], and robustness to small perturbations in node states [139] — of dynamical processes on networks. Across many domains, researchers have linked measures of differential entropy of a network to its ability to robustly perform a desired function [114, 139, 159, 178, 334, 467, 578].

Ornstein–Uhlenbeck processes [543] are simple and popular models for noisy coupled systems [1]. For example, they have been used as models for the dynamics for neuronal systems [31, 533], stock prices [325], and gene expression [453]. To give us a concrete setting, we use an Ornstein–Uhlenbeck (OU) process as a simple model of noisy information-propagation on a network and derive its differential entropy. We

[1] Bollobás (2013) distinguished between closed walks and cycles [60]. When using the terminology in Ref. [60], a cycle is not a special case of a closed walk. We use the terminology from Ref. [60] in this chapter. We explain the difference between a closed walk and a cycle in Section 4.4.1.

identify *circular walk graphs*[2] as the class of motifs that contribute to the differential entropy of a network with a multivariate OU process. A circular walk graph is either a cyclic[3] motif or a non-cyclic motif. Cyclic motifs occur, for example, on the 3-node feedback loop. Non-cyclic motifs occur, for example, on the 3-node feedforward loop. We find that a non-cyclic motif contributes more to differential entropy than a cyclic motif. However, cyclic motifs can have a greater total contribution to differential entropy than non-cyclic motifs if more cyclic motifs than non-cyclic motifs occur on a network. Our results show that the propagation of signals along cycles and redundant paths and trails are mechanisms by which network structure contributes to the network's differential entropy and network characteristics that are related to differential entropy[533–535]. We anticipate that our results are a possible starting point for connecting measures of global network characteristics to a network's local structure.

Several researchers have explored connections between network structure and information propagation of a multivariate OU process [31, 247, 533, 538]. Barnett et al. (2009) used the covariance matrix of a network with OU dynamics to derive 2-edge and 3-edge motifs that contribute to neural complexity [31], which is a function of mean differential entropy of subsystems [533]. Trousdale et al. (2012) reported that recurrent network structure, which allows a signal to propagate in cycles, influences cross-correlations in a neural network [538]. Hu et al. (2014) used motif counts to determine pairwise and higher-order correlations in networks with OU dynamics [247]. We compare our results to their findings in Section 5.7.

In contrast to Chapter 3, we consider directed, *weighted* networks with self-edges in this chapter. We consider a *directed weighted network* (V, E, W) to be an object that is defined by a set V of nodes (i.e., vertices), a set $E \subset V \times V$ of directed edges, and positive real edge weights $W : E \to \mathbb{R}_+$. We use the terms *network* and *graph* interchangeably. The systems of interest in this chapter are multivariate dynamical processes, for which a network encodes the coupling structure between variables. We consider a motif to be a small weakly connected *walk graph* (V_W, E_W, ℓ_W). For a network with node set V, an associated walk graph is a directed multigraph with a node set V_W that is a combination (with repetition allowed) of V. Edges in the edge set E_W correspond to walks in the associated network (V, E, W). Each edge in E_W has a non-negative integer *length* $\ell_W : E_W \to \mathbb{N}_0$ that is the length of the

[2]We define a *walk graph* as a directed multigraph in which each edge corresponds to a walk from one node to another. A *circular* walk graph is a walk graph such that if one replaces each directed edge by an undirected edge, the resulting graph is an undirected cycle. See Section 4.4 for details.

[3]A walk graph is *cyclic* if it is a directed cycle.

corresponding walk in (V, E, W). In Fig. 4.1, we give examples of walk graphs for a small network.

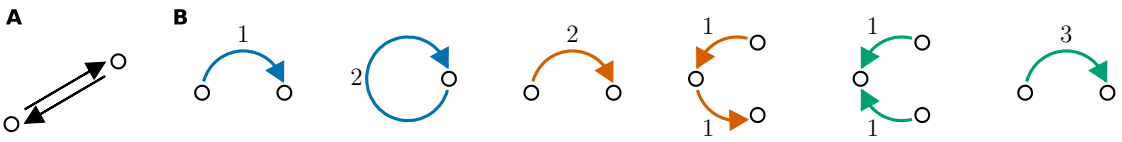


Figure 4.1: Examples of walk graphs. In (**A**), we show a small network (V, E, W). In (**B**), we show examples of associated walk graphs. Counting from left to right, the first and second walk graph (i.e., the walk graphs with blue edges) are examples of walk graphs that use each edge in E at most once and in which each node corresponds to a different node in V. The third and fourth walk graph (i.e., walk graphs with orange edges) are examples of walk graphs in which two nodes correspond to the same node in V. The fifth and sixth walk graph (i.e., walks graphs with green edges) are examples of walk graphs that use edges in E more than once.

This chapter proceeds as follows. In Section 4.2, we review notions of entropy in dynamical systems and network analysis. In Section 4.3, we give a brief introduction to the OU process and express covariance and maximal differential entropy of a network with OU dynamics as functions of the network's adjacency matrix. In Section 4.4, we introduce walk graphs and motivate their use as motifs in dynamics on networks. We also discuss the relationship between walk graphs and connected subgraphs, which are commonly associated with motifs in networks [17, 407, 450, 451, 486, 521, 542]. Using the expressions for covariance from Section 4.3, we characterize the motifs that contribute to covariance in a multivariate OU process on a network and identify the motifs that contribute the most to covariance in Section 4.5. In Section 4.6, we characterize the motifs that contribute to the differential entropy of a multivariate OU process on a network. In Section 5.7, we discuss our findings and compare our results to previous studies.

4.2 Background: Entropy and networks

Entropy is a central concept to this chapter. In this section, we briefly review common notions of entropy in dynamical systems and networks. The purpose of this discussion is to enable readers to connect the notion of entropy that we use in subsequent sections to other notions of entropy that are popular in dynamical systems and network analysis.

4.2.1 Shannon entropy and Kolmogorov–Sinai entropy

Two widely-used definitions of entropy are Shannon entropy (SE) and the Kolmogorov–Sinai entropy (KSE), which is an extension of SE to dynamical systems [148, 194].

Shannon (1949) considered an "information source" that emits a sequence (x_1, x_2, ...) of messages x_i, where x_i is an element of some finite or countable set X [482, 576]. Assuming that messages are independent of one another and follow a probability distribution $p : X \to \mathbb{R}$, one can consider this information source to be a random variable (X, p). Shannon defined its entropy to be the mean rate of information production:

$$H_{\mathrm{S}}(p) := -\sum_{x \in X} p(x) \log(p(x)) \geq 0 \,. \tag{4.1}$$

If the set X is uncountable, one often replaces H_{S} with the *differential entropy* [366, p. 27]

$$h_{\mathrm{S}}(p) := -\int_{x \in X} p(x) \log(p(x)) \,,$$

which, for a given countable partition of X, is an affine function of the negative Kullback–Leibler divergence from the uniform distribution on X to $p(x)$ [206, 480].

Sinai and Kolmogorov (1959) extended SE to dynamical systems [296, 491]. Instead of a random variable, one considers a dynamical system $(X, \mathcal{B}, p, T)$. This tuple consists of a phase space X with a Borel algebra $\mathcal{B}$, a probability measure $p : (X, \mathcal{B}) \to \mathbb{R}$, and a map $T : X \to X$ that describes the dynamics on X. Instead of taking the mean over all possible realizations of a random variable, one takes the mean over all possible trajectories in phase space and over time.

To formalize this idea, consider a partition $Q = \{Q_1, Q_2, \dots Q_k\}$ of the space X. The *T-pullback* of Q is [495]

$$T^{-1}Q := \{T^{-1}Q_1, T^{-1}Q_2, \dots, T^{-1}Q_k\} \,.$$

The *refinement* of two partitions Q and R is [495]

$$Q \vee R := \{Q_i \cap R_j \,|\, p(Q_i \cap R_j) > 0\} \,.$$

The refinement of a partition Q and its T-pullback consists of sets $Q_i \cap T^{-1}Q_j$ for some $i, j \leq k$. The probability $p(Q_i \cap T^{-1}Q_j)$ is the probability of observing $x \in Q_j$ at time $t = -1$ and $x \in Q_i$ at time $t = 0$. One can view the refinement $Q \vee T^{-1}Q$ as the set of possible 1-step trajectories on X with respect to the partition Q. Using

$\bigvee_{i=1}^{k} R_n$ to denote the refinement $R_1 \vee R_2 \vee \cdots \vee R_k$, one can write $\bigvee_{i=0}^{\ell} T^{-i}Q$ to describe the set of possible ℓ-step trajectories on X with respect to the partition Q.

The KSE is [495]

$$H_{\mathrm{KS}}(T) = \sup_{Q} h(T, Q)\,, \tag{4.2}$$

where

$$h(T, Q) = \lim_{k\to\infty} \frac{1}{\ell} H_{\mathrm{S}} \left(\bigvee_{i=0}^{\ell} T^{-i} Q \right) \tag{4.3}$$

is the time average of the SE of a stochastic process with respect to a partition Q of X. The division by the number ℓ of time steps in Eq. (4.3) has lead many researchers to interpret KSE as a "rate" of an SE [194, 199, 414]. Accordingly, many researchers refer to KSE by "Kolmogorov–Sinai entropy per unit time" [194], "Kolmogorov–Sinai entropy rate" [414], or simply "entropy rate" [199, 414].

For many dynamical systems, it is difficult to compute KSE [463]. For stochastic processes with limited memory, the system's state x_t is independent of previous states x_τ for all $\tau < t - \Delta$ (for some value of Δ). One can then sometimes simplify Eq. (4.2). For example, if T describes a first-order Markov process with a stationary distribution, one obtains [51]

$$H_{\mathrm{KS}}(T) = -\sum_{i,j=1}^{n} p_i^* p_{ij} \log(p_{ij})\,, \tag{4.4}$$

where n is the size (i.e., number of elements) of X, the probability distribution $p^* = (p_i^*)_{i=1,\ldots,n}$ is the stationary distribution of the Markov process, and p_{ij} is the transition probability from state x_i to state x_j. As another example, consider the case $\Delta = 0$. The state x_t is then independent of all previous states. At each time step, the state of the stochastic process is a sample from some probability distribution $p(x)$. The KSE of this stochastic process is equivalent to the Shannon entropy of the probability distribution $p(x)$. That is,

$$H_{\mathrm{KS}}(T) = H_{\mathrm{S}}(p(x))\,. \tag{4.5}$$

Equation (4.5) motivates the study of the SE of steady-state distributions of stochastic processes [31, 533–535].

4.2.2 Network entropy

Nazem (1979) was one of the first scholars to define a concept of a "network entropy" for the study of networks beyond those that have a lattice structure [388]. From the flow of goods in a freight transportation network, Nazem calculated transition probabilities for goods between cities. He used the KSE of the associated Markov process to characterize the associated U.S. freight transportation network.

During the last 50 years, network entropy has been used widely as a concept to describe the structure and dynamics of networks [137]. We distinguish two primary approaches to defining entropy on a network:

1. One can describe a network in terms of distributions of node, edge, or other subgraph properties. An example is its node degree distribution (i.e., graph entropy (see Section 2.4.4.4)). The SE of any of these distributions provides a notion of network entropy.

2. In the study of dynamical processes on networks, one often considers a stochastic process of N coupled elements with individual states $x_i \in X$ and pairwise coupling. One can use the KSE of the dynamics $T : X^N \to X^N$ of the coupled system to define a notion of network entropy.

Several scholars have proposed to use the SE of distributions of a variety of structural properties of nodes, edges, and subgraphs to classify networks. Wang et al. (2006) used SE of the degree distribution to study the robustness of a network's degree distribution to the removal of nodes [563]. Carpi et al. (2010) combined this SE with other information-theoretic measures on gene networks to develop indicators of Alzheimer's disease [84]. Solé and Valverde (2004) used the SE of the excess degree distribution[4] and the SE of the joint degree–degree distribution to study the evolution of networks under architectural constraints [499]. Zhao et al. (2011) used the distribution of cliques in a temporal network of phone calls to detect diurnal patterns in social interactions [608].

Kolmogorov–Sinai entropy is a popular measure for characterizing dynamical processes on networks. Monteforte and Wolf (2010) and Baptista et al. (2010) used KSE to study the effects of inhibitory neurons in a network with oscillatory neural dynamics [29, 374]. Kannathal et al. (2005) developed indicators for detecting epilepsy from electroencephalogram (EEG) data using KSE [271].

[4] A network's *excess degree distribution* is defined as $q(k) = \frac{k+1}{\langle k \rangle} p_{k+1}$, where p_k is the probability for a node in the network to have degree k and $\langle k \rangle = \sum_k k p_k$ is the mean degree of the network [398].

Many popular models for dynamics on networks are types of random walks [362]. For example, one can use random walks to derive centrality measures on networks [404] and to study flow and consensus dynamics on networks [131, 199, 304]. Random walks with the largest KSE for a given network structure are called *maximum-entropy random walks* and have been used in many studies of networks [199, 404, 425, 492].

For random walks on networks, one can use Eq. (4.4) to compute KSE from a walk's *stochastic matrix* $\mathbf{P} := (p_{ij})$. For a given network structure, one obtains the random walk that has the largest KSE by choosing the transitions probabilities $p_{ij} = a_{ij}\eta_j/(\lambda_1\eta_i)$, where $\mathbf{A} = (a_{ij})$ is the network's adjacency matrix, λ_1 is $\mathbf{A}$'s principal eigenvalue and $\boldsymbol{\eta} = (\eta_1, \eta_2, \ldots, \eta_N)^T$ is the normalized[5] principal eigenvector of $\mathbf{A}$ [75]. Demetrius and Manke (2005) used the KSE of a maximum-entropy random walk on a network to measure the robustness of dynamics on networks to small perturbations to the states of nodes [139]. Teschendorff et al. (2014) reported a positive correlation between drug resistance and the KSE of a maximum-entropy random walk on a protein-interaction network that corresponds to a drug's signaling pathway [528].

4.3 The Ornstein–Uhlenbeck process and its differential entropy

In this section, we give a short introduction to multivariate OU processes and derive expressions for their covariance matrix and differential entropy at steady state. We also discuss the conditions under which these expressions are valid. We propose an approximation $h_{[L]}$ for the differential entropy of an OU process at steady state and test its accuracy for multivariate OU processes on directed versions of $G(N, m)$ Erdős–Rényi random graphs (see Section 3.2.2).

4.3.1 The Ornstein–Uhlenbeck process

Uhlenbeck and Ornstein (1930) proposed a stochastic process to describe Brownian motion under the influence of friction [543]. Among the many extensions of their model, the multivariate OU process is a popular model for coupled noisy systems, including neuronal dynamics [31], stock prices [325], and the evolution of gene

[5] We normalize $\boldsymbol{\eta}$ such that $\sum_i \eta_i^2 = 1$.

expression [453]. In all of these studies, an n-variate OU process describes the dynamics on a network with n nodes, where the state of each node represents a neuron, stock, or gene-expression level.

Let's consider a system with a time-dependent n-variate state

$$\mathbf{x}_t := (x_{1,t}, x_{2,t}, \ldots, x_{N,t})\,.$$

One can describe the n-variate OU process as a map $T_{\mathrm{OU}} : \mathbf{x}_t \to \mathbf{x}_{t+dt}$, such that

$$d\mathbf{x}_{t+dt} = \theta(\epsilon\mathbf{A} - \mathbf{I})\mathbf{x}_t\, dt + \varsigma\, dW_t\,. \tag{4.6}$$

We describe the state of the process by a column vector $\mathbf{x}_t \in \mathbb{R}^N$. This process has an adjacency matrix $\mathbf{A}$, which can be directed and/or weighted, and a multivariate Wiener process W_t. The *reversion rate* $\theta > 0$, the *noise amplitude* ς^2, and the *coupling parameter* $\epsilon > 0$ are parameters of the OU process.

Consider a signal to be a temporary deviation of a node's state from its mean. The coupling parameter sets the rate at which a signal's amplitude increases or decreases when it is transmitted from one node to another. For sufficiently small ϵ, the eigenvalues of the matrix $\epsilon\mathbf{A} - \mathbf{I}$ have negative real parts[6]. The multivariate OU process then has a single stationary distribution, which is a 0-centered multivariate normal distribution.

The parameter θ is the rate at which a signal's amplitude increases or decreases over time. It thus determines the expected speed at which a node's state reverts to its mean. This connection between θ and the speed of signal decay in an OU process has led many researchers to refer to θ as the *reversion rate* of an OU process [405, 460, 539].

If all eigenvalues of $\epsilon\mathbf{A} - \mathbf{I}$ have negative real parts, the multivariate OU process has a single stationary distribution. We say that such an OU process is a process with *signal decay* because, in such a process, a signal's amplitude decreases with time. Denoting the spectral radius of a matrix by $\rho(\cdot)$, a sufficient condition for signal decay is $\rho(\epsilon\mathbf{A}) < 1$.

4.3.2 Variance and differential entropy of the Ornstein–Uhlenbeck process

The multivariate OU process with signal decay is a Markov process. Its stationary distribution is a multivariate normal distribution $\mathcal{N}(0, \mathbf{\Sigma})$ centered at $\langle\mathbf{x}\rangle = 0$ with

[6] A matrix that fulfills this condition is sometimes called a Hurwitz matrix or a stable matrix.

sample covariance matrix $\mathbf{\Sigma} := \langle \mathbf{x}_t \mathbf{x}_t^T \rangle$ [31]. In Appendix B.1, we show that an OU process with signal decay has covariance matrix

$$\mathbf{\Sigma} = \frac{\varsigma^2}{2\theta} \sum_{\ell=0}^{\infty} \sum_{r=0}^{\infty} 2^{-(\ell+r)} \binom{\ell+r}{\ell} (\epsilon \mathbf{A}^T)^\ell (\epsilon \mathbf{A})^r \,. \tag{4.7}$$

The differential entropy of the stationary distribution of the multivariate OU process is [112]

$$h_{\mathrm{S}} \left(\mathcal{N}(0, \mathbf{\Sigma}) \right) = \frac{1}{2} \left[N \log(2\pi e) + \log \left(\det(\mathbf{\Sigma}) \right) \right] \,.$$

For convenience, we introduce the shorthands $h_{[\infty]} := h_{\mathrm{S}} \left(\mathcal{N}(0, \mathbf{\Sigma}) \right)$ and

$$\mathbf{S}_{[a,b]} := \sum_{\ell=0}^{b} \sum_{\substack{r=0\,, \\ a \leq \ell + r \leq b}}^{b} \left(\frac{\epsilon}{2} \right)^{\ell+r} \binom{\ell+r}{\ell} (\mathbf{A}^T)^\ell \mathbf{A}^r \,.$$

We can then write

$$h_{[\infty]} = \frac{1}{2} \left[N \log \left(\frac{\pi e \varsigma^2}{\theta} \right) + \log \left(\det(\mathbf{S}_{[0,\infty]}) \right) \right] \,.$$

4.3.3 Short-range signal decay

When the positive-definite matrix $\mathbf{S}_{[0,\infty]}$ has sufficiently small eigenvalues, we can use the matrix identity $\det e^{\mathbf{A}} = e^{\mathrm{tr}\mathbf{A}}$ and a power-series expression for the logarithm of a matrix to obtain

$$\begin{aligned} h_{[\infty]} &= \frac{1}{2} \left[N \log \left(\frac{\pi e \varsigma^2}{\theta} \right) + \mathrm{tr} \left(\log(\mathbf{S}_{[0,\infty]}) \right) \right] \\ &= \frac{1}{2} \left[N \log \left(\frac{\pi e \varsigma^2}{\theta} \right) + \mathrm{tr} \left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (\mathbf{S}_{[1,\infty]})^k \right) \right] \,. \end{aligned} \tag{4.8}$$

This expression for the differential entropy of an OU process is valid whenever the eigenvalues of $\mathbf{S}_{[0,\infty]}$ are small enough for the sum in Eq. (4.8) to converge. If an OU process has adjacency matrix $\mathbf{A}$ and coupling parameter ϵ such that the sum in Eq. (4.8) converges, we say that an OU process has *short-range signal decay*[7]. In Appendix B.2, we show that a sufficient condition for short-range signal decay is that the Hilbert–Schmidt norm $\|\epsilon \mathbf{A}\|_2 < 1/2$. In that appendix, we also show that, when $\mathbf{A}$ is the adjacency matrix of a strongly connected network with non-negative edge weights, one can relax this condition to $\rho(\epsilon \mathbf{A}) < 1/2$.

[7] This is equivalent to the assumption of weak coupling in Ref. [31].

From Eq. (4.8), we see that a change in either ς or θ affects only the structure-independent offset $N\log(\pi e\varsigma^2/\theta)/2$. The two parameters thus do not affect the contribution of network structure and entropy. In the remainder of this chapter, we assume for convenience that $\varsigma = \theta = 1$ and write $h_{[\infty]}(\epsilon, \mathbf{A})$ to denote the differential entropy in Eq. (4.8) for an OU process with adjacency matrix $\mathbf{A}$ and coupling parameter ϵ.

4.3.4 Network structure matters

In this section, we investigate whether (and how much) one should expect network structure to influence the value of $h_{[\infty]}(\epsilon, \mathbf{A})$. Functions of powers of adjacency matrices have been the basis for the development of several measures of network structure [92, 164, 165, 168] and node centrality [44, 170]. Many of these measures tend to have a strong positive correlation to the number of edges in a network [92, 164, 168]. Depending on context, one can consider this to be either an advantage or a limitation of a measure. We are interested in the contribution of motifs in networks to $h_{[\infty]}(\epsilon, \mathbf{A})$. To confirm that network structure affects $h_{[\infty]}(\epsilon, \mathbf{A})$, we investigate $h_{[\infty]}(\epsilon, \mathbf{A})$ for random networks with fixed numbers of nodes and edges using a directed $G(N, m)$ random-graph model with self-edges. We show our results in Fig. 4.2.

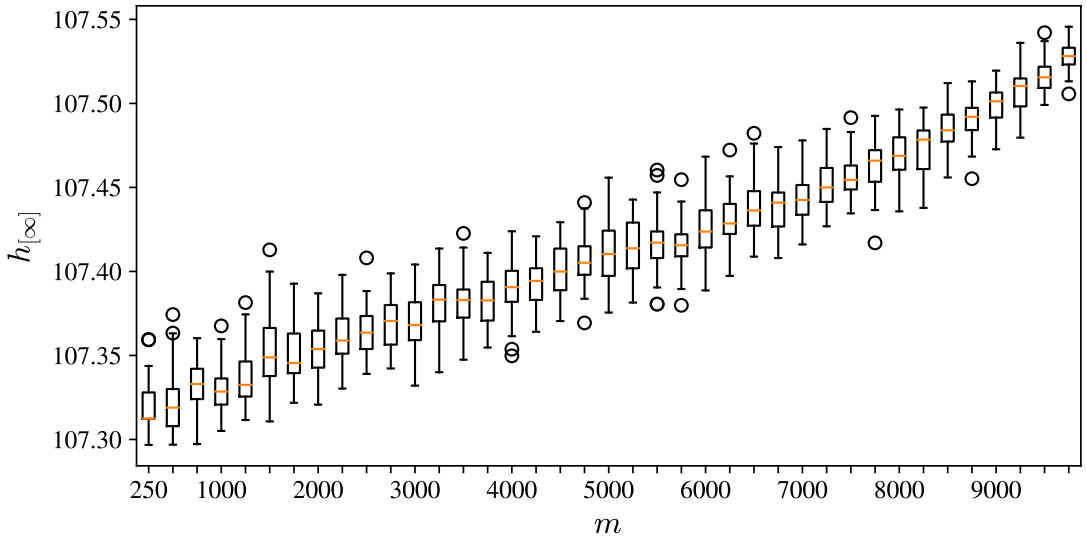


Figure 4.2: The differential entropy $h_{[\infty]}(\epsilon, \mathbf{A})$ for a multivariate OU process on a network drawn from a directed $G(N, m)$ Erdős–Rényi random-graph model that allows self-edges. We set $N = 100$ and consider 50 realizations of the $G(N, m)$ model for different values of m. For each m, we thus obtain a set of 50 values for $h_{[\infty]}(\epsilon, \mathbf{A})$, which we visualize using standard matplotlib box plots[9]. For each adjacency matrix, we choose ϵ such that $\|\epsilon\mathbf{A}\| = 0.49$.

We observe that $h_{[\infty]}(\epsilon, \mathbf{A})$ has a strong positive correlation with the number of edges. However, the values of $h_{[\infty]}(\epsilon, \mathbf{A})$ for two networks with the same number of nodes and edges can differ greatly (by up to 0.24 nats for the data in Fig. 4.2). We conclude that when one only has knowledge of the number of edges but not the network structure, one can make a very good estimate for the mean of $h_{[\infty]}(\epsilon, \mathbf{A})$ for an OU process on a directed $G(N, m)$ graph. For an individual network, it is not possible to obtain a similarly good estimate for $h_{[\infty]}(\epsilon, \mathbf{A})$ from the number of nodes and the number of edges. The differential entropy $h_{[\infty]}(\epsilon, \mathbf{A})$ of an individual network depends on the network's structure.

4.3.5 Finite-walk approximation

For an OU process with short-range signal decay, the amplitude of a signal decreases rapidly as the signal propagates along edges. The convergence of the sum in Eq. (4.8) indicates that signals that traverse many edges give only a small contribution to $h_{[\infty]}(\epsilon, \mathbf{A})$. More precisely, the contribution to $h_{[\infty]}(\epsilon, \mathbf{A})$ from signals that traverse L edges approaches 0 as $L \to \infty$. From $h_{[\infty]}(\epsilon, \mathbf{A})$, we construct an approximation $h_{[L]}(\epsilon, \mathbf{A})$ of the maximum differential entropy of an OU process by keeping only terms up to $O(\epsilon^L)$ in Eq. (4.8). The approximation $h_{[L]}(\epsilon, \mathbf{A})$ equals the sum of contributions to $h_{[\infty]}(\epsilon, \mathbf{A})$ by signals that traverse at most L edges. It does not include contributions of signals that traverse more than L edges. We write

$$h_{[L]}(\epsilon, \mathbf{A}) = H_0 + \sum_{i=1}^{n} \sum_{k=1}^{L} \frac{(-1)^{k+1}}{2k} \left(\frac{\epsilon}{2}\right)^L \sum_{\substack{\boldsymbol{\ell}, \mathbf{r} \in \mathbb{N}_0^k, \\ \sum \boldsymbol{\ell} + \sum \mathbf{r} = L, \\ \min_j (\ell_j + r_j) \geq 1}} \left[\left(\prod_{j=1}^{k} \binom{l_j + r_j}{l_j} \right) \nu_i^{\mathrm{EM}}(\mathbf{A}, \boldsymbol{\ell}, \mathbf{r}) \right], \tag{4.9}$$

where H_0 is a shorthand for the structure-independent term $N \log(\pi e)/2$. We use $\mathbb{N}_0^k$ to denote the set of k-tuples whose elements are non-negative integers and $\sum \boldsymbol{\ell}$ to denote the sum of all elements of a tuple $\boldsymbol{\ell}$. The number $\nu_i^{\mathrm{EM}}(\mathbf{A}, \boldsymbol{\ell}, \mathbf{r})$ is the ith diagonal element of the matrix

$$\left(\mathbf{A}^{\ell_1}\right)^T \mathbf{A}^{r_1} \left(\mathbf{A}^{\ell_2}\right)^T \mathbf{A}^{r_2} \ldots \left(\mathbf{A}^{\ell_k}\right)^T \mathbf{A}^{r_k}. \tag{4.10}$$

[9] Orange lines indicate the median q_2 of this set. The lower and upper edge of a box indicate the first quartile q_1 and third quartile q_3 of the set. If the smallest value in the set is larger than $q_2 - 1.5 \cdot (q_3 - q_1)$, the bottom whisker ends in the smallest value in the set. If the smallest value in the set is less than $q_2 - 1.5 \cdot (q_3 - q_1)$, the bottom whisker ends in $q_2 - 1.5 \cdot (q_3 - q_1)$. Black circles indicate all values that are smaller than the end of the bottom whisker. Similarly, the top whisker either ends in the largest value in the set or in $q_2 + 1.5 \cdot (q_3 - q_1)$.

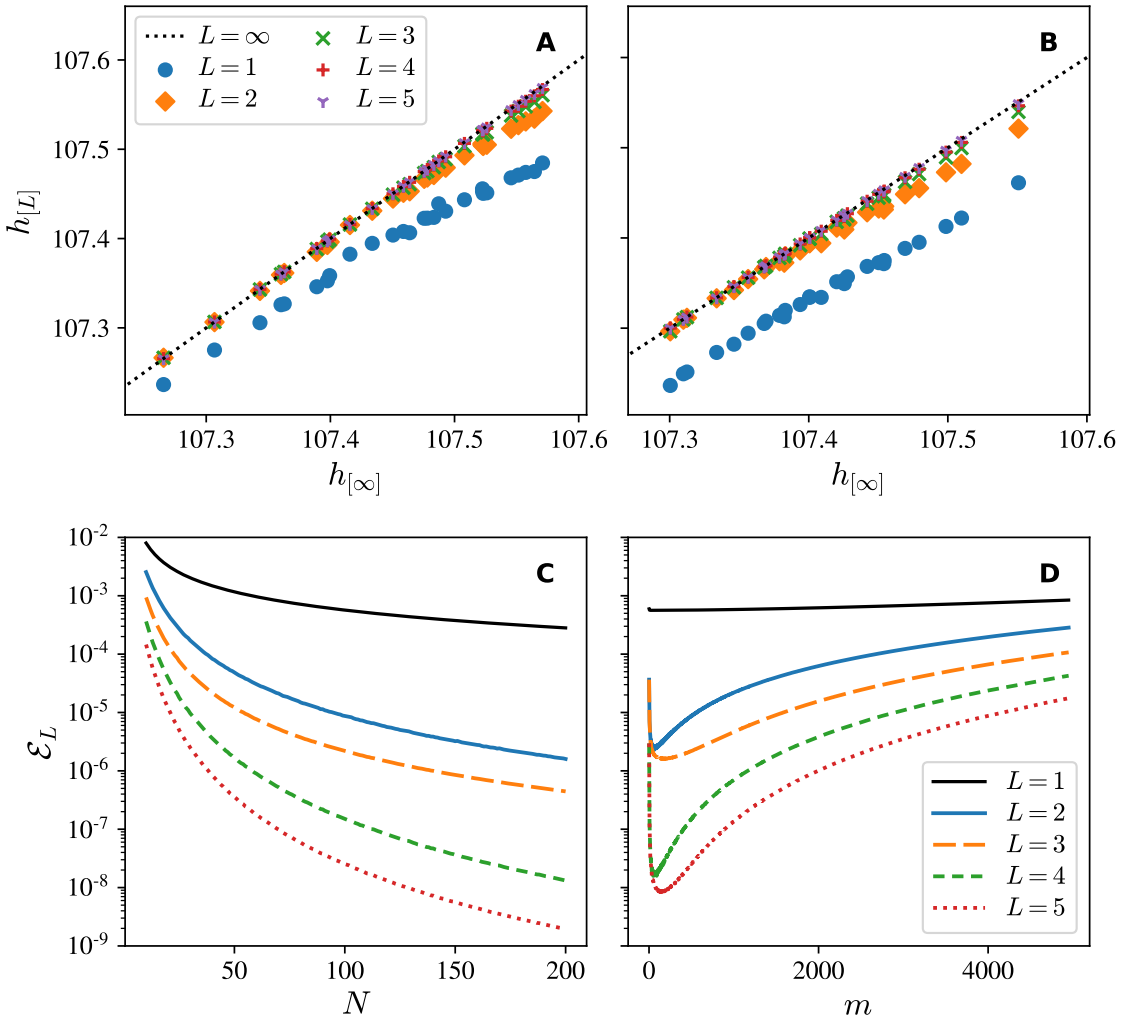


Figure 4.3: Accuracy of $h_{[L]}(\epsilon, \mathbf{A})$ as an approximation of $h_{[\infty]}(\epsilon, \mathbf{A})$ for $G(N, m)$ random graphs with self-edges. In (**A**) and (**B**), we show $h_{[L]}(\epsilon, \mathbf{A})$ and $h_{[\infty]}(\epsilon, \mathbf{A})$ for directed and undirected $G(N, m)$ random graphs, respectively. All graphs have $N = 100$ nodes. We choose the coupling parameter ϵ so that $\|\epsilon\mathbf{A}\| = 0.49$. The integer m takes equidistant values on $[1, n(n-1)]$ for the directed graphs and on $[1, N(N-1)/2]$ for the undirected graphs. In (**C**), we show the relative error $\mathcal{E}_L$ of $h_{[L]}(\epsilon, \mathbf{A})$ for undirected $G(N, m)$ graphs with mean degree 10. In (**D**), we show $\mathcal{E}_L$ for undirected $G(N, m)$ graphs with $N = 100$ nodes. All data in (**C**) and (**D**) are mean values of $\mathcal{E}_L$ over 100 $G(N, m)$ graphs.

In Fig. 4.3, we compare $h_{[\infty]}(\epsilon, \mathbf{A})$ and $h_{[L]}(\epsilon, \mathbf{A})$ OU processes on $G(N, m)$ random graphs with self-edges. In panels (**A**) and (**B**), we show results for a directed and undirected $G(N, m)$ random-graph model, respectively. We observe that $h_{[L]}(\epsilon, \mathbf{A})$ approximates $h_{[\infty]}(\epsilon, \mathbf{A})$ very well when $L \geq 3$. For progressively larger L, the difference between $h_{[\infty]}(\epsilon, \mathbf{A})$ and $h_{[L]}(\epsilon, \mathbf{A})$ is progressively smaller. In panels (**C**) and (**D**) of Fig. 4.3, we show the relative error $\mathcal{E}_L := |h_{[\infty]} - h_{[L]}|/h_{[\infty]}$ for different values

of the number n of nodes and the number m of edges in undirected $G(N, m)$ random graphs. We observe that, for constant m, the relative error decreases with n and that, for constant n, it increases with m. These observations indicate that $h_{[L]}(\epsilon, \mathbf{A})$ works best as an approximation for $h_{[\infty]}(\epsilon, \mathbf{A})$ when graphs are large and sparse. The relative error of the approximation tends to increase with the density[10] d of a network. For the data in panel (**C**), we find that $h_{[3]}(\epsilon, \mathbf{A})$ tends to have five or more significant digits when $n \geq 30$ (i.e., for networks with $d \lessapprox 0.33$). For the data in panel (**D**), we find that $h_{[3]}(\epsilon, \mathbf{A})$ tends to have five or more significant digits when $m \leq 2895$ (i.e., for networks with $d \lessapprox 0.58$). Increasing L increases the accuracy of $h_{[L]}(\epsilon, \mathbf{A})$ as an approximation $h_{[\infty]}(\epsilon, \mathbf{A})$ for all considered values of the model parameters n and m.

When one is interested solely in computing $h_{[\infty]}(\epsilon, \mathbf{A})$, the approximation $h_{[L]}(\epsilon, \mathbf{A})$ has no practical use. By using standard numerical solvers for Lyapunov or Sylvester equations [90], one can obtain a value of $h_{[\infty]}(\epsilon, \mathbf{A})$ that is accurate up to a user-prescribed numerical precision. (See, for example, the function `linalg.solve_lyap` in Python's `scipy` library.) For $L \geq 3$, the computation of $h_{[L]}(\epsilon, \mathbf{A})$ tends to take longer than a numerical computation of $h_{[\infty]}(\epsilon, \mathbf{A})$. Instead of increased computational efficiency, the approximation $h_{[L]}(\epsilon, \mathbf{A})$ can offer theoretical insights. It connects the differential entropy of a multivariate OU process to network structure. In Section 4.6, we show that one can use $h_{[L]}(\epsilon, \mathbf{A})$ to identify walks graphs that contribute to the differential entropy of a multivariate OU process.

4.4 Walk graphs as network motifs

In this section, we introduce walk graphs and motivate their use as motifs in dynamical processes on networks. In Section 4.4.1, we briefly review graph-theoretical concepts that are useful for the study of walk graphs. In Section 4.4.2, we define walk graphs and connect counts of walk graphs on a network to elements of powers of the network's adjacency matrix. We discuss the relationship between walk graphs and connected subgraphs in Section 4.4.3.

4.4.1 Some graph-theoretic concepts

In this chapter, we distinguish between *walks*, *trails*, and *paths* on directed networks. Consider a directed network (V, E). Recall from Chapter 2 that a *walk* in this

[10] The density d of a network is the ratio of the number of edges in a network to the number of possible edges in a network. An undirected network with n nodes and m edges has a density $d = 2m/n^2$ when one allows the network to have self-edges [398, p. 128].

network is a sequence

$$w = (v_{i_1}, e_{i_1,i_2}, v_{i_2}, e_{i_2,i_3}, \ldots, e_{i_{\ell-1},i_\ell}, v_{i_\ell}, e_{i_\ell,i_{\ell+1}}, v_{i_{\ell+1}})$$

of nodes $v_{i_1}, v_{i_2}, \ldots, v_{i_\ell}, v_{i_{\ell+1}} \in V$ and edges $e_{i_1,i_2}, e_{i_2,i_3}, \ldots, e_{i_{\ell-1},i_\ell}, e_{i_\ell,i_{\ell+1}} \in E$ such that each edge $e_{i,j}$ starts at the node v_i and ends at the node v_j [60, p. 4]. The number ℓ is the *length* of a walk. If no edge in E appears more than once in w, the walk w is also a *trail* [60]. If no node in V and no edge in E appears more than once in w, one can use the set of nodes in w and the set of edges in w to construct a *path*. Also recall from Chapter 3 that a *path* is a tuple (V_P, E_P) that consists of a node set $V_P \subseteq V$ and an edge set $E_P \subseteq E$ that one can combine to construct a sequence $(v_{i_1}, e_{i_1,i_2}, v_{i_2}, e_{i_2,i_3}, \ldots, e_{i_{\ell-1},i_\ell}, v_{i_\ell}, e_{i_\ell,i_{\ell+1}}, v_{i_{\ell+1}})$ of nodes and edges [60]. The number ℓ is the *length* of a path. A path is a subnetwork of a network. By contrast, a walk is a combination (with repetition allowed) of a network's nodes and edges.

A *closed walk* of length ℓ is a sequence

$$w_c = (v_{i_1}, e_{i_1 i_2}, v_{i_2}, e_{i_2 i_3}, \ldots, e_{i_{\ell-1} i_\ell}, v_{i_\ell}, e_{i_\ell i_1}, v_{i_1})$$

of nodes and edges [60]. A *cycle* of length ℓ is a tuple (V_C, E_C) that consists of a node set $V_C \subseteq V$ and an edge set $E_C \subseteq E$ that one can combine to construct a sequence $(v_{i_1}, e_{i_1,i_2}, v_{i_2}, e_{i_2,i_3}, \ldots, e_{i_{\ell-1},i_\ell}, v_{i_\ell}, e_{i_\ell,i_1})$ [60]. One can think of a cycle as a closed path.

We say that a graph is *cyclic* if it is a cycle. A graph is *non-cyclic* if it is not a cycle. It is *acyclic* if it is not a cycle and none of its subgraphs is a cycle. A directed *circular* graph is graph such that, if one replaces each directed edge by an undirected edge, the resulting graph is a cycle.

4.4.2 Characterizing and counting walk graphs

Consider a (possibly weighted) network (V, E, W). A *walk graph* on this network is a directed multigraph (V_W, E_W, ℓ_W), where V_W is a combination (with repetition allowed) of V and each edge in E_W corresponds to a walk in the associated network. The *length* $\ell_W(e)$ of an edge $e \in E_W$ is the length of the corresponding walk (i.e., the number of edges that a walk in (V, E, W) traverses). When studying walk graphs, it can be helpful to consider the number $|V_W|$ of nodes and the walk graph's *length* $L := \sum_{e \in E_W} \ell_W(e)$.

To briefly illustrate these concepts, we consider a walk graph with only one edge. A walk graph that consists of a single edge corresponds to a walk on the associated

network. If that edge in the walk graph is a self-edge, the walk graph is cyclic and its edge corresponds to a closed walk in the associated network.

On an unweighted network, the number of walks of length ℓ from a node i to a node j is given by the ijth element of the ℓth power of $\mathbf{A}$. Consequently, one can obtain the number of walk graphs on an unweighted network as a product of elements of powers of the network's adjacency matrix. For weighted networks, $(\mathbf{A}^{\ell})_{ii}$ is the sum of products of edge weights along length-ℓ closed walks that start and end at a node i. In the remainder of this chapter, we consider motifs in a network to be weakly connected walks graphs. We focus on small motifs with lengths $L \leq 8$. For a given motif structure — e.g., closed walks of length ℓ — we use the word *count* to refer to the number of motifs with this structure in unweighted networks or the sum of edge-weight products of motifs with this structure in weighted networks.

4.4.3 Walk graphs and subgraphs

Several researchers have used a walk-based approach to identify motifs in dynamical processes on networks [31, 247, 427, 538]. Importantly, this notion of motifs is conceptually different from the usual notion of motifs in biochemistry and network science, where many researchers consider motifs to be small connected subgraphs that occur more frequently in a network than one would expect from comparison with a suitable null model [17, 370, 486, 542]. Walk graphs and connected subgraphs depend only on a network's adjacency matrix. The two concepts thus provide alternative frameworks for characterizing a network's structure. In defining a motif to be a walk graph, we consider a motif to be a "small process" or a "small unit of movement" that occurs on a small connected subgraph in a network, because the notion of a walk implies that an agent or a signal visits nodes in a prescribed sequence, one node after the other. Previous work has demonstrated that these "small process" motifs provide an approach for connecting dynamics on networks to network structure [31, 247, 427, 538].

The counts of connected subgraphs in a network and the counts of occurrences of walk graphs are related. To illustrate this relationship, we consider two small example networks. Milo et al. (2002) identified small connected subgraphs that tend to be more frequent in real networks than in random networks drawn from several random-graph models [370]. They called one of these connected subgraphs the 3-node feedforward loop and another the 3-node feedback loop. We choose the 3-node feedforward loop and 3-node feedback loop as our example networks. In Fig. 4.4, we show all length-L walk graphs that can occur on a 3-node feedforward loop for

$L \leq 3$. In Fig. 4.5, we show all length-L walk graphs that can occur on a 3-node feedback loop for $L \leq 3$.

When comparing Figs. 4.4 and 4.5, we observe that some walk graphs can occur on the 3-node feedforward loop but not on the 3-node feedback loop, and vice versa. The difference in the sets of walk graphs for the 3-node feedforward loop and the 3-node feedback loop illustrates that the structure of a network constrains the structures of walk graphs that can occur on the network. The 3-node feedforward loop is an acyclic network with a maximum trail length of 2. Because the feedforward loop is a directed acyclic network, its associated walk graphs are also acyclic. Walk graphs on a directed acyclic network with a maximum trail length of 2 cannot have edges with $\ell_W(e) > 2$ and cannot have paths with edge lengths that sum to 3 or more.

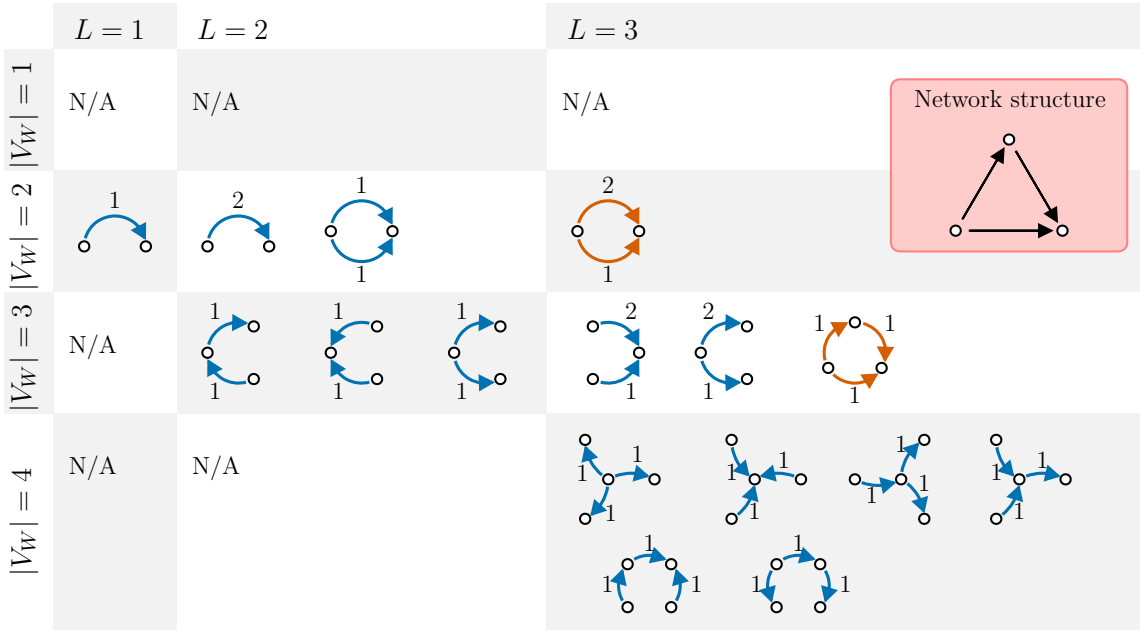


Figure 4.4: Walk graphs that can occur on a 3-node feedforward loop. We sort walk graphs according to their length L and their number $|V_W|$ of nodes. Numerical edge labels on walk graphs indicate the length of an edge. Walk graphs with orange edges are those that occur on the 3-node feedforward loop but not on the 3-node feedback loop. The pink inset shows the structure of the 3-node feedforward loop.

The structure of the 3-node feedback loop leads to other constraints on the structures of associated walk graphs. For example, a walk graph that can occur on the 3-node feedforward loop but not on the 3-node feedback loop is the circular walk graph with $|V_W| = 2$ and $L = 3$ in Fig. 4.4. It consists of a length-1 edge and a length-2 edge that share their start node and their end node. This walk graph

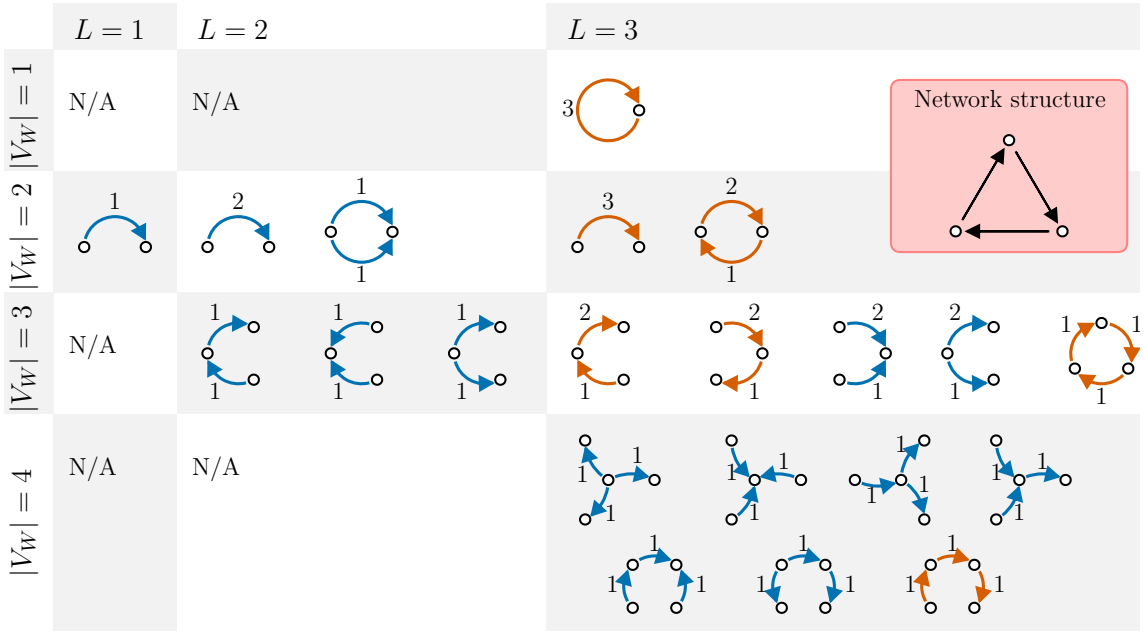


Figure 4.5: Walk graphs that can occur on a 3-node feedback loop. Walk graphs with orange edges are those that occur on the 3-node feedback loop but not on the 3-node feedforward loop. The pink inset shows the structure of the 3-node feedback loop.

cannot occur on the 3-node feedback loop, because for every pair (i, j) of nodes in the 3-node feedback loop, there is either a length-1 walk or a length-2 walk from i to j, but not both.

4.5 Covariance-inducing motifs of the Ornstein–Uhlenbeck process

In this section, we identify the motifs that contribute to covariance between two nodes i and j. Subsequently, we explore the magnitude of contributions of different motifs to covariance and identify the motifs that contribute the most to covariance of two nodes i and j.

4.5.1 Structure of covariance-inducing motifs

We can express the elements σ_{ij} of the covariance matrix $\mathbf{\Sigma}$ from Eq. (4.7) as a weighted sum of counts of covariance-inducing motifs (CMs):

$$\sigma_{ij} = \frac{1}{2} + \frac{1}{2} \sum_{\ell=0}^{\infty} \sum_{\substack{r=0, \\ \ell+r>0}}^{\infty} \omega_{\mathrm{CM}}(\epsilon, \ell, r) \nu_{ij}^{\mathrm{CM}}(\mathbf{A}, \ell, r)\,, \tag{4.11}$$

where $\nu_{ij}^{\mathrm{CM}}(\mathbf{A}, \ell, r)$ is the (i,j)th element of the matrix $\left(\mathbf{A}^{\mathbf{T}}\right)^{\ell} \mathbf{A}^{r}$ and

$$\omega_{\mathrm{CM}}(\epsilon, \ell, r) = \left(\frac{\epsilon}{2}\right)^{\ell+r} \binom{\ell + r}{\ell} \tag{4.12}$$

is the *weighting factor* for a contribution of a CM. We can interpret $\nu_{ij}^{\mathrm{CM}}(\mathbf{A}, \ell, r)$ as the count of motifs that consist of two walks that start at an arbitrary source node and end at i and j, respectively. The walk from the source node to i has length ℓ. The walk from the source node to j has length r. The parameters ℓ and r determine the lengths of all walks in a CM. We refer to a tuple that specifies the lengths of all walks in a motif as its *configuration*. A CM with configuration (ℓ, r) has length $L_{\mathrm{CM}} = \ell + r$. In Fig. 4.6, we show the structures and associated configurations of covariance-inducing, variance-inducing, and entropy-inducing motifs. A variance-inducing motif (VM) is a CM for which the nodes $i = j$. Other noteworthy special cases of CMs occur when either r or ℓ is equal to 0. A CM with $r = 0$ is a length-ℓ walk from node i to node j. A CM with $\ell = 0$ is a length-r walk from node j to node i.

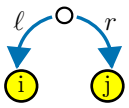
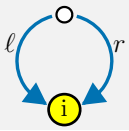
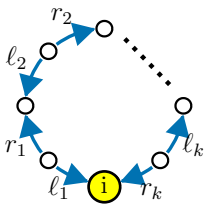
Motif class	Walk Diagram	Configuration
CM		(ℓ, r)
VM		(ℓ, r)
EM		$(\boldsymbol{\ell}, \mathbf{r})$ with $\boldsymbol{\ell} = (\ell_1, \ell_2, \ldots,$ $\mathbf{r} = (r_1, r_2, \ldots,$

Figure 4.6: Structures and configurations of covariance-inducing motifs (CMs), variance-inducing motifs (VMs), and entropy-inducing motifs (EMs).

The CM and VM structures are consistent with properties of covariation in systems of coupled random variables. A covariance σ_{ij} measures the joint "variability" of two random variables x_i and x_j [449, p. 138], where one takes variability to indicate a variable's deviation from its mean. This joint variability of x_i and x_j can arise from several causes:

1. Variability in x_i induces variability in x_j if there is a path from node i to node j.

2. Variability in x_j induces variability in x_i if there is a path from node j to node i.

3. Variability in a third variable x_s induces variability in both x_i and x_j if there are paths from s to i and from s to j.

4.5.2 Weighting factors for covariance-inducing motifs

We now examine the weighting factors $\omega_{\mathrm{CM}}(\epsilon, \ell, r)$ from Eq. (4.12). For any CM configuration, the weighting factor is positive. Therefore, all CMs have a positive contribution to covariances for the multivariate OU process. The value of a weighting factor depends on a CM's configuration (ℓ, r). Swapping ℓ and r leads to a CM with an identical weighting factor:

$$\omega_{\mathrm{CM}}(\epsilon, \ell, r) = \omega_{\mathrm{CM}}(\epsilon, r, \ell)\,.$$

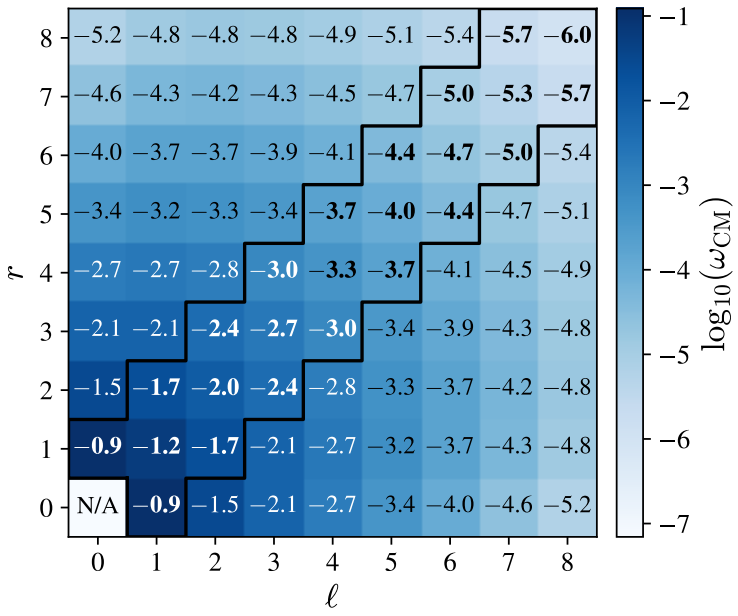


Figure 4.7: Weighting factors ω_{CM} for CMs with configurations (ℓ, r) with $\epsilon = 0.49$. The length $L_{\mathrm{CM}} = \ell + r$ increases along the diagonal from bottom left to top right. We indicate the CM configurations with largest weight for each value of L_{CM} by bold labels and delineate them with solid black line segments.

In Fig. 4.7, we show the weights ω_{CM}. The length L_{CM} increases along the diagonal from the bottom left to the top right. We indicate the CM configurations

with largest weighting factor for each value of L_{CM} by bold labels and delineate them with solid black line segments. For even L_{CM}, the weighting factor is maximal when $\ell = r = L_{\mathrm{CM}}/2$. For odd L_{CM}, the weighting factor is maximal when $\ell = (L_{\mathrm{CM}} \pm 1)/2$. Comparing the weights of CMs of different lengths, we find that short CMs (bottom left) tend to contribute more to covariances than long CMs. These observations are consistent with the notion that covariances and correlations should decay with the distance that a signal travels [36]. Because $\omega_{\mathrm{CM}}(\epsilon, r, \ell)$ is proportional to $\epsilon^{L_{\mathrm{CM}}}$ (see Eq. (4.12)), the weighting factors decrease sharply with increasing L_{CM} when ϵ is small (i.e., when signals decay over a short range).

4.5.3 Origins of variance in the Ornstein–Uhlenbeck process

By focusing on the diagonal elements of the covariance matrix $\mathbf{\Sigma}$, we can gain insights into the origins of variance in the multivariate OU process. We first note that a VM with $\ell, r > 0$ is a walk graph of two walks that share both their starting node and their destination. There are many such motifs when a network has many redundant paths or trails that connect the same pairs of nodes. A VM with $\ell = 0$ or $r = 0$ is a cycle that connects a target node to itself. Each diagonal element of $\mathbf{\Sigma}$ in Eq. (4.11) is the sum of a term that is independent of $\mathbf{A}$ and the counts of possibly many VMs. We interpret this equation as follows. Variance in state variable x_i arises via two mechanisms:

1. The term dW_t in Eq. (4.6) models intrinsic noise in each state variable x_i. It leads to the term $\mathbf{I}/2$ in Eq. (4.11). This term is independent of a network's structure.

2. The variance of a state variable $x_{i,t}$ exceeds its base value $\frac{1}{2}$ (i.e., $\frac{\varsigma^2}{2\theta}$, when $\theta, \varsigma \neq 1$) when cycles and redundant paths or trails in a network reinforce the input that the corresponding node i receives from other nodes or itself, because the presence of cycles and redundant paths or trails results in VMs in a network.

4.6 Entropy-inducing motifs of the Ornstein–Uhlenbeck process

In this section, we characterize the motifs that contribute to differential entropy in a multivariate OU process. In Section 4.6.1, we introduce entropy-inducing motifs

(EMs) as walk graphs that contribute to differential entropy. We study the weighting factor of an EM's contribution in Section 4.6.2. In Section 4.6.3, we link the contribution of EMs, which are walk graphs, to the contributions of connected subgraphs. For the 3-node feedforward loop and the 3-node feedback loop, we compare their contributions to differential entropy.

4.6.1 Structure of entropy-inducing motifs

Using $\mathbb{N}_0^*$ to denote the set of arbitrary-length tuples of non-negative integers and $|\cdot|$ to denote the length of a tuple, we rewrite Eq. (4.9) as a weighted sum of EMs:

$$h_{[L]}(\epsilon, \mathbf{A}) = \frac{N}{2}\log(\pi e) + \sum_{i=1}^{n} \sum_{\boldsymbol{\ell}\in\mathbb{N}_0^*} \sum_{\substack{\mathbf{r}\in\mathbb{N}_0^{|\boldsymbol{\ell}|}, \\ \sum\boldsymbol{\ell}+\sum\mathbf{r}=L, \\ \min_j(\ell_j+r_j)\geq 1}} \omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r})\nu_i^{\mathrm{EM}}(\mathbf{A}, \boldsymbol{\ell}, \mathbf{r}),$$

where $\nu_i^{\mathrm{EM}}(\mathbf{A}, \boldsymbol{\ell}, \mathbf{r})$ is the ith diagonal element of the matrix in Eq. (4.10) and

$$\omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r}) = \frac{(-1)^{|\boldsymbol{\ell}|+1}}{2|\boldsymbol{\ell}|}\left(\frac{\epsilon}{2}\right)^{\sum\boldsymbol{\ell}+\sum\mathbf{r}}\left(\prod_{j=1}^{|\boldsymbol{\ell}|}\binom{l_j+r_j}{l_j}\right).$$

From Eq. (4.10), we see that an EM is a concatenation of CMs. Only the diagonal elements of Eq. (4.10) contribute to $h_{[L]}(\epsilon, \mathbf{A})$. Therefore, the motifs that contribute to differential entropy are circular concatenations of CMs. We show the structure and configuration of an EM in Fig. 4.6.

4.6.2 Weighting factors for entropy-inducing motifs

The weighting factors for an EM are proportional to the product of the weighting factors of the CMs that form the EM. The order of the CMs in an EM does not change ω_{EM}. Moreover, if we apply the same permutation to the elements of $\boldsymbol{\ell}$ and the elements of $\mathbf{r}$, where we denote the permuted tuples $\boldsymbol{\ell}'$ and $\mathbf{r}'$, then $\omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r}) = \omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}', \mathbf{r}')$.

To identify the EMs that have the greatest positive or negative contribution to differential entropy, we focus on EMs that consist of CMs that have the maximal ω_{CM} for a given length L_{CM}. The configurations of these EMs have

$$r_i = \begin{cases} \ell_i - 1, & \text{if } \ell_i + r_i \text{ odd}, \\ \ell_i, & \text{if } \ell_i + r_i \text{ even}, \end{cases}$$

for $i \in \{1, 2, \ldots, |\boldsymbol{\ell}|\}$. For each integer partition[11] of L_{EM}, we can construct exactly one EM configuration with maximal CMs.

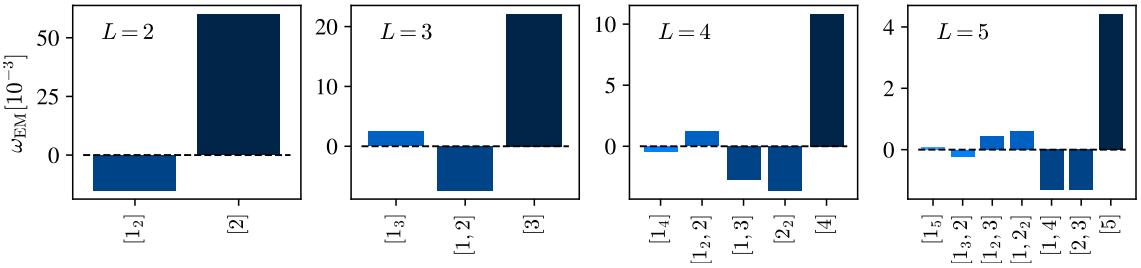


Figure 4.8: Weighting factors ω_{EM} for EMs with length L_{EM} and configurations $(\boldsymbol{\ell}, \mathbf{r})$. Labels on the horizontal axis indicate integer partitions of L_{EM}. Bar heights indicate $\omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r})$, where $\ell_i + r_i$ equals the ith number of the corresponding integer partitions and $r_i \in \{\ell_i, \ell_i - 1\}$. Bar colors indicate the absolute value of $\omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r})$. We use $\epsilon = 0.49$.

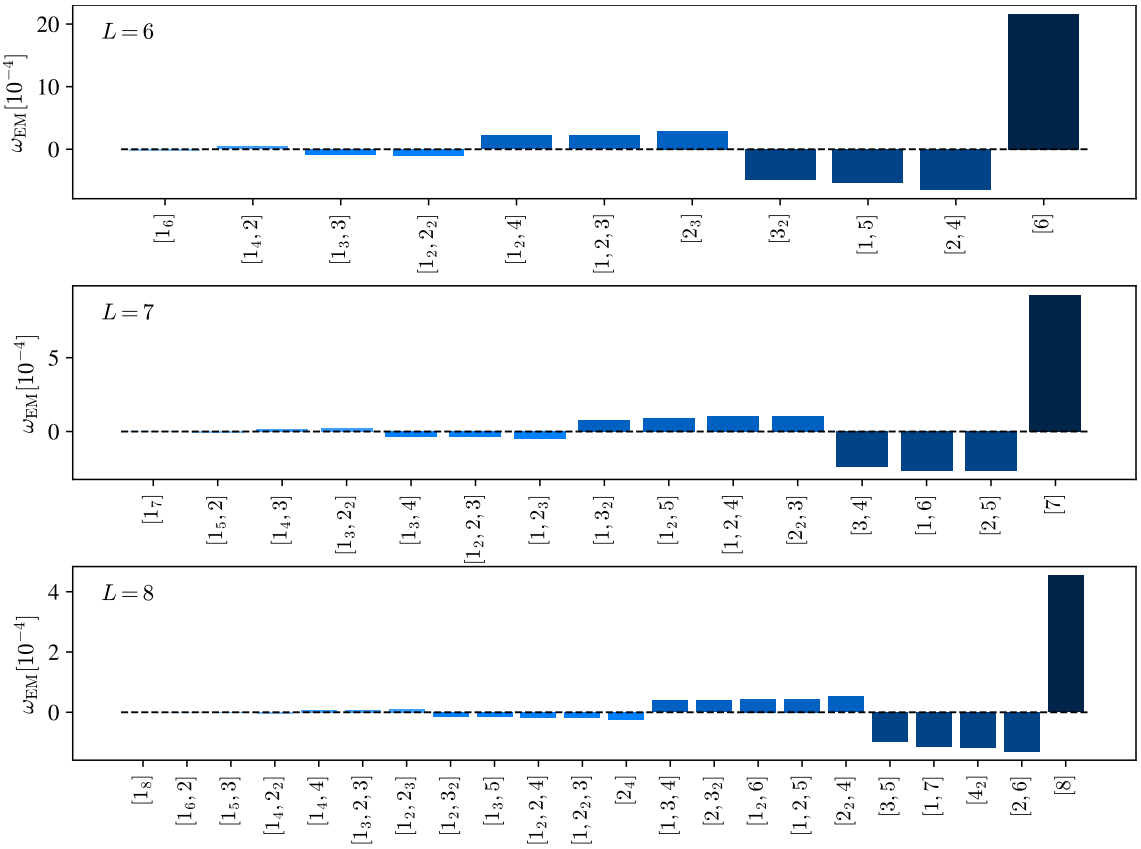


Figure 4.9: Weighting factors ω_{EM} for EMs with length L_{EM} and configurations $(\boldsymbol{\ell}, \mathbf{r})$. See the caption of Fig. 4.8 for further details.

In Figs. 4.8 and 4.9, we show the weighting factors of maximal EM configurations for all integer partitions of L_{EM}. We consider different values of L_{EM} in different panels. The absolute value of ω_{EM} decreases sharply as either the length L_{EM} or

[11] An integer partition of an integer n is a set $\{n_1, n_2, \ldots, n_k\}$ of positive integers such that $n = \sum_{i=1}^{k} n_i$ [21, p. 7].

the number of elements in an integer partition increases. Weighting factors that correspond to integer partitions with an odd number of elements are positive, and weighting factors that correspond to integer partitions with an even number of elements are negative. Our observations of ω_{EM} for EM lengths up to $L_{\mathrm{EM}} = 8$ suggest that, for a given L_{EM} and a given number of elements in an integer partition of L_{EM}, the weighting factor ω_{EM} has largest absolute value when many elements of the corresponding integer partition are equal to 2.

We now investigate the EMs that contribute with the largest weighting factor to differential entropy in a network with a multivariate OU process. In Fig. 4.10, we show the EMs with the largest positive contributions to differential entropy. We observe that short motifs tend to have the largest contributions to differential entropy. The first three EMs have lengths $L_{\mathrm{EM}} \leq 2$ and can occur on networks with a single edge. For any given motif length, a motif that is a pair of redundant walks of roughly equal length contributes more to differential entropy than a cyclic motif. Notably, the contribution of the EM with $((3),(0))$ is even less than that of the longer non-cyclic EM $((2),(2))$. We also observe that, for a given motif length, a motif that consists of one or two walks contributes more to differential entropy than a motif that consists of three walks.

The nine EMs that have the largest contributions are also VMs. This observation suggests that, under the assumption of short-range signal decay (see Section 4.3.3), the differential entropy of a multivariate OU process is closely related to the variances of its state variables $x_{i,t}$. It follows that the important mechanisms by which differential entropy arises in a network with OU dynamics are the same mechanisms that lead to variance in state variables. These mechanisms are (1) intrinsic noise, (2) signal propagation in non-cyclic motifs, and (3) signal propagation in cyclic motifs.

4.6.3 Entropy contribution of a 3-node feedforward loop and 3-node feedback loops

We can use the weighting factors in Fig. 4.10 to understand the differences in the contributions of different connected subgraphs to differential entropy. As example structures for connected subgraphs, we consider the 3-node feedforward loop and the 3-node feedback loop (see Figs. 4.4 and 4.5) which have respective adjacency matrices

$$\mathbf{A}_{\mathrm{FF}} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}_{\mathrm{FB}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

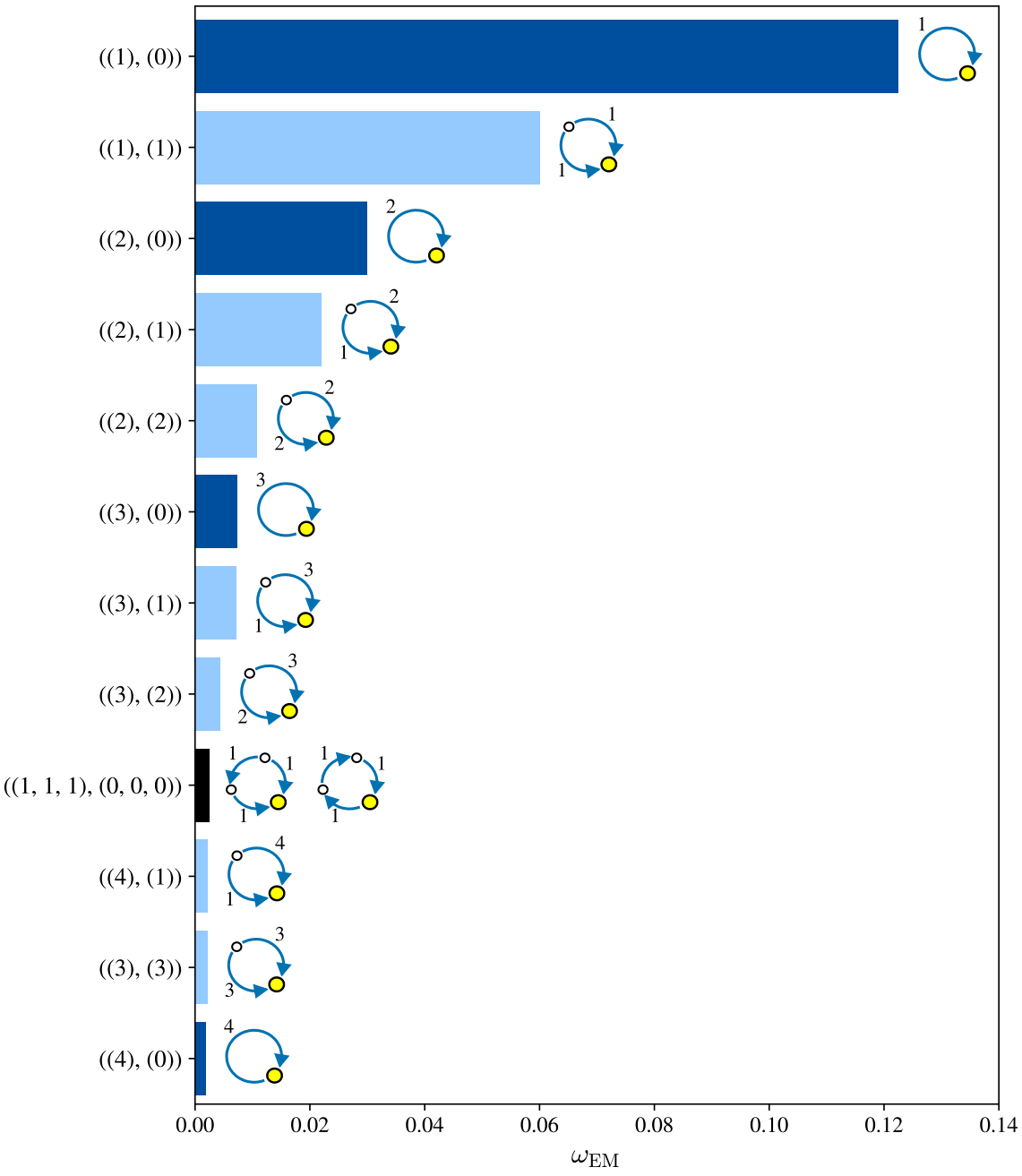


Figure 4.10: Largest weighting factors ω_{EM}. Bar labels (left) consists of EM configurations. To the right of each bar, we show corresponding walk-graph structures. Bars are light blue when the corresponding motifs are acyclic, dark blue when the corresponding motifs are cyclic, and black when the set of corresponding motifs includes cyclic and acyclic motifs. We use $\epsilon = 0.49$ in our computations of ω_{EM}.

To ensure short-range signal decay (see Section 4.3.3), we set $\epsilon = 0.49/\sqrt{3}$. We can obtain the contribution of a connected subgraph with adjacency matrix $\mathbf{A}$ to differential entropy by computing $h_{[\infty]}(\epsilon, \mathbf{A})$ (see Eq. (4.8)). For the 3-node feed-

forward loop and the 3-node feedback loop, we obtain $h_{[\infty]}(\epsilon, \mathbf{A}_{\mathrm{FF}}) \approx 3.249$ and $h_{[\infty]}(\epsilon, \mathbf{A}_{\mathrm{FB}}) \approx 3.251$, respectively. The 3-node feedback loop thus contributes more to the differential entropy of the multivariate OU process than the 3-node feedforward loop, although several EMs that occur on the feedfoward loop have larger weighting factors than the EMs that occur on the feedback loop. How is this possible? The EMs in Fig. 4.10 can help us answer this question. In a subgraph with adjacency matrix $\mathbf{A}$, the product $\nu_i^{\mathrm{EM}}(\mathbf{A}, \boldsymbol{\ell}, \mathbf{r})\omega_{\mathrm{EM}}(\epsilon, \boldsymbol{\ell}, \mathbf{r})$ is the entropy contribution of all EMs with configuration $(\boldsymbol{\ell}, \mathbf{r})$. For the 3-node feedforward loop and the 3-node feedback loop, we sum these products over the EM configurations in Fig. 4.10. For the 3-node feedforward loop, we obtain

$$3\,\omega_{\mathrm{EM}}((1),(1)) + 1\,\omega_{\mathrm{EM}}((2),(1)) + 1\,\omega_{\mathrm{EM}}((2),(2)) + 1\,\omega_{\mathrm{EM}}((1,1,1),(0,0,0)) \approx 0.215\,,$$

where, for convenience, we omit the dependence of weighting factors on ϵ. For the 3-node feedback loop, we obtain

$$3\,\omega_{\mathrm{EM}}((1),(1)) + 3\,\omega_{\mathrm{EM}}((3),(0)) + 3\,\omega_{\mathrm{EM}}((1,1,1),(0,0,0)) + 3\,\omega_{\mathrm{EM}}((3),(3)) \approx 0.216\,.$$

We observe that, even when only considering the contributions of the EMs that contribute most to differential entropy, we find that the 3-node feedback loop contributes more to differential entropy than the 3-node feedforward loop. The EMs that occur on the feedback loop tend to have smaller weighting factors than the EMs that occur on the feedforward loop, but the counts of EMs on the feedback loop tend to be higher than the counts of EMs on the feedforward loop. By "strength in numbers", the EMs on the feedback loop lead to a greater differential entropy contribution than the EMs on the feedforward loop.

4.7 Conclusions and discussion

We summarize our findings and compare them to previous work on the links between motifs and global network properties for a network with a multivariate Ornstein–Uhlenbeck process. We close with an outlook on possible generalizations and applications of our results.

4.7.1 Summary

Empirical and theoretical studies of networks in biology have used different notions of motifs to identify important functional units in networked systems [17, 247, 512]. For many of the proposed functional units, the mechanisms by which they contribute

to a system's function are unknown and frameworks to measure the relative importance of different motifs are not available. We examined a multivariate Ornstein–Uhlenbeck process as a simple model of noisy information propagation in a network. Using small, weakly connected "walk graphs" as motifs of dynamics on networks, we studied the relative contributions and mechanisms by which motifs contribute to variance, covariance, and differential entropy in this system. We found that, under the assumption of short-range signal decay, the mechanisms that lead to variance are also the important mechanisms that contribute to the system's differential entropy. These mechanisms are (1) node-intrinsic noise generation, (2) signal propagation along redundant paths or trails, and (3) signal propagation along cycles. Our comparison of the contributions of different motifs demonstrated that motifs that consist of two redundant walks contribute more than closed walks to the system's differential entropy. We also investigated the contributions of small connected subgraphs, which are a common notion of motifs in biochemistry and network science [17, 370, 486, 542] to differential entropy. Considering the 3-node feedforward loop and the 3-node feedback loop as examples of small connected subgraphs, we found that many entropy-inducing motifs with small weighting factors occur on the 3-node feedback loop and few entropy-inducing motifs with large weighting factors occur on the 3-node feedforward loop. By "strength in numbers", the motifs on the feedback loop lead to a greater combined contribution to differential entropy than the motifs on the feedforward loop.

4.7.2 An extended information-theoretic framework

In our work, we have used several ideas from Ref. [31], in which the authors expressed the covariance matrix of an Ornstein–Uhlenbeck process on a network as a function of the network's adjacency matrix and derived length-2 and length-3 walk graphs that contribute to neural complexity. We have expanded their framework in several ways:

1. We formalized the concept of walk graphs and discussed the differences between walk graphs and connected subgraphs.

2. We considered variance, covariance, and differential entropy of the multivariate Ornstein–Uhlenbeck process as information-theoretic properties of interest.

3. We showed that any change in the noise amplitude or the mean reversion rate affects all motif weights equally. All of our results regarding the rela-

tive importance of different motifs thus hold for any positive values of these parameters.

4. We computed explicit forms of the conditions that the coupling parameter ϵ and the adjacency matrix $\mathbf{A}$ need to fulfill for the system to have short-range signal decay.

5. We derived the structures and weighting factors of motifs that contribute to variance, covariance, and differential entropy in the multivariate Ornstein–Uhlenbeck process for arbitrary numbers of edges and showed numerical results for motifs with up to 8 edges.

Our results for motifs with lengths greater than 3 enabled us to discover a systematic difference between the contributions of cycles and motifs that include redundant paths or trails to the differential entropy of the multivariate Ornstein–Uhlenbeck process.

4.7.3 Disorder versus coherence

The foundation of the information-theoretic framework that we used and extended in this chapter is the analytical relationship between the covariance matrix and the adjacency matrix of the multivariate Ornstein–Uhlenbeck process (see Eq. (4.7)). Several researchers have found similar relationships for systems of coupled excitable units [411, 427, 538] and used these relationships to study the mechanisms by which network structure can contribute to synchrony or global coherence of excitable units [247, 411, 427, 538]. The motifs that they identified as contributors to synchrony or coherence are members of the class of motifs that we identified as covariance-inducing motifs.

Synchrony and global coherence describe states of order in a system. By contrast, entropy is a measure of a system's disorder. The finding of a class of motifs that contributes to global coherence and a class of motifs that provides positive and negative contributions to differential entropy suggests that different "network effects" lead to an increase of order or increase of disorder in a dynamical system. Which of these effects is dominant? In Section 4.3.4, we observed that the mean differential entropy of $G(N, m)$ graphs increases with the number of edges. This trend suggests that, when a system has short-range signal decay, adding an edge to a graph uniformly at random is likely to increase its differential entropy and that the network effects that increase system disorder are stronger than those that increase order. However, the large variability of differential entropy for $G(N, m)$ graphs with

the same number of edges suggests that it may be possible to observe positive or negative changes in system disorder when adding an edge to different graphs. The balance of network effects that increase order and disorder in a system is thus very sensitive to network structure. The discovery of network structures that are optimal for promoting order or disorder of units in a coupled system is an interesting avenue for future research.

4.7.4 Outlook

In this chapter, we have shown that, under the assumption of short-range signal decay, one can express the variance, covariance, and differential entropy of a multivariate Ornstein–Uhlenbeck process as a sum of counts of walk graphs. We anticipate that walk graphs that contribute to variance, covariance, and differential entropy in other processes on networks may have the same structure as the walk graphs that we identified for the multivariate Ornstein–Uhlenbeck process. Motivations for this anticipation include the following:

1. The Ornstein–Uhlenbeck process is the linear-response solution for various nonlinear models of excitable dynamics [210, 538].

2. The basis for the derivations that we presented in this chapter is the analytical relationship between adjacency and covariance matrices (see Eq. (4.7)). Similar relationships also hold for several models of excitable dynamics [210, 427, 538]. Some of these relationships do not require an assumption of weak coupling or signal decay.

3. Several previous studies of cross-correlations in systems with different dynamics have identified motifs that are examples of covariance-inducing motifs [247, 411, 427]. A model-independent reasoning about the difference between correlation and causation identifies the same class of motifs as those that contribute to covariance and correlation in a network in which information can propagate along edges. The structure of covariance-inducing motifs seems to reflect a fundamental aspect of information propagation in networks. This aspect is independent of the choice of a specific dynamical process. It may be that the structure of entropy-inducing motifs is also a fundamental aspect of information propagation in networks. The study of differential entropy of stochastic processes other than the multivariate Ornstein–Uhlenbeck process will be necessary to confirm or reject this hypothesis.

Our work furthers understanding of the connections between local network structure and information propagation. Barnett et al. (2009) connected walk graphs with neural complexity of networks [31]. Our work provides an important step towards extending their approach to many other network properties that are based on differential entropy. Examples of such properties include matching complexity [534], neural redundancy [535], and neural degeneracy [535]. These and other network properties that are based on differential entropy are relevant to the study of neuronal networks [31, 350] and beyond [266, 345, 444, 501, 616].

Chapter 5

Log-minor distributions and an application to estimating mean subsystem entropy

A common task in physics, information theory, and other fields is the analysis of properties of subsystems of a given system. Given the covariance matrix M of a system of n coupled variables, the covariance matrices of the subsystems are principal submatrices of M. The rapid growth with n of the set of principal submatrices makes it impractical to exhaustively study each submatrix for even modestly-sized systems. It is therefore of great interest to derive methods for approximating the distributions of important submatrix properties for a given matrix.

Motivated by the importance of differential entropy as a systemic measure of disorder, we study the distribution of log-determinants of principal $k \times k$ submatrices when the covariance matrix has bounded condition number. We derive upper bounds for the right tail and the variance of the distribution of minors, and we use these in turn to derive upper bounds on the standard error of the sample mean of subsystem entropy. Our results demonstrate that, despite the rapid growth of the set of subsystems with n, the number of samples that are needed to bound the sampling error is asymptotically independent of n. Instead, it is sufficient to increase the number of samples in linear proportion to k to achieve a desired sampling accuracy.

Symbol	Meaning
$\mathbb{R}$	the real numbers
$h(M)$	the differential entropy of a normal distribution with scale matrix M
M	a positive definite square matrix
D	a positive definite diagonal square matrix
Q	an orthogonal matrix
$\lambda_i(M)$	the ith largest eigenvalue of M
$\kappa(M)$	the condition number of M
$\tilde{\kappa}$	upper bound on a matrix' condition number
I	an index set
n	size of a square matrix
k	size of a princial submatrix
$\wedge_{n,k}$	minimum of $\{k, n-k\}$
$\mathcal{A}_k(M)$	the set of $k \times k$ principal submatrices of M
$A_k(M)$	a uniformly-random element of $\mathcal{A}_k(M)$
$\hat{A}_k(M)$	the top-left $k \times k$ principal submatrix of M
$Y_k(M)$	the log-determinant of $A_k(M)$
Pr	the probability of an event
$\mathbb{E}$	the expectation of a random variable
μ	a probability measure
$\mu_{M,k}$	probability measure of $Y_k(M)$
f	a function
$f_{\alpha,k}(s)$	a function of the $k \times k$ top-left principal submatrix of a matrix sM
$\mathcal{S}_n$	the set of permutations of n elements
s	a permutation
τ	a transposition
Π	kernel of the "random-transposition walk"
$\overline{\lambda}$	the spectral gap of the random transposition walk
S_Y	a sample mean of $Y_k(M)$
S_h	a sample mean of subsystem entropy
q	a number of samples
σ	the standard error of a sample mean
$\mathcal{E}$	the relative error of a sample mean
c_v	the coefficient of variation of a sample mean

Table 5.1: List of symbols for Chapter 5.

5.1 Introduction

In many fields of study, researchers use matrices to represent systems of interest. Statisticians and data scientists represent large tabular data sets as matrices [494]. In network science, it is common to use adjacency matrices to represent the structure of a network [398]. In dynamical systems, researchers use Jacobian matrices in the study of the linearized dynamics of a system of coupled variables [514]. For networks, dynamical systems, statistical analysis of large data sets, and other applications, it can be insightful (and even necessary) to examine their components (as subnetworks, subsystems, reduced data sets, and so on). Several researchers have used subsystem properties to characterize robustness and other salient properties of dynamical systems [136, 323, 324, 444, 533, 535]. Network scientists count and analyze motifs and other subgraphs in networks to characterize a network's structure [370, 486]. Several prominent tools in data science are based on linear sketching, an approach to data dimensionality reduction whereby one obtains a reduced data set via matrix multiplication [185, 587] or as a linear combination of submatrices [100, 303, 462]. An example of such a tool for dimensionality reduction is principal component analysis [326].

The various applications of submatrices motivate the mathematical study of their properties. In this chapter, we study the distribution of log-determinants of principal submatrices of a positive definite matrix and show that our results lead to controllable sampling guarantees for computing the mean differential entropy of subsystems for a dynamical system. Researchers have studied the differential entropy of subsystems in areas such as physics [413, 477], biology [323, 324], neuroscience [136, 533, 535], computer science [444], and coding theory [295]. For example, Tononi et al. (1999) computed a measure of network redundancy from the mean differential entropy of its subsystems of fixed size [535]. Teschendorff et al. (2014) [528] used differential entropy to define a measure of network robustness for protein-interaction networks .

For several symmetric multivariate distributions, estimates of differential entropy are affine functions of the log-determinant of a system's covariance matrix. Examples include the multivariate normal distribution [6], the multivariate t distribution [212, 384], and the multivariate Cauchy distribution [384]. For the n-variate normal distribution with covariance matrix M, for example, the differential entropy is [6]

$$h(M) = \frac{1}{2}\log(\det M) + \frac{n}{2}\left(1 + \log(2\pi)\right) . \tag{5.1}$$

The logarithm of the covariance matrix is thus sufficient to approximate the differential entropy of several multivariate distributions.

The principal submatrices of M are covariance matrices of subsystems that correspond to subsets of coupled variables. One can compute the differential entropy of a subsystem by computing h in Eq. (5.1) for a principal submatrix of M. A system of n coupled variables possesses $\binom{n}{k} \approx \frac{n^k}{k!}$ subsystems of k variables; each of these subsystems corresponds to one of the $\binom{n}{k}$ principal $k \times k$ submatrices of M. The exact computation of the distribution of differential subsystem entropy or its moments thus requires one to compute $O(n^k)$ distinct determinants, an infeasible task for large n and k. This task can be computationally prohibitive even for modestly-sized systems. To our knowledge, the largest system for which researchers have exactly computed the differential entropy of subsystems is a synthetic network with $n = 12$ variables and subsystems with $k \leq 12$ variables [535].

To address this problem, we study the distribution of log-determinants of principal $k \times k$ submatrices. We refer to these log-determinants as *log-minors* of *size* k. As we noted above, these log-determinants are sufficient to determine the subsystem entropy for many important multivariate distributions. Knowledge of the properties of this distribution thus enables the derivation of bounds on the sampling error when estimating subsystem entropy in many applications. We show that, given a bound on the condition number of M, the standard error of a sample mean of differential entropy is independent of n and sublinear in k, implying that one needs a sublinear number of samples in k to ensure a desired accuracy.

This chapter proceeds as follows. In Section 5.2, we introduce some notation that we use throughout this chapter. In Section 5.3, we give several upper bounds on the tail and variance of the distribution of log-minors of a positive-definite matrix with bounded condition number. We present proofs for these bounds in Section 5.4 and show numerical examples in Section 5.5. In Section 5.6, we apply our theorems to provide probabilistic guarantees on the sample mean and relative error, and we discuss implications for the design of practical schemes for estimating mean subsystem entropy. We conclude and discuss possible extensions in Section 5.7.

5.2 Notation

Let $M \in \mathbb{R}^{n \times n}$ be a positive-definite matrix. Let $\lambda_1(M) \geq \lambda_2(M) \cdots \geq \lambda_n(M) \geq 0$ be the eigenvalues of M. Because M is positive definite, it is also nonsingular; its condition number is $\kappa(M) = \lambda_1(M)/\lambda_n(M)$. For a given index set $I \in [n]^k := \{1, \ldots, n\}^k$, the matrix $M_I := [M_{i,j}]_{i,j \in I}$ is the corresponding principal submatrix of

M. For any fixed $k \leq n$, let $\mathcal{A}_k(M)$ denote the set of all such $k \times k$ submatrices of M, and let $A_k(M)$ denote a uniformly-random element of this set. We define a random variable $Y_k(M) := \log(\det A_k(M))$ and denote its empirical distribution by $\mu_{M,k}$. For convenience, we define $\wedge_{n,k} := \min\{k, n-k\}$.

5.3 Bounds on the distribution of log-minors

In this section, we state bounds on the distribution $\mu_{M,k}$ for a positive-definite matrix M with bounded condition number. We give upper bounds for the distribution's support, variance, and right tail. We also show that we can improve these bounds if M is diagonal.

Theorem 1 (Tail and variance bound for log-minors of a positive-definite matrix)**.** *Let M be a positive-definite $n \times n$ matrix with condition number $\kappa(M) \leq \tilde{\kappa}$. For every $r \geq 0$, we have*

$$\Pr\left(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r\right) \leq 6 \exp\left(-\frac{r}{\log \tilde{\kappa}} \sqrt{\frac{n}{k(n-k)}}\right). \tag{5.2}$$

Furthermore, the variance of $Y_k(M)$ satisfies

$$\operatorname{var}(Y_k(M)) \leq 12 \left(\frac{k(n-k)}{n}\right) (\log \tilde{\kappa})^2. \tag{5.3}$$

Remark 1. The tail bound in Eq. (5.2) does not guarantee that $Y_k(M)$ concentrates[1] on $\mathcal{A}_k(M)$. This is because the bound in Eq. (5.2) is increasing with respect to k and asymptotically constant with respect to n. Indeed, for the bound to approach 0 for a sequence $\{M_i\}$ of matrices, it is both necessary and sufficient that $\lim_{i\to\infty} \sqrt{k_i} \log \kappa_i(M) \to 0$. Because k cannot be smaller than 1, this condition requires the condition number $\kappa_i(M)$ to approach 1. The condition $\lim_{i\to\infty} \kappa_i(M) \to 1$ severely constrains the sequence $\{M_i\}$. In that limit, all eigenvalues of M are equal to each other and all log-minors are equal to $k \log \lambda_1$.

Theorem 2 (Support and variance bound for log-minors of a positive-definite matrix)**.** *Let M be a positive-definite $n \times n$ matrix with condition number $\kappa(M) \leq \tilde{\kappa}$. For any k, the random variable $Y_k(M)$ and its distribution $\mu_{M,k}$ satisfy the following properties:*

[1] Ledoux defined concentration of measure in Ref. [311] (on page 3) as follows. Let (X, d) be a metric space with probability measure μ on Borel sets of (X, d). The *concentration function* is defined as $\alpha_{(X,d,\mu)} := \sup\{1 - \mu(A_r); A \subset X, \mu(A) \geq \frac{1}{2}\}$, where $r > 0$ and $A_r := \{x \in X; d(x, A) < r\}$ is the open r-neighborhood of A. The measure μ has *normal concentration* on (X, d) if there are constants c and C such that $\alpha_{(X,d,\mu)} \leq C \exp(-cr^2)$ for every r.

1. *The distribution $\mu_{M,k}$ has bounded support that is contained in an interval whose length is no greater than $(\wedge_{n,k} \times \log \tilde{\kappa})$; and*

2. *the variance of $Y_k(M)$ satisfies*

$$\operatorname{var}(Y_k(M)) \leq \frac{1}{4} \left(\wedge_{n,k} \times \log \tilde{\kappa}\right)^2 . \tag{5.4}$$

Remark 2. The variance bound in Eq. (5.4) is much sharper than the one in Eq. (5.3). Both variance bounds are asymptotically constant with respect to n. For fixed k, the two variance bounds differ by a factor of 48 in the large-n limit.

Remark 3. For even n and $k \in \{1, n-1\}$, the bound on the variance in Eq. (5.4) is sharp when M is a $2\ell \times 2\ell$ (where $\ell = n/2$) diagonal matrix with entries $\lambda_1 = \cdots = \lambda_\ell = \tilde{\kappa}$ and $\lambda_{\ell+1} = \cdots = \lambda_{2\ell} = 1$.

When M is diagonal, we can derive a variance bound that is sharper than the bounds in Eqs. (5.3) and (5.4).

Theorem 3 (Variance bound for log-minors of a positive-definite diagonal matrix)**.** *Let D be a positive-definite $n \times n$ diagonal matrix with condition number $\kappa(D) \leq \tilde{\kappa}$. The variance of $Y_k(D)$ satisfies*

$$\operatorname{var}(Y_k(D)) \leq \frac{k}{4} \left(\frac{n-k}{n-1}\right) (\log \tilde{\kappa})^2 . \tag{5.5}$$

Remark 4. The two variance bounds in Eqs. (5.4) and (5.5) are asymptotically constant with respect to n and converge to the same limiting value of $k(\log \tilde{\kappa})^2/4$.

Remark 5. The variance bound for diagonal positive-definite matrices in Eq. (5.5) is sharper than the variance bound for general positive-definite matrices in Eq. (5.4). The former differs from the latter by a factor of $\max\{k, n-k\}/(n-1) \leq 1$.

Remark 6. For even n and any $k \leq n$, the bound on the variance in Eq. (5.5) is sharp when M is a $2\ell \times 2\ell$ (where $\ell = n/2$) diagonal matrix with entries $\lambda_1 = \cdots = \lambda_\ell = \tilde{\kappa}$ and $\lambda_{\ell+1} = \cdots = \lambda_{2\ell} = 1$. The sharpness of the bound for diagonal matrices indicates a limit to possible improvements for the variance bound for general positive-definite matrices. Specifically, one cannot hope to improve the variance bound in Eq. (5.4) by more than a factor of $\max\{k, n-k\}/(n-1)$.

The variance bound in Theorem 2 is sharp for a diagonal matrix. This observation and several examples in Section 5.5 motivate the following conjecture.

Conjecture 1 (Diagonal matrices maximize log-minor variance)**.** *Let $\mathcal{M}_\kappa$ be the set of positive-definite $n \times n$ matrices with condition number κ. For all $k < n$, κ, and $M \in \mathcal{M}_\kappa$, there exists a diagonal matrix $D \in \mathcal{M}_\kappa$ such that*

$$\mathrm{var}(Y_k(M)) \leq \mathrm{var}(Y_k(D))\,. \tag{5.6}$$

The variance bounds (see Eqs. (5.3)–(5.5)) have important implications for the accuracy of sample means of log-minors. We discuss these implications in Section 5.6.

5.4 Proofs of bounds on the distribution of log-minors

5.4.1 Proof of Theorem 1

To prove Theorem 1, we use Cauchy's interlacing theorem and results on Markov chains on countable sets. Chatterjee and Ledoux (2009) previously used this approach to prove a concentration result for empirical cumulative eigenvalue spectra of Hermitian matrices [87].

Proposition 1 (Cauchy's interlacing theorem [250])**.** *Let M be a Hermitian $n \times n$ matrix, and let A be a principal $(n-1)\times(n-1)$ submatrix of M. If M has eigenvalues $\lambda_1(M) \geq \lambda_2(M) \geq \cdots \geq \lambda_n(M)$ and A has eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_{n-1}(A)$, then*

$$\lambda_1(M) \geq \lambda_1(A) \geq \lambda_2(M) \geq \lambda_2(A) \geq \lambda_3(M) \geq \cdots \geq \lambda_{n-1}(M) \geq \lambda_{n-1}(A) \geq \lambda_n(M)\,.$$

Proposition 2 (Large-deviation inequality for functions on countable sets [311])**.** *Let $(\Pi, \mu_{\mathcal{A}})$ be a reversible Markov chain on a finite or countable set $\mathcal{A}$. Let $(\Pi, \mu_{\mathcal{A}})$ have a spectral gap[2] of $\overline{\lambda} > 0$. It follows that , whenever $f : \mathcal{A} \to \mathbb{R}$ is a function such that*

$$|||f|||_\infty^2 := \frac{1}{2} \sup_{A \in \mathcal{A}} \sum_{B \in \mathcal{A}} (f(A) - f(B))^2\, \Pi(A, B) \leq 1\,,$$

it is also true that f is integrable with respect to $\mu_{\mathcal{A}}$ and that, for every $r \geq 0$, the probability measure

$$\mu\left(\{f \geq \int_{\mathcal{A}} f\, d\mu_{\mathcal{A}} + r\}\right) \leq 3e^{-\frac{r}{2}\sqrt{\overline{\lambda}}}\,.$$

[2]The spectral gap (also called the "Poincaré constant") of a Markov chain $(\Pi, \mu_{\mathcal{A}})$ on a space $\mathcal{A}$ is the constant $\overline{\lambda}$ such that, for all functions f, we have $\overline{\lambda} \times \mathrm{var}_\mu(f) \leq \frac{1}{2}\sum_{A,B\in\mathcal{A}}[f(A) - f(B)]^2\Pi(A,B)\mu(\{A\})$. See, for example, Ref. [311, p. 50].

Remark 7. The expected squared distance in f between $A \in \mathcal{A}$ and its adjacent states in the Markov chain is $\sum_{B \in \mathcal{A}} (f(A) - f(B))^2 \, \Pi(A, B)$. One can thus think of $|||f|||_\infty^2$ as a measure of the expected squared distance between the greatest "outlier" A and adjacent states in $\mathcal{A}$. We thus refer to $|||f|||_\infty^2$ as the *squared outlier deviation* of f on $(\Pi, \mu_{\mathcal{A}})$.

Proposition 3 (Spectral gap of random-transposition walk [142])**.** *Let $\mathcal{S}_n$ be the set of permutations of n elements, and let $s \in \mathcal{S}_n$. Let the "random-transposition walk" be a reversible Markov chain $(\Pi, \mu_{\mathcal{S}_n})$ with kernel*

$$\Pi = \begin{cases} 1/n\,, & \text{if } s' = s\,, \\ 2/n^2\,, & \text{if } s' = \tau s \text{ for some transposition } \tau\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

The random-transposition walk has a spectral gap of $\overline{\lambda} = 2/n$.

Proof of Theorem 1. Every principal $k \times k$ submatrix of M is the top-left principal $k \times k$ submatrix of M after a permutation of its rows and columns. We denote the permutated matrix by sM and its top-left principal $k \times k$ submatrix by $\hat{A}_k(sM)$, where $s \in \mathcal{S}_n$ is a permutation of n elements.

For the top-left principal $k \times k$ submatrix, only the first k elements of s are relevant. There are $(n-k)!$ permutations $s \in \mathcal{S}_n$ that are identical in their first k elements, so there is a 1-to-$(n-k)!$ correspondence between $\mathcal{A}_k(M)$ and $\mathcal{S}_n$. Because of the correspondence between $\mathcal{A}_k(M)$ and $\mathcal{S}_n$, we obtain the same distribution for a function $f(A_k(M))$, where we choose $A_k(M)$ uniformly at random from $\mathcal{A}_k(M)$, and for $f(\hat{A}_k(sM))$, where we choose s uniformly at random from $\mathcal{S}_n$.

Let $f_{\alpha,k} : \mathcal{S}_n \to \mathbb{R}$ be such that

$$f_{\alpha,k}(s) := \alpha \log(\det \hat{A}_k(sM)) \tag{5.7}$$

for some $\alpha \in \mathbb{R}$. To find an upper bound on the squared outlier deviation for $f_{\alpha,k}$ on the random-transposition walk, we make two observations:

1. Consider two permutations, s and s', that are adjacent in the random-transposition walk; that is, $s' = \tau s$ for some transposition τ. The determinant is invariant under basis transformation, so the value of $f_{\alpha,k}(s')$ can differ from $f_{\alpha,k}(s)$ only if τ is a transposition that swaps one of the first k elements in s with one of the last $n-k$ elements in s. There are n^2 possible transpositions for a sequence of n elements; and $2k(n-k)$ of these transpositions swap one of the first k elements of the sequence with one of the last $n-k$ elements of the sequence.

Consequently, the fraction of transpositions that change the value of $f_{\alpha,k}$ has an upper bound of

$$b_1 := \frac{2k}{n^2}(n-k)\,.$$

2. Using Cauchy's interlacing theorem (see Proposition 1), one can find an upper bound b_2 for $|f_{\alpha,k}(s) - f_{\alpha,k}(s')|$. For any $k < n$ and any pair s, s' of adjacent permutations, there exists a matrix $C \in A_{k+1}(M)$ such that $\hat{A}_k(sM)$ and $\hat{A}_k(s'M)$ are principal submatrices of C. Cauchy's interlacing theorem implies that

   (a) $\lambda_1(M)$ is an upper bound on the largest eigenvalue of C;

   (b) $\lambda_n(M)$ is a lower bound on the smallest eigenvalue of C;

   (c) $\sum_{i=1}^{k} \log \lambda_i(C)$ is an upper bound on $\log(\det \hat{A}_k(sM))$ and $\log(\det \hat{A}_k(s'M))$; and

   (d) $\sum_{i=2}^{k+1} \log \lambda_i(C)$ is a lower bound on $\log(\det \hat{A}_k(sM))$ and $\log(\det \hat{A}_k(s'M))$.

   Therefore,

$$\begin{aligned}|f_{\alpha,k}(s) - f_{\alpha,k}(s')| &\leq \sum_{i=1}^{k} \log \lambda_i(C) - \sum_{i=2}^{k+1} \log \lambda_i(C) \\ &\leq \log \lambda_1(M) - \log \lambda_n(M) \leq \log \tilde{\kappa}\,.\end{aligned}$$

   This upper bound for $|f_{\alpha,k}(s) - f_{\alpha,k}(s')|$ holds for all adjacent pairs $s, s' \in \mathcal{S}_n$. We can thus set the upper bound to be $b_2 = \alpha \log \tilde{\kappa}$.

We obtain an upper bound for the squared outlier deviation of $f_{\alpha,k}$ of

$$\begin{aligned}|||f_{\alpha,k}|||_\infty^2 &\leq \frac{1}{2} b_1 b_2^2 \\ &= k(n-k)\left(\frac{\alpha \log \tilde{\kappa}}{n}\right)^2 .\end{aligned}$$

Let $\alpha' = n \log \tilde{\kappa} / \sqrt{k(n-k)}$. The function $f_{\alpha'}$ on $(\Pi, \mu_{\mathcal{S}_n})$ has a squared outlier deviation of $|||f_{\alpha'}|||_\infty^2 \leq 1$. We can thus use the tail bound for functions on countable sets (see Proposition 2) for $f_{\alpha'}$. Therefore,

$$\Pr(|f_{\alpha'}(\hat{A}_k(sM)) - \mathbb{E}[f_{\alpha'}(A_k(sM)]| \geq \alpha' r) \leq 6e^{-\frac{\alpha' r}{2}\sqrt{\tilde{\lambda}}}\,.$$

We can substitute $f_{\alpha'}(\hat{A}_k(sM))$ in Eq. (5.7) by $\alpha' Y_k(M)$, because of the correspondence between $\mathcal{A}_k(M)$ and $\mathcal{S}_n$. Applying Proposition 3, we obtain

$$\begin{aligned}&\Pr(|\alpha' Y_k(M) - \alpha' \mathbb{E}[Y_k(M)] \geq \alpha' r) \leq 6e^{-\frac{\alpha' r}{\sqrt{2n}}} \\ &\implies \Pr(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r) \leq 6 \exp\left(-\frac{r}{\log \tilde{\kappa}} \sqrt{\frac{n}{k(n-k)}}\right) .\end{aligned}$$

This proves the first statement of Theorem 1 (see Eq. (5.2)).

We derive a bound on the variance of $\log(\det A)$ from Eq. (5.2) from a direct calculation. First, we write

$$\begin{aligned}\operatorname{var}(Y_k(M)) &= \mathbb{E}[(Y_k(M) - \mathbb{E}[Y_k(M)])^2] \\ &= \int_0^\infty \Pr(Y_k(M) - \mathbb{E}[Y_k(M)])^2 \geq u) du \\ &= \int_0^\infty \Pr(Y_k(M) - \mathbb{E}[Y_k(M)]) \geq \sqrt{u}\} du\,.\end{aligned}$$

Using the tail bound in Eq. (5.2), it follows that

$$\begin{aligned}\operatorname{var}(Y_k(M)) &\leq \int_0^\infty 6 \exp\left(-\frac{\sqrt{u}}{\log\tilde{\kappa}} \sqrt{\frac{n}{k(n-k)}}\right) du \\ &= 12 \left(\frac{k(n-k)}{n}\right) (\log\tilde{\kappa})^2\end{aligned}$$

□

5.4.2 Proof of Theorem 2

We prove Theorem 2 using Cauchy's interlacing theorem and Popoviciu's inequality.

Proposition 4 (Popoviciu's inequality [433, 485])**.** *Let X be a real-valued random variable supported on the interval $[x_{\min}, x_{\max}]$. It then follows that X has variance*

$$\operatorname{var}(X) \leq (x_{\max} - x_{\min})^2/4\,.$$

For a proof of this version of Popoviciu's inequality, see Ref. [354].

Proof of Theorem 2. For any finite n and k, the set $\mathcal{A}_k(M)$ of principal $k \times k$ submatrices of an $n \times n$ matrix M has finite cardinality $\binom{n}{k}$. It follows that the distribution of any function of $A_k(M)$ has finite support. We define an interval $[r_1, r_2]$ with $r_1 := \min Y_k(M)$ and $r_2 := \max Y_k(M)$, such that the support of $\mu_{M,k}$ is a finite subset of $[r_1, r_2]$.

We can obtain any principal $k \times k$ submatrix A of M by removing $n - k$ row–column pairs from M. Successive applications of Cauchy's interlacing theorem show that $\lambda_1(A) \leq \lambda_1(M)$ and $\lambda_k(A) \geq \lambda_n(M)$. It follows that

$$r_1 = k \log \lambda_n(M)\,, \qquad r_2 = k \log \lambda_1(M)\,.$$

Therefore,

$$r_2 - r_1 \leq k\,(\log\tilde{\kappa} + \log\lambda_n(M)) - k \log\lambda_n(M) = k \log\tilde{\kappa}\,.$$

If $k > n/2$, any two principal $k \times k$ submatrices share $2k - n$ rows and columns. They can thus differ in at most $n - k$ rows and columns. It follows that one can refine the lower and upper bounds on the support of $\mu_{M,k}$ so that $[r_1, r_2] \leq \wedge_{n,k} \times \log \tilde{\kappa}$. We have thus proven the first part of Theorem 1. Applying Popiviciu's inequality to $X = Y_k(M)$ with $x_{\max} - x_{\min} \leq r_2 - r_1$ yields the variance bound in Theorem 2. $\square$

5.4.3 Proof of Theorem 3

For our proof of Theorem 3, we maximize $\mathrm{var}(Y_k(D))$ with respect to the eigenvalues of D.

Proof of Theorem 3. Let D be a positive-definite diagonal matrix with entries $\lambda_1(D) \geq \lambda_2(D) \geq \cdots \geq \lambda_n(D) > 0$. Define $x_i = \log \lambda_i(D)$ for each $i \in [n]$; and let $I \in [n]^k$. It then follows that

$$\log(\det D_I) = \sum_{i \in I} x_i \,. \tag{5.8}$$

We now consider the function $v(x_1, x_2, \ldots, x_n) := \mathrm{var}(Y_k(D))$. From Eq. (5.8), we see that every value of $Y_k(D)$ is a sum of a subset of the variables x_i. Therefore, the function $v(x_1, x_2, \ldots, x_n)$ is convex (i.e., concave up) in the variables $x_1, x_2, \ldots, x_n$. Furthermore, the variance is translation-invariant. We may therefore, without loss of generality, suppose that $x_n = 0$ (corresponding to $\lambda_n(D) = 1$) and $x_1 = \log \kappa(D)$ (corresponding to $\lambda_n(D) = \kappa(D)$). Consequently, the maximization of the variance $\mathrm{var}(Y_k(D))$ amounts to the maximization of v over the volume $[0, \log \kappa(D)]^n$ associated with an n-dimensional hypercube with edge length $\log \kappa(D)$. The solutions lie at the vertices of this hypercube. Therefore,

$$x_i = \begin{cases} \log \kappa(D) \,, & i \leq \ell \,, \\ 0 \,, & \text{otherwise} \,, \end{cases}$$

for some $\ell \in [n]$. We may now view $Y_k(D)/(\log \kappa(D))$ as a hypergeometric random variable on a population of size n for which ℓ elements have the value 1 and $n - l$ elements have the value 0. The variance of this hypergeometric random variable is

$$\mathrm{var}(Y_k(D)) = \frac{k(n-k)}{n^2(n-1)} \, (n - \ell) \, \ell (\log \kappa(D))^2 \,,$$

which is maximal at

$$\ell^* = \begin{cases} \frac{n}{2} \,, & n \text{ even} \,, \\ \frac{n \pm 1}{2} \,, & n \text{ odd} \,. \end{cases}$$

The maximal value of $\mathrm{var}(Y_k(D))$ for even n leads to the variance bound

$$\mathrm{var}(Y_k(D)) \leq \frac{k(n-k)}{4(n-1)}(\log \kappa(D))^2 \tag{5.9}$$

$$\leq \frac{k(n-k)}{4(n-1)}(\log \tilde{\kappa})^2 \,. \tag{5.10}$$

Comparing the maximal values of $\mathrm{var}(Y_k(D))$ for even n and for odd n shows that Eq. (5.9) is a variance bound for all n. □

5.5 Examples

In this section, we compare the tail of the distribution $\mu_{M,k}$ for several example matrices M to the bounds in Theorems 1 and 3.

We consider four examples of positive-definite $n \times n$ matrices with $n = 20$ and fixed condition number $\kappa = 3$.

Example E1. Consider the diagonal matrix M_{E1} that maximizes the variance of $Y_k(M_{\mathrm{E1}})$. (See the proof of Theorem 3.) For even n, this matrix has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, where $\lambda_1, \ldots, \lambda_{n/2} = \tilde{\kappa}$ and $\lambda_{n/2+1}, \ldots, \lambda_n = 1$.

Example E2. Consider a diagonal matrix M_{E2} with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. We set $\lambda_1 = \kappa$ and $\lambda_n = 1$. We draw $\lambda_2, \ldots, \lambda_{n-1}$ from a uniform distribution on $[1, \kappa]$.

Example E3. We obtain a non-diagonal positive-definite matrix M_{E3} with condition number κ via an orthogonal transformation of M_{E1}. That is,

$$M_{\mathrm{E3}} := Q^{-1} M_{\mathrm{E1}} Q \,,$$

where Q is an orthogonal matrix that we choose from the Haar measure over the group of orthogonal matrices. We use Stewart's algorithm [511] to generate Q.

Example E4. We again generate a random orthogonal matrix Q using Stewart's algorithm. We obtain another non-diagonal positive-definite matrix $M_{\mathrm{E4}} := Q^{-1} M_{\mathrm{E2}} Q$ via an orthogonal transformation of M_{E2}.

In Figure 5.1, we show the empirical probability densities of $Y_k(M)$ for Examples E1, E2, E3, and E4 using four different values of k. For all four examples, we observe that the interval on which $\mu_{M,k}$ is supported shifts to the right for progressively

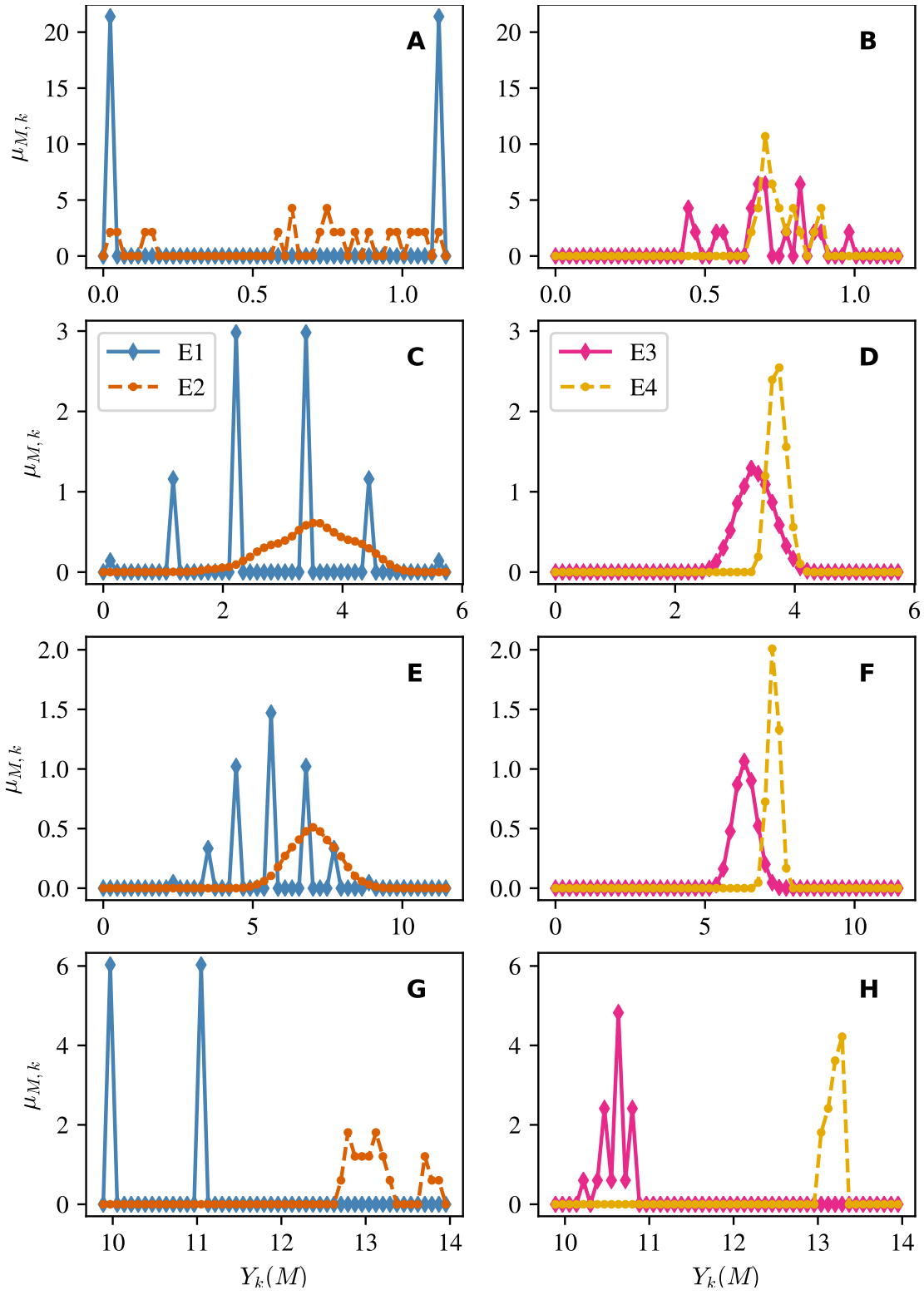


Figure 5.1: Approximate probability density functions for log-minors. We show $\mu_{M,k}$ for Examples E1, E2, E3, and E4 with $n = 20$ and condition number $\kappa = 3$. We show results for $k = 1$ in panels (**A**) and (**B**), for $k = 5$ in panels (**C**) and (**D**), for $k = 10$ in panels (**E**) and (**F**), and for $k = 19$ in panels (**G**) and (**H**).

larger k. The length of the supported interval increases with $\wedge_{n,k}$. For $k = 5$ and $k = 10$ — the cases in which $\wedge_{n,k}$ is larger than 1 — the distribution $\mu_{M,k}$ are almost symmetric about $\mathbb{E}[Y_k(M)]$ for all four examples. For Example E1, the distribution $\mu_{M,k}$ is symmetric about its mean for all examined values of k. Its density is nonzero at $\wedge_{n,k} + 1$ equidistant points.

k	Example	$\mathbb{E}[Y_k(M)]$	$\mathrm{var}(Y_k(M))$	Variance bound from		
				Theorem 1	Theorem 2	Theorem 3
1	E1	0.549	0.302	13.760	0.302	0.302
	E2	0.689	0.115	13.760	0.302	0.302
	E3	0.683	0.021	13.760	0.302	N/A
	E4	0.739	0.005	13.760	0.302	N/A
5	E1	2.747	1.191	54.312	1.509	1.191
	E2	3.446	0.454	54.312	1.509	1.191
	E3	3.283	0.091	54.312	1.509	N/A
	E4	3.649	0.020	54.312	1.509	N/A
10	E1	5.493	1.588	72.416	3.017	1.588
	E2	6.893	0.605	72.416	3.017	1.588
	E3	6.213	0.128	72.416	3.017	N/A
	E4	7.176	0.031	72.416	3.017	N/A
19	E1	10.437	0.302	13.760	0.302	0.302
	E2	13.096	0.115	13.760	0.302	0.302
	E3	10.570	0.022	13.760	0.302	N/A
	E4	13.155	0.008	13.760	0.302	N/A

Table 5.2: Expectation and variance for $Y_k(M)$ for Examples E1, E2, E3, and E4. For comparison, we show the numerical values of the variance bounds from Theorem 1 (see Eq. (5.3)) and Theorem 2 (see Eq. (5.4)). For the examples with diagonal matrices (E1 and E2), we also show the numerical value of the variance bound from Theorem 3 (see Eq. (5.5)).

In Table 5.2, we show $\mathbb{E}[Y_k(M)]$ and $\mathrm{var}(Y_k(M))$ for the distributions in Figure 5.1. We first consider the expectation of $Y_k(M)$. For all four examples, we observe that $\mathbb{E}[Y_k(M)]$ increases with k. For all examined values of k, we see that $\mathbb{E}[Y_k(M_{\mathrm{E4}})] > \mathbb{E}[Y_k(M_{\mathrm{E2}})] > \mathbb{E}[Y_k(M_{\mathrm{E3}})] > \mathbb{E}[Y_k(M_{\mathrm{E1}})]$. Our observations thus suggest that the expectation of $Y_k(M)$ is large when we choose eigenvalues of M uniformly at random from the interval $[1, \kappa]$ and small when we set half of the eigenvalues of M to 1 and the other half to κ.

We now give several observations about the variance of $Y_k(M)$. For all examined values of k, we see that $\mathrm{var}(Y_k(M_{\mathrm{E1}})) > \mathrm{var}(Y_k(M_{\mathrm{E2}})) > \mathrm{var}(Y_k(M_{\mathrm{E3}})) > \mathrm{var}(Y_k(M_{\mathrm{E4}}))$. Our observation of larger $\mathrm{var}(Y_k(M))$ for the examples with diagonal matrices (Examples E1 and E2) than for the examples with non-diagonal matrices (Examples E3 and E4) gives intuitive support for Conjecture 1. Our observation that $\mathrm{var}(Y_k(M_{\mathrm{E1}})) > \mathrm{var}(Y_k(M_{\mathrm{E2}}))$ reflects the fact that Example E1 maximizes the variance in this case (see Theorem 3).

For all examined k, the value of the variance bound in Theorem 1 (see Eq. (5.3)) is at least 24 times larger than the value of the variance bound in Theorem 2 (see Eq. (5.4)). For $k = 1$ and $k = 19$, the cases in which $\wedge_{n,k} = 1$, the value of the variance bound in Theorem 2 is equal to the value of the variance bound for diagonal positive-definite matrices (Theorem 3). Additionally, it is sharp in Example E1.

In Fig. 5.2, we show the empirical tails $\Pr(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r)$ for our four examples. We also show the tail bound B1 from Theorem 1 and two Chebyshev bounds[3], B2 and B3, which we obtain from the variance bounds in Theorems 2 and 3, respectively. Consistent with our observations in Table 5.2 on $\mathrm{var}(Y_k(M))$, we observe that the tail probability tends to be larger for the examples with diagonal matrices (Examples E1 and E2) than for the examples with non-diagonal matrices (Examples E3 and E4).

The difference in functional form guarantees that the bound B1 intersects with the Chebyshev bound B2 at two values of r. If we denote these values by r' and $r'' > r'$, the bound B1 is sharper than B2 on $[0, r']$ and $[r'', \infty)$. In our observations, both bounds exceed the trivial bound $\Pr(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r) \leq 1$ on $[0, r']$. The value r'' lies outside the support of $\Pr(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r)$. We thus see that B1 is sharper than B2 only for values of r for which neither bound is informative.

For $k = 1$ and $k = 19$, the bounds B2 and B3 coincide and are sharp at $r = (\log \kappa)/2$ when $M = M_{\mathrm{E1}}$. For $k = 5$ and $k = 10$, the bound (B3) for diagonal positive-definite matrices is sharper than the bound (B2) for general positive-definite matrices. The difference between the two bounds is most visible for $k = 10$, which is the case that maximizes $\wedge_{n,k}$.

5.6 Estimating mean subsystem entropy

We now consider the implications of our results in Section 5.3 for the problem of estimating the mean subsystem entropy of a given system of coupled variables.

[3] We can obtain a tail bound from a variance bound by using Chebyshev's inequality [54, p. 429], $\Pr(|X - \mathbb{E}[X]| \geq r) \leq \frac{\mathrm{var}(X)}{r^2}$, for an integrable random variable X and $r \in \mathbb{R}_+$.

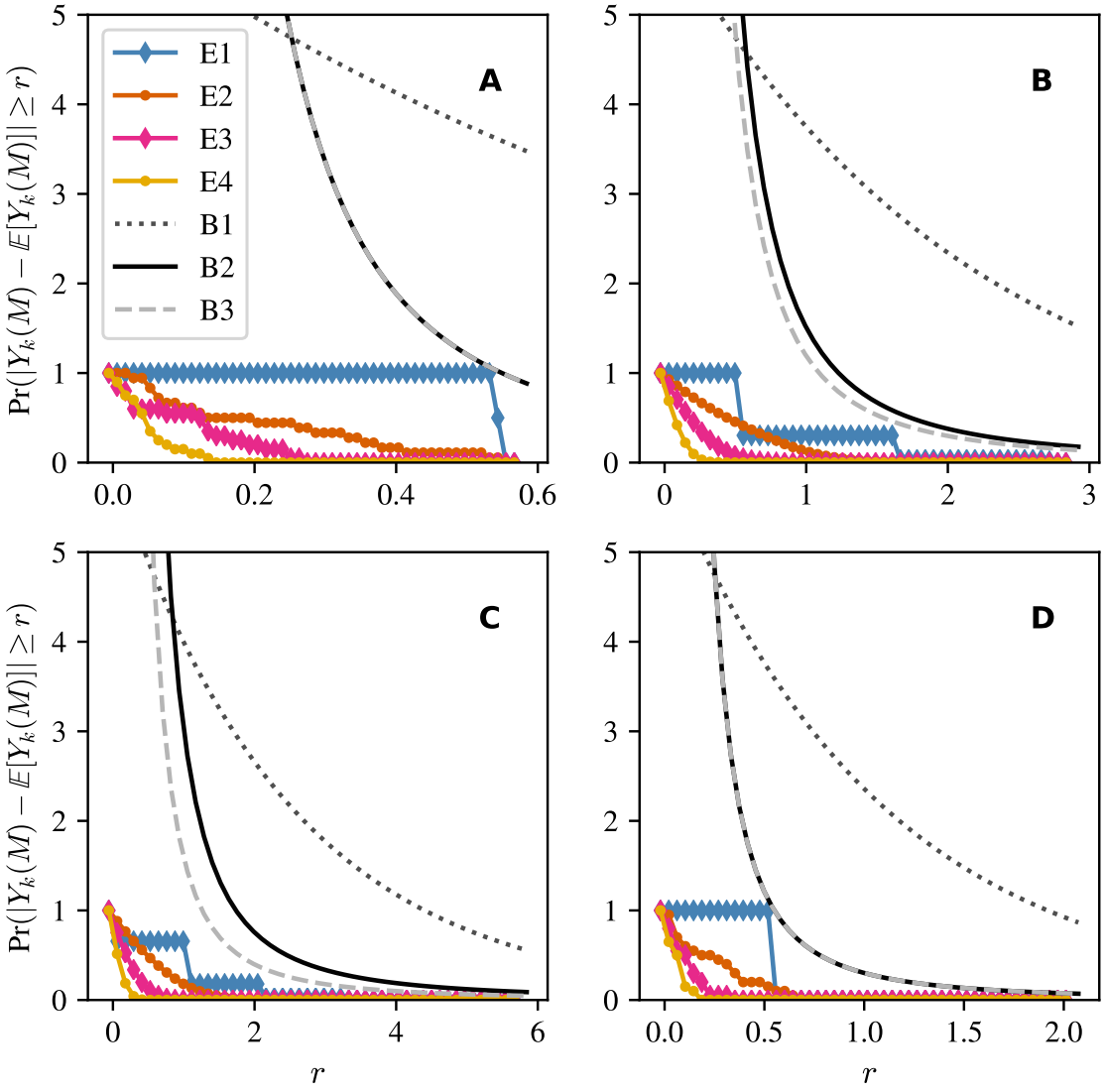


Figure 5.2: Empirical tails of the distribution of log-minors. We show $\Pr(|Y_k(M) - \mathbb{E}[Y_k(M)]| \geq r)$ for positive-definite $n \times n$ matrices from examples E1, E2, E3, and E4 with $n = 20$ and condition number $\kappa = 3$. We show results for $k \times k$ submatrices with (**A**) $k = 1$, (**B**) $k = 5$, (**C**) $k = 10$, and (**D**) $k = 19$. The curve B1 represents the tail bound from Theorem 1. The curves B2 and B3 visualize Chebyshev bounds that we obtain from the variance bounds in Theorems 2 and 3, respectively.

When the joint distribution of variables is a multivariate normal distribution, one can compute the differential entropy of a subsystem by applying Eq. (5.1) to the corresponding sub-covariance matrix. We are interested in the *mean subsystem entropy* $\mathbb{E}[h(A_k(M))]$ for subsystems of k variables. As we noted previously, the large number of subsystems for even modest values of n and k render it prohibitive to exactly compute $\mathbb{E}[h(A_k(M))]$. Fortunately, the tail and variance bounds in

Section 5.3 allow us to instead provide sampling guarantees, through which one can achieve a prescribed sampling accuracy. We give upper bounds on the standard error and on the coefficient of variation for both a sample mean of $Y_k(M)$ and a sample mean of subsystem entropy.

Fix a subsystem size k and sample size $q \geq 1$. The q-sample mean of $Y_k(M)$ is

$$S_Y := \frac{1}{q} \sum_{i=1}^{q} \log(\det A_i)\,,$$

where we choose each A_i uniformly at random from $\mathcal{A}_k(M)$. The q-sample mean of subsystem entropy is

$$S_h := \frac{1}{q} \sum_{i=1}^{q} h(A_i)\,.$$

We use S_Y and S_h as estimators of the population means $\mathbb{E}[Y_k(M)]$ and $\mathbb{E}[h(A_k(M))]$, respectively. These estimators are unbiased, as $\mathbb{E}[S_Y] = \mathbb{E}[Y_k(M)]$ and $\mathbb{E}[S_h] = \mathbb{E}[h(A_k(M))]$. A measure of reliability of an estimator is the *standard error*, which one computes as the estimator's standard deviation. Because $h(A_k(M))$ differs from $Y_k(M)/2$ by a constant, the sample mean S_h has the standard error

$$\sigma(S_h) = \frac{1}{2}\sigma(S_Y) = \frac{1}{2}\sqrt{\frac{\operatorname{var}\left(Y_k(M)\right)}{q}}\,.$$

We may therefore use the bounds of Theorems 1–3 to derive bounds on the standard error for S_Y and S_h.

Corollary 1 (Standard error of the sample mean subsystem entropy)**.** *Let M be a covariance matrix of an n-variate normal distribution, and suppose that the condition number of M satisfies $\kappa(M) \leq \tilde{\kappa}$. Let S_h be the q-sample mean of the entropy of subsets of k variables; and let S_Y be the q-sample mean of log-determinants of $k \times k$ principal submatrices of M. It then follows, for any subsystem size k, that the standard error of the mean subsystem entropy is $\sigma(S_h) = \sigma(S_Y)/2$ and that $\sigma(S_Y)$ satisfies*

$$\sigma(S_Y) \leq \sqrt{\frac{12k(n-k)}{qn}} \log \tilde{\kappa}\,, \tag{5.11}$$

and

$$\sigma(S_Y) \leq \frac{1}{2}\sqrt{\frac{\wedge_{n,k}}{q}} \log \tilde{\kappa}\,. \tag{5.12}$$

Furthermore, if M is diagonal,

$$\sigma(S_Y) \leq \frac{1}{2}\sqrt{\frac{k(n-k)}{q(n-1)}} \log \tilde{\kappa}\,. \tag{5.13}$$

The *coefficient of variation* $c_v(S)$ is another measure of reliability for estimators. It measures the size of the typical error of an estimator S as a fraction of the magnitude of $\mathbb{E}[S]$. As a formula, it is given by

$$c_v(S) := \frac{\sigma(S)}{\mathbb{E}[S]}\,. \tag{5.14}$$

The coefficient of variation for S_Y arises from the standard deviation of the relative error

$$\mathcal{E} := \left|\frac{Y_k(M) - \mathbb{E}[Y_k(M)]}{\mathbb{E}[Y_k(M)]}\right|$$

of $Y_k(M)$ because

$$\frac{\sigma(S_Y)}{\mathbb{E}[S_Y]} = \frac{\sigma(Y_k(M)) - \mathbb{E}[Y_k(M)]}{\mathbb{E}[Y_k(M)]\sqrt{q}} = \sigma\left(\left|\frac{Y_k(M) - \mathbb{E}[Y_k(M)]}{\mathbb{E}[Y_k(M)]}\right|\right)/\sqrt{q} = \sigma(\mathcal{E})/\sqrt{q}\,.$$

For a multivariate Gaussian distribution, the following corollaries give bounds on the coefficient of variation for the sample mean of log-minors and for the sample mean of subsystem entropy.

Corollary 2 (Coefficient of variation for a sample mean of log-minors)**.** *Let $\lambda_1(M), \ldots, \lambda_n(M)$ be the eigenvalues of M; we order them from largest to smallest. Let*

$$\ell(M) = \min\{|\log \lambda_1(M)|\,, |\log \lambda_n(M)|\}\,.$$

If $\ell(M) \neq 0$, the coefficient of variation for a q-sample mean S_Y of $Y_k(M)$ satisfies

$$c_v(S_Y) \leq \frac{\log \hat{\kappa}}{\ell(M)}\sqrt{\frac{12(n-k)}{qkn}} \tag{5.15}$$

and

$$c_v(S_Y) \leq \frac{\log \hat{\kappa}}{2\ell(M)}\sqrt{\frac{\wedge_{n,k}}{qk^2}}\,. \tag{5.16}$$

Proof. This corollary follows from Eq. (5.14). We use Eqs. (5.3) and (5.4) as upper bounds on the numerator. For all k, a lower bound on the denominator is $|\mathbb{E}[Y_k(M)]| \geq k\ell(M)$. □

Corollary 3 (Coefficient of variation for mean subsystem entropy)**.** *For an n-variate Gaussian distribution with covariance matrix M, the coefficient of variation $c_v(S_h)$ for a q-sample mean S_h of subsystem entropy satisfies*

$$c_v(S_h) \leq \frac{2\log \hat{\kappa}}{\ell(M) + \log 2e\pi}\sqrt{\frac{12(n-k)}{qkn}} \tag{5.17}$$

and

$$c_v(S_h) \leq \frac{\log \hat{\kappa}}{\ell(M) + \log 2e\pi}\sqrt{\frac{\wedge_{n,k}}{qk^2}}\,. \tag{5.18}$$

Proof. We derive this result from Eq. (5.14); we use Eqs. (5.3) and (5.4) to bound the numerator, and we use Eq. (5.1) to bound the expectation in the denominator. □

The bounds in Eqs. (5.16) and (5.18) are sharper bounds than Eqs. (5.15) and (5.17). From Eqs. (5.16) and (5.18), we see that both $c_v(S_Y)$ and $c_v(S_h)$ decay in proportion to $\sqrt{k}$. Indeed, under a certain regularity condition (which we specify in Corollary 4), the coefficient of variation decays to 0 in the limit of large n and large k.

Corollary 4 (Concentration of the relative error)**.** *Let $\{M_i\}$ be a sequence of positive-definite matrices of dimension $n(i)$. Let $k = k(i) \leq n(i)$ be a function of i. Suppose that the sequence*

$$a_i := \sqrt{k(i)}\ell(M_i) \tag{5.19}$$

is nondecreasing and unbounded. It then follows that $c_v(S_Y) \to 0$ and $\mathcal{E}$ converges in probability to 0 as i becomes large.

Remark 8. A sufficient condition for the concentration of $\mathcal{E}$ is that the sequence $\{M_i\}$ has fixed condition number and the smallest eigenvalue λ_n is bounded away from both 0 and 1. Formally, the latter condition is

$$\text{there exists } \delta > 0 \text{ such that, for all } i\,, \text{ we have } \lambda_{n(i)}(M_i) \in [\delta, 1-\delta] \cap [1+\delta, \infty)\ . \tag{5.20}$$

Remark 9. A popular model for sample covariance matrices is the Wishart ensemble[4]. A sequence $\{M_i\}$ of Wishart matrices can satisfy the condition in Eq. (5.20) if the ratio $c := n/n_f$ of the number n of variables and the number n_f of degrees of freedom is $c \notin \{1/4, 1\}$ [276, 348].

One can use these bounds on the standard error to choose a sample size q that guarantees a desired accuracy of a sample mean. In Fig. 5.3, we show our bounds on the standard error and the coefficient of variation of S_Y and S_h with $q = 2000k$ and $\ell(M) = 1$. In the left panels, we show the bounds B1 from Eq. (5.11) and B2 from Eq. (5.12) on the standard error of S_Y and S_h. In the right panels, we show the bounds B1' (see Eqs. (5.15) and (5.17)) and B2' (see Eqs. (5.16) and (5.18)) on the coefficient of variation of S_Y and S_h. In panels (**A**) and (**B**), we vary the system size n for fixed subsystem size k. We observe for $n \leq 2k$ that the values of the bounds

[4]The Wishart ensemble $W_n(V, n_f)$ with scale matrix V and n_f degrees of freedom is the ensemble of random matrices $M := n_f^{-1} \sum_{i=1}^{n_f} X_i^T X_i$, where the $X_1, X_2, \ldots, X_{n_f}$ are n_f realizations of an n-variate random variable with 0-mean Gaussian distribution $N_n(0, V)$ [202, 276].

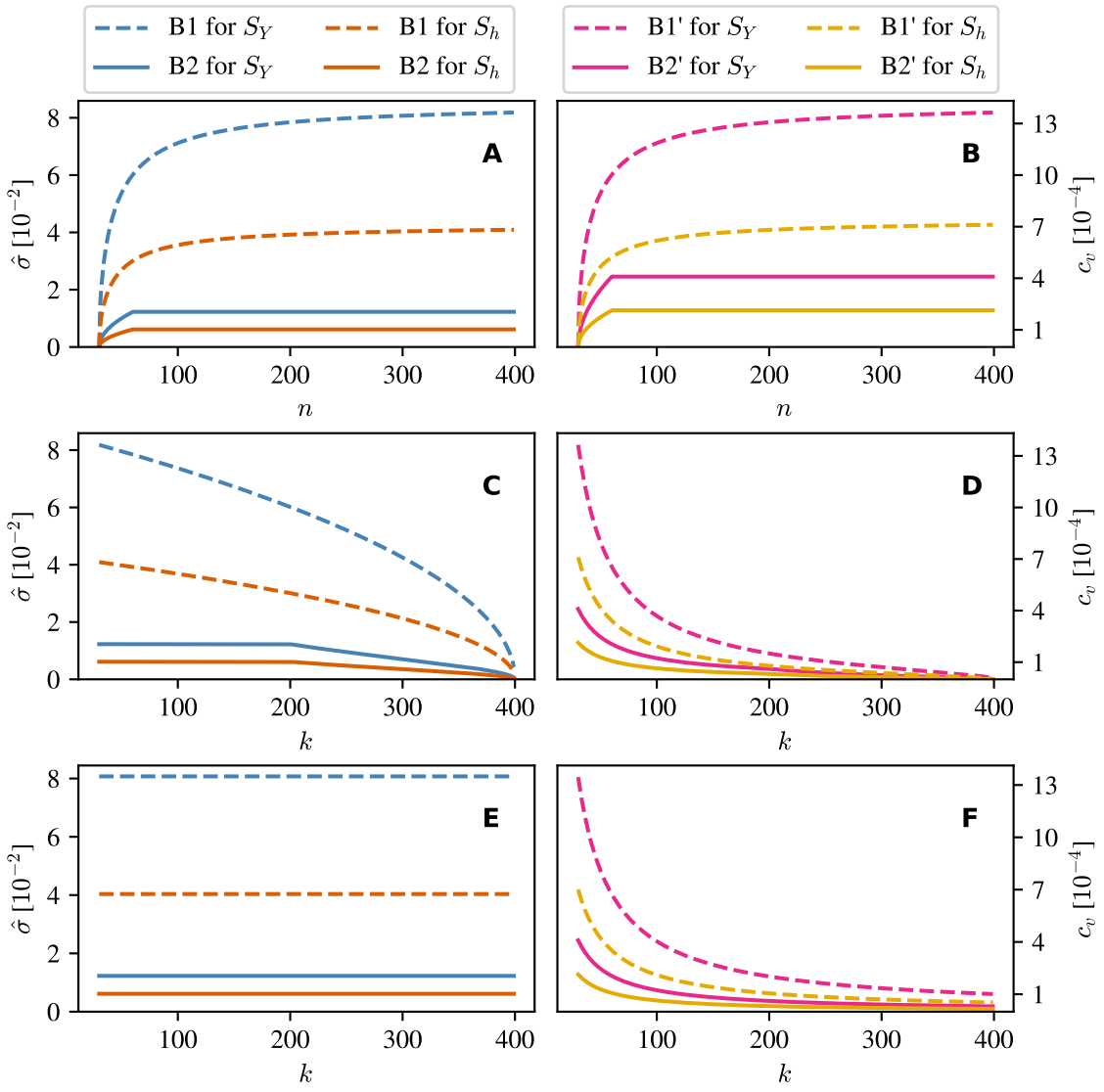


Figure 5.3: Bounds on the standard error and coefficient of variation for a sample mean S_Y of log-minors and a sample mean S_h of subsystem entropy. In the left panels, we show the bounds B1 from Eq. (5.11) and B2 from Eq. (5.12) on the standard error of S_Y and S_h. In the right panels, we show the bounds B1' (see Eqs. (5.15) and (5.17)) and B2' (see Eqs. (5.16) and (5.18)) on the coefficient of variation of S_Y and S_h. We show results for progressively larger n with fixed $k = 30$ in panels (**A**) and (**B**); results for progressively larger k with fixed $n = 400$ in panels (**C**) and (**D**); and results for progressively larger k with a fixed ratio $k/n = 0.1$ in panels (**E**) and (**F**). We compute bounds using $q = 2000k$, $\tilde{\kappa} = 3$, and $\ell(M) = 1$.

B2 and B2' increase with n. For $n > 2k$, the bounds B2 and B2' are independent of n. The bounds B1 and B1' are less sharp than the bounds B2 and B2'. The values of B1 and B1' increase with n and approach their asymptotic values from below. For example, the bound B1 for $\sigma(S_Y)$ has a limiting value of $\sqrt{12k/q} \times \log \tilde{\kappa} \approx 0.08$. In panels (**C**) and (**D**), we vary k for fixed n. We observe for $k \leq n/2$ that the value of the bound B2 on the standard error is independent of k. For $k > n/2$, the value of B2 decreases with increasing k and is 0 for $k = n$. The bound B1 is less sharp than B2. Its value decreases with increasing k for any $k \leq n$. The values of the bounds B1' and B2' on the coefficient of variation decrease with increasing subsystem size and are 0 for $k = n$. In panels (**E**) and (**F**), we vary k for fixed ratio k/n, and we observe that the bounds on the standard error are independent of k if the ratio k/n is constant. The values of the bounds B1' and B2' decrease with increasing k. This is consistent with our previous observation that $c_v(S_Y)$ vanishes if the sequence a_i (see Eq. (5.19)) becomes unbounded.

It is important to note that all of our bounds on the standard error and the coefficient of variation of S_Y and S_h are asymptotically constant in n. It is thus not necessary to sample proportionally more minors from a larger matrix. Instead, to guarantee a desired accuracy of a sample mean of log-minors or subsystem entropy, one can choose q to be a function of k. To ensure that the standard error is constant or decreases with growing n and k, it is sufficient to choose q in linear proportion to k. When the smallest and largest eigenvalues of a system's correlation matrix are fixed, one can ensure that the coefficient of variation is constant or decreasing with growing n and k by choosing q in linear proportion to k^{-1}.

5.7 Conclusions

We examined the problem of estimating the mean subsystem entropy of a system of n coupled variables with covariance matrix M. When the joint distribution of a system's variables is an n-variate Gaussian, t, or Cauchy distribution, the mean differential entropy of subsystems is an affine function of the log-minors of the covariance matrix [6, 384]. We derived tail and variance bounds on the distribution of log-minors of fixed size of a positive-definite matrix with bounded condition number. Using our variance bounds, we provided upper bounds on the standard error and on the coefficient of variation of both the sample mean of log-minors and the sample mean of subsystem entropy. Our results indicate that, despite the rapid growth of the number of subsystems with n, the accuracy of these sample means is asymptotically independent of a system's size. Instead, it is sufficient to increase

the number of samples in linear proportion to the size of subsystems to achieve a desired sampling accuracy.

Our results are salient to studies that use mean subsystem entropy to examine systems of coupled variables [295, 533, 535]. Even for a system with as few as 50 variables, sampling just 0.001% of its subsystem entropies can require the computation of over a billion log-determinants. Using the largest and smallest eigenvalues of a system's covariance matrix to determine the number of samples that are needed to achieve a prescribed accuracy for a sample mean can thus facilitate a quantitative study of mean subsystem entropy when it would otherwise be impossible.

Throughout this chapter, we relied only on knowledge of the largest and smallest eigenvalues of a system's covariance matrix. We expect that it is possible to derive sharper bounds than our current results when one knows the complete spectrum of a system's covariance matrix, likely by relying on Cauchy's interlacing theorem (Proposition 1) to control the log-minors.

We presented two bounds on the variance of a log-minor that we choose uniformly at random from the set of log-minors of size k of an $n \times n$ positive-definite matrix. The variance bound in Theorem 2 is sharper than the one in Theorem 1, but either bound is sufficient to deduce that the accuracy of a sample mean of subsystem entropy is asymptotically independent of a system's size and that one can achieve a prescribed accuracy by choosing the number of samples in linear proportion to the size of subsystems.

The proof of our first bound (see Section 5.4.1) relies on the existence of an upper bound for the difference between $\log(\det A)$ for two different principal submatrices $A \in \mathcal{A}_k(M)$ and the invariance of $\log(\det A)$ under a basis transformation of A. The proof of our second bound (see Section 5.4.2) relies on the existence of an upper bound and a lower bound for the support of the distribution $\mu_{M,k}$.

Similar bounds and the invariance under basis transformation hold for several other matrix properties, including the largest and smallest eigenvalues. It is thus plausible that one can derive similar results for the standard error and coefficient of variation for many spectral properties of principal submatrices. For example, Chatterjee and Ledoux (2009) proved a large-deviation inequality for the empirical cumulative eigenvalue distribution of principal submatrices of Hermitian matrices [87]. These and other variance and tail bounds on submatrix properties offer welcoming possibilities to enhance computational studies that characterize complex systems based on the mean properties of their subsystems. For example, they can provide

guarantees for linear sketching techniques, which are relevant for data dimensionality reduction. They can also facilitate the use of methods of spectral graph analysis in the study of subgraphs and some notions of motifs in networks.

Chapter 6

Conclusions and outlook

In this dissertation, we studied several aspects of robustness and entropy of dynamics on networks. We anticipate that our results are relevant contributions to the study of network robustness for biomedical applications, motif-based investigations of dynamical processes on networks, and subsystem-based investigations of complex systems.

6.1 Robustness of networks to structural changes

In Chapter 2, we reviewed the study of robustness of networks to structural changes. Many studies of network robustness lead to different (and sometimes seemingly contradictory) conclusions on the characteristics of robust network structures. Alderson and Doyle (2010) proposed to distinguish different notions of robustness by several specifications [14]. Their framework helps resolve apparent conflicts between results of different studies of robustness, because it helps recognize that many studies consider different notions of robustness that do not need to be correlated with one another. However, it leads to the problem of choosing relevant specifications for a robustness problem. In our review, we considered various choices for specifying a model of a system, a measure of insensitivity, a set of perturbations, and a performance measure. We categorized different measures of insensitivity and gave examples of studies that demonstrated that the choices of a network model, a set of perturbations, and a performance measure crucially affect the outcome of a study of robustness. We surveyed the study of network robustness to identify network models, targeting strategies for sequential node or edge removal, and performance measures that many researchers have used to define notions of network robustness. Some notions of robustness are very relevant for one application, but not relevant

for others. We discussed the relevance of different notions of robustness to different applications of network models.

When choosing a targeting strategy for sequential node or edge removal, we distinguished between three cases: (1) When a researcher aims to demonstrate that different targeting strategies for sequential node or edge removal can lead to different impacts on a system's performance, it is usually sufficient to compare the impact of removing nodes or edges with the largest degree to the impact of removing them uniformly at random; (2) for studies of adversarial attacks on a system, we advised that the choice of targeting strategies should reflect a researcher's expectation of an adversary's behavior; and (3) a researcher can be interested in determining a good heuristic method for finding a sequence of nodes or edges that, if removed, lead to a large impact on a system's performance. In such cases, it can be useful to consider several centrality measures for selecting nodes or edges for removal, because, for some performance measures, betweenness centrality and other centrality measures identify sequences of nodes or edges that, if removed, lead to larger impacts than targeting nodes or edges based on their largest degree.

When choosing a performance measure, we argued that researchers should use node labels, dynamics on networks, and other features of network models to guide their search for performance measures that are relevant for a given application. If such features are not available, it can help to consider the types of walks or paths on a network that one expects to be relevant to a system function. In this review, we have aimed to demonstrate the breadth of approaches to the study of network robustness and to link relevant notions of robustness to applications. We hope that our review can be a helpful starting point for researchers who are interested in using network models to study the robustness of a real-world system.

6.2 Plausibility of performance measures for protein-interaction networks

The choice of a relevant performance measure for protein-interaction networks (PINs) is a challenging task. Methods for detecting protein–protein interactions have high false-positive rates [214, 504]. Some protein–protein interactions are specific stages of an organism's early development and are unlikely to contribute to the organism's viability in its adult stage [214]. The protein–protein interactions that contribute to a function in a cell or an organism can do so in many different ways (see Section 1.1). To our knowledge, there do not exist mechanistic models of dynamics on PINs.

One can use knowledge about proteins and their functions in biological systems to construct node labels for nodes in a PIN. However, for many proteins, one knows none or only a subset of their functions in biological systems [141, 291, 369].

In Chapter 3, we examined several notions of robustness in PINs. As a network model, we introduced function-specific PINs (fPINs), which are samples of a protein interactome. They include pathway-associated proteins and proteins that are "close to" (specifically, in the neighborhood of) pathway-associated proteins or "between" (specifically, on short paths between) pairs of pathway-associated proteins. We considered several targeting strategies and several performance measures. We examined the structural properties of fPINs, and we found that fPINs have very small mean local clustering coefficients and global clustering coefficients. When comparing the impact of node removal on fPINs, we found that impacts depend strongly on the choice of a performance measure and weakly on the choice of a targeting strategy for node removal and the choice of a sampling method for constructing fPINs. We anticipate that these observations are helpful for building intuition about the effects that different specifications for studies of robustness of PINs may have on the results of such a study. However, it is important to keep in mind that there exist exceptions to these observations.

To identify relevant notions of robustness for fPINs, we constructed a plausibility test for performance measures that assigns criterion scores for selective robustness, structural sensitivity, and perturbation-size monotonicity of structural properties under node removal. We compared the plausibility of ten structural properties as performance measures and found, for almost all sampling methods, that the relative LCC size F and algebraic connectivity a are the most plausible performance measures of the ten structural properties.

By introducing weighting factors for criterion scores or introducing further criteria for plausibility, one can adapt our plausibility test to generalize our study of fPINs robustness or to examine network robustness for other applications. We expect that an adapted plausibility test can lead to different rankings of the plausibility of performance measures than the ranking that we obtained in our study of robustness in fPINs. A comparison of plausibility scores for different plausibility criteria or different weighting factors for criterion scores may be able to demonstrate that there is no structural property of networks that is the most plausible performance measure for all notions of plausibility. However, it is possible that some structural properties are plausible for many different applications. Comparisons of plausibility scores for different plausibility criteria and network models can help (1)

identify such structural properties and (2) highlight that the plausibility of performance measures depends in general on the criteria that a researcher has chosen for a plausibility test. Plausibility tests can thereby be a useful tool for identifying relevant notions of robustness for a specific application when researchers cannot use dynamics on networks or other features of network models to guide their choice of a performance measure.

6.3 Walk graphs in a multivariate Ornstein–Uhlenbeck process on a network

Tononi et al. (1999) proposed a model that links robustness and redundancy in biological systems to the mean differential entropy of subsystems of a stochastic dynamical process on a network [535]. In Chapter 4, motivated by this link, we studied the differential entropy of a multivariate Ornstein–Uhlenbeck process on a network. We proposed walk graphs as a type of motif that can help identify important building blocks in dynamical processes on networks. We derived the contributions of different walk graphs to variance, covariance, and differential entropy of a multivariate Ornstein–Uhlenbeck process and discussed the mechanisms by which walk graphs contribute to these system properties. We found that, under the assumption of short-range signal decay, the mechanisms that lead to variance are also the important mechanisms that contribute to a system's differential entropy. These mechanisms are (1) node-intrinsic noise generation, (2) signal propagation along redundant paths, and (3) signal propagation along cycles. Our comparison of the contributions of different motifs demonstrated that motifs that consist of two redundant walks contribute more than closed walks to a system's differential entropy. We investigated the entropy contributions of the 3-node feedforward loop and the 3-node feedback loop, which many researchers consider to be important building blocks in biological systems [17, 370, 486, 542] and engineering systems [20]. We found that many entropy-inducing walk graphs with small weighting factors occur on the 3-node feedback loop and few entropy-inducing walk graphs with large weighting factors occur on the 3-node feedforward loop. By "strength in numbers", the motifs on the feedback loop lead a greater total contribution to differential entropy than the motifs on the feedforward loop.

Our results in Chapter 4 extend a framework that Barnett et al. (2009) used to identify length-2 and length-3 walk graphs that contribute to neural complexity [31]. We have extended their framework in several ways: (1) We formalized

the concept of walk graphs and discussed the differences between walk graphs and connected subgraphs. (2) We showed that relative differences between walk-graph contributions are independent of the choice of the mean reversion rate and the noise amplitude of the Ornstein–Uhlenbeck process. Results on the relative importance of different walk graphs thus hold for any positive values of these parameters. (3) We derived the conditions that the coupling parameter ϵ and the adjacency matrix $\mathbf{A}$ need to fulfill for an Ornstein–Uhlenbeck process on a network to have short-range signal decay. (4) We derived the structures and weighting factors of walk graphs that contribute to variance, covariance, and differential entropy in the multivariate Ornstein–Uhlenbeck process for arbitrary numbers of edges and showed numerical results for motifs with up to 8 edges.

We anticipate that these extensions can facilitate the discovery of walk graphs that contribute to further insights on entropy-based properties of dynamics on networks (e.g., matching complexity [534], neural redundancy [535], and neural degeneracy [535]), which are relevant to the study of neuronal networks [31, 350] and many others systems [266, 345, 444, 501, 616]. More generally, our results in this chapter demonstrate that it can be useful to consider dynamics on a network (instead of just a network's structure) as a composite entity that one can decompose into many small parts. This decomposition of dynamics on networks provides a framework for studying the mechanisms by which dynamics on networks and network structure contribute to various properties of dynamical systems.

6.4 Sampling guarantees for the mean differential entropy of subsystems

In Chapter 5, we considered the problem of estimating the mean differential entropy of subsystems of a system of n coupled variables with covariance matrix M. When the joint distribution of a system's variables is an n-variate Gaussian, t, or Cauchy distribution, the mean differential entropy of subsystems is an affine function of the log-minors of the covariance matrix [6, 384]. We derived upper bounds for the right tail and the variance of the distribution of minors, and we used these to derive upper bounds on the standard error of the sample mean of a subsystem's differential entropy of subsystems. Our results indicate that, despite the rapid growth of the number of subsystems with n, the accuracy of these sample means is asymptotically independent of a system's size. Instead, to achieve a desired sampling accuracy, it

is sufficient to increase the number of samples in linear proportion to the size of subsystems and the logarithm of M's condition number.

Our results have important implications for the study of mean subsystem entropy of systems of coupled variables. Even for a system with as few as 50 variables, sampling the differential entropy of just 0.001% of its subsystems with 25 variables requires the computation of over a billion log-determinants. Using the condition number of a system's covariance matrix to determine the number of samples that are needed to achieve a prescribed accuracy for a sample mean can thus facilitate a quantitative study of mean differential entropy of subsystems when it would otherwise be impossible.

Our derivation of sampling guarantees for the mean differential entropy of subsystems relies on the bounds of the support of the distribution of differential entropy for subsystems. We found that the sampling guarantees are asymptotically independent of n, because the bounds of the support of log-minors of M are independent of n. We expect that it is possible to use our approach to derive sampling guarantees for other submatrix functions that have a bounded support with bounds independent of n. The resulting sampling guarantees and our sampling guarantees for the mean differential entropy of subsystems can enhance computational studies that characterize complex systems based on the mean properties of their subsystems. For example, they can facilitate the use of methods of spectral graph analysis in the study of subgraphs, graphlets, and motifs in networks.

6.5 Outlook

The results in this dissertation indicate several new and promising avenues of research on robustness and entropy of dynamics on networks. We proposed fPINs as a type of network models for biological functions that aim to capture biological pathways that a biological system uses to ensure its viability, as well as alternative biological pathways that the system can use when environmental changes or therapeutic interventions inhibit the use of the default pathway. It would be interesting to test the use of fPINs for identifying effective interventions for biological functions and, in particular, identifying proteins as potential drug targets for tumor drugs that block default pathways and alternative pathways for a tumor's survival.

Our plausibility-based framework for identifying salient notions of robustness can help researchers find relevant performance measures when they do not possess knowledge of additional model features that can guide their search. Moreover, a

comparison of plausibility scores for different plausibility criteria and network models can (1) help identify structural network properties that tend to be plausible performance measures for many applications and (2) enable researchers to study how the choice of criteria affects the plausibility scores of performance measures.

We proposed walk graphs as a type of motifs that facilitates a motif-based study of dynamics on networks. We expect that the study of walk graphs can uncover mechanisms by which dynamics on networks and network structure contribute to various properties of dynamical systems. In particular, we anticipate that further study of walk graphs in multivariate Ornstein–Uhlenbeck processes can identify walk graphs that contribute to a system's structural and functional redundancy.

Our derivation of sampling guarantees for the mean differential entropy of subsystems demonstrates that it is possible to estimate the mean differential entropy of subsystems from a very small fraction of a system's subsystems. Many studies of links between a system and its subsystems have used the counts of connected subnetworks to characterize a system [17, 247, 512]. However, counting connected subnetworks is a computationally expensive problem [182]. Our sampling guarantees do not improve methods for counting connected subnetworks but they indicate that it is possible to compute accurate estimates of some properties of subsystems (for either connected or fragmented subnetworks) from very few samples. We anticipate that, for uncovering links between a system and its subsystems, the study of mean subsystem properties can be a computationally cheap and potentially useful alternative to the study of connected subnetworks.

Appendix A:

Comparison of structural properties for different sampling methods

A.1 Comparison of degree distributions and local clustering coefficients

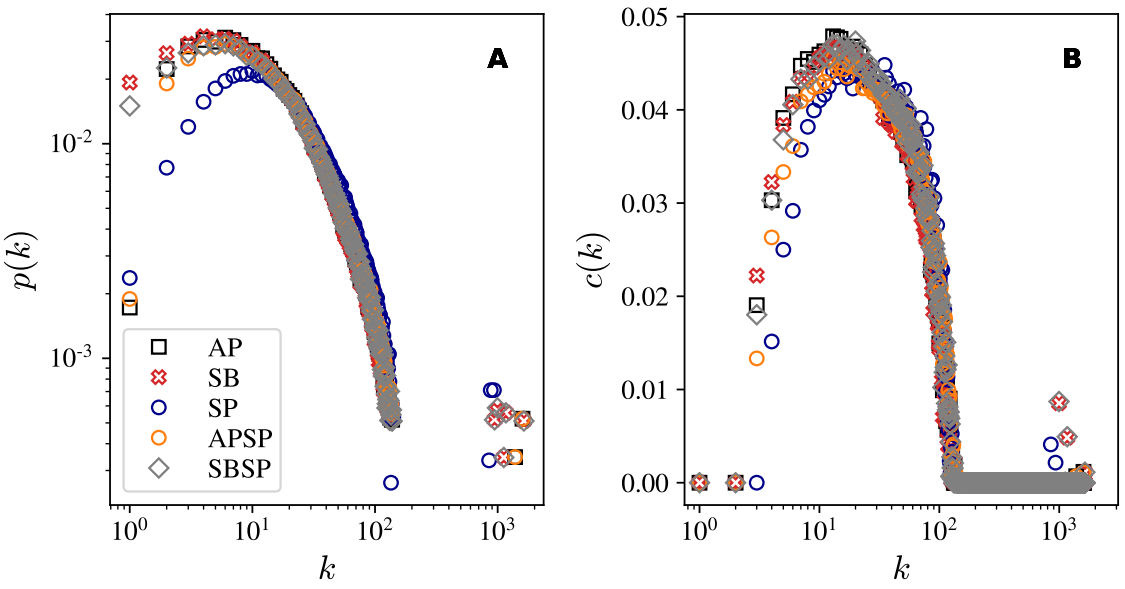


Figure A.1: Degree distribution and mean local clustering coefficient for fPINs from different sampling methods. In (**A**), we show the degree distribution $p(k)$ for sets of fPINs that we obtained using five different sampling methods: 3-hop all-path (AP) sampling , 1-hop snowball (SB) sampling, shortest-path (SP) sampling and all-path–shortest-path (APSP) sampling, and snowball–shortest-path (SBSP) sampling. In (**B**), we show the mean local clustering coefficient $c(k)$ for fPINs from the five samlping methods.

A.2 Comparison of structural properties for different network sizes

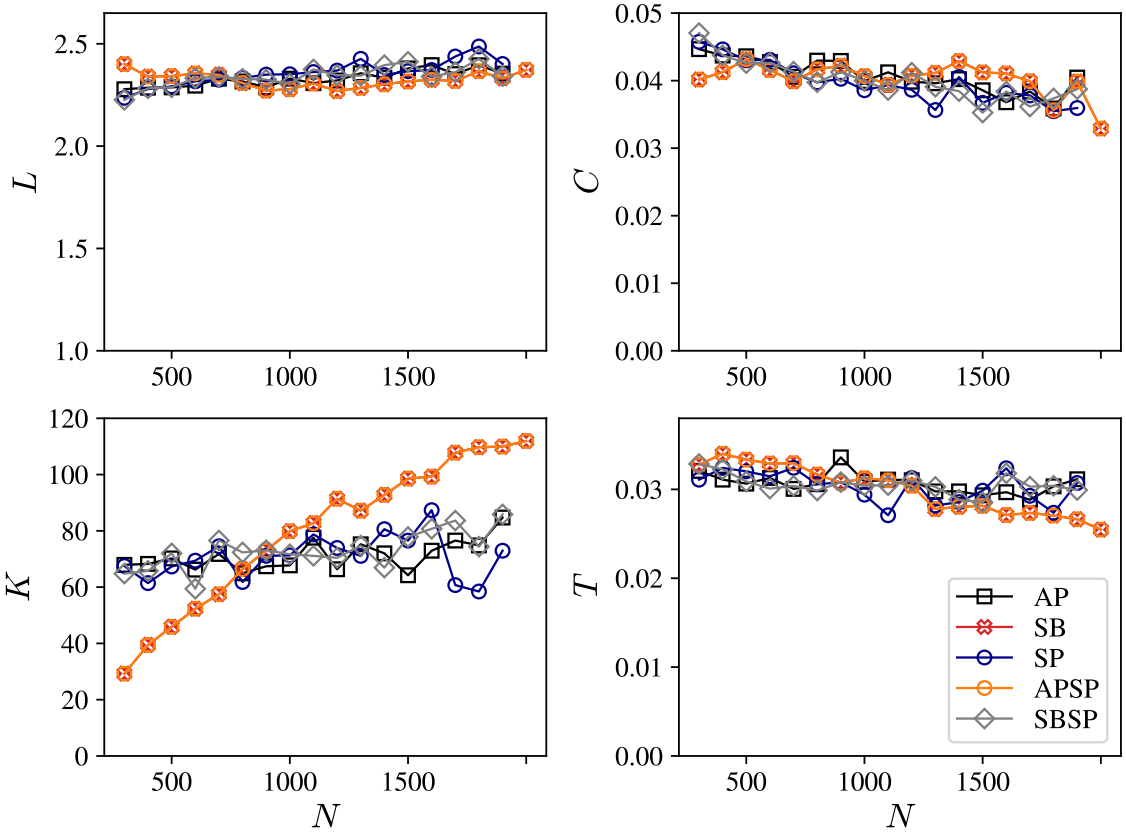


Figure A.2: Structural properties for fPINs with different sizes. We show the mean shortest-path length L, the natural connectivity K, and mean local clustering coefficient C, and the global clustering coefficient T for fPINs from five different sampling methods (see Fig. A.1 for details). For each sampling method, markers indicate indicate the mean of L, K, C, and T over fPINs with sizes in $\{N, N + 99\}$.

Appendix B:

Covariance matrix of an Ornstein–Uhlenbeck process with signal decay

B.1 Derivation of the covariance matrix

At time $t+dt$, the state vector of an OU process with adjacency matrix $\mathbf{A}$, coupling parameter ϵ, noise amplitude ς^2, and rate of reversion θ is

$$\mathbf{x}_{t+dt} = \mathbf{K}\mathbf{x}_t + \varsigma\, dW_t\,, \tag{B.1}$$

where $\mathbf{K} = \mathbf{I} + \theta(\epsilon\mathbf{A} - \mathbf{I})dt$. An OU process with a stationary distribution has a covariance matrix

$$\mathbf{\Sigma} = \langle \mathbf{x}_t^T \mathbf{x}_t \rangle = \langle \mathbf{x}_{t+dt}^T \mathbf{x}_{t+dt} \rangle\,, \tag{B.2}$$

where $\langle\cdot\rangle$ denotes an ensemble average. Using Eq. (B.1) to substitute x_{t+dt} into Eq. (B.2), we obtain

$$\begin{aligned}\mathbf{\Sigma} &= \langle (\mathbf{K}x_t + \varsigma\, dW_t)^T (\mathbf{K}x_t + \varsigma\, dW_t) \rangle \\ &= \langle \mathbf{K}^T x_t^T x_t \mathbf{K} + \varsigma^2 \mathbf{I} dt \rangle\,,\end{aligned} \tag{B.3}$$

because dW_t is a mean-0 unit-variance stochastic process that is independent of x_t. Executing the ensemble average in Eq. (B.3), we obtain

$$\begin{aligned}\mathbf{\Sigma} &= \mathbf{K}^T \mathbf{\Sigma} \mathbf{K} + \varsigma^2 \mathbf{I}\, dt \\ &= [\mathbf{I} + \theta(\epsilon\mathbf{A} - \mathbf{I})\, dtdt]^T \mathbf{\Sigma} [\mathbf{I} + \theta(\epsilon\mathbf{A} - \mathbf{I})\, dtdt] + \varsigma^2 \mathbf{I}\, dtdt \\ &= \mathbf{\Sigma} + \theta \left[(\epsilon\mathbf{A} - \mathbf{I})^T \mathbf{\Sigma} + \mathbf{\Sigma}(\epsilon\mathbf{A} - \mathbf{I}) + \frac{\varsigma^2}{\theta}\mathbf{I} \right] dt + O(dt^2)\,.\end{aligned}$$

To first order in dt, we thus have

$$0 = (\epsilon\mathbf{A} - \mathbf{I})^T \mathbf{\Sigma} + \mathbf{\Sigma}(\epsilon\mathbf{A} - \mathbf{I}) + \frac{\varsigma^2}{\theta}\mathbf{I}\,. \tag{B.4}$$

Equation (B.4) is a Lyapunov equation [90, 175]. For an OU process with signal decay, the solution of Eq. (B.4) is [175]

$$\mathbf{\Sigma} = \frac{\varsigma^2}{\theta}\int_0^\infty e^{(\epsilon\mathbf{A}^T-\mathbf{I})t}e^{(\epsilon\mathbf{A}-\mathbf{I})t}dt = \frac{\varsigma^2}{\theta}\mathbf{\Sigma}_0\,, \tag{B.5}$$

where

$$\mathbf{\Sigma}_0 := \int_0^\infty e^{(\epsilon\mathbf{A}^T-\mathbf{I})t}e^{(\epsilon\mathbf{A}-\mathbf{I})t}dt$$

is the covariance matrix of the OU process when $\varsigma = \theta = 1$. One can express $\mathbf{\Sigma}_0$ as a sum of products of powers of the adjacency matrix and its transpose [31, 32]:

$$\mathbf{\Sigma}_0 = \sum_{L=0}^{\infty} 2^{-(L+1)} \sum_{\ell=0}^{l} \binom{L}{\ell} (\epsilon\mathbf{A}^T)^\ell (\epsilon\mathbf{A})^{L-\ell}\,. \tag{B.6}$$

Therefore,

$$\mathbf{\Sigma} = \frac{\varsigma^2}{\theta}\sum_{L=0}^{\infty} 2^{-(L+1)} \sum_{\ell=0}^{l} \binom{L}{\ell} (\epsilon\mathbf{A}^T)^l (\epsilon\mathbf{A})^{L-\ell}\,. \tag{B.7}$$

Whenever $\epsilon\mathbf{A}-\mathbf{I}$ is stable, the sum in Eq. (B.7) is absolutely convergent. Therefore, we may write

$$\mathbf{\Sigma} = \frac{\varsigma^2}{\theta}\sum_{ell=0}^{\infty}\sum_{ell=r}^{\infty} 2^{-(\ell+r)} \sum_{l=0}^{l} \binom{\ell+r}{\ell} (\epsilon\mathbf{A}^T)^\ell (\epsilon\mathbf{A})^{r}\,.$$

B.2 Conditions for short-range signal decay

The sum in Eq. (4.8) converges if the matrix $\mathbf{\Sigma}_0 = \frac{\theta}{\varsigma^2}\mathbf{\Sigma}$ has eigenvalues [216, p. 38]

$$\nu_{i=1,\ldots,n} \in (0,1)\,.$$

From Eq. (B.5), we see that $\mathbf{\Sigma}_0$ is the positive-definite controllability Gramian of the controllable linear time-invariant system $\dot{x}(t) = (\mathbf{I}-\epsilon\mathbf{A})x(t) + \mathbf{I}u(t)$ with state vector $x(t)$ and input vector $u(t)$ [90, p. 145]. A sufficient condition for short-range signal decay thus only needs to constrain the largest eigenvalue of $\mathbf{\Sigma}_0$.

First, we show that a sufficient condition for short-range signal decay is $\|\epsilon\mathbf{A}\|_2 < 1/2$. Applying the Hilbert–Schmidt norm to both sides of Eq. (B.6), we obtain

$$\begin{aligned}
\|\mathbf{\Sigma}_0\|_2 &= \|\sum_{L=0}^{\infty} 2^{-(L+1)} \sum_{\ell=0}^{l} \binom{L}{\ell} (\epsilon\mathbf{A}^T)^\ell (\epsilon\mathbf{A})^{L-\ell}\| \\
&\leq \sum_{L=0}^{\infty} 2^{-(L+1)} \sum_{\ell=0}^{l} \binom{L}{\ell} \|\epsilon\mathbf{A}\|_2^L \\
&= \frac{1}{2}\sum_{L=0}^{\infty} \|\epsilon\mathbf{A}\|_2^L\,,
\end{aligned}$$

where we used the identity $\|\mathbf{A}^T\|_2 = \|\mathbf{A}\|_2$ and subadditivity and submultiplicativity of the Hilbert–Schmidt norm.

When $\|\epsilon\mathbf{A}\|_2 < 1/2$, it follows that $\|\mathbf{\Sigma}_0\|_2 < 1$, so $\nu_{i=1,\dots,n} \in (0,1)$ for the positive-definite matrix $\mathbf{\Sigma}_0$. It follows that the sum in Eq. (4.8) converges.

If $\mathbf{A}$ is the adjacency matrix of a strongly connected graph with non-negative edge weights, one can relax the condition for short-range signal decay. The adjacency matrix of a strongly connected graph is irreducible [70]. For an irreducible matrix with non-negative entries, the Perron–Frobenius theorem guarantees the existence of a simple positive real leading eigenvalue $\lambda_{\max} = \rho(\mathbf{A})$ [70]. The transpose of $\mathbf{A}$ is also an adjacency matrix of a strongly connected graph with non-negative edge weights, so $\mathbf{A}^T$ also has a simple positive real leading eigenvalue $\lambda_{\max}$. Ortega (1990) proved the existence of a submultiplicative matrix norm $\|\mathbf{M}\|$, such that $\rho(\mathbf{M}) = \|\mathbf{M}\|$ for all complex square matrices $\mathbf{M}$ with simple max-modulus eigenvalues[1] [410, p. 24]. The matrices $\mathbf{A}$ and $\mathbf{A}^T$ are matrices with a single simple max-modulus eigenvalue. We thus write

$$\rho(\mathbf{\Sigma}_0) \leq \|\mathbf{\Sigma}_0\| = \|\sum_{L=0}^{\infty} 2^{-(L+1)} \sum_{\ell=0}^{l} \binom{L}{\ell} (\epsilon\mathbf{A}^T)^{\ell} (\epsilon\mathbf{A})^{L-\ell}\|$$

and use the subadditivity and submulitplicativity of $\|\cdot\|$ to obtain

$$\rho(\mathbf{\Sigma}) \leq \frac{1}{2}\sum_{L=0}^{\infty} \|\epsilon\mathbf{A}\|^L = \frac{1}{2}\sum_{L=0}^{\infty} (\rho(\epsilon\mathbf{A}))^L \ .$$

For $\rho(\epsilon\mathbf{A}) < 1/2$, it follows that $\rho(\mathbf{\Sigma}_0) < 1$. We thus have $\nu_{i=1,\dots,n} \in (0,1)$, so the sum in Eq. (4.8) converges.

[1] A max-modulus eigenvalue μ_i of a matrix $\mathbf{M}$ is an eigenvalue that satisfies $\mu_i = \rho(\mathbf{M})$.

Bibliography

[1] O. O. Aalen and H. K. Gjessing. Survival models based on the Ornstein–Uhlenbeck process. *Lifetime Data Analysis*, 10(4):407–423, 2004.

[2] M. L. Acencio and N. Lemke. Towards the prediction of essential genes by integration of network topology, cellular localization and biological process information. *BMC Bioinformatics*, 10(1):290, 2009.

[3] C. Adami. Information theory in molecular biology. *Physics of Life Reviews*, 1(1):3–22, 2004.

[4] S. Agreste, S. Catanese, P. De Meo, E. Ferrara, and G. Fiumara. Network structure and resilience of mafia syndicates. *Information Sciences*, 351:30–47, 2016.

[5] M. Ahmadi, R. Jafari, S.-A. Marashi, and A. Farazmand. Evidence for the relationship between the regulatory effects of microRNAs and attack robustness of biological networks. *Computers in Biology and Medicine*, 63:83–91, 2015.

[6] N. A. Ahmed and D. V. Gokhale. Entropy expressions and their estimators for multivariate distributions. *IEEE Transactions on Information Theory*, 35(3):688–692, 1989.

[7] J. Akiyama and M. Kano. *Factors and Factorizations of Graphs: Proof Techniques in Factor Theory*. Springer, Berlin, Germany, 2011.

[8] L. Akoglu, P. O. V. de Melo, and C. Faloutsos. Quantifying reciprocity in large weighted communication networks. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 85–96. Springer, 2012.

[9] R. Albert, I. Albert, and G. L. Nakarado. Structural vulnerability of the North American power grid. *Physical Review E*, 69(2):025103, 2004.

[10] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002.

[11] R. Albert, H. Jeong, and A.-L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382, 2000.

[12] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, and P. Wal-

ter. *Molecular Biology of the Cell*. Taylor & Francis, New York, NY, USA, 6th edition, 2014.

[13] M. Aldana and P. Cluzel. A natural class of robust networks. *Proceedings of the National Academy of Sciences of the United States of America*, 100(15):8710–8714, 2003.

[14] D. L. Alderson and J. C. Doyle. Contrasting views of complexity and their implications for network-centric infrastructures. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(4):839–852, 2010.

[15] W. Ali, C. Deane, and G. Reinert. Protein interaction networks and their statistical analysis. *Handbook of Statistical Systems Biology*, pages 200–234, 2011.

[16] S. Allesina, A. Bodini, and M. Pascual. Functional links and robustness in food webs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1524):1701–1709, 2009.

[17] U. Alon. Network motifs: Theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461, 2007.

[18] M. Altar. Complexity, Uncertainty and Systemic Risk: A Survey. *Central Bank Journal of Law and Finance*, 2(2):65–105, 2015.

[19] G. Altay, M. Asim, F. Markowetz, and D. E. Neal. Differential C3NET reveals disease networks of direct physical interactions. *BMC Bioinformatics*, 12(1):296, 2011.

[20] W. Altmann. *Practical Process Control for Engineers and Technicians*. Elsevier, Burlington, MA, USA, 2005.

[21] G. E. Andrews and K. Eriksson. *Integer Partitions*. Cambridge University Press, Cambridge, United Kingdom, 2nd edition, 2004.

[22] F. Arrigo and M. Benzi. Updating and downdating techniques for optimizing network communicability. *SIAM Journal on Scientific Computing*, 38(1):B25–B49, 2016.

[23] H. Ashrafian, W. J. McKenna, and H. Watkins. Disease pathways and novel therapeutic targets in hypertrophic cardiomyopathy. *Circulation Research*, 109(1):86–96, 2011.

[24] D. Ayyildiz, E. Gov, R. Sinha, and K. Y. Arga. Ovarian cancer differential interactome and network entropy analysis reveal new candidate biomarkers. *Omics: A Journal of Integrative Biology*, 21(5):285–294, 2017.

[25] A. S. Azmi. Network pharmacology for cancer drug discovery: Are we there yet? *Future Medicinal Chemistry*, 4(8):939–941, 2012.

[26] A. S. Azmi and R. M. Mohammad. Rectifying cancer drug discovery through network pharmacology. *Future Medicinal Chemistry*, 6(5):529–539, 2014.

[27] J. A. Baggio, S. B. BurnSilver, A. Arenas, J. S. Magdanz, G. P. Kofinas, and M. De Domenico. Multiplex social ecological network analysis reveals how social changes affect community robustness more than resource depletion. *Proceedings of the National Academy of Sciences of the United States of America*, 113(48):13708–13713, 2016.

[28] K. Bähner, K. Zweig, I. Leal, and R. Wirth. Robustness of plant–insect herbivore interaction networks to climate change in a fragmented temperate forest landscape. *Bulletin of Entomological Research*, 107(5):563–572, 2017.

[29] M. Baptista, F. M. Kakmeni, and C. Grebogi. Combined effect of chemical and electrical synapses in Hindmarsh–Rose neural networks on synchronization and the rate of information. *Physical Review E*, 82(3):036203, 2010.

[30] L. Barnett, A. B. Barrett, and A. K. Seth. Granger causality and transfer entropy are equivalent for Gaussian variables. *Physical Review Letters*, 103(23):238701, 2009.

[31] L. Barnett, C. L. Buckley, and S. Bullock. Neural complexity and structural connectivity. *Physical Review E*, 79(5):051914, 2009.

[32] L. Barnett, C. L. Buckley, and S. Bullock. Neural complexity: A graph theoretic interpretation. *Physical Review E*, 83(4):041906, 2011.

[33] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3747–3752, 2004.

[34] M. Barthélemy. *Spatial Networks*. Springer, Berlin, Germany, 2018.

[35] S. M. Bartram, G. W. Brown, and J. E. Hund. Estimating systemic risk in the international financial system. *Journal of Financial Economics*, 86(3):835–869, 2007.

[36] B. Barzel and O. Biham. Quantifying the connectivity of a network: The network correlation function method. *Physical Review E*, 80(4):046104, 2009.

[37] J. Bascompte. Structure and dynamics of ecological networks. *Science*, 329(5993):765–766, 2010.

[38] Q. Bashir, S. Scanu, and M. Ubbink. Dynamics in electron transfer protein complexes. *The FEBS Journal*, 278(9):1391–1400, 2011.

[39] K. E. Bassler, C. I. Del Genio, P. L. Erdős, I. Miklós, and Z. Toroczkai. Exact sampling of graphs with prescribed degree correlations. *New Journal of Physics*, 17(8):083052, 2015.

[40] G. Bebek and J. Yang. PathFinder: Mining signal transduction pathway segments from protein-protein interaction networks. *BMC Bioinformatics*, 8(1):335, 2007.

[41] J. L. Beck. Bayesian system identification based on probability logic. *Structural Control and Health Monitoring*, 17(7):825–847, 2010.

[42] J. Behre, T. Wilhelm, A. von Kamp, E. Ruppin, and S. Schuster. Structural robustness of metabolic networks with respect to multiple knockouts. *Journal of Theoretical Biology*, 252(3):433–441, 2008.

[43] S. Benoit, J.-E. Colliard, C. Hurlin, and C. Pérignon. Where the risks lie: A survey on systemic risk. *Review of Finance*, 21(1):109–152, 2017.

[44] M. Benzi and C. Klymko. Total communicability as a centrality measure. *Journal of Complex Networks*, 1(2):124–149, 2013.

[45] F. Berkes and H. Ross. Community resilience: Toward an integrated approach. *Society & Natural Resources*, 26(1):5–20, 2013.

[46] A. Beygelzimer, G. Grinstein, R. Linsker, and I. Rish. Improving network robustness by edge modification. *Physica A: Statistical Mechanics and its Applications*, 357(3-4):593–612, 2005.

[47] R. Bhamra, S. Dani, and K. Burnard. Resilience: The concept, a literature review and future directions. *International Journal of Production Research*, 49(18):5375–5393, 2011.

[48] C. Biely, K. Dragosits, and S. Thurner. The prisoner's dilemma on co-evolving networks under perfect rationality. *Physica D: Nonlinear Phenomena*, 228(1):40–48, 2007.

[49] A. Bigdeli, A. Tizghadam, and A. Leon-Garcia. Comparison of network criticality, algebraic connectivity, and other graph metrics. In *Proceedings of the 1st Annual Workshop on Simplifying Complex Network for Practitioners*, page 4. ACM, 2009.

[50] K. Bilal, M. Manzano, S. U. Khan, E. Calle, K. Li, and A. Y. Zomaya. On the characterization of the structural robustness of data center networks. *IEEE Transactions on Cloud Computing*, 1(1):1–1, 2013.

[51] P. Billingsley. *Ergodic Theory and Information*. John Wiley & Sons, New York City, NY, USA, 1965.

[52] D. Bisias, M. Flood, A. W. Lo, and S. Valavanis. A survey of systemic risk analytics. *Annual Review of Financial Economics*, 4(1):255–296, 2012.

[53] F. Blanchini and E. Franco. Structurally robust biological networks. *BMC Systems Biology*, 5(1):74, 2011.

[54] J. K. Blitzstein and J. Hwang. *Introduction to Probability*. CRC Press, Boca Raton, FL, USA, 2014.

[55] F. Bloch, M. O. Jackson, and P. Tebaldi. Centrality measures in networks. *Available at SSRN*, page 2749124, 2017.

[56] 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. *Physics Reports*, 544(1):1–122, 2014.

[57] M. Boguná and R. Pastor-Satorras. Epidemic spreading in correlated complex networks. *Physical Review E*, 66(4):047104, 2002.

[58] M. Boguná and M. Á. Serrano. Generalized percolation in random directed networks. *Physical Review E*, 72(1):016106, 2005.

[59] P. Boldi, M. Rosa, and S. Vigna. Robustness of social networks: Comparative results based on distance distributions. In *International Conference on Social Informatics*, pages 8–21. Springer, 2011.

[60] B. Bollobás. *Modern Graph Theory*, volume 184. Springer-Verlag, Berlin, Germany, 2013.

[61] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Mathematics*, 1(1):1–35, 2004.

[62] T. A. Boltz, P. Devkota, and S. Wuchty. Collective influencers in protein interaction networks. *Scientific Reports*, 9(1):3948, 2019.

[63] P. Bonacich. Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, 2(1):113–120, 1972.

[64] P. Bonacich, A. C. Holdren, and M. Johnston. Hyper-edges and multidimensional centrality. *Social Networks*, 26(3):189–203, 2004.

[65] S. P. Borgatti, K. M. Carley, and D. Krackhardt. On the robustness of centrality measures under conditions of imperfect data. *Social Networks*, 28(2):124–136, 2006.

[66] J. Borge-Holthoefer, R. A. Banos, S. González-Bailón, and Y. Moreno. Cascading behaviour in complex socio-technical networks. *Journal of Complex Networks*, 1(1):3–24, 2013.

[67] J. Borge-Holthoefer, Y. Moreno, and A. Arenas. Topological versus dynamical robustness in a lexical network. *International Journal of Bifurcation and Chaos*, 22(07):1250157, 2012.

[68] K. Borkovich, F. Farrelly, D. Finkelstein, J. Taulien, and S. Lindquist. hsp82 is an essential protein that is required in higher concentrations for growth of cells at higher temperatures. *Molecular and Cellular Biology*, 9(9):3919–3930, 1989.

[69] A. Bossi and B. Lehner. Tissue specificity and the human protein interaction network. *Molecular Systems Biology*, 5(1), 2009.

[70] M. Boyle. Notes on the Perron–Frobenius theory of nonnegative matrices, 2015.

[71] M. Bradonjić, A. Hagberg, and F. Pan. Performance of wireless sensor networks under random node failures. In *MILCOM Military Communications Conference*, pages 1784–1789. IEEE, 2011.

[72] A. Braunstein, L. Dall'Asta, G. Semerjian, and L. Zdeborová. Network dismantling. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44):12368–12373, 2016.

[73] S. L. Bressler and A. K. Seth. Wiener–Granger causality: A well established methodology. *Neuroimage*, 58(2):323–329, 2011.

[74] C. Brun, F. Chevenet, D. Martin, J. Wojcik, A. Guénoche, and B. Jacq. Functional classification of proteins for the prediction of cellular function from a protein-protein interaction network. *Genome Biology*, 5(1):R6, 2003.

[75] Z. Burda, J. Duda, J.-M. Luck, and B. Waclaw. Localization of the maximal entropy random walk. *Physical Review Letters*, 102(16):160602, 2009.

[76] F. Caccioli, P. Barucca, and T. Kobayashi. Network models of financial systemic risk: A review. *Journal of Computational Social Science*, 1(1):81–114, 2018.

[77] D. S. Callaway, M. E. Newman, S. H. Strogatz, and D. J. Watts. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters*, 85(25):5468, 2000.

[78] L. Cantini, E. Medico, S. Fortunato, and M. Caselle. Detection of gene communities in multi-networks reveals cancer drivers. *Scientific Reports*, 5:17386, 2015.

[79] G. Cantwell, Y. Liu, B. F. Maier, A. C. Schwarze, C. A. Serván, J. Snyder, and G. St-Onge. Thresholding normally distributed data creates complex networks. *arXiv preprint arXiv:1902.08278*, 2019.

[80] J. Cárdenas, M. Mouronte, L. Moyano, M. Vargas, and R. Benito. On the robustness of spanish telecommunication networks. *Physica A: Statistical Mechanics and its Applications*, 389(19):4209–4216, 2010.

[81] J. M. Carlson and J. Doyle. Highly optimized tolerance: A mechanism for power laws in designed systems. *Physical Review E*, 60(2):1412, 1999.

[82] J. M. Carlson and J. Doyle. Highly optimized tolerance: Robustness and design in complex systems. *Physical Review Letters*, 84(11):2529, 2000.

[83] J. M. Carlson and J. Doyle. Complexity and robustness. *Proceedings of the National Academy of Sciences of the United States of America*, 99(suppl 1):2538–2545, 2002.

[84] L. C. Carpi, O. A. Rosso, P. M. Saco, and M. G. Ravetti. Analyzing complex networks evolution through information theory quantifiers. *Physics Letters A*, 375(4):801–804, 2011.

[85] S. Catanese, P. De Meo, and G. Fiumara. Resilience in criminal networks. *Atti Della Accademia Peloritana dei Pericolanti-Classe di Scienze Fisiche, Matematiche e Naturali*, 94(2):1, 2016.

[86] A. Chatr-Aryamontri, B.-J. Breitkreutz, R. Oughtred, L. Boucher, S. Heinicke, D. Chen, C. Stark, A. Breitkreutz, N. Kolas, L. O'Donnell, et al. The biogrid interaction database: 2015 update. *Nucleic Acids Research*, 43(D1):D470–D478, 2015.

[87] S. Chatterjee and M. Ledoux. An observation about submatrices. *Electronic Communications in Probability*, 14(48):495–500, 2009.

[88] G. Chaverri. Comparative social network analysis in a leaf-roosting bat. *Behavioral Ecology and Sociobiology*, 64(10):1619–1630, 2010.

[89] B.-S. Chen and C.-W. Li. On the interplay between entropy and robustness of gene regulatory networks. *Entropy*, 12(5):1071–1101, 2010.

[90] C.-T. Chen. *Linear System Theory and Design*. Oxford University Press, Inc., Oxford, United Kingdom, 4th edition, 2013.

[91] L. Chen, R.-S. Wang, and X.-S. Zhang. *Biomolecular networks: Methods and applications in systems biology*, volume 10. John Wiley & Sons, New York City, NY, USA, 2009.

[92] Y. Chen, G. Paul, S. Havlin, F. Liljeros, and H. E. Stanley. Finding a better immunization strategy. *Physical Review Letters*, 101(5):058701, 2008.

[93] Z. Chen, J. Wu, Y. Xia, and X. Zhang. Robustness of interdependent power grids and communication networks: A complex network perspective. *Ieee Transactions on Circuits and Systems II: Express Briefs*, 65(1):115–119, 2017.

[94] C.-S. Chin and M. P. Samanta. Global snapshot of a protein interaction network — a percolation based approach. *Bioinformatics*, 19(18):2413–2419, 2003.

[95] H.-C. Chiu, A. Kovacs, D. A. Ford, F.-F. Hsu, R. Garcia, P. Herrero, J. E. Saffitz, and J. E. Schaffer. A novel mouse model of lipotoxic cardiomyopathy. *The Journal of Clinical Investigation*, 107(7):813–822, 2001.

[96] L.-H. Chu and B.-S. Chen. Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets. *BMC Systems Biology*, 2(1):56, 2008.

[97] N. N. Chung, L. Y. Chew, J. Zhou, and C. H. Lai. Impact of edge removal on the centrality betweenness of the best spreaders. *Europhysics Letters*, 98(5):58004, 2012.

[98] S. Ciliberti, O. C. Martin, and A. Wagner. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Computational Biology*, 3(2):e15, 2007.

[99] E. Clark, A. Nellis, S. J. Hickinbotham, S. Stepney, T. Clarke, M. Pay, and P. Young. Degeneracy enriches artificial chemistry binding systems. In *European Conference on Artificial Life, Paris*, pages 133–140, 2011.

[100] K. L. Clarkson and D. P. Woodruff. Numerical linear algebra in the streaming model. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 205–214. ACM, 2009.

[101] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. Resilience of the internet to random breakdowns. *Physical Review Letters*, 85(21):4626, 2000.

[102] R. Cohen and S. Havlin. *Complex networks: Structure, robustness and function*. Cambridge University Press, Cambridge, United Kingdom, 2010.

[103] C. J. Colbourn. Network resilience. *Siam Journal on Algebraic Discrete Methods*, 8(3):404–409, 1987.

[104] T. Coletta and P. Jacquod. Performance measures in electric power networks under line contingencies. *IEEE Transactions on Control of Network Systems*, 2019.

[105] A. F. Colladon and F. Vagaggini. Robustness and stability of enterprise intranet social networks: The impact of moderators. *Information Processing & Management*, 53(6):1287–1298, 2017.

[106] J. R. Correa and N. E. Stier-Moses. Wardrop equilibria. *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

[107] M. Costanzo, A. Baryshnikova, J. Bellay, Y. Kim, E. D. Spear, C. S. Sevier, H. Ding, J. L. Koh, K. Toufighi, S. Mostafavi, et al. The genetic landscape of a cell. *Science*, 327(5964):425–431, 2010.

[108] M. Costanzo, A. Baryshnikova, C. L. Myers, B. Andrews, and C. Boone. Charting the genetic interaction map of a cell. *Current Opinion in Biotechnology*, 22(1):66–74, 2011.

[109] M. Costanzo, B. VanderSluis, E. N. Koch, A. Baryshnikova, C. Pons, G. Tan, W. Wang, M. Usaj, J. Hanchard, S. D. Lee, et al. A global genetic interaction network maps a wiring diagram of cellular function. *Science*, 353(6306):aaf1420, 2016.

[110] E. Costenbader and T. W. Valente. The stability of centrality measures when networks are sampled. *Social Networks*, 25(4):283–307, 2003.

[111] T. Couffinhal, M. Silver, L. P. Zheng, M. Kearney, B. Witzenbichler, and J. M. Isner. Mouse model of angiogenesis. *The American Journal of Pathology*, 152(6):1667–1679, 1998.

[112] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley & Sons, New York City, NY, USA, 2nd edition, 2006.

[113] L. Cowen, T. Ideker, B. J. Raphael, and R. Sharan. Network propagation: a universal amplifier of genetic associations. *Nature Reviews Genetics*, 18(9):551–562, 2017.

[114] E. Creaco, A. Fortunato, M. Franchini, and M. Mazzola. Comparison between entropy and resilience as indirect measures of reliability in the framework of water distribution network design. *Procedia Engineering*, 70:379–388, 2014.

[115] P. Crucitti, V. Latora, and M. Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4):045104, 2004.

[116] P. Crucitti, V. Latora, M. Marchiori, and A. Rapisarda. Error and attack tolerance of complex networks. *Physica A: Statistical Mechanics and its Applications*, 340(1-3):388–394, 2004.

[117] P. Csermely, T. Korcsmáros, H. J. Kiss, G. London, and R. Nussinov. Structure and dynamics of molecular networks: A novel paradigm of drug discovery: A comprehensive review. *Pharmacology & Therapeutics*, 138(3):333–408, 2013.

[118] L. Cuadra, S. Salcedo-Sanz, J. Del Ser, S. Jiménez-Fernández, and Z. Geem. A critical review of robustness in power grids using complex networks concepts. *Energies*, 8(9):9211–9265, 2015.

[119] P. Curry, R. Curry, and J. Churchman. Mini Metro, 2015. https://dinopoloclub.com/minimetro/.

[120] S. L. Cutter, L. Barnes, M. Berry, C. Burton, E. Evans, E. Tate, and J. Webb. A place-based model for understanding community resilience to natural disasters. *Global Environmental Change*, 18(4):598–606, 2008.

[121] D. M. Cvetkovic, M. Doob, H. Sachs, et al. *Spectra of Graphs*, volume 10. Academic Press, New York City, NY, USA, 1980.

[122] B. R. da Cunha and S. Gonçalves. Performance of attack strategies on modular networks. *Journal of Complex Networks*, 5(6):913–923, 2017.

[123] L. Dall'Asta, A. Barrat, M. Barthélemy, and A. Vespignani. Vulnerability of weighted networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(04):P04006, 2006.

[124] M. M. Danziger, A. Bashan, Y. Berezin, and S. Havlin. Percolation and cascade dynamics of spatial networks with partial dependency. *Journal of Complex Networks*, 2(4):460–474, 2014.

[125] L. Dartnell, E. Simeonidis, M. Hubank, S. Tsoka, I. D. L. Bogle, and L. G. Papageorgiou. Robustness of the p53 network and biological hackers. *Febs Letters*, 579(14):3037–3042, 2005.

[126] W. Dáttilo, N. Lara-Rodríguez, P. Jordano, P. R. Guimarães Jr, J. N. Thompson, R. J. Marquis, L. P. Medeiros, R. Ortiz-Pulido, M. A. Marcos-García, and V. Rico-Gray. Unravelling Darwin's entangled bank: Architecture and robustness of mutualistic networks with multiple interaction types. *Proceedings of the Royal Society B: Biological Sciences*, 283(1843):20161564, 2016.

[127] T. N. Davis, M. S. Urdea, F. R. Masiarz, and J. Thorner. Isolation of the yeast calmodulin gene: Calmodulin is an essential protein. *Cell*, 47(3):423–431, 1986.

[128] T. M. Dawson and V. L. Dawson. Molecular pathways of neurodegeneration in Parkinson's disease. *Science*, 302(5646):819–822, 2003.

[129] C. de Abreu, B. R. da Cunha, and S. Gonçalves. Empirical determination of the optimum attack for fragmentation of modular networks. *Arxiv Preprint Arxiv: 1810.05962*, 2018.

[130] N. M. M. De Abreu. Old and new results on algebraic connectivity of graphs. *Linear Algebra and its Applications*, 423(1):53–73, 2007.

[131] M. H. de Badyn, A. Chapman, and M. Mesbahi. Network entropy: A system-theoretic perspective. In *54th Annual Conference on Decision and Control*, pages 5512–5517. IEEE, 2015.

[132] O. De Bandt and P. Hartmann. Systemic risk: A survey. 2000.

[133] M. De Domenico and A. Arenas. Modeling structure and resilience of the dark network. *Physical Review E*, 95(2):022313, 2017.

[134] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas. Ranking in interconnected multilayer networks reveals versatile nodes. *Nature Communications*, 6:6868, 2015.

[135] A. De la Fuente. From 'differential expression' to 'differential networking' — identification of dysfunctional regulatory networks in diseases. *Trends in Genetics*, 26(7):326–333, 2010.

[136] M. De Lucia, M. Bottaccio, M. Montuori, and L. Pietronero. Topological approach to neural complexity. *Physical Review E*, 71(1):016114, 2005.

[137] M. Dehmer and A. Mowshowitz. A history of graph entropy measures. *Information Sciences*, 181(1):57–78, 2011.

[138] G. del Rio, D. Koschützki, and G. Coello. How to identify essential genes from molecular networks? *BMC Systems Biology*, 3(1):102, 2009.

[139] L. Demetrius and T. Manke. Robustness and network evolution — An entropic principle. *Physica A: Statistical Mechanics and its Applications*, 346(3):682–696, 2005.

[140] Y. Deng, J. Wu, M. Qi, and Y. Tan. Optimal disintegration strategy in spatial networks with disintegration circle model. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(6):061102, 2019.

[141] K. Dhanyalakshmi, M. B. Naika, R. Sajeevan, O. K. Mathew, K. M. Shafi, R. Sowdhamini, and K. N. Nataraja. An approach to function annotation for Proteins of Unknown Function (PUFs) in the transcriptome of Indian mulberry. *PLoS One*, 11(3):e0151323, 2016.

[142] P. Diaconis and M. Shahshahani. Generating a random permutation with random transpositions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57(2):159–179, 1981.

[143] P. J. Dickinson, H. Bunke, A. Dadej, and M. Kraetzl. On graphs with unique node labels. In *International Workshop on Graph-Based Representations in Pattern Recognition*, pages 13–23. Springer, 2003.

[144] K. M. Dipple, J. K. Phelan, and E. R. McCabe. Consequences of complexity within biological networks: Robustness and health, or vulnerability and disease. *Molecular Genetics and Metabolism*, 74(1–2):45–50, 2001.

[145] P. S. Dodds, D. J. Watts, and C. F. Sabel. Information exchange and the robustness of organizational networks. *Proceedings of the National Academy of Sciences of the United States of America*, 100(21):12516–12521, 2003.

[146] C. Donnat, S. Holmes, et al. Tracking network dynamics: A survey using graph distances. *The Annals of Applied Statistics*, 12(2):971–1012, 2018.

[147] P. Dorato. A historical review of robust control. *IEEE Control Systems Magazine*, 7(2):44–47, 1987.

[148] T. Downarowicz. *Entropy in Dynamical Systems*, volume 18. Cambridge University Press, Cambridge, United Kingdom, 2011.

[149] J. Doyle and J. M. Carlson. Power laws, highly optimized tolerance, and generalized source coding. *Physical Review Letters*, 84(24):5656, 2000.

[150] J. C. Doyle, D. L. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka, and W. Willinger. The "robust yet fragile" nature of the internet. *Proceedings of the National Academy of Sciences of the United States of America*, 102(41):14497–14502, 2005.

[151] D. J. Du Plessis, N. Nouwen, and A. J. Driessen. The sec translocase. *Biochimica et Biophysica Acta – Biomembranes*, 1808(3):851–865, 2011.

[152] J. A. Dunne, R. J. Williams, and N. D. Martinez. Network structure and biodiversity loss in food webs: Robustness increases with connectance. *Ecology Letters*, 5(4):558–567, 2002.

[153] J. A. Dunne, R. J. Williams, and N. D. Martinez. Network structure and robustness of marine food webs. *Marine Ecology Progress Series*, 273:291–302, 2004.

[154] G. M. Edelman and J. A. Gally. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24):13763–13768, 2001.

[155] T. Efferth, V. B. Konkimalla, Y.-F. Wang, A. Sauerbrey, S. Meinhardt, F. Zintl, J. Mattern, and M. Volm. Prediction of broad spectrum resistance of tumors towards anticancer drugs. *Clinical Cancer Research*, 14(8):2405–2412, 2008.

[156] H. A. Eiselt and C.-L. Sandblom. *Integer Programming and Network Models*. Springer, Berlin, Germany, 2000.

[157] L. Eisenberg and T. H. Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

[158] W. Ellens, F. Spieksma, P. Van Mieghem, A. Jamakovic, and R. Kooij. Effective graph resistance. *Linear Algebra and its Applications*, 435(10):2491–2506, 2011.

[159] F. Emmert-Streib and M. Dehmer. Information processing in the transcriptional regulatory network of yeast: Functional robustness. *BMC Systems Biology*, 3(1):35, 2009.

[160] T. Erlebach, A. Hall, L. Moonen, A. Panconesi, F. Spieksma, and D. Vukadinović. Robustness of the internet at the topology and routing level. In *Dependable Systems: Software, Computing, Networks*, pages 260–274. Springer, 2006.

[161] E. Estrada. Network robustness to targeted attacks. The interplay of expansibility and degree distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, 52(4):563–574, 2006.

[162] E. Estrada. Protein bipartivity and essentiality in the yeast protein- protein interaction network. *Journal of Proteome Research*, 5(9):2177–2184, 2006.

[163] E. Estrada. Food webs robustness to biodiversity loss: The roles of connectance, expansibility and degree distribution. *Journal of Theoretical Biology*, 244(2):296–307, 2007.

[164] E. Estrada. The communicability distance in graphs. *Linear Algebra and its Applications*, 436(11):4317–4328, 2012.

[165] E. Estrada. Returnability as a criterion of disequilibrium in atmospheric reactions networks. *Journal of Mathematical Chemistry*, 50(6):1363–1372, 2012.

[166] E. Estrada and N. Hatano. Communicability in complex networks. *Physical Review E*, 77(3):036111, 2008.

[167] E. Estrada and N. Hatano. Communicability angle and the spatial efficiency of networks. *SIAM Review*, 58(4):692–715, 2016.

[168] E. Estrada and D. J. Higham. Network properties revealed through matrix functions. *SIAM Review*, 52(4):696–714, 2010.

[169] E. Estrada and P. A. Knight. *A First Course in Network Theory*. Oxford University Press, Oxford, United Kingdom, 2015.

[170] E. Estrada and J. A. Rodriguez-Velazquez. Subgraph centrality in complex networks. *Physical Review E*, 71(5):056103, 2005.

[171] E. Estrada and J. A. Rodríguez-Velázquez. Subgraph centrality and clustering in complex hyper-networks. *Physica A: Statistical Mechanics and its Applications*, 364:581–594, 2006.

[172] E. Estrada, E. Vargas-Estrada, and H. Ando. Communicability angles reveal critical edges for network consensus dynamics. *Physical Review E*, 92(5):052809, 2015.

[173] A. Fabregat, K. Sidiropoulos, P. Garapati, M. Gillespie, K. Hausmann, R. Haw, B. Jassal, S. Jupe, F. Korninger, S. McKay, et al. The reactome pathway knowledgebase. *Nucleic Acids Research*, 44(D1):D481–D487, 2016.

[174] G. Fagiolo. Clustering in complex directed networks. *Physical Review E*, 76(2):026107, 2007.

[175] F. W. Fairman. *Linear Control Theory: The State Space Approach*. John Wiley & Sons, New York City, NY, USA, 1998.

[176] R. Z. Farahani, E. Miandoabchi, W. Y. Szeto, and H. Rashidi. A review of urban transportation network design problems. *European Journal of Operational Research*, 229(2):281–302, 2013.

[177] M.-A. Félix and M. Barkoulas. Pervasive robustness in biological systems. *Nature Reviews Genetics*, 16(8):483–496, 2015.

[178] F. Feng and L. Wang. Robustness measure of China's railway network topology using relative entropy. *Discrete Dynamics in Nature and Society*, 2013:391709, 2013.

[179] P. B. Fernandes. Technological advances in high-throughput screening. *Current Opinion in Chemical Biology*, 2(5):597–603, 1998.

[180] T. Feyessa and M. Bikdash. Measuring nodal contribution to global network robustness. In *2011 Proceedings of IEEE Southeastcon*, pages 131–135. IEEE, 2011.

[181] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23(2):298–305, 1973.

[182] F. V. Fomin, D. Lokshtanov, V. Raman, S. Saurabh, and B. Raghavendra Rao. Faster algorithms for finding and counting subgraphs. *Journal of Computer and System Sciences*, 78(3):698–706, 2012.

[183] K. A. Ford. Refinement, reduction, and replacement of animal toxicity tests by computational methods. *Ilar Journal*, 57(2):226–233, 2016.

[184] B. K. Fosdick, D. B. Larremore, J. Nishimura, and J. Ugander. Configuring random graph models with fixed degree sequences. *SIAM Review*, 60(2):315–355, 2018.

[185] D. P. Francis and K. Raimond. A practical streaming approximate matrix multiplication algorithm. *Journal of King Saud University-Computer and Information Sciences*, 2018.

[186] T. L. Frantz, M. Cataldo, and K. M. Carley. Robustness of centrality measures under uncertainty Examining the role of network topology. *Computational and Mathematical Organization Theory*, 15(4):303, 2009.

[187] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40:35–41, 1977.

[188] V. P. Freire and D. R. Figueiredo. Ranking in collaboration networks using a group based metric. *Journal of the Brazilian Computer Society*, 17(4):255–266, 2011.

[189] R. Fritsche-Guenther, F. Witzel, A. Sieber, R. Herr, N. Schmidt, S. Braun, T. Brummer, C. Sers, and N. Blüthgen. Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. *Molecular Systems Biology*, 7:489, 2011.

[190] C. Fu, S. Deng, G. Jin, X. Wang, and Z.-G. Yu. Bayesian network model for identification of pathways by integrating protein interaction with genetic interaction data. *BMC Systems Biology*, 11(4):81, 2017.

[191] Z. Galias. Probabilistic model for studying blackouts in power networks. *Ieee Journal on Emerging and Selected Topics in Circuits and Systems*, 7(2):218–227, 2017.

[192] J. Gao, X. Liu, D. Li, and S. Havlin. Recent progress on the resilience of complex networks. *Energies*, 8(10):12187–12210, 2015.

[193] X. Gao, K. Li, and B. Chen. Invulnerability measure of a military heterogeneous network based on network structure entropy. *IEEE Access*, 6:6700–6708, 2017.

[194] P. Gaspard. Brownian motion, dynamical randomness and irreversibility. *New Journal of Physics*, 7(1):77, 2005.

[195] L. F. Gay and S. K. Sinha. Resilience of civil infrastructure systems: Literature review for improved asset management. *International Journal of Critical Infrastructures*, 9(4):330–350, 2013.

[196] P. M. Gleiss, P. F. Stadler, A. Wagner, and D. A. Fell. Relevant cycles in chemical reaction networks. *Advances in Complex Systems*, 4(02n03):207–226, 2001.

[197] A. Golan, G. Judge, and L. Karp. A maximum entropy approach to estimation and inference in dynamic models *or* counting fish in the sea using maximum entropy. *Journal of Economic Dynamics and Control*, 20(4):559–582, 1996.

[198] E. Goles, M. Montalva, and G. A. Ruz. Deconstruction and dynamical robustness of regulatory networks: Application to the yeast cell cycle networks. *Bulletin of Mathematical Biology*, 75(6):939–966, 2013.

[199] J. Gómez-Gardenes and V. Latora. Entropy rate of diffusion processes on complex networks. *Physical Review E*, 78(6):065102, 2008.

[200] J. Goñi, A. Avena-Koenigsberger, N. V. de Mendizabal, M. P. van den Heuvel, R. F. Betzel, and O. Sporns. Exploring the morphospace of communication efficiency in complex networks. *PLoS One*, 8(3):e58070, 2013.

[201] M. W. Gonzalez and M. G. Kann. Chapter 4: Protein Interactions and Disease. *PLoS Computational Biology*, 8(12):002819, 2012.

[202] N. Goodman. The distribution of the determinant of a complex wishart distributed matrix. *The Annals of Mathematical Statistics*, 34(1):178–180, 1963.

[203] A. N. Gorban and O. Radulescu. Dynamical robustness of biological networks with hierarchical distribution of time scales. *IET Systems Biology*, 1(4):238–246, 2007.

[204] W. Gou and Z. Jin. How heterogeneous susceptibility and recovery rates affect the spread of epidemics on networks. *Infectious Disease Modelling*, 2(3):353–367, 2017.

[205] C. W. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438, 1969.

[206] R. M. Gray. *Entropy and Information Theory*. Springer-Verlag, Berlin, Germany, 2011.

[207] R. Greco, A. Di Nardo, and G. Santonastaso. Resilience and entropy as indices of robustness of water distribution networks. *Journal of Hydroinformatics*, 14(3):761–771, 2012.

[208] T. Gross and B. Blasius. Adaptive coevolutionary networks: A review. *Journal of the Royal Society Interface*, 5(20):259–271, 2007.

[209] T. W. Grunberg and D. F. Gayme. Performance measures for linear oscillator networks over arbitrary graphs. *IEEE Transactions on Control of Network Systems*, 5(1):456–468, 2016.

[210] D. Grytskyy, T. Tetzlaff, M. Diesmann, and M. Helias. A unified view on weakly correlated recurrent networks. *Frontiers in Computational Neuroscience*, 7:131, 2013.

[211] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamics Systems, and Bifurcations of Vector Fields*. Springer, Berlin, Germany, 6th edition, 2002.

[212] J.-L. Guerrero-Cusumano. A measure of total variability for the multivariate t distribution with applications to finance. *Information Sciences*, 92(1–4):47–63, 1996.

[213] H. Gui, M. Li, P. C. Sham, and S. S. Cherny. Comparisons of seven algorithms for pathway analysis using the WTCCC Crohn's disease dataset. *BMC Research Notes*, 4(1):386, 2011.

[214] L. Hakes, J. W. Pinney, D. L. Robertson, and S. C. Lovell. Protein–protein interaction networks and biology — What's the connection? *Nature Biotechnology*, 26(1):69–72, 2008.

[215] A. G. Haldane and R. M. May. Systemic risk in banking ecosystems. *Nature*, 469(7330):351–355, 2011.

[216] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, volume 222. Springer-Verlag, Berlin, Germany, 2015.

[217] C. Han and L. Liu. Topological vulnerability of subway networks in China. In *International Conference on Management and Service Science*, pages 1–4. IEEE, 2009.

[218] J. T. Hancock. *Cell Signalling*. Oxford University Press, Oxford, United Kingdom, 2017.

[219] I. Hanukoglu. Electron transfer proteins of cytochrome P450 systems. In *Advances in molecular and cell biology*, volume 14, pages 29–56. Elsevier, 1996.

[220] Y.-H. Hao, J.-L. Song, Y. Lin, and J.-H. Han. Vulnerability of Complex Networks under Approximate Longest Chain Attack Strategies. In *ITM Web of Conferences*, volume 7, page 01011. EDP Sciences, 2016.

[221] T. Hase, H. Tanaka, Y. Suzuki, S. Nakagawa, and H. Kitano. Structure of protein interaction networks and their implications on drug design. *PLoS Computational Biology*, 5(10):e1000550, 2009.

[222] A. Hasman. A critical review of contagion risk in banking. *Journal of Economic Surveys*, 27(5):978–995, 2013.

[223] S. He, S. Li, and H. Ma. Effect of edge removal on topological and functional robustness of complex networks. *Physica A: Statistical Mechanics and its Applications*, 388(11):2243–2253, 2009.

[224] X. He and J. Zhang. Why do hubs tend to be essential in protein networks? *PLoS Genetics*, 2(6):e88, 2006.

[225] Z. He, S. Liu, and M. Zhan. Dynamical robustness analysis of weighted complex networks. *Physica A: Statistical Mechanics and its Applications*, 392(18):4181–4191, 2013.

[226] L. S. Heath and N. Parikh. Generating random graphs with tunable clustering coefficients. *Physica A: Statistical Mechanics and its Applications*, 390(23-24):4577–4587, 2011.

[227] M. R. Heath, D. C. Speirs, and J. H. Steele. Understanding patterns and processes in models of trophic cascades. *Ecology Letters*, 17(1):101–114, 2014.

[228] D. Helbing. Globally networked risks and how to respond. *Nature*, 497(7447):51–59, 2013.

[229] J. P. Hendrick and F.-U. Hartl. Molecular chaperone functions of heat-shock proteins. *Annual Review of Biochemistry*, 62(1):349–384, 1993.

[230] H. Hermjakob, L. Montecchi-Palazzi, C. Lewington, S. Mudali, S. Kerrien, S. Orchard, M. Vingron, B. Roechert, P. Roepstorff, A. Valencia, et al. IntAct: An open source molecular interaction database. *Nucleic Acids Research*, 32(suppl 1):D452–D455, 2004.

[231] H. J. Herrmann, C. M. Schneider, A. A. Moreira, J. S. Andrade Jr, and S. Havlin. Onion-like network topology enhances robustness against malicious attacks. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(01):P01027, 2011.

[232] R. P. Hertzberg and A. J. Pope. High-throughput screening: New technology for the 21st century. *Current Opinion in Chemical Biology*, 4(4):445–451, 2000.

[233] A. M. Hinsby, L. Kiemer, E. O. Karlberg, K. Lage, A. Fausbøll, A. S. Juncker, J. S. Andersen, M. Mann, and S. Brunak. A wiring of the human nucleolus. *Molecular Cell*, 22(2):285–295, 2006.

[234] R. R. Hoffman and P. A. Hancock. Measuring resilience. *Human Factors*, 59(4):564–581, 2017.

[235] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han. Attack vulnerability of complex networks. *Physical Review E*, 65(5):056109, 2002.

[236] P. Holme and J. Saramäki. Temporal networks. *Physics Reports*, 519(3):97–125, 2012.

[237] Å. J. Holmgren. Using graph models to analyze the vulnerability of electric power networks. *Risk Analysis*, 26(4):955–969, 2006.

[238] C. Hong, N. He, O. Lordan, B.-Y. Liang, and N.-Y. Yin. Efficient calculation of the robustness measure R for complex networks. *Physica A: Statistical Mechanics and its Applications*, 478:63–68, 2017.

[239] X. Hong-lin, Y. Han-bing, G. Cui-fang, and Z. Ping. Social network analysis based on network motifs. *Journal of Applied Mathematics*, 2014, 2014.

[240] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[241] A. L. Hopkins. Network pharmacology. *Nature Biotechnology*, 25(10):1110–1111, 2007.

[242] A. L. Hopkins. Network pharmacology: The next paradigm in drug discovery. *Nature Chemical Biology*, 4(11):682–690, 2008.

[243] M. Hossain, S. Alam, T. Rees, and H. Abbass. Australian airport network robustness analysis: A complex network approach. In *Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia*, 2013.

[244] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez. A review of definitions and measures of system resilience. *Reliability Engineering & System Safety*, 145:47–61, 2016.

[245] H. M. Hsieh-Li, J.-G. Chang, Y.-J. Jong, M.-H. Wu, N. M. Wang, C. H. Tsai, and H. Li. A mouse model for spinal muscular atrophy. *Nature Genetics*, 24(1):66–70, 2000.

[246] J. X. Hu, C. E. Thomas, and S. Brunak. Network biology concepts in complex disease comorbidities. *Nature Reviews Genetics*, 17(10):615–629, 2016.

[247] Y. Hu, J. Trousdale, K. Josić, and E. Shea-Brown. Local paths to global coherence: Cutting networks down to size. *Physical Review E*, 89(3):032802, 2014.

[248] Y. Huang, J.-F. Martínez, J. Sendra, and L. López. Resilientwireless sensor networks using topology control: A review. *Sensors*, 15(10):24735–24770, 2015.

[249] Z.-Y. Huang, M.-K. Kim, T.-H. Kim-Han, Z. K. Indik, and A. D. Schreiber. Effect of locally administered Syk siRNA on allergen-induced arthritis and asthma. *Molecular immunology*, 53(1-2):52–59, 2013.

[250] S.-G. Hwang. Cauchy's interlace theorem for eigenvalues of Hermitian matrices. *The American Mathematical Monthly*, 111(2):157–159, 2004.

[251] T. Ideker and N. J. Krogan. Differential network biology. *Molecular Systems Biology*, 8(1), 2012.

[252] M. Isalan, C. Lemerle, K. Michalodimitrakis, C. Horn, P. Beltrao, E. Raineri, M. Garriga-Canut, and L. Serrano. Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, 452(7189):840, 2008.

[253] S. Iyer, T. Killingback, B. Sundaram, and Z. Wang. Attack robustness and centrality of complex networks. *PLoS One*, 8(4):e59613, 2013.

[254] M. O. Jackson. Centrality, gossip, and diffusion of information in networks. In *Proceedings of the fifteenth ACM conference on Economics and computation*, pages 1–1. ACM, 2014.

[255] E. W. Jacobsen and G. Cedersund. Structural robustness of biochemical network models-with application to the oscillatory metabolism of activated neutrophils. *Iet Systems Biology*, 2(1):39–47, 2008.

[256] T. Jacquemin and R. Jiang. Walking on a tissue-specific disease-protein-complex heterogeneous network for the discovery of disease-related protein complexes. *BioMed Research International*, 2013, 2013.

[257] M. Jaffe, A. Dziulko, J. D. Smith, R. P. St Onge, S. F. Levy, and G. Sherlock. Improved discovery of genetic interactions using CRISPRiSeq across multiple environments. *Genome Research*, 29(4):668–681, 2019.

[258] A. Jamakovic, R. Kooij, P. Van Mieghem, and E. R. van Dam. Robustness of networks against viruses: The role of the spectral radius. In *2006 Symposium on Communications and Vehicular Technology*, pages 35–38. IEEE, 2006.

[259] A. Jamakovic and S. Uhlig. Influence of the network structure on robustness. In *15th IEEE International Conference on Networks*, pages 278–283. IEEE, 2007.

[260] A. Jamakovic and S. Uhlig. On the relationship between the algebraic connectivity and graph's robustness to node and link failures. In *3rd EuroNGI Conference on Next Generation Internet Networks*, pages 96–102. IEEE, 2007.

[261] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, 2001.

[262] H. Jeong and B.-J. Yoon. SEQUOIA: significance enhanced network querying through context-sensitive random walk and minimization of network conductance. *BMC Systems Biology*, 11(3):20, 2017.

[263] C. Jernberg, S. Löfmark, C. Edlund, and J. K. Jansson. Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *The ISME Journal*, 1(1):56–66, 2007.

[264] L. G. Jeub, P. Balachandran, M. A. Porter, P. J. Mucha, and M. W. Mahoney. Think locally, act locally: Detection of small, medium-sized, and large communities in large networks. *Physical Review E*, 91(1):012821, 2015.

[265] C. Johansen, J. Horney, and I. Tien. Metrics for evaluating and improving community resilience. *Journal of Infrastructure Systems*, 23(2):04016032, 2016.

[266] N. J. Joshi, G. Tononi, and C. Koch. The minimal complexity of adapting agents increases with fitness. *PLoS Computational Biology*, 9(7):e1003111, 2013.

[267] K. Juszczyszyn, K. Musiał, P. Kazienko, and B. Gabrys. Temporal changes in local topology of an email-based social network. *Computing and Informatics*, 28(6):763–779, 2012.

[268] M. Kaiser and C. C. Hilgetag. Edge vulnerability in neural and metabolic networks. *Biological Cybernetics*, 90(5):311–317, 2004.

[269] C. N. Kaiser-Bunbury, S. Muff, J. Memmott, C. B. Müller, and A. Caflisch. The robustness of pollination networks to the loss of species and interactions: A quantitative approach incorporating pollinator behaviour. *Ecology Letters*, 13(4):442–452, 2010.

[270] M. Kamalahmadi and M. M. Parast. A review of the literature on the principles of enterprise and supply chain resilience: Major findings and directions for future research. *International Journal of Production Economics*, 171:116–133, 2016.

[271] N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan. Entropies for detection of epilepsy in EEG. *Computer Methods and Programs in Biomedicine*, 80(3):187–194, 2005.

[272] K. Kapoor, D. Sharma, and J. Srivastava. Weighted node degree centrality for hypergraphs. In *IEEE Network Science Workshop*, pages 152–155. IEEE, 2013.

[273] F. Karimi and P. Holme. Threshold model of cascades in empirical temporal networks. *Physica A: Statistical Mechanics and its Applications*, 392(16):3476–3483, 2013.

[274] B. Karrer, E. Levina, and M. E. Newman. Robustness of community structure in networks. *Physical Review E*, 77(4):046119, 2008.

[275] B. Karrer and M. E. Newman. Random graphs containing arbitrary distributions of subgraphs. *Physical Review E*, 82(6):066118, 2010.

[276] E. Katzav and I. P. Castillo. Large deviations of the smallest eigenvalue of the Wishart-Laguerre ensemble. *Physical Review E*, 82(4):040104, 2010.

[277] M. Keeling. The implications of network structure for epidemic dynamics. *Theoretical Population Biology*, 67(1):1–8, 2005.

[278] F. Keller and N. Di Stefano. The Robustness of Musical Language: A Perspective from Complex Systems Theory. In *Biological Robustness*, pages 207–217. Springer, 2018.

[279] H. Kesten. *Percolation Theory for Mathematicians*, volume 423. Springer, Berlin, Germany, 1982.

[280] M. Khansari, A. Kaveh, Z. Heshmati, and M. A. Motlaq. Centrality measures for immunization of weighted networks. *Network Biology*, 6(1):12–27, 2016.

[281] A. Kikuchi. Roles of Axin in the Wnt signalling pathway. *Cellular Signalling*, 11(11):777–788, 1999.

[282] S.-W. Kim, W. Suda, S. Kim, K. Oshima, S. Fukuda, H. Ohno, H. Morita, and M. Hattori. Robustness of gut microbiota of healthy adults in response to probiotic intervention revealed by high-throughput pyrosequencing. *DNA Research*, 20(3):241–253, 2013.

[283] D. Kirouac and M. Onsum. Using network biology to bridge pharmacokinetics and pharmacodynamics in oncology. *Cpt: Pharmacometrics & Systems Pharmacology*, 2(9):1–7, 2013.

[284] H. Kitano. Computational systems biology. *Nature*, 420(6912):206–210, 2002.

[285] H. Kitano. Biological robustness. *Nature Reviews Genetics*, 5(11):826–837, 2004.

[286] H. Kitano. Cancer as a robust system: Implications for anticancer therapy. *Nature Reviews Cancer*, 4(3):227–235, 2004.

[287] H. Kitano. A robustness-based approach to systems-oriented drug design. *Nature Reviews Drug Discovery*, 6(3):202–210, 2007.

[288] H. Kitano. Towards a theory of biological robustness. *Molecular Systems Biology*, 3(1), 2007.

[289] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. Multilayer networks. *Journal of Complex Networks*, 2(3):203–271, 2014.

[290] S. Klamt, U.-U. Haus, and F. Theis. Hypergraphs and cellular networks. *PLoS Computational Biology*, 5(5):e1000385, 2009.

[291] T. Kleffmann, D. Russenberger, A. von Zychlinski, W. Christopher, K. Sjölander, W. Gruissem, and S. Baginsky. The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. *Current Biology*, 14(5):354–362, 2004.

[292] D. J. Klein and M. Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993.

[293] J. M. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845, 2000.

[294] Y. Koç, M. Warnier, R. E. Kooij, and F. M. Brazier. An entropy-based metric to quantify the robustness of power grids against cascading failures. *Safety Science*, 59:126–134, 2013.

[295] R. Koetter and M. Médard. An algebraic approach to network coding. *IEEE/ACM Transactions on Networking (TON)*, 11(5):782–795, 2003.

[296] A. N. Kolmogorov. Entropy per unit time as a metric invariant of automorphism. *Doklady of Russian Academy of Sciences*, 124:754–755, 1959.

[297] K. Komurov, M. A. White, and P. T. Ram. Use of data-biased random walks on graphs for the retrieval of context-specific networks from genomic data. *PLoS Computational Biology*, 6(8):e1000889, 2010.

[298] D. C. Krakauer and J. B. Plotkin. Redundancy, antiredundancy, and the robustness of genomes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(3):1405–1409, 2002.

[299] A. Kumar, M. Gopalswamy, A. Wolf, D. J. Brockwell, M. Hatzfeld, and J. Balbach. Phosphorylation-induced unfolding regulates p19INK4d during the human cell cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 115(13):3344–3349, 2018.

[300] K. Kümpornsin, C. Modchang, A. Heinberg, E. H. Ekland, P. Jirawatcharadech, P. Chobson, N. Suwanakitti, S. Chaotheing, P. Wilairat, K. W. Deitsch, et al. Origin of robustness in generating drug-resistant malaria parasites. *Molecular Biology and Evolution*, 31(7):1649–1660, 2014.

[301] Y.-K. Kwon and K.-H. Cho. Analysis of feedback loops and robustness in network evolution based on boolean models. *BMC Bioinformatics*, 8(1):430, 2007.

[302] Y.-K. Kwon and K.-H. Cho. Quantitative analysis of robustness and fragility in biological networks based on feedback dynamics. *Bioinformatics*, 24(7):987–994, 2008.

[303] A. Kyrillidis, M. Vlachos, and A. Zouzias. Approximate matrix multiplication with application to linear embeddings. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2182–2186. IEEE, 2014.

[304] R. Lambiotte, R. Sinatra, J.-C. Delvenne, T. S. Evans, M. Barahona, and V. Latora. Flow graphs: Interweaving dynamics and structure. *Physical Review E*, 84(1):017102, 2011.

[305] R. F. Lance, P. Bailey, D. L. Lindsay, and N. S. Cobb. Precipitation and the robustness of a plant and flower-visiting insect network in a xeric ecosystem. *Journal of Arid Environments*, 144:48–59, 2017.

[306] A. Larhlimi, S. Blachon, J. Selbig, and Z. Nikoloski. Robustness of metabolic networks: A review of existing definitions. *Biosystems*, 106(1):1–8, 2011.

[307] S. LaRocca, J. Johansson, H. Hassel, and S. Guikema. Topological performance measures as surrogates for physical flow models for risk and vulnerability analysis for electric power systems. *Risk Analysis*, 35(4):608–623, 2015.

[308] V. Latora and M. Marchiori. Vulnerability and protection of infrastructure networks. *Physical Review E*, 71(1):015103, 2005.

[309] J. H. Lawton and V. K. Brown. Redundancy in ecosystems. In *Biodiversity and ecosystem function*, pages 255–270. Springer, 1994.

[310] P. Lax. *Functional Analysis*. John Wiley & Sons, New York City, NY, USA, 2002.

[311] M. Ledoux. *The Concentration of Measure Phenomenon*. American Mathematical Society, Providence, RI, USA, 2001.

[312] A. Lee, J. O'Rourke, M. C. De Ungria, B. Robertson, G. Daskalopoulos, and M. F. Dixon. A standardized mouse model of Helicobacter pylori infection: Introducing the Sydney strain. *Gastroenterology*, 112(4):1386–1397, 1997.

[313] S. H. Lee and P. Holme. Exploring maps with greedy navigators. *Physical Review Letters*, 108(12):128701, 2012.

[314] S. H. Lee, P.-J. Kim, and H. Jeong. Statistical properties of sampled networks. *Physical Review E*, 73(1):016102, 2006.

[315] J. Lehár, A. Krueger, G. Zimmermann, and A. Borisy. High-order combination effects and biological robustness. *Molecular Systems Biology*, 4(1), 2008.

[316] M. D. Leiserson, F. Vandin, H.-T. Wu, J. R. Dobson, J. V. Eldridge, J. L. Thomas, A. Papoutsaki, Y. Kim, B. Niu, M. McLellan, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nature Genetics*, 47(2):106–114, 2015.

[317] A. Lesne. Robustness: Confronting lessons from physics and biology. *Biological Reviews*, 83(4):509–532, 2008.

[318] C. Li, H. Wang, and P. Van Mieghem. Epidemic threshold in directed networks. *Physical Review E*, 88(6):062802, 2013.

[319] D. Li, B. Fu, Y. Wang, G. Lu, Y. Berezin, H. E. Stanley, and S. Havlin. Percolation transition in dynamical traffic network with evolving critical bottlenecks. *Proceedings of the National Academy of Sciences of the United States of America*, 112(3):669–672, 2015.

[320] D. Li, J. Li, S. Ouyang, J. Wang, S. Wu, P. Wan, Y. Zhu, X. Xu, and F. He. Protein interaction networks of Saccharomyces cerevisiae, Caenorhabditis elegans and Drosophila melanogaster: Large-scale organization and robustness. *Proteomics*, 6(2):456–461, 2006.

[321] H.-J. Li, H. Wang, and L. Chen. Measuring robustness of community structure in complex networks. *Europhysics Letters*, 108(6):68009, 2015.

[322] M. Li, H. Zhang, J.-x. Wang, and Y. Pan. A new essential protein discovery method based on the integration of protein-protein interaction and gene expression data. *BMC Systems Biology*, 6(1):15, 2012.

[323] Y. Li, G. Dwivedi, W. Huang, M. L. Kemp, and Y. Yi. Quantification of degeneracy in biological systems for characterization of functional interactions between modules. *Journal of Theoretical biology*, 302:29–38, 2012.

[324] Y. Li and Y. Yi. Systematic measures of biological networks II: Degeneracy, complexity, and robustness. *Communications on Pure and Applied Mathematics*, 69(10):1952–1983, 2016.

[325] Z. Liang, K. C. Yuen, and J. Guo. Optimal proportional reinsurance and investment in a stock market with Ornstein–Uhlenbeck process. *Insurance: Mathematics and Economics*, 49(2):207–215, 2011.

[326] E. Liberty. Simple and deterministic matrix sketching. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 581–588. ACM, 2013.

[327] L. Licata, L. Briganti, D. Peluso, L. Perfetto, M. Iannuccelli, E. Galeota, F. Sacco, A. Palma, A. P. Nardozza, E. Santonico, et al. MINT, the molecular interaction database: 2012 update. *Nucleic Acids Research*, 40(D1):D857–D861, 2012.

[328] G. Lima-Mendez and J. van Helden. The powerful law of the power law and other myths in network biology. *Molecular BioSystems*, 5(12):1482, 2009.

[329] Y.-y. Lin, Y. Qi, J.-y. Lu, X. Pan, D. S. Yuan, Y. Zhao, J. S. Bader, and J. D. Boeke. A comprehensive synthetic genetic interaction network governing yeast histone acetylation and deacetylation. *Genes & Development*, 22(15):2062–2074, 2008.

[330] R. G. Little. Toward more robust infrastructure: Observations on improving the resilience and reliability of critical systems. In *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the*, page 9. IEEE, 2003.

[331] B. Liu, Z. Li, X. Chen, Y. Huang, and X. Liu. Recognition and vulnerability analysis of key nodes in power grid based on complex network centrality. *Ieee Transactions on Circuits and Systems II: Express Briefs*, 65(3):346–350, 2017.

[332] J.-G. Liu, J.-H. Lin, Q. Guo, and T. Zhou. Locating influential nodes via dynamics-sensitive centrality. *Scientific Reports*, 6:21380, 2016.

[333] S.-Y. Liu, C. Li, Y.-P. Feng, and G. Rong. Clustering structure and logistics: A new framework for supply network analysis. *Chemical Engineering Research and Design*, 91(8):1383–1389, 2013.

[334] J. Lloyd-Price, A. Gupta, and A. S. Ribeiro. Robustness and information propagation in attractors of random boolean networks. *PloS One*, 7(7):e42018, 2012.

[335] O. Lordan, J. M. Sallan, P. Simo, and D. Gonzalez-Prieto. Robustness of the air transport network. *Transportation Research Part E: Logistics and Transportation Review*, 68:155–163, 2014.

[336] M. Loreau. Does functional redundancy exist? *Oikos*, 104(3):606–611, 2004.

[337] V. H. Louzada, F. Daolio, H. J. Herrmann, and M. Tomassini. Smart rewiring for network robustness. *Journal of Complex Networks*, 1(2):150–159, 2013.

[338] L. Lü, D. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou. Vital nodes identification in complex networks. *Physics Reports*, 650:1–63, 2016.

[339] Z.-M. Lu and X.-F. Li. Attack vulnerability of network controllability. *PLoS One*, 11(9):e0162289, 2016.

[340] M. A. Lzquierdo, A. G. v. d. Zee, J. B. Vermorken, P. v. d. Valk, J. A. Beliën, G. Giaccone, G. L. Scheffer, M. J. Flens, H. M. Pinedo, P. Kenemans, et al. Drug resistance-associated marker Lrp for prediction of response to chemotherapy and prognoses in advanced ovarian carcinoma. *Journal of the National Cancer Institute*, 87(16):1230–1237, 1995.

[341] H. Ma and H. Zhao. Drug target inference through pathway analysis of genomics data. *Advanced Drug Delivery Reviews*, 65(7):966–972, 2013.

[342] L. Ma, J. Liu, B. Duan, and M. Zhou. A theoretical estimation for the optimal network robustness measure R against malicious node attacks. *Europhysics Letters*, 111(2):28003, 2015.

[343] W. Ma, L. Lai, Q. Ouyang, and C. Tang. Robustness and modular design of the Drosophila segment polarity network. *Molecular Systems Biology*, 2(1), 2006.

[344] J. Macia and R. V. Solé. Distributed robustness in cellular networks: Insights from synthetic evolved circuits. *Journal of the Royal Society Interface*, 6(33):393–400, 2008.

[345] J. Macía, R. V. Solé, and S. F. Elena. The causes of epistasis in genetic networks. *Evolution: International Journal of Organic Evolution*, 66(2):586–596, 2012.

[346] E. Mackin and S. Patterson. Optimizing the coherence of composite networks. In *American Control Conference (ACC)*, pages 4334–4340. IEEE, 2017.

[347] K. Magis. Community resilience: An indicator of social sustainability. *Society and Natural Resources*, 23(5):401–416, 2010.

[348] S. N. Majumdar. Extreme eigenvalues of Wishart matrices: Application to entangled bipartite system. In G. Akemann, J. Baik, and P. Di Francesco, editors, *The Oxford Handbook of Random Matrix Theory*. Oxford University Press, Oxford, United Kingdom, 2011.

[349] M. Man, S. Liu, X. Chang, and M. Lu. The Biological Property of Synthetic Evolved Digital Circuits with ESD Immunity–Redundancy or Degeneracy? *Journal of Bionic Engineering*, 10(3):396–403, 2013.

[350] M. Man, Y. Zhang, G. Ma, K. Friston, and S. Liu. Quantification of degeneracy in hodgkin–huxley neurons on Newman–Watts small world network. *Journal of Theoretical Biology*, 402:62–74, 2016.

[351] R. Mani, R. P. St Onge, J. L. Hartman, G. Giaever, and F. P. Roth. Defining genetic interaction. *Proceedings of the National Academy of Sciences of the United States of America*, 105(9):3461–3466, 2008.

[352] T. Manke, L. Demetrius, and M. Vingron. Lethality and entropy of protein interaction networks. *Genome Informatics*, 16(1):159–163, 2005.

[353] T. Manke, L. Demetrius, and M. Vingron. An entropic characterization of protein interaction networks and cellular robustness. *Journal of The Royal Society Interface*, 3(11):843–850, 2006.

[354] P. C. Marques. Variance of a bounded random variable. Cross Validated. (version: 2016-04-23).

[355] J. Martín-Hernández, H. Wang, P. Van Mieghem, and G. D'Agostino. Algebraic connectivity of interdependent networks. *Physica A: Statistical Mechanics and its Applications*, 404:92–105, 2014.

[356] J. Masel and M. L. Siegal. Robustness: Mechanisms and consequences. *Trends in Genetics*, 25(9):395–403, 2009.

[357] J. Masel and M. V. Trotter. Robustness and evolvability. *PTrends in Genetics*.

[358] A. R. Mashaghi, A. Ramezanpour, and V. Karimipour. Investigation of a protein complex network. *The European Physical Journal B*, 41(1):113–121, 2004.

[359] P. H. Mason. Degeneracy at multiple levels of complexity. *Biological Theory*, 5(3):277–288, 2010.

[360] P. H. Mason, B. Winter, A. Grignolio, et al. Hidden in plain view: Degeneracy in complex systems. *Biosystems*, 128:1–8, 2015.

[361] M. Mastri, S. Rosario, A. Tracz, R. E Frink, R. A Brekken, and J. ML Ebos. The challenges of modeling drug resistance to antiangiogenic therapy. *Current Drug Targets*, 17(15):1747–1754, 2016.

[362] N. Masuda, M. A. Porter, and R. Lambiotte. Random walks and diffusion on networks. *Physics Reports*, 716:1–58, 2017.

[363] Y. Meepetchdee and N. Shah. Logistical network design with robustness and complexity considerations. *International Journal of Physical Distribution & Logistics Management*, 37(3):201–222, 2007.

[364] M. A. R. Mello, F. M. D. Marquitti, P. R. Guimarães, E. K. V. Kalko, P. Jordano, and M. A. M. de Aguiar. The modularity of seed dispersal: Differences in structure and robustness between bat–and bird–fruit networks. *Oecologia*, 167(1):131, 2011.

[365] M. A. R. Mello, F. M. D. Marquitti, P. R. Guimaraes Jr, E. K. V. Kalko, P. Jordano, and M. A. M. de Aguiar. The missing part of seed dispersal networks: Structure and robustness of bat-fruit interactions. *PLoS One*, 6(2):e17395, 2011.

[366] J. V. Michalowicz, J. M. Nichols, and F. Bucholtz. *Handbook of Differential Entropy*. CRC Press, Boca Raton, FL, USA, 2013.

[367] S. Milgram. The small world problem. *Psychology Today*, 2(1):60–67, 1967.

[368] J. C. Miller. Percolation and epidemics in random clustered networks. *Physical Review E*, 80(2):020901, 2009.

[369] C. L. Mills, P. J. Beuning, and M. J. Ondrechen. Biochemical functional predictions for protein structures of unknown or uncertain function. *Computational and Structural Biotechnology Journal*, 13:182–191, 2015.

[370] 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.

[371] I. Mirzaev and J. Gunawardena. Laplacian dynamics on general graphs. *Bulletin of Mathematical Biology*, 75(11):2118–2149, 2013.

[372] B. Mohar, Y. Alavi, G. Chartrand, and O. Oellermann. The laplacian spectrum of graphs. *Graph Theory, Combinatorics, and Applications*, 2(871-898):12, 1991.

[373] J. Monostori. Supply chains robustness: Challenges and opportunities. *Procedia CIRP*, 67:110–115, 2018.

[374] M. Monteforte and F. Wolf. Dynamical entropy production in spiking neuron networks in the balanced state. *Physical Review Letters*, 105(26):268104, 2010.

[375] A. N. Monteiro. BRCA1: Exploring the links to transcription. *Trends in Biochemical Sciences*, 25(10):469–474, 2000.

[376] K. Morino, G. Tanaka, and K. Aihara. Robustness of multilayer oscillator networks. *Physical Review E*, 83(5):056208, 2011.

[377] F. Morone and H. A. Makse. Influence maximization in complex networks through optimal percolation. *Nature*, 524(7563):65–68, 2015.

[378] F. Morone, B. Min, L. Bo, R. Mari, and H. A. Makse. Collective influence algorithm to find influencers via optimal percolation in massively large social media. *Scientific Reports*, 6:30062, 2016.

[379] A. E. Motter and Y.-C. Lai. Cascade-based attacks on complex networks. *Physical Review E*, 66(6):065102, 2002.

[380] J. Mourier, C. Brown, and S. Planes. Learning and robustness to catch-and-release fishing in a shark social network. *Biology Letters*, 13(3):20160824, 2017.

[381] S. Mugisha and H.-J. Zhou. Identifying optimal targets of network attack by belief propagation. *Physical Review E*, 94(1):012305, 2016.

[382] P. Müller, D. Kuttenkeuler, V. Gesellchen, M. P. Zeidler, and M. Boutros. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature*, 436(7052):871–875, 2005.

[383] M. W. Nachman and S. L. Crowell. Estimate of the mutation rate per nucleotide in humans. *Genetics*, 156(1):297–304, 2000.

[384] S. Nadarajah and S. Kotz. Mathematical properties of the multivariate t distribution. *Acta Applicandae Mathematica*, 89(1–3):53–84, 2005.

[385] A. Nagurney and Q. Qiang. Robustness of transportation networks subject to degradable links. *Europhysics Letters*, 80(6):68001, 2007.

[386] A. Nair and J. M. Vidal. Supply network topology and robustness against disruptions–an investigation using multi-agent model. *International Journal of Production Research*, 49(5):1391–1404, 2011.

[387] M. Nakayama, S. Imamura, T. Taniguchi, and R. Tateno. Does conversion from natural forest to plantation affect fungal and bacterial biodiversity, community structure, and co-occurrence networks in the organic horizon and mineral soil? *Forest Ecology and Management*, 446:238–250, 2019.

[388] S. M. Nazem. Entropy of freight transportation and price variation. *European Journal of Operational Research*, 3(5):413–416, 1979.

[389] S. Neumayer and E. Modiano. Network reliability with geographically correlated failures. In *Proceedings IEEE INFOCOM*, pages 1–9. IEEE, 2010.

[390] A. R. Neveu. A survey of network-based analysis and systemic risk measurement. *Journal of Economic Interaction and Coordination*, 13(2):241–281, 2018.

[391] M. Newman and G. Ghoshal. Bicomponents and the robustness of networks to failure. *Physical Review Letters*, 100(13):138701, 2008.

[392] M. E. Newman. Assortative mixing in networks. *Physical Review Letters*, 89(20):208701, 2002.

[393] M. E. Newman. Mixing patterns in networks. *Physical Review E*, 67(2):026126, 2003.

[394] M. E. Newman. Analysis of weighted networks. *Physical Review E*, 70(5):056131, 2004.

[395] M. E. Newman. Random graphs with clustering. *Physical Review Letters*, 103(5):058701, 2009.

[396] M. E. Newman, A.-L. E. Barabási, and D. J. Watts. *The structure and dynamics of networks*. Princeton University Press, Princeton, NJ, USA, 2006.

[397] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[398] M. E. J. Newman. *Networks*. Oxford University Press, Oxford, United Kingdom, 2nd edition, 2018.

[399] T. Nie, Z. Guo, K. Zhao, and Z.-M. Lu. New attack strategies for complex networks. *Physica A: Statistical Mechanics and its Applications*, 424:248–253, 2015.

[400] T. Nie, Z. Guo, K. Zhao, and Z.-M. Lu. The dynamic correlation between degree and betweenness of complex network under attack. *Physica A: Statistical Mechanics and Its Applications*, 457:129–137, 2016.

[401] T. Nishikawa, J. Sun, and A. E. Motter. Sensitive dependence of optimal network dynamics on network structure. *Physical Review X*, 7(4):041044, 2017.

[402] Q. Niu, A. Zeng, Y. Fan, and Z. Di. Robustness of centrality measures against network manipulation. *Physica A Statistical Mechanics and its Applications*, 438:124–131, 2015.

[403] E. P.-B. Noon, F. P. Davis, and D. L. Stern. Evolved repression overcomes enhancer robustness. *Developmental Cell*, 39(5):572–584, 2016.

[404] J. K. Ochab. Maximal-entropy random walk unifies centrality measures. *Physical Review E*, 86(6):066109, 2012.

[405] C. P. Ogbogbo. Modeling crude oil spot price as an Ornstein–Uhlenbeck process. *International Journal of Mathematical Analysis and Optimization: Theory and Applications*, 2018:261–275, 2018.

[406] U. Y. Ogras and R. Marculescu. "it's a small world after all": NoC performance optimization via long-range link insertion. *IEEE Transactions on Very Large Scale Integration Systems*, 14(7):693–706, 2006.

[407] T. Ohnishi, H. Takayasu, and M. Takayasu. Network motifs in an inter-firm network. *Journal of Economic Interaction and Coordination*, 5(2):171–180, 2010.

[408] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkevičiute, C. Suo, and A. Fornito. Consistency and differences between centrality measures across distinct classes of networks. *PLoS One*, 14(7):1–23, 07 2019.

[409] T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

[410] J. M. Ortega. *Numerical Analysis: A Second Course*. SIAM, Philadelphia, PA, USA, 1990.

[411] S. Ostojic, N. Brunel, and V. Hakim. How connectivity, background activity, and synaptic properties shape the cross-correlation between spike trains. *Journal of Neuroscience*, 29(33):10234–10253, 2009.

[412] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(1):17, 2017.

[413] D. N. Page. Average entropy of a subsystem. *Physical Review Letters*, 71(9):1291, 1993.

[414] M. Paluš. On entropy rates of dynamical systems and Gaussian processes. *Physics Letters A*, 227(5-6):301–308, 1997.

[415] K. Park and D. Kim. Localized network centrality and essentiality in the yeast–protein interaction network. *Proteomics*, 9(22):5143–5154, 2009.

[416] M. Pascual, J. A. Dunne, et al. *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford University Press, Oxford, United Kingdom, 2006.

[417] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani. Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3):925, 2015.

[418] R. Pastor-Satorras and A. Vespignani. Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200, 2001.

[419] G. Paul, T. Tanizawa, S. Havlin, and H. E. Stanley. Optimization of robustness of complex networks. *The European Physical Journal B*, 38(2):187–191, 2004.

[420] S. Pei, J. Wang, F. Morone, and H. A. Makse. Influencer identification in dynamical complex systems. *arXiv preprint arXiv:1907.13017*, 2019.

[421] X. Pei and S. Zhu. Measurements of Financial Contagion: A Primary Review from the Perspective of Structural Break. In *Optimization and Control for Systems in the Big-Data Era*, pages 61–84. Springer, 2017.

[422] T. P. Peixoto. The graph-tool python library. *figshare*, 2014.

[423] T. P. Peixoto and S. Bornholdt. Evolution of robust network topologies: Emergence of central backbones. *Physical Review Letters*, 109(11):118703, 2012.

[424] G.-S. Peng, S.-Y. Tan, J. Wu, and P. Holme. Trade-offs between robustness and small-world effect in complex networks. *Scientific Reports*, 6:37317, 2016.

[425] X. Peng and Z. Zhang. Maximal entropy random walk improves efficiency of trapping in dendrimers. *The Journal of Chemical Physics*, 140(23):234104, 2014.

[426] R. Penrose. A generalized inverse for matrices. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 51, pages 406–413. Cambridge University Press, 1955.

[427] V. Pernice, B. Staude, S. Cardanobile, and S. Rotter. How structure determines correlations in neuronal networks. *PLoS Computational Biology*, 7(5):e1002059, 2011.

[428] B. R. Phillips and R. S. Middleton. SimWIND: A geospatial infrastructure model for optimizing wind power generation and transmission. *Energy Policy*, 43:291–302, 2012.

[429] P. C. Phillips, S. P. Otto, and M. C. Whitlock. Beyond the average. In *Epistasis and the evolutionary process*, pages 20–38. Oxford University Press, Oxford, United Kingdom, 2000.

[430] S. Pinnaka, R. Yarlagadda, and E. K. Çetinkaya. Modelling robustness of critical infrastructure networks. In *2015 11th International Conference on the Design of Reliable Communication Networks (DRCN)*, pages 95–98. IEEE, 2015.

[431] A. Plough, J. E. Fielding, A. Chandra, M. Williams, D. Eisenman, K. B. Wells, G. Y. Law, S. Fogleman, and A. Magaña. Building community disaster resilience: Perspectives from a large urban county department of public health. *American Journal of Public Health*, 103(7):1190–1197, 2013.

[432] P. Poornima, J. D. Kumar, Q. Zhao, M. Blunder, and T. Efferth. Network pharmacology of cancer: From understanding of complex interactomes to the design of multi-target specific therapeutics from nature. *Pharmacological Research*, 111:290–302, 2016.

[433] T. Popoviciu. Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica*, 9:129–145, 1935.

[434] M. A. Porter and J. P. Gleeson. Dynamical Systems on Networks. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials*, 4, 2016.

[435] M. Pósfai, A. Fekete, and G. Vattay. Shortest-path sampling of dense homogeneous networks. *Europhysics Letters*, 89(1):18007, 2010.

[436] T. K. Prasad, R. Goel, K. Kandasamy, S. Keerthikumar, S. Kumar, S. Mathivanan, D. Telikicherla, R. Raju, B. Shafreen, A. Venugopal, et al. Human protein reference database — 2009 update. *Nucleic Acids Research*, 37(suppl 1):D767–D772, 2009.

[437] C. J. Price and K. J. Friston. Degeneracy and cognitive anatomy. *Trends in Cognitive Sciences*, 6(10):416–421, 2002.

[438] C.-L. Pu, W.-J. Pei, and A. Michaelson. Robustness analysis of network controllability. *Physica A: Statistical Mechanics and its Applications*, 391(18):4420–4425, 2012.

[439] X. Qi, E. Fuller, Q. Wu, Y. Wu, and C.-Q. Zhang. Laplacian centrality: A new centrality measure for weighted networks. *Information Sciences*, 194:240–253, 2012.

[440] F. Radicchi. Percolation in real interdependent networks. *Nature Physics*, 11(7):597–602, 2015.

[441] F. Radicchi and G. Bianconi. Redundant interdependencies boost the robustness of multiplex networks. *Physical Review X*, 7(1):011013, 2017.

[442] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi. Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(9):2658–2663, 2004.

[443] S. Rai and D. P. Agrawal. Distributed computing network reliability. 1990.

[444] M. Randles, D. Lamb, E. Odat, and A. Taleb-Bendiab. Distributed redundancy and robustness in complex systems. *Journal of Computer and System Sciences*, 77(2):293–304, 2011.

[445] M. Rausand and A. Høyland. *System Reliability Theory: Models, Statistical Methods, and Applications*, volume 396. John Wiley & Sons, New York City, NY, USA, 2004.

[446] J. Ren, J. Wang, M. Li, H. Wang, and B. Liu. Prediction of essential proteins by integration of PPI network topology and protein complexes information. In *International Symposium on Bioinformatics Research and Applications*, pages 12–24. Springer, 2011.

[447] T. Ren, Y.-F. Wang, M.-M. Liu, and Y.-J. Xu. Analysis of robustness of urban bus network. *Chinese Physics B*, 25(2):020101, 2016.

[448] X.-L. Ren, N. Gleinig, D. Helbing, and N. Antulov-Fantulin. Generalized network dismantling. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14):6554–6559, 2019.

[449] J. Rice. *Mathematical Statistics and Data Analysis*. Thomson Higher Education, Belmont, CA, USA, 3rd edition, 2006.

[450] J. M. Rip, K. S. McCann, D. H. Lynn, and S. Fawcett. An experimental test of a fundamental food web motif. *Proceedings of the Royal Society of London B: Biological Sciences*, pages 1743–1749, 2010.

[451] K. Ristl, S. J. Plitzko, and B. Drossel. Complex response of a food-web module to symmetric and asymmetric migration between several patches. *Journal of Theoretical Biology*, 354:54–59, 2014.

[452] E. Rodríguez-Núñez and J. C. García-Palomares. Measuring the vulnerability of public transport networks. *Journal of Transport Geography*, 35:50–63, 2014.

[453] R. V. Rohlfs, P. Harrigan, and R. Nielsen. Modeling gene expression evolution with an extended Ornstein–Uhlenbeck process accounting for within-species variation. *Molecular Biology and Evolution*, 31(1):201–211, 2013.

[454] M. Rosas-Casals, S. Valverde, and R. V. Solé. Topological vulnerability of the European power grid under errors and attacks. *International Journal of Bifurcation and Chaos*, 17(07):2465–2475, 2007.

[455] J. S. Rosenfeld. Functional redundancy in ecology and conservation. *Oikos*, 98(1):156–162, 2002.

[456] S. Roy and B. Ravindran. Measuring network centrality using hypergraphs. In *Proceedings of the 2nd ACM IKDD Conference on Data Sciences*, pages 59–68. ACM, 2015.

[457] A. S Azmi. Adopting network pharmacology for cancer drug discovery. *Current Drug Discovery Technologies*, 10(2):95–105, 2013.

[458] M. Sahini and M. Sahimi. *Applications of percolation theory*. CRC Press, Boca Raton, FL, USA, 1994.

[459] A. Sajjad, S. Saeed, M. Ali, F. Z. A. Khan, Y. J. Kwon, and M. Devoto. Effect of temporal data aggregation on the perceived structure of a quantitative plant–floral visitor network. *Entomological Research*, 47(6):380–387, 2017.

[460] T. Sanderson, G. Hertzler, T. Capon, and P. Hayman. A real options analysis of australian wheat production under climate change. *Australian Journal of Agricultural and Resource Economics*, 60(1):79–96, 2016.

[461] S. Santamaría, J. Galeano, J. M. Pastor, and M. Méndez. Removing interactions, rather than species, casts doubt on the high robustness of pollination networks. *Oikos*, 125(4):526–534, 2016.

[462] T. Sarlos. Improved approximation algorithms for large matrices via random projections. In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 143–152. IEEE, 2006.

[463] S.-I. Sasa and K. Hayashi. Computation of the Kolmogorov–Sinai entropy using statistitical mechanics: Application of an exchange Monte Carlo method. *Europhysics Letters*, 74(1):156, 2006.

[464] H. Sayama, I. Pestov, J. Schmidt, B. J. Bush, C. Wong, J. Yamanoi, and T. Gross. Modeling complex systems with adaptive networks. *Computers & Mathematics With Applications*, 65(10):1645–1664, 2013.

[465] M. H. Schaefer, J.-F. Fontaine, A. Vinayagam, P. Porras, E. E. Wanker, and M. A. Andrade-Navarro. HIPPIE: Integrating protein interaction networks with experiment based quality scores. *PLoS One*, 7(2):1–8, 2012.

[466] M. T. Schaub, J.-C. Delvenne, R. Lambiotte, and M. Barahona. Multiscale dynamical embeddings of complex networks. *Physical Review E*, 99(6):062308, 2019.

[467] T. Schieber, M. Ravetti, and P. M. Pardalos. A review on network robustness from an information theory perspective. In *International Conference on Discrete Optimization and Operations Research*, pages 50–60. Springer, 2016.

[468] T. A. Schieber, L. Carpi, A. C. Frery, O. A. Rosso, P. M. Pardalos, and M. G. Ravetti. Information theory perspective on network robustness. *Physics Letters A*, 380(3):359–364, 2016.

[469] M. Schleuning, J. Fruend, O. Schweiger, E. Welk, J. Albrecht, M. Albrecht, M. Beil, G. Benadi, N. Bluethgen, H. Bruelheide, et al. Ecological networks are more sensitive to plant than to animal extinction under climate change. *Nature Communications*, 7:13965, 2016.

[470] C. M. Schneider, A. A. Moreira, J. S. Andrade, S. Havlin, and H. J. Herrmann. Mitigation of malicious attacks on networks. *Proceedings of the National Academy of Sciences of the United States of America*, 108(10):3838–3841, 2011.

[471] S. L. Schreiber and B. E. Bernstein. Signaling network model of chromatin. *Cell*, 111(6):771–778, 2002.

[472] T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461, 2000.

[473] S. Schuster, D. A. Fell, and T. Dandekar. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nature Biotechnology*, 18(3):326–332, 2000.

[474] A. C. Schwarze, P. S. Chodrow, and M. A. Porter. Log-minor distributions and an application to estimating mean subsystem entropy. *arXiv preprint arXiv:1901.09456*, 2019.

[475] D. M. Scott, D. C. Novak, L. Aultman-Hall, and F. Guo. Network robustness index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14(3):215–227, 2006.

[476] S. Segarra and A. Ribeiro. Stability and continuity of centrality measures in weighted graphs. *IEEE Transactions on Signal Processing*, 64(3):543–555, 2015.

[477] S. Sen. Average entropy of a quantum subsystem. *Physical Review Letters*, 77(1):1, 1996.

[478] Z. Sha and J. H. Panchal. Towards the design of complex evolving networks with high robustness and resilience. *Procedia Computer Science*, 16:522–531, 2013.

[479] A. Shahrokni and R. Feldt. A systematic review of software robustness. *Information and Software Technology*, 55(1):1–17, 2013.

[480] C. Shalizi. 36-754, Advanced Probability II or Almost None of Stochastic Processes. Lecture Notes at Carnegie Mellon University Department of Statistics, 2006.

[481] Y. Shang. Random lifts of graphs: Network robustness based on the estrada index. *Applied Mathematics E-Notes*, 12:53–61, 2012.

[482] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, 1949.

[483] R. Sharan, I. Ulitsky, and R. Shamir. Network-based prediction of protein function. *Molecular Systems Biology*, 3(1), 2007.

[484] B. Shargel, H. Sayama, I. R. Epstein, and Y. Bar-Yam. Optimization of robustness and connectivity in complex networks. *Physical Review Letters*, 90(6):068701, 2003.

[485] R. Sharma, M. Gupta, and G. Kapoor. Some better bounds on the variance with applications. *Journal of Mathematical Inequalities*, 4(3):355–363, 2010.

[486] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics*, 31(1):64–68, 2002.

[487] D. Shizuka and D. R. Farine. Measuring the robustness of network community structure using assortativity. *Animal Behaviour*, 112:237–246, 2016.

[488] T. Shlomi, D. Segal, E. Ruppin, and R. Sharan. QPath: A method for querying pathways in a protein-protein interaction network. *BMC Bioinformatics*, 7(1):199, 2006.

[489] H. A. Simon. The architecture of complexity. In *Facets of systems science*, pages 457–476. Springer, 1991.

[490] I. Simonsen, L. Buzna, K. Peters, S. Bornholdt, and D. Helbing. Transient dynamics increasing network vulnerability to cascading failures. *Physical Review Letters*, 100(21):218701, 2008.

[491] Y. G. Sinai. On the notion of entropy of a dynamical system. *Doklady of Russian Academy of Sciences*, 124:768–771, 1959.

[492] R. Sinatra, J. Gómez-Gardenes, R. Lambiotte, V. Nicosia, and V. Latora. Maximal-entropy random walks in complex networks with limited information. *Physical Review E*, 83(3):030103, 2011.

[493] M. Singh and N. Ferrara. Modeling and predicting clinical efficacy for drugs targeting the tumor milieu. *Nature Biotechnology*, 30(7):648–657, 2012.

[494] S. S. Skiena. *The Data Science Design Manual*. Springer, Cham, Switzerland, 2017.

[495] A. B. Sletsjœ. The entropy of a dynamical system. The Abel Prize 2014 Brochure.

[496] J. A. Smith and J. Moody. Structural effects of network sampling coverage I: Nodes missing at random. *Social Networks*, 35(4):652–668, 2013.

[497] J. M. Smith. The concept of information in biology. *Philosophy of Science*, 67(2):177–194, 2000.

[498] R. V. Solé, M. Rosas-Casals, B. Corominas-Murtra, and S. Valverde. Robustness of the European power grids under intentional attack. *Physical Review E*, 77(2):026102, 2008.

[499] R. V. Solé and S. Valverde. Information theory of complex networks: On evolution and architectural constraints. In *Complex Networks*, pages 189–207. Springer, 2004.

[500] A. Solé-Ribalta, M. De Domenico, S. Gómez, and A. Arenas. Random walk centrality in interconnected multilayer networks. *Physica D: Nonlinear Phenomena*, 323:73–79, 2016.

[501] V. Sperati, V. Trianni, and S. Nolfi. Evolving coordinated group behaviours through maximisation of mean mutual information. *Swarm Intelligence*, 2(2-4):73–95, 2008.

[502] D. A. Spielman. Spectral graph theory and its applications. In *48th Annual IEEE Symposium on Foundations of Computer Science*, pages 29–38. IEEE, 2007.

[503] O. Sporns, G. Tononi, and G. M. Edelman. Theoretical neuroanatomy: Relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex*, 10(2):127–141, 2000.

[504] E. Sprinzak, S. Sattath, and H. Margalit. How reliable are experimental protein–protein interaction data? *Journal of Molecular Biology*, 327(5):919–923, 2003.

[505] R. P. St Onge, R. Mani, J. Oh, M. Proctor, E. Fung, R. W. Davis, C. Nislow, F. P. Roth, and G. Giaever. Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. *Nature Genetics*, 39(2):199–206, 2007.

[506] Z. Stanic. Graphs with small spectral gap. *Electronic Journal of Linear Algebra*, 26(1):28, 2013.

[507] M. Staniek and K. Lehnertz. Symbolic transfer entropy. *Physical Review Letters*, 100(15):158101, 2008.

[508] D. Stauffer and A. Aharony. *Introduction to Percolation Theory*. Taylor & Francis, Oxford, United Kingdom, 2018.

[509] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle III, and J. Doyle. Robustness of cellular functions. *Cell*, 118(6):675–685, 2004.

[510] U. Stelzl, U. Worm, M. Lalowski, C. Haenig, F. H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen, et al. A human protein-protein interaction network: A resource for annotating the proteome. *Cell*, 122(6):957–968, 2005.

[511] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM Journal on Numerical Analysis*, 17(3):403–409, 1980.

[512] L. Stone, D. Simberloff, and Y. Artzy-Randrup. Network motifs and their origins. *PLoS Computational Biology*, 15(4):e1006749, 2019.

[513] K. Strimbu and J. A. Tavel. What are biomarkers? *Current Opinion in HIV and AIDS*, 5(6):463, 2010.

[514] S. H. Strogatz. *Nonlinear Dynamics and Chaos*. CRC Press, Boca Raton, FL, USA, 2018.

[515] M. P. Stumpf and P. J. Ingram. Probability models for degree distributions of protein interaction networks. *Europhysics Letters*, 71(1):152, 2005.

[516] M. P. Stumpf, C. Wiuf, and R. M. May. Subnets of scale-free networks are not scale-free: Sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12):4221–4224, 2005.

[517] L. Šubelj and M. Bajec. Robust network community detection using balanced propagation. *The European Physical Journal B*, 81(3):353–362, 2011.

[518] X. Sun, V. Gollnick, and S. Wandelt. Robustness analysis metrics for worldwide airport network: A comprehensive study. *Chinese Journal of Aeronautics*, 30(2):500–512, 2017.

[519] M. Szell, R. Lambiotte, and S. Thurner. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences of the United States of America*, 107(31):13636–13641, 2010.

[520] K. Takemoto, T. Tamura, Y. Cong, W.-K. Ching, J.-P. Vert, and T. Akutsu. Analysis of the impact degree distribution in metabolic networks using branching process approximation. *Physica A: Statistical Mechanics and its Applications*, 391(1-2):379–387, 2012.

[521] F. W. Takes, W. A. Kosters, B. Witte, and E. M. Heemskerk. Multiplex network motifs as building blocks of corporate networks. *Applied Network Science*, 3(1):39, 2018.

[522] G. Tanaka, K. Morino, and K. Aihara. Dynamical robustness in complex networks: The crucial role of low-degree nodes. *Scientific Reports*, 2:232, 2012.

[523] J. Tang and T. Aittokallio. Network pharmacology strategies toward multi-target anticancer therapies: From computational models to experimental design principles. *Current Pharmaceutical Design*, 20(1):23–36, 2014.

[524] L. Tang, K. Jing, J. He, and H. E. Stanley. Robustness of assembly supply chain networks by considering risk propagation and cascading failure. *Physica A: Statistical Mechanics and its Applications*, 459:129–139, 2016.

[525] C. M. Taniguchi, B. Emanuelli, and C. R. Kahn. Critical nodes in signalling pathways: Insights into insulin action. *Nature Reviews Molecular Cell Biology*, 7(2):85–96, 2006.

[526] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, and H. E. Stanley. Optimization of network robustness to waves of targeted and random attacks. *Physical Review E*, 71(4):047101, 2005.

[527] M. Tanowitz and M. von Zastrow. Identification of protein interactions by yeast two-hybrid screening and coimmunoprecipitation. In *Receptor Signal Transduction Protocols*, volume 259, pages 353–369. Springer, 2004.

[528] A. E. Teschendorff, P. Sollich, and R. Kuehn. Signalling entropy: A novel network-theoretical framework for systems analysis and interpretation of functional omic data. *Methods*, 67(3):282–293, 2014.

[529] K. Thulasiraman and M. N. Swamy. *Graphs: Theory and Algorithms*. John Wiley & Sons, New York City, NY, USA, 2011.

[530] P. Tieri, A. Grignolio, A. Zaikin, M. Mishto, D. Remondini, G. C. Castellani, and C. Franceschi. Network, degeneracy and bow tie. Integrating paradigms and architectures to grasp the complexity of the immune system. *Theoretical Biology and Medical Modelling*, 7(1):32, 2010.

[531] A. Tizghadam and A. Leon-Garcia. On iterative calculation of moore-penrose laplacian and resistance distance. In *Sigmetrics MAMA Workshop*, 2015.

[532] G. A. Tobin. Sustainability and community resilience: The holy grail of hazards planning? *Global Environmental Change Part B: Environmental Hazards*, 1(1):13–25, 1999.

[533] G. Tononi, O. Sporns, and G. M. Edelman. A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11):5033–5037, 1994.

[534] G. Tononi, O. Sporns, and G. M. Edelman. A complexity measure for selective matching of signals by the brain. *Proceedings of the National Academy of Sciences of the United States of America*, 93(8):3422–3427, 1996.

[535] G. Tononi, O. Sporns, and G. M. Edelman. Measures of degeneracy and redundancy in biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 96(6):3257–3262, 1999.

[536] J. T. Tou and J. T. Tou. *Advances in Information Systems Science*, volume 6. Springer, Berlin, Germany, 2012.

[537] S. Trajanovski, J. Martín-Hernández, W. Winterbach, and P. Van Mieghem. Robustness envelopes of networks. *Journal of Complex Networks*, 1(1):44–62, 2013.

[538] J. Trousdale, Y. Hu, E. Shea-Brown, and K. Josić. Impact of network structure and cellular response on spike time correlations. *PLoS Computational Biology*, 8(3):e1002408, 2012.

[539] M.-T. Tsai, J.-D. Saphores, and A. Regan. Valuation of freight transportation contracts under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 47(6):920–932, 2011.

[540] B. R. Tukamuhabwa, M. Stevenson, J. Busby, and M. Zorzini. Supply chain resilience: Definition, review and theoretical foundations for further study. *International Journal of Production Research*, 53(18):5592–5623, 2015.

[541] N. Tuncbag, S. J. Gosline, A. Kedaigle, A. R. Soltis, A. Gitter, and E. Fraenkel. Network-based interpretation of diverse high-throughput datasets through the omics integrator software package. *PLoS Computational Biology*, 12(4):e1004879, 2016.

[542] J. J. Tyson and B. Novák. Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry*, 61:219–240, 2010.

[543] G. E. Uhlenbeck and L. S. Ornstein. On the theory of the Brownian motion. *Physical Review*, 36(5):823, 1930.

[544] N. Urruty, D. Tailliez-Lefebvre, and C. Huyghe. Stability, robustness, vulnerability and resilience of agricultural systems. A review. *Agronomy for Sustainable Development*, 36(1):15, 2016.

[545] M. Usaj, Y. Tan, W. Wang, B. VanderSluis, A. Zou, C. L. Myers, M. Costanzo, B. Andrews, and C. Boone. Thecellmap. org: A web-accessible database for visualizing and mining the global yeast genetic interaction network. *G3: Genes, Genomes, Genetics*, 7(5):1539–1549, 2017.

[546] M. M. Usaj, E. B. Styles, A. J. Verster, H. Friesen, C. Boone, and B. J. Andrews. High-content screening for quantitative cell biology. *Trends in Cell Biology*, 26(8):598–611, 2016.

[547] E. Valdano, M. R. Fiorentin, C. Poletto, and V. Colizza. Epidemic threshold in continuous-time evolving networks. *Physical Review Letters*, 120(6):068302, 2018.

[548] T. W. Valente, K. Coronges, C. Lakon, and E. Costenbader. How correlated are network centrality measures? *Connections (Toronto, Ont.)*, 28(1):16–26, 2008.

[549] P. Van Mieghem. Robustness of large networks. In *2005 IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2372–2377. IEEE, 2005.

[550] P. Van Mieghem. *Graph Spectra for Complex Networks*. Cambridge University Press, Cambridge, United Kingdom, 2010.

[551] P. Van Mieghem, J. Omic, and R. Kooij. Virus spread in networks. *IEEE/ACM Transactions on Networking (TON)*, 17(1):1–14, 2009.

[552] M. Ventresca and D. Aleman. Network robustness versus multi-strategy sequential attack. *Journal of Complex Networks*, 3(1):126–146, 2014.

[553] R. Vicente, M. Wibral, M. Lindner, and G. Pipa. Transfer entropy — a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience*, 30(1):45–67, 2011.

[554] M. Vidal, M. E. Cusick, and A.-L. Barabási. Interactome networks and human disease. *Cell*, 144(6):986–998, 2011.

[555] M. C. Vieira and M. Almeida-Neto. A simple stochastic model for complex coextinctions in mutualistic networks: Robustness decreases with connectance. *Ecology Letters*, 18(2):144–152, 2015.

[556] A. Vinayagam, T. E. Gibson, H.-J. Lee, B. Yilmazel, C. Roesel, Y. Hu, Y. Kwon, A. Sharma, Y.-Y. Liu, N. Perrimon, et al. Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proceedings of the National Academy of Sciences of the United States of America*, 113(18):4976–4981, 2016.

[557] C. Von Mering, R. Krause, B. Snel, M. Cornell, S. G. Oliver, S. Fields, and P. Bork. Comparative assessment of large-scale data sets of protein–protein interactions. *Nature*, 417(6887):399–403, 2002.

[558] A. Wagner. Robustness against mutations in genetic networks of yeast. *Nature Genetics*, 24(4):355–361, 2000.

[559] A. Wagner. Distributed robustness versus redundancy as causes of mutational robustness. *Bioessays*, 27(2):176–188, 2005.

[560] A. Wagner. *Robustness and Evolvability in Living Systems*, volume 24. Princeton University Press, Princeton, NJ, USA, 2007.

[561] W. C. Wah. *Synchronization in complex networks of nonlinear dynamical systems*. World scientific, 2007.

[562] M. M. Walsh, E. H. Einstein, and K. A. Gluck. A quantification of robustness. *Journal of Applied Research in Memory and Cognition*, 2(3):137–148, 2013.

[563] B. Wang, H. Tang, C. Guo, and Z. Xiu. Entropy optimization of scale-free networks' robustness to random failures. *Physica A: Statistical Mechanics and its Applications*, 363(2):591–596, 2006.

[564] H. Wang, J. Huang, X. Xu, and Y. Xiao. Damage attack on complex networks. *Physica A: Statistical Mechanics and its Applications*, 408:134–148, 2014.

[565] J. Wang, L. Cai, B. Xu, and Y. Wu. Abnormal behavior in cascading dynamics with node weight. *PLoS One*, 10(10):e0139941, 2015.

[566] J. Wang, C. Zhang, Y. Huang, and C. Xin. Attack robustness of cascading model with node weight. *Nonlinear Dynamics*, 78(1):37–48, 2014.

[567] S.-J. Wang and C. Zhou. Hierarchical modular structure enhances the robustness of self-organized criticality in neural networks. *New Journal of Physics*, 14(2):023005, 2012.

[568] W.-X. Wang and G. Chen. Universal robustness characteristic of weighted networks against cascading failure. *Physical Review E*, 77(2):026101, 2008.

[569] X. Wang, L. Feng, R. E. Kooij, and J. L. Marzo. Inconsistencies among spectral robustness metrics. In *International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, pages 119–136. Springer, 2018.

[570] X. Wang, E. Pournaras, R. E. Kooij, and P. Van Mieghem. Improving robustness of complex networks via the effective graph resistance. *The European Physical Journal B*, 87(9):221, 2014.

[571] Z. Wang, C. Zang, K. Cui, D. E. Schones, A. Barski, W. Peng, and K. Zhao. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell*, 138(5):1019–1031, 2009.

[572] J. G. Wardrop. Some theoretical aspects of road traffic research. *Proceedings of the Institution of Civil Engineers*, 1(3):325–362, 1952.

[573] S. Watterson, S. Marshall, and P. Ghazal. Logic models of pathway biology. *Drug Discovery Today*, 13(9-10):447–456, 2008.

[574] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393(6684):440–442, 1998.

[575] C. Wawra, M. Kühl, and H. A. Kestler. Extended analyses of the Wnt/β-catenin pathway: Robustness and oscillatory behaviour. *FEBS Letters*, 581(21):4043–4048, 2007.

[576] W. Weaver. The mathematics of communication. *Scientific American*, 181(1):11–15, 1949.

[577] J. West, S. Beck, X. Wang, and A. E. Teschendorff. An integrative network algorithm identifies age-associated differential methylation interactome hotspots targeting stem-cell differentiation pathways. *Scientific Reports*, 3:1630, 2013.

[578] J. West, G. Bianconi, S. Severini, and A. E. Teschendorff. Differential network entropy reveals cancer system hallmarks. *Scientific Reports*, 2:802, 2012.

[579] J. Whitacre and A. Bender. Degeneracy: A design principle for achieving robustness and evolvability. *Journal of Theoretical Biology*, 263(1):143–153, 2010.

[580] J. M. Whitacre. Degeneracy: A link between evolvability, robustness and complexity in biological systems. *Theoretical Biology and Medical Modelling*, 7(1):6, 2010.

[581] J. M. Whitacre. Biological robustness: Paradigms, mechanisms, and systems principles. *Frontiers in Genetics*, 3:67, 2012.

[582] J. M. Whitacre and A. Bender. Networked buffering: A basic mechanism for distributed robustness in complex adaptive systems. *Theoretical Biology and Medical Modelling*, 7(1):20, 2010.

[583] T. Wilhelm, J. Behre, and S. Schuster. Analysis of structural robustness of metabolic networks. *Systems Biology*, 1(1):114–120, 2004.

[584] M. J. Williams and M. Musolesi. Spatio-temporal networks: Reachability, centrality and robustness. *Royal Society Open Science*, 3(6):160196, 2016.

[585] W. Winterbach, H. Wang, M. Reinders, P. Van Mieghem, and D. de Ridder. Metabolic network destruction: Relating topology to robustness. *Nano Communication Networks*, 2(2-3):88–98, 2011.

[586] D. Witthaut and M. Timme. Braess's paradox in oscillator networks, desynchronization and power outage. *New Journal of Physics*, 14(8):083036, 2012.

[587] D. P. Woodruff. Sketching as a tool for numerical linear algebra. *Foundations and Trends® in Theoretical Computer Science*, 10(1–2):1–157, 2014.

[588] J. Wu, M. Barahona, T. Yue-Jin, and D. Hong-Zhong. Natural connectivity of complex networks. *Chinese Physics Letters*, 27(7):078902, 2010.

[589] J. Wu, H.-z. Deng, and Y.-j. Tan. Spectral measure of robustness for Internet topology. In *2010 3rd International Conference on Computer Science and Information Technology*, volume 6, pages 50–54. IEEE, 2010.

[590] Y. Wu, X. Zhang, J. Yu, and Q. Ouyang. Identification of a topological characteristic responsible for the biological robustness of regulatory networks. *PLoS Computational Biology*, 5(7):e1000442, 2009.

[591] Z.-X. Wu and P. Holme. Onion structure and network robustness. *Physical Review E*, 84(2):026106, 2011.

[592] Z. Wunderlich and L. A. Mirny. Using the topology of metabolic networks to predict viability of mutant strains. *Biophysical Journal*, 91(6):2304–2311, 2006.

[593] J. Xia, D. M. Rocke, G. Perry, and M. Ray. Differential network analyses of Alzheimer's disease identify early events in Alzheimer's disease pathology. *International Journal of Alzheimer's Disease*, 2014, 2014.

[594] Z. Xu, Q. Zhang, D. Chen, and Y. He. Characterizing the Connectivity of Railway Networks. *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[595] K. Yamashita, Y. Yasuda, R. Nakamura, and H. Ohsaki. On the Predictability of Network Robustness from Spectral Measures. In *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, volume 2, pages 24–29. IEEE, 2019.

[596] Y. Yang, Z. Li, Y. Chen, X. Zhang, and S. Wang. Improving the robustness of complex networks with preserving community structure. *PLoS One*, 10(2):e0116551, 2015.

[597] A. Yazdani and P. Jeffrey. Applying network theory to quantify the redundancy and structural robustness of water distribution systems. *Journal of Water Resources Planning and Management*, 138(2):153–161, 2011.

[598] M. A. Yıldırım, K.-I. Goh, M. E. Cusick, A.-L. Barabási, and M. Vidal. Drug–target network. *Nature Biotechnology*, 25(10):1119, 2007.

[599] J.-G. Young, G. Petri, F. Vaccarino, and A. Patania. Construction of and efficient sampling from the simplicial configuration model. *Physical Review E*, 96(3):032312, 2017.

[600] M. Young, S. Zimmer, and A. Whitmore. Drug molecules and biology: Network and systems aspects. In *Designing Multi-Target Drugs*. RSC Publishing, Cambridge, United Kingdom, 2012.

[601] M. Youssef, R. Kooij, and C. Scoglio. Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics. *Journal of Computational Science*, 2(3):286–298, 2011.

[602] J. Zanghellini, D. E. Ruckerbauer, M. Hanscho, and C. Jungreuthmayer. Elementary flux modes in a nutshell: Properties, calculation and applications. *Biotechnology Journal*, 8(9):1009–1016, 2013.

[603] L. Zdeborová, P. Zhang, and H.-J. Zhou. Fast and simple decycling and dismantling of networks. *Scientific Reports*, 6:37954, 2016.

[604] A. Zeng and W. Liu. Enhancing network robustness against malicious attacks. *Physical Review E*, 85(6):066130, 2012.

[605] D. Zhang, E. K. Cetinkaya, and J. P. Sterbenz. Robustness of mobile ad hoc networks under centrality-based attacks. In *2013 5th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, pages 229–235. IEEE, 2013.

[606] H. Zhang, E. Fata, and S. Sundaram. A notion of robustness in complex networks. *Ieee Transactions on Control of Network Systems*, 2(3):310–320, 2015.

[607] S. Zhang, G. Jin, X.-S. Zhang, and L. Chen. Discovering functions and revealing mechanisms at molecular level from biological networks. *Proteomics*, 7(16):2856–2869, 2007.

[608] K. Zhao, M. Karsai, and G. Bianconi. Entropy of dynamical social networks. *PloS One*, 6(12):e28116, 2011.

[609] D. Zhou, H. E. Stanley, G. D'Agostino, and A. Scala. Assortativity decreases the robustness of interdependent networks. *Physical Review E*, 86(6):066103, 2012.

[610] F. Zhou, Y. Yuan, and M. Zhang. Robustness Analysis of Interdependent Urban Critical Infrastructure Networks Against Cascade Failures. *Arabian Journal for Science and Engineering*, 44(3):2837–2851, 2019.

[611] M. Zhou and J. Liu. A memetic algorithm for enhancing the robustness of scale-free networks against malicious attacks. *Physica A: Statistical Mechanics and its Applications*, 410:131–143, 2014.

[612] S. Zhou and R. J. Mondragón. Redundancy and robustness of AS-level Internet topology and its models. *Electronics Letters*, 40(2):151–152, 2004.

[613] F. Zhu. Improved collective influence of finding most influential nodes based on disjoint-set reinsertion. *Scientific Reports*, 8(1):14503, 2018.

[614] M. Zhu, L. Gao, X. Li, Z. Liu, C. Xu, Y. Yan, E. Walker, W. Jiang, B. Su, X. Chen, et al. The analysis of the drug–targets based on the topological properties in the human protein–protein interaction network. *Journal of Drug Targeting*, 17(7):524–532, 2009.

[615] E. Zotenko, J. Mestre, D. P. O'Leary, and T. M. Przytycka. Why do hubs in the yeast protein interaction network tend to be essential: Reexamining the connection between the network topology and essentiality. *PLoS Computational Biology*, 4(8):e1000140, 2008.

[616] D. Zubillaga, G. Cruz, L. Aguilar, J. Zapotécatl, N. Fernández, J. Aguilar, D. Rosenblueth, and C. Gershenson. Measuring the complexity of self-organizing traffic lights. *Entropy*, 16(5):2384–2407, 2014.