Understanding Complex Systems

Springer : COMPLEXITY

Federico Battiston
Giovanni Petri *Editors*

Higher-Order Systems

Springer

Springer Complexity

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

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

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

The three major book publication platforms of the Springer Complexity program are the monograph series "Understanding Complex Systems" focusing on the various applications of complexity, the "Springer Series in Synergetics", which is devoted to the quantitative theoretical and methodological foundations, and the "Springer Briefs in Complexity" which are concise and topical working reports, case studies, surveys, essays and lecture notes of relevance to the field. In addition to the books in these two core series, the program also incorporates individual titles ranging from textbooks to major reference works.

Indexed by SCOPUS, INSPEC, zbMATH, SCImago.

Series Editors

Understanding Complex Systems

Founding Editor: Scott Kelso

More information about this series at https://link.springer.com/bookseries/5394

Federico Battiston · Giovanni Petri
Editors

Higher-Order Systems

Springer

Editors
Federico Battiston
Department of Network and Data Science
Central European University
Vienna, Austria

Giovanni Petri
ISI Foundation
Turin, Italy

ISSN 1860-0832 ISSN 1860-0840 (electronic)
Springer Complexity
ISBN 978-3-030-91373-1 ISBN 978-3-030-91374-8 (eBook)
https://doi.org/10.1007/978-3-030-91374-8

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Abstract Over the last 20 years networks have emerged as the paradigmatic framework to model complex systems. Yet, as simple collections of nodes and links, they are intrinsically limited to pairwise interactions, limiting our ability to describe, understand and predict complex phenomena which arise from higher-order interactions. Here we introduce the new modeling framework of higher-order systems, where hypergraphs and simplicial complexes are used to describe complex patterns of interactions among any number of agents. We discuss its potential to model real-world systems, and how considering their higher-order organization can lead to the emergence of novel dynamical behavior. We believe that in the future years the framework of higher-order systems will become an important reference for network scientists interested in better modeling the interconnected world we live in.

Every day we witness the complexity of our world at many scales and across many domains. Over the last few decades, the study of such complexity has overtaken more traditional reductionist approaches, significantly affecting the way we look at the reality around us [1]. We have learned that no matter how good our knowledge of the basic units is, this is often not enough to understand system-level emerging phenomena. Epileptic seizures can not be explained by the most precise description of how a single neuron works, nor the viral spread of rumours, diseases or cooperation in a human population can be properly described by considering individual behavior in isolation. In all such cases, the complex patterns of dependencies, feedbacks and interactions among the system constituents play a fundamental role to determine the emergence of complex collective behavior.

In this context, networks have risen as the primary tool to model complex systems [2, 3]. Indeed, they are the simplest mathematical object able to explicitly encode relational information: system units are described by nodes, and pairwise interactions are encoded by links among them. Following the increasing availability of large-scale datasets, networks expanded beyond traditional graph theory [4, 5] and are now widely used to characterize the heterogeneous and multiscale structure of many social, economic, political, transportation, biological and ecological

systems [6–9]. In addition, the growth of network science has had a huge impact on research on dynamical systems and stochastic processes. Indeed, we now know that, instead of more traditional lattice-based or well-mixed coupling schemes, considering complex patterns of interactions among coupled agents, from the small-world phenomenon to high clustering and degree heterogeneity [10, 11], can dramatically impact even the simplest dynamical models. To this day, networks are the natural substrate to describe complex emerging phenomena such as diffusion [12], synchronization [13], spreading [14], social dynamics [15] or cooperation [16] in complex systems.

However, despite their success, networks can only provide a simplified description of reality. Indeed, links encode connections among single pairs of nodes, inherently constraining a faithful system representation to dyadic relationships only. Yet, from biological to ecological and social systems, the basic units of a system routinely interact in groups of three of more nodes at a time. For this reason, more comprehensive mathematical frameworks are needed, which are able to fully capture the richness stemming from such higher-order interactions [17, 18]. Hypergraphs and simplicial complexes are the two most commonly adopted representations that capture by construction such group interactions.

In their simplest definition, hypergraphs [19] are natural extensions of networks with the additional property that (hyper)edges can contain more than two nodes and hence describe higher-order interactions. Just like for networks, many different flavours of hypergraphs exist: hyperedges can be weighted, directed, or both; hypergraphs can be multilayer, temporal and so on. In fact, their extreme generality as a framework to encode interactions justifies their widespread use in applications ranging from computer science to biology. At the same time, they are not fully free from certain difficulties in interpretation. For example, there are many different Laplacian operators compatible with the same hypergraph, which in turn complicates the understanding of their effects on dynamical processes [17].

Simplicial complexes [20] find themselves at the other extreme of this spectrum of flexibility. They are able to describe groups as well as hypergraphs in terms of simplices (collections of vertices), but are much more constrained in which simplices must be present in a complex. In fact, simplicial complexes have an internal closure condition, meaning that for each simplex in the complex all its sub-simplices must be contained in the complex too. This induces a certain rigidity in modeling efforts. For instance, it is not possible to describe the situation in which three nodes are part of a group as a triple, but are not part of other groups in pairs, which is instead trivial to represent in hypergraphs through a single hyper-edges with three nodes. A positive side of this limitation is, however, that simplicial complexes (and related representations, e.g. cell complexes) provide access to powerful mathematical tools that bridge between geometry, topology and dynamical systems [21]. Indeed, a large fraction of the recent development in computational approaches to data spaces and their topology is based on simplicial descriptions [22] .

Against this background, this book introduces the most common higher-order descriptions and their properties, with the explicit aim to provide a comprehensive overview of the state-of-the-art for higher-order systems [17]. Its chapters are contributed by leading researchers and research teams from a variety of backgrounds and disciplines, from nonlinear dynamics and topological data analyses to biology and neuroscience. Our target has been to cover such emerging field both in breadth and in depth. For this reason , on the one hand, some chapters are organised as reviews on central key topics, such as models of higher-order systems, topological data analysis, or social contagions. On the other hand, others provide detailed accounts of specific findings, from explosive synchronization and chaos in higher-order synchronization to collective games. More in detail, the book is organized in three main parts:

- The first part of the book is devoted to the mathematical characterization of higher-order structures. In Chap. 1 Mulas, Horak and Jost begin with a thorough introduction on the mathematical formalisms of hypergraphs and simplicial complexes and their corresponding spectral theories. In Chap. 2 Bobrowski and Krioukov discuss geometric and topological models of higher-order structures. In Chap. 3 Vaccarino, Fugacci and Scaramuccia presents an overview of algebraic and topological tools for the analysis of real-world data. In Chap. 4 Eriksson, Carletti, Lambiotte, Rojas and Rosvall focus on the presence of complex mesoscale structures known as communities, and on flow-based methods to extract them in systems with higher-order interactions.
- The second part of the book discusses dynamical processes unfolding on higher-order systems. In Chap. 5 Carletti and Fanelli overview reaction-diffusion schemes and pattern formation in hypergraphs. In Chap. 6 Pikovsky and Rosenblum introduce a theory of higher-order mean-field phase coupling, and discuss harmonics, phase reduction and higher-order network reconstruction from dynamics. In Chap. 7 Ashwin, Bick and Rodrigues focus on heteroclinic dynamics and chaos in coupled phase oscillators with higher-order interactions. In Chap. 8 Skardal and Arenas review explosive phenomena and multistability in systems with higher-order interactions. In Chap. 9 Lucas, Cencetti and Battiston introduce a multiorder Laplacian operator, which allows for a spectral characterization of Kuramoto dynamics with higher-order interactions. In Chap. 10, Gambuzza, Di Patti, Gallo, Lepri, Romance, Criado, Frasca, Latora and Boccaletti exploit the same operator to extend the master stability function approach to simplicial complexes. In Chap. 11, Millán, Restrepo, Torres and Bianconi introduce simplicial synchronization, where state variables are placed not only on the nodes, but also on the links of a higher-order system. In Chap. 12 Schaub, Seby, Frantzen, Roddenberry, Zhu and Segarra overview signal processing in the presence of higher-order interactions. In Chap. 13 Barrat, de Arruda, Iacopini and Moreno survey contagion processes in higher-order systems. In Chap. 14 Neuhäuser, Lambiotte and Schaub discuss opinion dynamics with multibody interactions. Finally, in Chap. 15 Alvarez-Rodriguez, Battiston, de Arruda, Moreno, Perc and Latora discuss collective games with higher-order interactions

- The third part of the book discusses applications of the higher-order network framework. In particular, in Chap. 16 Feng, Hickok and Porter focus on application of topological data analyses to spatial systems. In Chap. 17 Expert and Petri summarise recent results and open challenges in the higher-order organization of the human brain. In Chap. 18 Ogbunugafor and Scarpino discuss the role of higher-order interactions in biological systems and epistasis in particular.

Overall, we hope that this book will provide, as a single resource, a useful guide to navigate the most important findings of the emergent field of higher-order systems, which we believe will become an important reference for complexity scientists in coming years.

Vienna, Austria — Federico Battiston
Turin, Italy — Giovanni Petri

References

1. P.W. Anderson, Science **177**(4047), 393 (1972)
2. A.L. Barabási, Nat. Phys. **8**(1), 14 (2011)
3. A. Vespignani, Nat. Phys. **8**(1), 32 (2012)
4. B. Bollobás, B. Béla, *Random graphs*, vol. 73 (Cambridge University Press, 2001)
5. L. Lovász, *Large Networks and Graph Limits*, vol. 60 (American Mathematical Society, 2012)
6. R. Albert, A.L. Barabási, Rev. Mod. Phys. **74**(1), 47 (2002)
7. S.N. Dorogovtsev, J.F.F. Mendes, Adv. Phys. **51**(4), 1079 (2002)
8. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
9. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Phys. Rep. **424**(4–5), 175 (2006). (Fervier)
10. D.J. Watts, S.H. Strogatz, Nature **393**(6684), 440 (1998)
11. A.L. Barabási, R. Albert, Science **286**(5439), 509 (1999)
12. N. Masuda, M.A. Porter, R. Lambiotte, Phys. Rep. (2017)
13. A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Nat. Phys. **17**(10), 1093–1098 (2022)
14. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Rev. Mod. Phys. **87**(3), 925 (2015)
15. C. Castellano, S. Fortunato, V. Loreto, Rev. Mod. Phys. **81**(2), 591 (2009)
16. G. Szabó, G. Fath, Phys. Rep. **446**(4–6), 97 (2007)
17. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, Phys. Rep. **874**, 1 (2020)
18. F. Battiston, E. Amico, A. Barrat, G. Bianconi, G.F. de Arruda, B. Franceschiello, I. Iacopini, S. Kefi, V. Latora, Y. Moreno, M. Murray, T. Peixoto, F. Vaccarino, G. Petri, Nat. Phys. (2021)
19. C. Berge, (1973), *Graph and Hypergraphs*. North-Holland Publishing Company, Amsterdam
20. P.S. Aleksandrov, *Combinatorial Topology*, vol. 1 (Courier Corporation, 1998)
21. L. Wasserman, Ann. Rev. Stat. Appl. **5**, 501 (2018)
22. G. Carlsson, Nat. Rev. Phys. **2**(12), 697 (2020)

Contents

Chapter 1
Graphs, Simplicial Complexes and Hypergraphs: Spectral Theory and Topology

Raffaella Mulas, Danijela Horak, and Jürgen Jost

Abstract In this chapter we discuss the spectral theory of discrete structures such as graphs, simplicial complexes and hypergraphs. We focus, in particular, on the corresponding Laplace operators. We present the theoretical foundations, but we also discuss the motivation to model and study real data with these tools.

Keywords Simplicial complexes · Hypergraphs · Laplace operators · Eigenvalues · Topology

1.1 Introduction

1.1.1 Motivating Examples

Example 1 Consider scientists A, B, C, D that work in the same field and assume that there exists a joint publication of A, B, C, a single author paper of A, a joint paper of B, C and one of C, D. What are formal structures that model these relations?

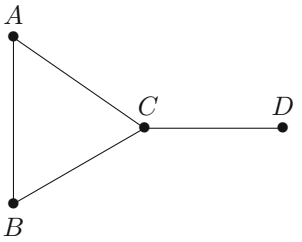
(a) A **graph**: We take A, B, C, D as the vertices of a graph and connect vertices by an edge when the two scientists are coauthors. Thus, there are edges $(A, B), (B, C), (A, C), (C, D)$.

R. Mulas
Alan Turing Institute of London and the University of Southampton, 2QR, 96 Euston Rd, London NW1 2DB, UK

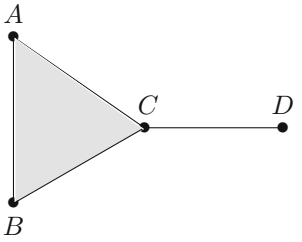
R. Mulas · J. Jost (✉)
Max Planck Institute for Mathematics in the Sciences, lnselstr. 22, 04103 Leipzig, Germany
e-mail: juergen.jost@mis.mpg.de

D. Horak
AIG London, The AIG Building, 58 Fenchurch Street, London EC3M 4AB, UK

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_1

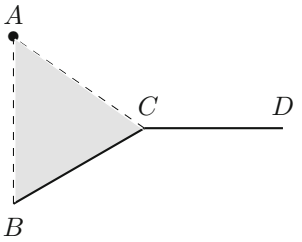


(b) A **simplicial complex**: The preceding model does not distinguish the fact that there is a 3-author paper between A, B, C from the possibility that there might be three 2-author papers between A, B, between B, C and between A, C. The latter would yield the same graph as the former. We therefore add a 2-dimensional simplex (A, B, C) to represent the 3-author paper.



dummy

(c) A **hypergraph**: Still, this does not account for the facts that while there is a joint paper between B, C, there are no 2-author papers involving A, B only or A, C only. Nor does it account for the fact that only A has a single author paper. We therefore consider the hypergraph with hyperedges (A), (B, C), (C, D), (A, B, C) that represents the full information about the collaboration pattern.



Example 2 We consider a metric space (X, d). For simplicity, we assume that it is finite, although that is not needed for the construction. Thus, we have points $x_1, \ldots, x_N$ with mutual distances $d(x_i, x_j)$ that are positive for $x_i \neq x_j$, symmetric $(d(x_i, x_j) = d(x_j, x_i))$ and satisfy the triangle inequality $(d(x_i, x_j) \leq d(x_i, x_k) + d(x_k, x_j))$. For $x \in X$ and $\rho \geq 0$, we consider the closed ball $B(x, \rho) := \{y \in X : d(x, y) \leq \rho\}$ of radius ρ. For radii $r_i > 0, i = 1, \ldots .N$, we then form a complex

by associating a $(d-1)$-dimensional simplex to every family $(x_{i_1}, r_{i_1}), \ldots, (x_{i_d}, r_{i_d})$ that satisfies

$$\bigcap_{j=1,\ldots,d} B(x_{i_j}, r_{i_j}) \neq \emptyset, \tag{1.1}$$

that is, whenever these d balls have some point in common, for $d = 0, \ldots, N$. This is a simplicial complex (see the definition below), because whenever some collections of balls has a nonempty intersection, then this also holds for any subcollection. Such a complex is called a *Čech complex*, and it will play an important role in our discussion of topology. When we take all $r_i \equiv r$ and then let r vary, we get the family of complexes whose topology yields the bar codes of topological data analysis (TDA).

Example 3 We consider a system of chemical reactions with substances A, B, C, D, E. Reaction 1 transforms A, B into C, with E acting as a catalyst. Reaction 2 has B, C as inputs and A, D as outputs.

(a) We represent this by a **directed hypergraph** with one hyperedge $(A, B, E; C, E)$ going from the inputs A, B, E to the outputs C, E (note that the catalyst E appears both as an input and as an output) and another hyperedge $(B, C; A, D)$.
(b) Alternatively, we construct a **(directed) bipartite graph** with vertices A, B, C, D, E, 1, 2, with edges from A, from B and from E to 1, from 1 to C and to E, from B and from C to 2, and from 2 to A and to D.

1.1.2 How Can We Handle the Data?

While the preceding describes some modelling options, we should keep in mind that the real data sets are much larger. There are probably millions of scientists, with intricate patterns of collaborations. And by now, more than 20 million chemical substances and 40 million reactions have been reported in the chemical literature (see [1, 2]).

Therefore, we need not only mathematical concepts to model the data, but also mathematical tools to extract qualitative or quantitative information from the models about the data. Since (hyper)graphs and simplicial complexes are combinatorial objects, we may have to face the danger of combinatorial explosion, that the effort needed to analyze them grows exponentially with their size. For instance, the question whether two (unlabelled) graphs are isomorphic to each other is known to be in the complexity class NP [3]. That is, we cannot easily decide whether two data structures modelled as graphs are abstractly the same or different. And, of course, since simplicial complexes or hypergraphs are more complicated structures than graphs, the difficulties become even more severe.

But when we look at the global structures, for instance the collaboration pattern of all scientists in some field, or the chemical hypergraph consisting of all known substances and reactions, perhaps we are not so much interested in local details, but

rather wish to extract some global qualitative information that identifies characteristic features of the data sets. And the methods to extract that kind of information should scale nicely with the data size.

Here, we describe one such method, the spectral analysis (for networks modelled as graphs, this has been systematically applied in [4–6]). This can be seen as some kind of Fourier transform. We define a linear operator (or several in the case of simplicial complexes) acting on functions on the vertices (or more generally, the simplices) and consider its spectrum, that is, the collection of its eigenvalues. There exist quick and robust numerical schemes to compute these eigenvalues (or possibly only those that are most interesting for our purposes). And as we shall explain in this contribution, from these eigenvalues we can read off many important qualitative properties of the underlying structure. This requires some mathematical theory, as we shall describe. The rewards are high, as we shall see.

Of course, the spectrum cannot provide full information about the underlying structure. That is, two (hyper)graphs or two simplicial complexes with the same spectra need not be isomorphic. In fact, it is not surprising that graphs cannot always be distinguished by their spectra, since, as mentioned, the graph isomorphism problem is in the complexity class NP, and therefore, according to current belief, is unlikely to be possible to solve it in polynomial time. However, spectra provide important information about many qualitative properties.

We should point out that there are also other geometric quantities that are on one hand readily computed and on the other yield useful qualitative information about the underlying structure. We should mention in particular the so-called *Ricci curvatures* whose statistics reveal important patterns (see for instance [7]). The name *curvature*, while perhaps not being very appropriate for a discrete structure, reveals its origin in Riemannian geometry (see [8]). In fact, also the spectral theory of Laplace operators was originally explored in Riemannian geometry (see [9]), and many ideas developed there are also useful here.

In any case, in this contribution, in line with the purpose of the present volume, we only discuss the spectral theory of discrete structures.

1.1.3 Definitions of Structures

We shall begin with the simplest structures before we shall subsequently enrich or decorate them. Keeping real data in mind, all objects will be assumed *finite*.

The starting point is a finite collection V of elements. As such, such a collection is amorphous. But we want to assume that these elements stand in certain discrete relations, and these relations then provide us with the structure to work with. The basic definition shall now identify some such structures that we shall explore in this contribution.

Definition 1.1.1

- A *graph* consists of a set V of *vertices* and a set E of *edges* which are unordered pairs of different vertices. When $e = (i, j)$ is an edge between the vertices $i, j \in V$, we say that i and j are *neighbors* and write $i \sim j$. The *degree* $\deg i$ is the number of neighbors of i.
- A *hypergraph* consists of a set V of *vertices* and a set H of *hyperedges* which are nonempty sets of vertices. An *oriented hypergraph* has hyperedges consisting of two disjoint sets of vertices, called the inputs and the outputs of the corresponding hyperedge. Either of them, but not both might be empty.
- A *simplicial complex* is defined on a set of vertices V as a subset of its power set, i.e. $\mathcal{C} \subset \mathcal{P}(V)$ that is closed under taking subsets, i.e. for a *simplex* $C \in \mathcal{C}$, any of its subsets $C' \subset C$ is also a simplex in $\mathcal{C}$. In contrast to the previous two definitions, the empty set is a simplex. A simplex c with $d + 1$ vertices is called a d-simplex (with d being considered as its dimension), its subsimplices are called its *faces*, and its $(d - 1)$-dimensional faces are called its *facets*. The *down degree* of c is the number of its facets, and its *up degree* is the number of $(d + 1)$-simplices of which c is a facet.

Every simplicial complex is a hypergraph when we consider every simplex (except the empty one) as a hyperedge. As it turns out, however, this is not a very useful perspective, and it is much better to develop the theories for these two concepts differently.

We can enrich this definition by providing additional structure.

Definition 1.1.2

- The graph (V, E) is *directed* when the edge set E contains ordered pairs of vertices. Likewise, the oriented hypergraph (V, H) is directed when we consider each hyperedge $h = (h_1; h_2)$ as going from the input set h_1 to the output set h_2.
- We may allow the graph (V, E) to contain self-loops, that is, edges of the form $e = (i, i)$ with $i \in V$.
- The (hyper)graph or simplicial complex is *weighted* when each vertex, (hyper)edge, simplex is allowed to carry some real number as its weight. When, for instance, the edges $e = (i, j)$ of a graph (V, E) carry weights $w_e = w_{ij}$, we put $\deg i = \sum_j w_{ij}$. Here, $w_{ij} = 0$ may simply express the fact that i and j are not connected by an edge.

In most cases, the weights are assumed to be nonnegative (so as to make also the vertex degrees nonnegative), but in certain cases, also negative weights might be permitted. Since subsequently, we want to normalize by the degree, we may want to assume that $\deg i \neq 0$ for all i.

1.1.4 Conventions

The conventions for the sign and the eigenvalues of Laplacians vary in the literature, and even in our own papers. This is partly due to the fact that in different circumstances, different conventions seem to be most natural and convenient.

Here, we make the Laplacian into a positive operator. Since the Laplacian is also selfadjoint, its spectrum is real and nonnegative. We denote its smallest eigenvalue by λ_1, and not by λ_0, as often done in the literature, as in the basic case of a graph, that smallest eigenvalue is $= 0$. For a simplicial complex or a hypergraph, that smallest eigenvalue, however, in general is no longer 0, or in contrast, there might be several eigenvalues $= 0$ while for a connected graph, there is only one such eigenvalue.

1.2 Graphs

While this contribution is not about graphs, we nevertheless develop their theory first, because both simplicial complexes and hypergraphs can be advantageously treated as generalizations of graphs. This is particularly true for the Laplacians and their spectra.

We shall develop the theory in such a manner that the generalizations appear most natural. We should also point out that the theory to be developed here is inspired by the corresponding theory in Riemannian geometry, see [9]. Systematic presentations of the spectral theory of graphs can be found in [10, 11]. While we shall partly use those references, we develop the theory here somewhat differently, in order to motivate and to facilitate the generalizations to simplicial complexes [12, 13] and hypergraphs [14] which is the main interest of this contribution.

1.2.1 The Laplacian

We start with the formula for the Laplace operator on an unweighted and undirected graph.

Definition 1.2.1 Let Γ be a graph with vertex set V and edge set E without isolated vertices (i.e., $\deg v > 0$ for every $v \in V$). Its *(vertex) Laplace operator* (or *Laplacian* for short) operates on functions $f : V \to \mathbb{R}$ via

$$L^0 f(v) := f(v) - \frac{1}{\deg v} \sum_{w \sim v} f(w) \quad \text{for } v \in V. \tag{1.2}$$

This definition is easy to understand: The Laplacian takes the value of the function f at the vertex v in question and subtracts from it the average of the values of the neighbors of v. It can also be interpreted as the negative of the generator of a random walk on the vertex set, where the random walker, when it finds herself at the vertex v, randomly selects one of the neighbors of v for her next position.

Warning: The Laplacian L^0 is **not** the *algebraic Laplacian* used in classical graph theory. That latter operator is given by

$$\mathcal{L}f(v) := \deg v \; f(v) - \sum_{w \sim v} f(w) \quad \text{for } v \in V. \tag{1.3}$$

When the graph is *regular*, that is $\deg v \equiv const$, then L^0 and $\mathcal{L}$ differ only by a constant factor. For general graphs, their theory is different, however. L^0 has better mathematical properties than $\mathcal{L}$ and, in particular, generalizes much more readily to both simplicial complexes and hypergraphs. That is why we take L^0 as our basic object.

We are interested in the spectrum, that is, the eigenvalues of L^0. In order to derive properties of that spectrum, we need to introduce some more mathematical structure. Perhaps somewhat surprisingly, we shall begin by introducing another operator that operates on functions on oriented edges. Here, the oriented edge $e = [v, w]$ is considered as going from the tail or input v to the head or output w. A change of orientation corresponds to reversing the ordering and considering $e^- = [w, v]$. That is, $e^+ = [v, w]$ and $e^- = [w, v]$ carry opposite orientations. We shall understand below why this is a natural construction. From now on, we let E denote the set of *oriented* edges of our graph. We then consider functions γ defined on oriented edges that are required to satisfy

$$\gamma(e^-) = -\gamma(e^+) \tag{1.4}$$

for all oriented edges. (A reader knowledgable in Riemannian geometry [9] might be reminded of exterior 1-forms.) We then define a Laplacian operating on such functions.

Definition 1.2.2 The Laplacian for functions on oriented edges is

$$L^1\gamma(e) := \frac{1}{\deg v_0} \cdot \sum_{v_0 \in e' = [v_0, w]} \gamma(e') - \frac{1}{\deg v_1} \cdot \sum_{v_1 \in e'' = [v_1, w]} \gamma(e''), \tag{1.5}$$

where $e = [v_0, v_1]$, and γ has to satisfy 1.4.

Again, this Laplacian is easy to understand. When we consider γ as a flow along oriented edges, then it compares the difference between out- and inflow at the tail with the difference between in- and outflow at the head. That is, it compares the flow in the direction of the edge, what comes in at the tail and what goes out at the head, with that in the opposite direction, what goes out at the tail and what comes in at the head. That is, $\gamma(e)$ measures the net flow of γ through the oriented edge e. Changing the orientation of e changes the sign of this Laplacian, that is,

$$L^1\gamma(e^-) = -L^1\gamma(e^+), \tag{1.6}$$

in accordance with 1.4.

As we shall see in a moment, the two operators L^0 and L^1 are intimately related. In particular, they have the same spectrum, apart from possibly the multiplicity of the eigenvalue 0.

We first define the boundary operator for a function $f : V \to \mathbb{R}$. Let $e = [v_0, v_1]$ be an oriented edge, then

$$\delta f(e) := f(v_0) - f(v_1). \tag{1.7}$$

In order to define an adjoint of δ, we shall utilize a scalar product for functions $f, g : V \to \mathbb{R}$ on vertices, defined by

$$(f, g)_V := \sum_{v \in V} \deg v \cdot f(v) \cdot g(v) \tag{1.8}$$

and a scalar product for functions $\omega, \gamma : E \to \mathbb{R}$ on oriented edges

$$(\omega, \gamma)_E := \sum_{e \in E} \omega(e) \cdot \gamma(e). \tag{1.9}$$

Lemma 1.2.1 *The operator*

$$\delta^* : \{\gamma : E \to \mathbb{R}\} \longrightarrow \{f : V \to \mathbb{R}\}$$

defined as

$$\delta^*(\gamma)(v) := \frac{\sum_{e'=[v,w]} \gamma(e') - \sum_{e''=[w,v]} \gamma(e'')}{2 \deg v} \tag{1.10}$$

is the adjoint of δ in the sense that

$$(\delta f, \gamma)_E = (f, \delta^* \gamma)_V \text{ for all } f, \gamma. \tag{1.11}$$

Proof

$$\begin{aligned}
(\delta f, \gamma)_E &= \sum_{e=[v_0,v_1]} \gamma(e) \cdot (f(v_0) - f(v_1)) \\
&= \frac{1}{2} \sum_{v \in V} f(v) \cdot \left(\sum_{e'=[v,w]} \gamma(e') - \sum_{e''=[w,v]} \gamma(e'') \right) \\
&= \sum_{v \in V} \deg v \cdot f(v) \cdot \frac{\sum_{e'=[v,w]} \gamma(e') - \sum_{e''=[w,v]} \gamma(e'')}{2 \deg v} \\
&= \sum_{v \in V} \deg v \cdot f(v) \cdot \delta^*(\gamma)(v) \\
&= (f, \delta^* \gamma)_V .
\end{aligned}$$

□

Lemma 1.2.2 *We have*

$$L^0 = \delta^*\delta \tag{1.12}$$

and

$$L^1 = \delta\delta^*. \tag{1.13}$$

Proof

$$\begin{aligned}\delta^*(\delta f)(v) =& \frac{\sum_{e'=[v,w']} \delta f(e') - \sum_{e''=[w'',v]} \delta f(e'')}{2\deg v} \\ =& \frac{\sum_{e'=[v,w']}(f(v) - f(w'))}{2\deg v} + \\ & - \frac{\sum_{e''=[w'',v]}(f(w'') - f(v))}{2\deg v} \\ =& L^0 f(v)\end{aligned}$$

and for $e = [v_0, v_1]$

$$\begin{aligned}\delta(\delta^*\gamma)(e) =& \delta^*\gamma(v_0) - \delta^*\gamma(v_1) \\ =& \frac{\sum_{e'=[v_0,w]} \gamma(e') - \sum_{e''=[w,v_0]} \gamma(e'')}{2\deg v_0} \\ & - \frac{\sum_{f'=[v_1,w]} \gamma(f') - \sum_{f''=[w,v_1]} \gamma(f'')}{2\deg v_1} \\ =& L^1\gamma(e).\end{aligned}$$

□

Corollary 1.2.1 *We have*

$$(f, L^0 f)_V = (\delta f, \delta f)_E = (L^0 f, f)_V \tag{1.14}$$

and

$$(\gamma, L^1\gamma)_E = (\delta^*\gamma, \delta^*\gamma)_V = (L^1\gamma, \gamma)_E \tag{1.15}$$

for all f, γ.
In particular, the operators L^0 and L^1 are self-adjoint and nonnegative, and all their eigenvalues are real and nonnegative.
L^0 and L^1 have the same spectrum, except possibly for the multiplicity of the eigenvalue 0.

Proof 1.14 and 1.15 follow from Lemmas 1.2.1 and 1.2.2, and these relations then imply the next claim. Finally, two operators AB and BA have the same eigenvalues except possibly for the multiplicity of the eigenvalue 0. □

Remark The algebraic Laplacian $\mathcal{L}$ from 1.3 is self-adjoint w.r.t. the product

$$(f, g)_{\text{alg}} := \sum_v f(v)g(v).$$

1.2.2 The Spectrum

We consider a graph $\Gamma = (V, E)$ with N vertices.

Definition 1.2.3 We say that Γ is *connected* if for any $v', v'' \in V$, there exist vertices $v_1, \ldots, v_k$ and (unoriented) edges $e_0 = (v', v_1), e_1 = (v_1, v_2), \ldots, e_{k-1} = (v_{k-1}, v_k)$, $e_k = (v_k, v'')$, that is, if v' and v'' can be connected by a sequence of edges. Such a sequences of edges is called a *path* from v' to v''.

We shall usually assume that our graphs are connected, even though we may not always say so explicitly. For a graph that is not connected, we can simply treat its connected components individually. We shall also usually assume that our graph (or, if not connected, every connected component) has more than one vertex. Thus, every vertex is assumed to support at least one edge, and hence have positive degree.

Lemma 1.2.3 *Let $v_0 \in V$ be a local maximum of $f : V \to \mathbb{R}$, that is $f(v) \geq f(w)$ for all $w \sim v$. Then*

$$L^0 f(v) \geq 0, \tag{1.16}$$

and in fact $L^0 f(v) > 0$ unless $f(w) = f(v)$ for all $w \sim v$.

Proof If $f(v) \geq f(w)$ for all $w \sim v$, then also $f(v) \geq \frac{1}{\deg v} \sum_{w \sim v} f(w)$. The last claim then is obvious. □

Corollary 1.2.2 *If Γ is connected, then $L^0 f = 0$ implies $f \equiv const$.*

Proof Let v_0 be a local maximum of f. Then by the Lemma, $f(w) = f(v)$ for all $w \sim v$. Thus, all neighbors of v are also local maxima, with the same value of f. Iterating the argument, also all the neighbors of those neighbors have the same (maximal) value of f, and so on. Since Γ is connected, we can reach each vertex from v through a path connecting local neighbors, and therefore f is constant. □

Corollary 1.2.3 *If Γ is connected, then $\lambda = 0$ is a simple eigenvalue of L^0, and it is an eigenvalue of L^1 with multiplicity $|E| - |V| + 1$.*

Proof By Corollary 1.2.2, the only solutions of $L^0 f = 0 \cdot f$ are the constants. Therefore, the corresponding Eigenspace is 1-dimensional. The vector space on which L^0

operates is $|V|$-dimensional, and that of L^1 is $|E|$-dimensional. Hence they have $|V|$ and $|E|$ eigenvalues counted with multiplicity. In particular, by what we have already proved, L^0 has $|V|-1$ nonzero eigenvalues. By Corollary 1.2.1, these are also the nonzero eigenvalues of L^1. Hence the eigenvalue 0 of L^1 must have multiplicity $|E|-|V|+1$. □

Lemma 1.2.4 *All eigenvalues of L^0 and L^1 satisfy*

$$0 \leq \lambda \leq 2. \tag{1.17}$$

Proof The eigenvalues are nonnegative by Corollary 1.2.1. Let f_λ be an eigenfunction for the eigenvalue $\lambda > 0$. Then

$$\lambda f(v) = f(v) - \frac{1}{\deg v} \sum_{w \sim v} f(w). \tag{1.18}$$

Let v be a vertex where $|f|$ assumes its maximum. Then

$$\lambda |f(v)| \leq |f(v)| + \frac{1}{\deg v} \sum_{w \sim v} |f(w)| \leq 2|f(v)| \tag{1.19}$$

which implies $\lambda \leq 2$. □

Corollary 1.2.4 *The eigenvalue $\lambda = 2$ is attained for an eigenfunction that satisfies*

$$f(w) = -f(v) \text{ whenever } w \sim v. \tag{1.20}$$

Such an eigenfunction exists if and only if the graph is bipartite, that is, has two classes V_1, V_2 of vertices and allows only connections between vertices from different classes.

Proof 1.20 is the condition for equality in 1.19. And we can construct such an f only in the bipartite case, where we can take $f(v) = 1$ for $v \in V_1$, $f(v) = -1$ for $v \in V_2$. □

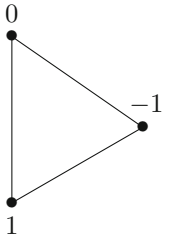
It is often convenient to reformulate the eigenvalue Eq. 1.18 as

$$(1-\lambda) f(v) = \frac{1}{\deg v} \sum_{w \sim v} f(w). \tag{1.21}$$

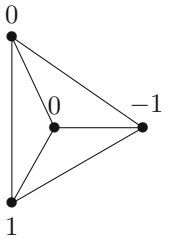
Thus, not only are the eigenvalues 0 and 2 extremal, but also the eigenvalue 1 is special as in that case, the right hand side of 1.21 vanishes, that is, the average of the values of f over the neighbors of any vertex vanishes.

Let us consider some simple **examples** where we can determine the eigenvalues of L^0 (and hence also infer those of L^1 by Corollary 1.2.3). Further examples can be found in Sect. 1.3.4.

is an eigenfunction for the eigenvalue 2 for the two-vertex graph K_2.



is an eigenfunction for the complete graph K_3 with three vertices, for the eigenvalue $\frac{3}{2}$. By permuting the vertices, we obtain another linearly independent such eigenfunction with the same eigenvalue. Hence the spectrum of K_3 is $(0, \frac{3}{2}, \frac{3}{2})$.



is an eigenfunction for the complete graph K_4 with four vertices, for the eigenvalue $\frac{4}{3}$. By permuting the vertices, we obtain two further linearly independent such eigenfunction with the same eigenvalue. Hence the spectrum of K_4 is $(0, \frac{4}{3}, \frac{4}{3}, \frac{4}{3})$.

Inductively, we see that the complete graph K_N has the eigenvalue 0 with multiplicity 1 and the eigenvalue $\frac{N}{N-1}$ with multiplicity $N-1$.

We can also see a principle here that will also be useful for determining the spectra of simplicial complexes and hypergraphs. Whenever we have two vertices v_1, v_{-1} with the same other neighbors, we obtain an eigenfunction when we put

$$f(v) := \begin{cases} 1 & \text{for } v = v_1 \\ -1 & \text{for } v = v_{-1} \\ 0 & \text{else} \end{cases} \tag{1.22}$$

as is readily checked from 1.2.

Let us explore this observation a bit more. First of all, such a graph possesses an automorphism ι that exchanges v_1 and v_{-1} and leaves all other vertices fixed. Secondly, as will become clear below, we may assume that all other eigenfunctions f' satisfy

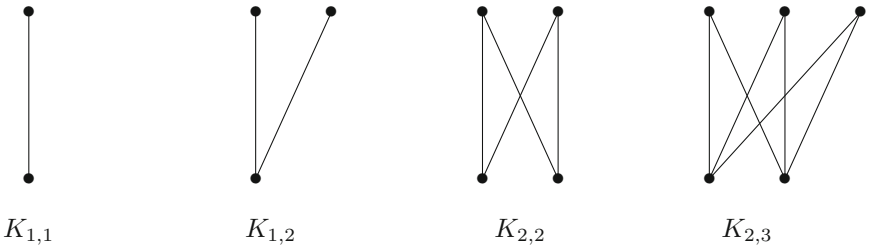
$$f'(v_1) = f'(v_{-1}). \tag{1.23}$$

This is useful for an iterative determination of the eigenvalues and eigenfunctions. Third, when v_1 and v_{-1} are not neighbors, then the eigenvalue for f from 1.22 is $= 1$, because then, by 1.21, for every vertex u

$$\sum_{w \sim u} f(w) = 0. \tag{1.24}$$

We may then say that v_1 and v_{-1} are *duplicates* of each other.

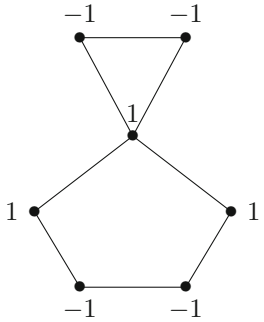
With this principle [15], we can iteratively determine the entire spectrum of the *complete bipartite graph* $K_{m,n}$. That graph has two groups V_1, V_{-1} of m and n vertices, resp., such that every vertex in one group is connected to every vertex in the other group, but there are no connections inside either group. This graph is obtained from $K_{1,1} = K_2$ by the repeated duplication of nodes



$K_{m,n}$ then has the eigenvalue 0 with multiplicity 1 and the eigenvalue 1 with multiplicity $m + n - 2$. The remaining eigenvalue is 2. In fact, every graph that is *bipartite*, meaning that it has two groups V_1, V_{-1} of vertices without any connections inside either group, has that eigenvalue, with an eigenfunction that is $= 1$ on V_1 and $= -1$ on V_{-1}. In fact, only bipartite graphs can carry the eigenvalue 2, as the condition 1.20 of Corollary 1.2.4 can only be satisfied on such graphs.

An example of a complete bipartite graph is the *star graph* $K_{1,n}$ that has one central vertex connected to n peripheral ones.

Remark While the presence of the eigenvalue 1, as in 1.24, can be seen as an indication of vertex duplication in empirical spectra [16], we should point out that this eigenvalue can also arise in other graphs:



Here, the eigenfunction for the eigenvalue 1 is invariant under the automorphisms of the graph.

We now develop a more abstract perspective on symmetries and spectra, following [17].

Definition 1.2.4 An *automorphism* of the graph $\Gamma = (V, E)$ consists of bijections $\sigma : V \to V$ and $\sigma : E \to E$ such that $\sigma(v) \in \sigma(e)$ if and only if $v \in e$.
For $f : V \to \mathbb{R}$, we then put $\sigma_* f(v) = f(\sigma(v))$.

The Laplace operator commutes with automorphisms.

Lemma 1.2.5 *If σ is an automorphism of $\Gamma = (V, E)$, then*

$$L^0(\sigma_* f)(v) = \sigma_*(L^0 f)(v) \tag{1.25}$$

for all $v \in V$ and $f : V \to \mathbb{R}$.

The *proof* is obvious.

We can use Lemma 1.2.5 to decompose the spectrum of L^0. Let τ be an automorphism of $\Gamma = (V, E)$ with

$$\tau^2 = \mathrm{id}. \tag{1.26}$$

Then τ has two possible eigenvalues, ± 1, on the space of functions $f : V \to \mathbb{R}$, and L^0 leaves those two Eigenspaces $L_\pm$ invariant. Also, $V = V_0 \cup V_1$ where V_0 is the set of those vertices that are fixed by τ. Thus $\tau(v_0) = v_0$ if and only if $v_0 \in V_0$. We write $V_1 = V' \cup V''$ where V', V'' are disjoint and $\tau(V') = V''$. Since also $\tau(V'') = V'$ because of (1.26), V' and V'' play symmetric roles.

Without loss of generality, we assume that V' (and hence also V'') is connected, as otherwise we can rearrange the decomposition of V_1 and/or write τ as the composition of several such automorphisms.

Lemma 1.2.6 *We can decompose V into spaces generated by symmetric and antisymmetric eigenfunctions. More precisely, we have a $|V'|$-dimensional space of functions $f : V \to \mathbb{R}$ generated by eigenfunctions of L^0 that vanish on V_0 and that are antisymmetric on V' and V'', $f(v'') = -f(v')$ if $v'' = \tau(v') \in V''$ for $v' \in V'$. The remaining $(|V'| + |V_0|)$-dimensional space is generated by eigenfunctions that are symmetric on V' and V'', that is, $f(v'') = f(v')$.*

Proof The first class of functions are those that are eigenfunctions of τ for the eigenvalue -1, and the second class has eigenvalue 1. By Lemma 1.2.5, these are unions of Eigenspaces of L^0. Since $|V''| = |V'|$ and $V = V_0 \cup V' \cup V''$, this generates the space of all functions on V.

Definition 1.2.5 Let $\Gamma = (V, E)$ be a graph. An *induced subgraph*, also called a *motif*, $\hat{\Gamma}$ has some nonempty vertex set $\hat{V} \subset V$ and an edge set $\hat{E} \subset E$ such that any two $v_1, v_2 \in \hat{V}$ are contained in an edge $e \in \hat{E}$ whenever they are contained in e in Γ.

Let $\hat{\Gamma}$ be a motif in Γ. We then have the induced Laplacian

$$L^0_{\Gamma,\hat{\Gamma}} f(v) = f(v) - \frac{1}{\deg_\Gamma v} \cdot \left(\sum_{v' \in \hat{V}, v' \sim v} f(v') \right) \tag{1.27}$$

where $\deg_\Gamma v$ denotes the degree of v in Γ.

Definition 1.2.6 We say that the motif Γ' with vertex set V' is a *duplicated motif* if V' and V'' are disconnected, that is, when there is no edge containing elements from both V' and V''.
We say that Γ' and Γ'' with vertex sets V' and V'' are *twin motives* if for every $e \in E$ we have that $v' \in e$ if and only if $v'' = \tau(v') \in e$.

Lemma 1.2.7 *Let Γ' be a duplicated motif in Γ, and let $v_0 \in V_0$ be a neighbor of some $v' \in V'$. Then v_0 is also a neighbor of $v'' = \tau(v') \in V''$.*

Proof Since $v_0 \in V_0$ is fixed by the automorphism τ, and since τ maps the edge e containing v_0 and v' onto an edge $\tau(e)$ containing $v'' = \tau(v')$ and $v_0 = \tau(v_0)$, the claim follows.

Lemma 1.2.8 *Let Γ' be a duplicated motif in Γ. Then we find a basis of eigenfunctions of the Laplacian L^0 of Γ of functions f satisfying either*

1.

$$L^0_{\Gamma,\Gamma'} f(v) = \begin{cases} \lambda f(v) & \text{for } v \in V' \cup V'' \\ -\lambda f(\tau(v)) & \text{for } v \in V'' \\ 0 & \text{for } v \in V_0 \end{cases} \tag{1.28}$$

2. or

$$f(\tau(v)) = f(v) \text{ for } v \in V'. \tag{1.29}$$

The latter eigenfunctions are those of the graph Γ^τ obtained as the quotient of Γ by τ, that is, the graph with vertex set $V_0 \cup V'$ and all edges induced by Γ.

Proof If $v_0 \in V_0$ and $f(v_0) = 0$ and if f is antisymmetric, then also $L^0 f(v_0) = 0$, since by Lemma 1.2.7 the contributions from its neighbors v' and $v'' = \tau(v')$ cancel in $L^0 f(v_0)$. The result then follows from Lemma 1.2.6, since a neighbor w of $v \in V'$ is contained either in V' or in V_0, in which case for an antisymmetric f, $f(w) = 0$, and therefore, we can restrict the computation in (1.28) to the induced Laplacian, that is, consider only the vertices in V'.

Examples:

1. We had already looked at the example of a duplicated vertex, that is, where v' and v'' are not connected, but have the same neighbors.

2. Let V' consist of a single vertex v' connected to $v'' = \tau(v')$ by an edge. Then 1.28 becomes
$$f(v') - \frac{1}{\deg_\Gamma v'} f(v'') = \lambda f(v'),$$
that is, since $f(v'') = -f(v')$,
$$\lambda = 1 + \frac{1}{\deg_\Gamma v'}.$$

3. We next duplicate an edge $e = (v_1', v_2')$. Then two eigenvalues and eigenfunctions of L^0 are obtained by solving
$$\begin{aligned} f(v_1') - \frac{1}{\deg_\Gamma v_1'} f(v_2') &= \lambda f(v_1') \\ f(v_2') - \frac{1}{\deg_\Gamma v_2'} f(v_1') &= \lambda f(v_2') \\ f(v) &= 0 \qquad \text{for all other } v. \end{aligned}$$
This yields [15]
$$\lambda = 1 \pm \frac{1}{\sqrt{\deg_\Gamma v_1' \deg_\Gamma v_2'}}.$$

4. It should now be clear how to analyze the duplication or twinning of other motives.

1.2.3 Rayleigh Quotients and the Courant-Fischer-Weyl Scheme

We now want to develop a more systematic approach for studying Laplacian spectra. We shall employ the fundamental

Theorem 1.2.1 *(Courant-Fischer-Weyl min-max principle) Let H be an N-dimensional vector space with a positive definite scalar product $(.,.)$. Let $\mathcal{H}_k$ be the family of all k-dimensional subspaces of H. Let $A : H \to H$ be a self adjoint linear operator. Then the eigenvalues $\lambda_1 \le \ldots \le \lambda_N$ of A can be obtained by*
$$\lambda_k = \min_{H_k \in \mathcal{H}_k} \max_{g(\neq 0) \in H_k} \frac{(Ag, g)}{(g, g)} = \max_{H_{N-k+1} \in \mathcal{H}_{N-k+1}} \min_{g(\neq 0) \in H_{N-k+1}} \frac{(Ag, g)}{(g, g)}. \tag{1.30}$$
The vectors g_k realizing such a min-max or max-min then are corresponding Eigenvectors, and the min-max spaces $\mathcal{H}_k$ are spanned by the Eigenvectors for the eigenvalues $\lambda_1, \ldots, \lambda_k$, and analogously, the max-min spaces $\mathcal{H}_{N-k+1}$ are spanned by the Eigenvectors for the eigenvalues $\lambda_k, \ldots, \lambda_N$.

Thus, we also have

$$\begin{aligned}\lambda_k &= \min_{g(\neq 0)\in H,(g,g_j)=0 \text{ for } j=1,\ldots,k-1} \frac{(Ag,g)}{(g,g)} \\ &= \max_{g(\neq 0)\in H,(g,g_\ell)=0 \text{ for } \ell=k+1,\ldots,N} \frac{(Ag,g)}{(g,g)}. \end{aligned} \tag{1.31}$$

In particular,

$$\lambda_1 = \min_{g(\neq 0)\in H} \frac{(Ag,g)}{(g,g)}, \qquad \lambda_N = \max_{g(\neq 0)\in H} \frac{(Ag,g)}{(g,g)}. \tag{1.32}$$

A *prf* can for instance be found in [18].

Definition 1.2.7 $\frac{(Ag,g)}{(g,g)}$ is called the *Rayleigh quotient* of g.

According to Theorem 1.2.1, the eigenvalues of L^0 are given by minimax values of

$$\frac{(L^0 g,g)}{(g,g)} = \frac{(\delta g,\delta g)}{(g,g)} = \frac{\sum_{e=(v,w)}(g(v)-g(w))(g(v)-g(w))}{\sum_v \deg v\ g(v)^2}, \tag{1.33}$$

that is, when we denote an eigenfunction for the eigenvalue λ_k by f_k,

$$\lambda_k = \min_{f:(f,f_j)=0 \text{ for } j=1,\ldots,k-1} \frac{\sum_{e=(v,w)}(f(v)-f(w))^2}{\sum_v \deg v\ f(v)^2}. \tag{1.34}$$

It is not difficult, for instance, to read off Lemma 1.2.4 from this formula.

Theorem 1.2.1 will play a fundamental role in our analysis of the spectra of simplicial complexes and hypergraphs below. In order to see its usefulness, let us look here at the following result.

Corollary 1.2.5 *A graph with N vertices is complete if and only if its spectrum consists of* 0 *as a simple eigenvalue and* $\frac{N}{N-1}$ *with multiplicity* $N-1$.
For a graph with N vertices that is not complete, we have

$$\lambda_2 \leq 1 \text{ and } \lambda_N > \frac{N}{N-1}. \tag{1.35}$$

Proof By Theorem 1.2.1,

$$\lambda_2 = \min_{g:\ \sum_v \deg v\ g(v)=0} \frac{(L^0 g,g)}{(g,g)}. \tag{1.36}$$

When Γ is not complete, we can find two vertices v_1, v_2 that are not connected by an edge, and take g with $g(v_1), g(v_2) \neq 0$, but

$$\deg v_1\ g(v_1) + \deg v_2\ g(v_2) = 0 \text{ and } g(v) = 0 \text{ for all other } v. \tag{1.37}$$

Inserting this into the Rayleigh quotient 1.36 makes that expression 1, and the minimum therefore is ≤ 1. We leave the second inequality of 1.35 as an exercise.

From Corollary 1.2.5 we see that the complete graphs K_N are completely determined by their spectrum. Thus, there are no other graphs that are isospectral with K_N. There do not even exist graphs whose spectrum is very close to that of K_N. In fact, not only do we have $\lambda_2 \leq 1 = \frac{N-1}{N-1}$ for non-complete graphs, but Das and Sun [19] proved that for all non-complete graphs we also have

$$\lambda_N \geq \frac{N+1}{N-1}, \tag{1.38}$$

with equality if and only if the complement graph (that is, the graph that connects precisely those vertices that are not neighbors in the graph under consideration) is a single edge or a complete bipartite graph with both parts of size $\frac{N-1}{2}$. More precise results in this direction can be found in [20].

Corollary 1.2.6 *Eigenfunctions for different eigenvalues of L^0 are orthogonal to each other w.r.t. $(.,.)_V$. In particular, all the eigenfunctions f for λ_k with $k \geq 2$ are orthogonal to the constants (the eigenfunctions for $\lambda_1 = 0$) and satisfy therefore*

$$\sum_v \deg v f(v) = 0. \tag{1.39}$$

1.2.4 Cheeger-Type Estimates

In this section, we only provide a survey of results, but no proofs. As we have seen, the spectrum of a graph can tell us its number of connected components (the multiplicity of the eigenvalue 0), whether it is bipartite (if and only if it has the eigenvalue 2) or complete (see Corollary 1.2.5), and it can indicate node duplications [15]. In fact, the spectrum reflects the general symmetries of a graph [17].

But the spectrum controls many further qualitative properties of a graph by inequalities relating eigenvalues to other quantities characterizing graphs. In particular, we have the so-called Cheeger-type estimates, the first of which were discovered by Dodziuk [21] and Alon and Milman [22].

In fact, these estimates concern a quantity that had already been introduced by Pólya and Szegö [23] into graph theory, but the estimate was inspired by Cheeger's estimate for an analogous constant in Riemannian geometry. The *(Polya)-Cheeger constant* is defined as

$$h := \min_S \frac{|E(S, \overline{S})|}{\min(\mathrm{vol}(S), \mathrm{vol}(\overline{S}))}, \tag{1.40}$$

where $|E(S, \overline{S})|$ denotes the number of edges between the two complementary subsets S, $\overline{S}$ of the vertex set V, and

$$\mathrm{vol}(S) := \sum_{v \in S} \deg v. \tag{1.41}$$

Thus, the aim is to cut the graph into two large vertex sets with few connections between them. That is, we want to *cluster* the vertex set. The result of [21, 22] then is

Theorem 1.2.2 *The eigenvalue λ_2 of a connected graph satisfies*

$$\frac{1}{2}h^2 \leq 1 - \sqrt{1-h^2} \leq \lambda_2 \leq 2\,h. \tag{1.42}$$

We do not provide the proof here which can be found, for instance, also in [10, 11]. A systematic general treatment will be given in [24].

The inequality 1.42 tells us that the first nonzero eigenvalue λ_2 of a connected graph measures how different that graph is from a disconnected one (where both that eigenvalue and h would be 0). We also recall that the largest eigenvalue λ_N is $= 2$ if and only if the graph is bipartite. We may therefore ask whether in general $2 - \lambda_N$ could also tell us how different a graph is from being bipartite. That works, indeed, and there is a *dual Cheeger constant* that bounds the largest eigenvalue [25, 26] (or a bipartiteness ratio [27]). It is defined as

$$\bar{h} := \max_{\text{partitions } V = V_1 \sqcup V_2 \sqcup V_3} \frac{|E(V_1, V_2)|}{\mathrm{vol}(V_1) + \mathrm{vol}(V_2)}. \tag{1.43}$$

This constant is $= 1$ if and only if the graph is bipartite (we take V_1, V_2 as the two classes of the bipartite graph and $V_3 = \emptyset$). It satisfies an analogue of (1.42),

$$2\bar{h} \leq \lambda_N \leq 1 + \sqrt{1 - (1-\bar{h})^2}.$$

The two constants h and $\bar{h}$ are related to each other [26].

We also have another characterization of h.

Theorem 1.2.3 *For every connected graph,*

$$h = \min_{f: V \to \mathbb{R} \text{ non constant}} \max_{t \in \mathbb{R}} \frac{\sum_{v \sim w} \big|f(v) - f(w)\big|}{\sum_{v \in V} \deg v \cdot \big|f(v) - t\big|}$$

and

$$\frac{1}{2}h \leq \min_{f: V \to \mathbb{R} \text{ s.t. } \sum_{v \in V} \deg v \cdot f(v) = 0} \frac{\sum_{v \sim w} \big|f(v) - f(w)\big|}{\sum_{v \in V} \deg v \cdot \big|f(v)\big|} \leq h.$$

For background on this result and the relation of the Cheeger constant with Rayleigh quotients, we refer to [28–35]. The nodal sets of an eigenfunction f_2 for λ_2, that is $V_\pm := \{v \in V : \pm f_2 > 0\}$ are connected and, according to the characterization in Theorem 1.2.3, there are only few edges between V_+ and V_-. Thus, they provide

natural clusters for the graph. A general survey of spectral clustering of graphs is found in [36].

There is, however, no analogue for $\bar{h}$ of Theorem 1.2.3, although optimal sets V_1, V_2 in 1.43 should have few internal connections and therefore yield some structure that approaches a bipartite one. Therefore, in [37], another quantity was introduced,

$$Q := \max_{e=(v,w)} \left(\frac{1}{\deg v} + \frac{1}{\deg w} \right).$$

We also need

$$\tau := \max_{e=(v,w):\deg w \geq \deg v} \left(\frac{(\deg w - \deg v + N) \cdot \deg v}{\deg v + \deg w} \right).$$

We then have the result of [37] which draws upon the duality between L^0 and L^1.

Theorem 1.2.4 *For every graph,*

$$Q = \max_{\gamma: E \to \mathbb{R}} \frac{\sum_{v \in V} \frac{1}{\deg v} \cdot \left| \sum_{e_{in}: v \text{ input}} \gamma(e_{in}) - \sum_{e_{out}: v \text{ output}} \gamma(e_{out}) \right|}{\sum_{e \in E} |\gamma(e)|}$$

and

$$Q \leq \lambda_N \leq Q \cdot \tau.$$

We also have

$$Q = \max_{\hat{\Gamma} \subset \Gamma \text{ bipartite}} \frac{\sum_{v \in V} \frac{\deg_{\hat{\Gamma}}(v)}{\deg_{\Gamma}(v)}}{|E(\hat{\Gamma})|}.$$

The last equation shows why Q is related to bipartiteness.

1.2.5 Generalizations and Extensions

1.2.5.1 Weighted Graphs

We can put weights on the vertices and/or the edges of a graph. When the vertex weights are denoted by w_i and the edge weights by w_e or w_{ij} for $e = (i, j)$, and we then require the symmetry

$$w_{ij} = w_{ji}, \tag{1.44}$$

$$(f,g)_V := \sum_i w_i f(i)g(i) \quad \text{for } f,g : V \to \mathbb{R} \tag{1.45}$$

$$(\gamma,\omega)_E := \sum_e w_e \gamma(e)\omega(e) \quad \text{for } \gamma,\omega : E \to \mathbb{R}. \tag{1.46}$$

In order to avoid problems, it might be good to require that all weights are nonnegative and every vertex supports at least one edge with a positive weight. (Formally, we could put $w_e = 0$ to indicate that the edge e is not present, that is, $e \notin E$. Thus, an ordinary graph has weight 1 for all $e \in E$ and weight 0 for all $e \notin E$.) In order to define an adjoint δ^* of the boundary operator δ as above, these two products have to be compatible in the sense that

$$w_i = \sum_j w_{ij}. \tag{1.47}$$

In fact, from this perspective, for a graph without edge weights, as considered above, $\deg i$ are the natural weights for the vertices i. In some situations, like neural networks, it might be natural to also admit negative edge weights (for inhibitory connections). In that case, we need to assume that $w_i \neq 0$ in 1.47 for all vertices, in order to be able to define our Laplacian.

That definition then is a simple and natural generalization of 1.2:

$$L^0 f(i) := f(i) - \frac{1}{w_i} \sum_j w_{ij} f(j). \tag{1.48}$$

And the preceding scheme can then be used to define L^1 analogously. We don't spell out the details here as such a scheme will come up again when we discuss simplicial complexes.

In summary, the spectral theory for weighted graphs (with the restrictions that we have imposed here) is not principally different from that of unweighted graphs.

1.2.5.2 Directed Graphs

When we allow for directed edges, going for instance from i to j, but not back from j to i, or more generally, give up the symmetry 1.44 in the weighted case, the theory becomes very different. The Laplacians, defined as in 1.2 or in 1.48, then are no longer self-adjoint, and consequently, their spectrum need no longer be real. Complex eigenvalues may occur. A corresponding theory has been developed in [38]. So far, however, this has not yet been pursued much, and systematic applications to real data have not yet been carried out, although many networks are naturally directed. Examples range from neural networks to weblinks or citations.

Importantly, chemical reaction networks should be modelled by *directed hypergraphs*. This calls for the development of the corresponding theory.

1.2.5.3 Signed Graphs

There is another version that, in contrast to directed graphs, supports a Laplacian with a real spectrum.

Definition 1.2.8 A *signed graph* Γ consists of a vertex set V and a set E of undirected edges with a sign function

$$s : E \to \{+1, -1\}. \tag{1.49}$$

A reference on the spectral theory of signed graphs is [39].
Such a signed graph may also carry a weight function

$$w : E \to \mathbb{R}_+, \tag{1.50}$$

but as we discussed the easy incorporation of weights into the theory already above, we neglect that possibility here.

The sign distinguishes between positive and negative relations, like friendship vs. hostility in a social network.

Definition 1.2.9 Let Γ be a signed graph. Its Laplacian is defined by

$$L_s^0 f(v) := f(v) - \frac{1}{\deg v} \sum_{v' \sim v} s(vv') f(v') = \frac{1}{\deg v} \sum_{v' \sim v} (f(v) - s(vv') f(v')) \tag{1.51}$$

with $\deg v$ defined as before as the number of neighbors of v, for functions $f : V \to \mathbb{R}$.

Lemma 1.2.9 *L_s^0 is selfadjoint w.r.t. the product 1.8 $(f, g) = \sum_v \deg v \ f(v)g(v)$, and*

$$(L_s^0 f, g) = \sum_{v \sim v'} (f(v) - s(vv') f(v'))(g(v) - s(vv') g(v')) = (f, L_s^0 g). \tag{1.52}$$

1.2.5.4 Self-Loops

We had excluded self-loops, that is, edges of the form $e = (i, i)$ for some vertex i. There is no deeper reason for such an exclusion, except perhaps historical contingencies in graph theory. The theory works as before when we allow for self-loops at some or all vertices.

To see what happens, let us consider the graph K_N^0 where each vertex i is connected with all vertices j, including itself. We then have

$$L_{K_N^0}^0 f(i) = f(i) - \frac{1}{N} \sum_j f(j) = \left(1 - \frac{1}{N}\right) f(i) - \frac{1}{N} \sum_{j \neq i} f(j), \tag{1.53}$$

which we may compare to the Laplacian on the complete graph K_N,

$$L^0_{K_N} f(i) = f(i) - \frac{1}{N-1} \sum_{j \neq i} f(j). \tag{1.54}$$

The spectrum of $L^0_{K^0_N}$ has the eigenvalue 0 and the eigenvalue 1, the latter with multiplicity $N-1$, whereas the corresponding eigenvalue of $L^0_{K_N}$ was $\frac{N}{N-1}$.

1.3 Simplicial Complexes

1.3.1 Homology of Simplicial Complexes

Definition 1.3.1 Let $\Sigma \subset \mathcal{P}(V)$ be a simplicial complex with vertex set $V = \{v_1, \ldots, v_N\}$. A collection of subsets of V, with $\emptyset \in \Sigma$, $S \in \Sigma$ is called a *q-simplex* if it contains precisely $q+1$ vertices.
When $S = \{v_{\sigma_0}, \ldots, v_{\sigma_q}\}$ is a q-simplex, then the ordered set $[v_{\sigma_0}, \ldots, v_{\sigma_q}]$ is called an *oriented q-simplex*. Changing the ordering by an odd permutation of the vertices induces the opposite orientation.
Let G be an abelian group. A *q-chain* is a formal linear combination

$$c_q = \sum_{i=1}^{m} g_i \sigma_q^i \tag{1.55}$$

for elements g_i of G and q-simplices σ_q^i.

Definition 1.3.2 The *boundary* of an oriented q-simplex $\sigma_q = [v_0, v_1, \ldots, v_q]$ is the $(q-1)$-chain

$$\partial \sigma_q := \sum_{i=0}^{q} (-1)^i (v_0, \ldots, \hat{v}_i, \ldots, v_q) \text{ for } q > 0, \tag{1.56}$$

and, of course, $\partial \sigma_0 = 0$ for a 0-chain. Here, $\hat{v}_i$ means that the vertex v_i is omitted. The boundary of the q-chain $c_q = \sum_{i=1}^{m} g_i \sigma_q^i$ then is, by linearity,

$$\partial c_q := \sum_{i=1}^{m} g_i \partial \sigma_q^i. \tag{1.57}$$

When we want to emphasize that ∂ operates on q-chains, we shall write ∂_q.

Definition 1.3.3 The q-chain c_q is called *closed* or, equivalently, a cycle, if

$$\partial_q c_q = 0. \tag{1.58}$$

The q-chain c_q is called a *boundary* if there exists some $(q+1)$-chain γ_{q+1} with

$$\partial_{q+1}\gamma_{q+1} = c_q. \tag{1.59}$$

Theorem 1.3.1

$$\partial_{q-1}\partial_q = 0 \textit{ for all } q. \tag{1.60}$$

We shall usually abbreviate this fundamental relation as

$$\partial^2 = 0. \tag{1.61}$$

Proof Because of 1.57, it suffices to show that

$$\partial\partial\sigma_q = 0 \tag{1.62}$$

for any oriented q-simplex. Since $C_s = 0$ for $s < 0$, we only need to consider the case $q \geq 2$. For $\sigma_q = [v_0, \ldots, v_q]$, we have

$$\begin{aligned}
\partial\partial\sigma_q &= \partial \sum_{i=0}^{q} (-1)^i (v_0, \ldots, \hat{v}_i, \ldots, v_q) \\
&= \sum_{i=0}^{q} (-1)^i \partial (v_0, \ldots, \hat{v}_i, \ldots, v_q) \\
&= \sum_{i=0}^{q} (-1)^i \left(\sum_{j=0}^{i-1} (-1)^j (v_0, \ldots, \hat{v}_j, \ldots, \hat{v}_i, \ldots, v_q) \right. \\
&\quad \left. + \sum_{j=i+1}^{q} (-1)^{j-1} (v_0, \ldots, \hat{v}_i, \ldots, \hat{v}_j, \ldots, v_q) \right) \\
&= \sum_{j<i} (-1)^{i+j} (v_0, \ldots, \hat{v}_j, \ldots, \hat{v}_i, \ldots, v_q) \\
&\quad + \sum_{j>i} (-1)^{i+j-1} (v_0, \ldots, \hat{v}_i, \ldots, \hat{v}_j, \ldots, v_q),
\end{aligned}$$

and exchanging i and j in the last sum gives the result.

Definition 1.3.4 The quotient group

$$H_q(\Sigma, G) := \ker \partial_q / \mathrm{im}\partial_{q+1} \tag{1.63}$$

is called the *qth homology group* (with coefficients in G) of the simplicial complex Σ.

When Σ is a simplicial complex and G is an abelian group G, we can define the group of q-cochains

$$C^q(\Sigma, G) := \mathrm{Hom}(C_q(\Sigma), G). \tag{1.64}$$

We then get the coboundary operators

$$\delta^q : C^q(\Sigma, G) \to C^{q+1}(\Sigma, G)$$
$$\phi \mapsto \phi \circ \partial_q.$$

Explicitly,

$$(\delta^q f)([v_0, \ldots, v_{q+1}]) = \sum_{j=0}^{q+1} (-1)^j f([v_0, \ldots, \hat{v}_j \ldots v_{q+1}]). \tag{1.65}$$

From Theorem 1.3.1, it follows that

$$\delta^q \circ \delta^{q-1} = 0. \tag{1.66}$$

We can therefore proceed to

Definition 1.3.5 The qth cohomology group of the simplicial complex Σ with coefficients in the abelian group G is

$$H^q(\Sigma, G) := \ker \delta^q / \mathrm{im} \delta^{q-1}. \tag{1.67}$$

1.3.2 Laplace Operators on Simplicial Complexes

In this section, we develop the theory of [13]. Σ is a simplicial complex. $[F]$ will denote an oriented simplex, that is, an ordered set $[v_0, \ldots, v_q]$ of vertices. Also, for the abelian group G of the previous section, we take the real field $\mathbb{R}$. The cochain groups $C^q(\Sigma, \mathbb{R})$ then are vector spaces over $\mathbb{R}$.

Definition 1.3.6 The dimension $b_q(\Sigma)$ of $H^q(\Sigma, \mathbb{R})$ is called the qth Betti number of Σ.

Also, concerning the orientations, we have for any $\phi \in C^q(\Sigma, \mathbb{R})$,

$$\phi(-\sigma_q) = -\phi(\sigma_q), \tag{1.68}$$

that is, changing the orientation yields a minus sign.

Definition 1.3.7 The adjoint $(\delta^q)^* : C^{q+1}(\Sigma, \mathbb{R}) \to C^q(\Sigma, \mathbb{R})$ of the coboundary operator δ^q is defined by

$$(\delta^q f, g)_{C^{q+1}} = (f, (\delta^q)^* g)_{C^q},$$

for $f \in C^q(\Sigma, \mathbb{R})$ and $g \in C^{q+1}(\Sigma, \mathbb{R})$, where $(\ ,\ )_{C^q}$ denotes the inner product on the $C^q(\Sigma, \mathbb{R}) = C^q$ for short.

We then have the arrows

$$C^{q-1}(\Sigma,\mathbb{R}) \underset{\delta^{q-1*}}{\overset{\delta^{q-1}}{\rightleftarrows}} C^q(\Sigma,\mathbb{R}) \underset{\delta^{q*}}{\overset{\delta^q}{\rightleftarrows}} C^{q+1}(\Sigma,\mathbb{R}), \tag{1.69}$$

enabling us to define the following three operators on $C^q(\Sigma,\mathbb{R})$:

Definition 1.3.8 (i) The *q-dimensional combinatorial up Laplace operator* or simply q-up Laplace operator of the simplicial complex Σ is

$$L^q{}_{up} := (\delta^q)^*\delta^q,$$

(ii) the *q-dimensional combinatorial down Laplace operator* or q-down Laplace operator

$$L^q{}_{down} := \delta^{q-1}(\delta^{q-1})^*,$$

(iii) the *q-dimensional combinatorial Laplace operator* or q-Laplace operator

$$L^q := L^q{}_{up} + L^q{}_{down} = (\delta^q)^*\delta^q + \delta^{q-1}(\delta^{q-1})^*.$$

For $q = 0$, that is, when we look at the operators on the vertices of a simplicial complex, we have $L^0{}_{down} = 0$, and hence

$$L^0 = L^0{}_{up}. \tag{1.70}$$

Similarly, for $q = \dim\Sigma$, the up-Laplacian vanishes, and

$$L^{\dim\Sigma} = L^{\dim\Sigma}{}_{down}. \tag{1.71}$$

The operators L^q_{up}, L^q_{down} and L^q are obviously self-adjoint. Also

Lemma 1.3.1 *The operators $L = L^q_{up}, L^q_{down}, L^q$ are nonnegative, that is, they satisfy*

$$(Lf, f) \geq 0 \text{ for all } f \in C^q. \tag{1.72}$$

Proof We have, generalizing (1.14),

$$(L^q_{up} f, f) = ((\delta^q)^*\delta^q f, f) = (\delta^q f, \delta^q f) \geq 0 \tag{1.73}$$

$$(L^q_{down} f, f) = (\delta^{q-1}(\delta^{q-1})^* f, f) = ((\delta^{q-1})^* f, (\delta^{q-1})^* f) \geq 0 \tag{1.74}$$

$$(L^q f, f) = (\delta^q f, \delta^q f) + ((\delta^{q-1})^* f, (\delta^{q-1})^* f) \geq 0. \tag{1.75}$$

In particular, from 1.73-1.75

Corollary 1.3.1

$$L^q_{up} f = 0 \ \textit{if and only if} \ \delta^q f = 0 \tag{1.76}$$

$$L^q_{down} f = 0 \ \textit{if and only if} \ (\delta^{q-1})^* f = 0 \tag{1.77}$$

$$L^q f = 0 \ \textit{if and only if} \ \delta^q f = 0 \ \textit{and} \ (\delta^{q-1})^* f = 0. \tag{1.78}$$

Since the operators L^q_{up}, L^q_{down} and L^q are self-adjoint, nonnegative operators on finite-dimensional Hilbert spaces, we have

Theorem 1.3.2 *The eigenvalues of the operators $L^q_{up}(\Sigma)$, $L^q_{down}(\Sigma)$ and $L^q(\Sigma)$ are real and nonnegative.*

Corollary 1.3.1 characterizes the eigenvalue 0. The other eigenvalues then are positive. Furthermore, Theorem 1.2.1 tells us that the eigenvalues admit a variational characterization

We can easily prove Eckmann's theorem [40], which is a discrete version of the Hodge theorem.

Theorem 1.3.3 *For a simplicial complex Σ,*

$$\ker L^q(\Sigma) \cong H^q(\Sigma, \mathbb{R}).$$

Thus, the multiplicity of the eigenvalue 0 *of the operator $L^q(\Sigma)$ equals the dimension of $H^q(\Sigma, \mathbb{R})$, that is, the Betti number b_q.*

Proof By 1.78,

$$\begin{aligned} \ker L^q(\Sigma) &= \ker \delta^q \cap \ker \delta^{q-1^*} \\ &= \ker \delta^q \cap (\mathrm{im}\delta^{q-1})^{\perp} \\ &\cong H^q(\Sigma, \mathbb{R}). \end{aligned}$$

Also, one readily checks that $\dim H^q(\Sigma, \mathbb{R}) = \dim H_q(\Sigma, \mathbb{R}) = b_q$.

While cohomology groups, like homology groups, were defined as quotients, that is, as equivalence classes of elements of C^q, Theorem 1.3.3 provides us with concrete representatives in C^q of those equivalence classes, the so-called harmonic cocycles.

We note that Eckmann's Theorem does not depend on the choice of scalar products on the spaces C^q (although the harmonic cocycles do). That theorem is concerned with the eigenvalue 0 of the Laplacian. We shall now investigate the nonzero part of the spectrum.

Since $\delta^q \delta^{q-1} = 0$ and $\delta^{q-1^*} \delta^{q*} = 0$ (recall 1.69),

$$\mathrm{im} L^q_{down}(\Sigma) \subset \ker L^q_{up}(\Sigma), \tag{1.79}$$

$$\mathrm{im} L^q_{up}(\Sigma) \subset \ker L^q_{down}(\Sigma). \tag{1.80}$$

Therefore, λ is a nonzero eigenvalue of $L_i(\Sigma)$ if and only if it is a nonzero eigenvalue of either $L^q_{up}(\Sigma)$ or $L^q_{down}(\Sigma)$. Therefore, the nonzero parts of the spectra satisfy

$$\mathrm{spec}_{\neq 0}(L^q(\Sigma)) = \mathrm{spec}_{\neq 0}(L_i^{up}(\Sigma)) \cup \mathrm{spec}_{\neq 0}(L_i^{down}(\Sigma)). \tag{1.81}$$

The multiplicity of the eigenvalue 0 may be different, however.

Since $\mathrm{spec}_{\neq 0}(AB) = \mathrm{spec}_{\neq 0}(BA)$, for linear operators A and B on Hilbert spaces, we get the following equality.

$$\mathrm{spec}_{\neq 0}(L^q_{up}(\Sigma)) = \mathrm{spec}_{\neq 0}(L^{q+1}_{down}(\Sigma)). \tag{1.82}$$

From (1.81) and (1.82) we conclude that each of the three families of multisets

$$\{\mathrm{spec}_{\neq 0}(L^q(\Sigma)) \mid 0 \leq q \leq m\}, \{\mathrm{spec}_{\neq 0}(L^q_{up}(\Sigma)) \mid 0 \leq q \leq m-1\}$$
$$\text{or } \{\mathrm{spec}_{\neq 0}(L^q_{down}(\Sigma)) \mid 1 \leq q \leq m\}$$

determines the other two. Therefore, it suffices to consider only one of them. In the sequel, we shall often omit the argument Σ from our Laplace operators.

We now look at scalar products on the spaces of cochains, as needed for the definition of the Laplace operators. Here, we only consider positive inner products, and when we shall speak about a scalar product in the sequel, we shall always assume that it be positive definite.

Each simplex generates a cochain, consisting of its real multiples. We assume that the cochains generated by different simplices are orthogonal to each other. This restricts the possible scalar products. A scalar product with this property can be obtained from a *weight function* w that associates to every simplex σ a positive real number. In fact, any such positive inner product on the space $C^q(\Sigma, \mathbb{R})$ can be written in terms of a weight function w as

$$(f, g)_{C^q} = \sum_{\sigma \in S_q} w(\sigma) f([\sigma]) g([\sigma]), \tag{1.83}$$

where S_q is the space of q-simplices by S_q. In the sequel, we shall write $\mathrm{sgn}([\sigma], [\sigma']) = \pm 1$ for two orientations of a simplex when those orientations coincide/differ. We also write $\partial\sigma$ for the cellular boundary of a simplex, that is, for the collection of its facets.

By simple linear algebra, the q-up Laplace operator is then given by

$$\begin{aligned}(L^q_{up} f)([\sigma]) = &\sum_{\substack{\rho \in S_{q+1}: \\ \sigma \in \partial\rho}} \frac{w(\rho)}{w(\sigma)} f([\sigma]) \\ &+ \sum_{\substack{\sigma' \in S_q : \sigma \neq \sigma', \\ \sigma, \sigma' \in \partial\rho}} \frac{w(\rho)}{w(\sigma)} \mathrm{sgn}([\sigma], \partial[\rho]) \mathrm{sgn}([\sigma'], \partial[\rho]) f([\sigma']),\end{aligned} \tag{1.84}$$

and the q-down Laplace operator is

$$(L^q_{down} f)([\sigma]) = \sum_{\tau \in \partial\sigma} \frac{w(\sigma)}{w(\tau)} f([\sigma]) + \sum_{\sigma' : \sigma \cap \sigma' = \tau} \frac{w(\sigma')}{w(\tau)} \mathrm{sgn}([\tau], \partial[\sigma]) \mathrm{sgn}([\tau], \partial[\sigma']) f([\sigma']). \tag{1.85}$$

For our purposes, however, we need some relation between the weights in different dimensions.

Definition 1.3.9 *The degree* of a q-simplex σ of Σ is

$$\deg \sigma := \sum_{\rho \in S_{q+1}(\Sigma) : \sigma \in \partial\rho} w(\rho). \tag{1.86}$$

Definition 1.3.10 If the weight function w on Σ satisfies

$$w(\sigma) = \deg \sigma, \tag{1.87}$$

for every $\sigma \in S_q(\Sigma)$, we call the Laplace operator defined on the cochain complex of Σ the *weighted normalized combinatorial Laplace operator*. If in addition the weights of the facets of Σ are equal to 1, then the Laplace operator is called the *normalized combinatorial Laplace operator*.

When 1.87 holds, 1.84 simplifies to become

$$(L^q_{up} f)([\sigma]) = f([\sigma]) + \frac{1}{\deg \sigma} \sum_{\substack{\sigma' \in S_q : \sigma \neq \sigma', \\ \sigma, \sigma' \in \partial\rho}} w(\rho) \mathrm{sgn}([\sigma], \partial[\rho]) \mathrm{sgn}([\sigma'], \partial[\rho]) f([\sigma']). \tag{1.88}$$

In the following, we discuss only the results related to the normalized combinatorial Laplace operator; the results for a more general case can be found in [12, 13].

1.3.3 Spectra of Simplicial Complexes

In the following, we restrict our analysis of $L^q_{up}(\Sigma)$ to pure, $(q+1)$-dimensional simplicial complexes. It follows from the definition that simplices of dimension lower than $q-1$ and higher than $q+1$ will have no influence on the non-zero spectrum of $L^q_{up}(\Sigma)$.

We could refine the class of relevant simplicial complexes even further, without any loss of generality , by limiting our attention to q-path connected components.

Definition 1.3.11 A simplicial complex Σ is q-path connected if and only if for any two q-faces E, F of Σ there exists a sequence of q-simplices $E = F_0, F_1, \ldots, F_n = F$, such that every two neighbouring simplices intersect in a $(q-1)$-face, i.e., they are $(q-1)$-down neighbours. The maximal q-path connected subcomplexes of Σ are called *q-path connected components*.

Any two q-path connected components share faces of dimension $q-2$, at most, therefore $\mathrm{spec}_{\neq 0}(L^q_{up}(\Sigma))$ is a multiset union of its q-path connected components. Hence, without loss of generality we shall assume that simplicial complexes in the subsequent analysis are pure and $q+1$- path connected.

We first generalize Lemma 1.2.4

Lemma 1.3.2 *All eigenvalues of L^q_{up} satisfy*

$$0 \leq \lambda \leq q+2. \tag{1.89}$$

Proof The spectrum of any bounded symmetric operator on a Hilbert space is real. Operator L^q_{up} is self-adjoint, thus symmetric, hence its eigenvalues are real. The non-negativity follows from the Courant-Fischer-Weyl theorem and non-negative weights.

Furthermore, for every $f \in C^q(\Sigma, \mathbb{R})$

$$\begin{aligned}
(L^q_{up} f, f) &= (\delta f, \delta f) \\
&= \Big(\sum_{\bar{F} \in S_{q+1}(\Sigma)} f(\partial[\bar{F}])\, \mathrm{e}_{[\bar{F}]}, \sum_{\bar{F} \in S_{q+1}(K)} f(\partial[\bar{F}])\, \mathrm{e}_{[\bar{F}]}\Big) \\
&= \sum_{\bar{F} \in S_{q+1}(K)} f(\partial[\bar{F}])^2 w(\bar{F}) \\
&\leq (q+2) \sum_{F \in S_q(K)} f([F])^2 \sum_{\bar{F} \in S_{q+1}(K): F \in \partial \bar{F}} w(\bar{F}) \qquad (1.90) \\
&= (q+2) \sum_{F \in S_q(K)} f([F])^2 \deg F \qquad (1.91) \\
&= (q+2)(f, f). \qquad (1.92)
\end{aligned}$$

Here $\mathrm{e}_{\bar{F}}$ denotes the elementary functional and (1.90) is obtained by using the Cauchy-Schwarz inequality. From (1.92) and the Courant-Fischer-Weyl min-max principle (Theorem 1.2.1) it follows that $\lambda \leq q+2$ for all $\lambda \in \mathrm{spec}(L^q_{up})$

The exact number of zero eigenvalues in the spectrum of L^q_{up} is given in the following theorem.

Theorem 1.3.4 *The multiplicity of the eigenvalue zero in* $\mathrm{spec}(L^q_{up})$ *is*

$$\dim C^q - \sum_{i=0}^{q} (-1)^{i+q} (\dim C^i - \dim H^i), \tag{1.93}$$

or equivalently

$$\dim C^q + \sum_{i=1}^{\dim K - q} (-1)^i (\dim C^{q+i} - \dim H^{q+i}). \tag{1.94}$$

Proof The following are short exact sequences that split

$$0 \to \ker \delta_q \to C^q \to \operatorname{im} \delta_q \to 0,$$

$$0 \to \operatorname{im} \delta_{q-1} \to \ker \delta_q \to H^q \to 0.$$

This is a direct consequence of the fact that $\operatorname{im} \delta_q$ and H^q are projective modules. Therefore,

$$\dim C^q = \dim \ker \delta_q + \dim \operatorname{im} \delta_q, \tag{1.95}$$

and

$$\dim \ker \delta_q = H^q + \dim \operatorname{im} \delta_{q-1}. \tag{1.96}$$

From (1.95) and (1.96)

$$\dim \operatorname{im} \delta_q = \sum_{i=0}^{q} (-1)^{q+i} (\dim C^i - \dim H^i).$$

The number of zeros in the spectrum of L^q_{up} is equal to the dimension of its kernel, thus

$$\begin{aligned} \dim \ker L^q_{up} &= \dim \ker \delta_q \\ &= \dim C^q - \sum_{i=0}^{q} (-1)^{q+i} (\dim C^i - \dim H^i). \end{aligned}$$

The expression (1.94) for the number of zeros in $\operatorname{spec}(L^q_{up})$ is easily obtained by using the Euler characteristic and the equality $\chi = \sum_{i=0}^{\dim \Sigma} (-1)^i \dim C^i = \sum_{i=0}^{\dim \Sigma} (-1)^i \dim H^i$.

Corollary 1.3.2 *Let Σ be a pure simplicial complex of dimension $q+1$, then the number of zero eigenvalues in the spectrum of $L^q_{up}(\Sigma)$ is $\dim C^q - \dim C^{q+1} + \dim H^{q+1}$.*

Using the above results, we can derive some lower bounds for the maximal eigenvalue of $L^q_{up}(\Sigma)$.

Theorem 1.3.5 *Let Σ be a pure simplicial complex of dimension $q+1$; let λ_m be the maximum eigenvalue in the spectrum of L^q_{up}, then*

$$\frac{\dim C^q}{(\dim C^{q+1} - \dim H^{q+1})} \leq \lambda_m. \tag{1.97}$$

Proof The sum of all eigenvalues is equal to the trace of the Laplace matrix, and in the case of the normalized Laplacian, the trace of the q-dimensional upper Laplacian equals the number of q-simplices.

The number of zero eigenvalues in the spectrum of L^q_{up} according to Corollary 1.3.2 is $\dim C^q - \dim C^{q+1} + \dim H^{q+1}$. Thus, the number of non-zero eigenvalues in L^q_{up} is exactly $\dim C^{q+1} - \dim H^{q+1}$. Hence,

$$\frac{\dim C^q}{(\dim C^{q+1} - \dim H^{q+1})} \leq \lambda_m,$$

which proves the theorem. □

Note that for $q = 0$ (i.e., for graphs) inequality (1.97) reduces to $\frac{V}{(E - \dim H^1)} = \frac{V}{(V-1)} \leq \lambda_m$, and it attains the lower bound when the underlying graph is complete. Interestingly, when Σ is a $q+1$-dimensional simplex, then from 1.97 it follows that $q + 2 \leq \lambda_m$; together with the inequality from Lemma 1.3.2 we conclude $\lambda_m = q + 2$.

The upper bound for the spectrum of the normalized graph Laplacian ($q = 0$) is 2 and is attained for bipartite graphs. There are many possible characterisations of bipartite graphs, the one that we will consider in this section is

Definition 1.3.12 A 1-dimensional simplicial complex (a graph) is bipartite if and only if it has no cycles of odd length.

We shall then see that we can generalize the characterization of bipartite graphs, as attaining the upper bound 2 for the largest eigenvalue, for the upper bound $q + 2$ of $\mathrm{spec}(L^q_{up}(\Sigma))$. We shall start with the definition of high dimensional cycles, which we shall refer to as circuits, to avoid the confusion with co-chain cycles.

Definition 1.3.13 A pure simplicial complex Σ of dimension q is called a *q-path of length m* if there is an ordering of its q-simplices $F_1 < F_2 < \ldots < F_m$, such that F_i and F_j are $(q-1)$-down neighbours if and only if $\mid j - l \mid = 1$; Σ is an *i-circuit* of length $(m-1)$ when $F_m = F_1$.

Theorem 1.3.6 *Let Σ be a pure q-connected simplicial complex, then the following statements are equivalent:*

(1) $q + 2$ is an eigenvalue of $L^{up}_q(\Sigma)$,
(2) There are no $(q+1)$-orientable circuits of odd length nor $(q+1)$-non orientable circuits of even length in Σ.

The following holds

$$(L^q_{up}(\Sigma)f, f) = \sum_{\bar{F}\in S_{q+1}(\Sigma)} f(\partial[\bar{F}])^2 w(\bar{F})$$
$$\leq (q+2) \sum_{F\in S_q(\Sigma)} f([F])^2 \deg(F). \tag{1.98}$$

The equality in (1.98) is reached if and only if there exists a function $f \in C^q(\Sigma, \mathbb{R})$, such that

$$\operatorname{sgn}([F_i], \partial[\bar{F}])\, f([F_i]) = \operatorname{sgn}([F_j], \partial[\bar{F}])\, f([F_j]),$$

for every $\bar{F}$ in S_{q+1} and every $F_i, F_j \in \partial\bar{F}$. Thus, $|f([F])|$ is a constant for every $F \in S_q(\Sigma)$; without loss of generality we shall assume further that $|f([F])| = 1$, then $f([F])$ is equal either to $\operatorname{sgn}([F], \partial[\bar{F}])$ or $-\operatorname{sgn}([F], \partial[\bar{F}])$, for every $F \in \partial\bar{F}$. Therefore, f could be viewed as a choice of orientation on the $(q+1)$-skeleton of Σ.

Theorem 1.3.7 *The existence of a function f satisfying the equality in (1.98) is equivalent to the existence of an orientation on the $(q+1)$-skeleton of Σ, for which any two $(q+1)$-simplices intersecting in a common q-face induce the same orientation on the intersecting simplex (This condition is opposite to the condition of coherently oriented simplices).*

Proof (**Proof of Theorem** 1.3.6) (1) $\Rightarrow$ (2) proceeds by contradiction. Assume that there exists a $(q+1)$-orientable circuit of odd length, whose q-simplices $F_1, \ldots, F_{2m+1}$ are ordered as suggested in Definition 1.3.13. Then it is possible to orient these simplices in such a way that every two neighbouring simplices induce different orientations on their intersecting face. Denote these oriented simplices by $[F_1], \ldots, [F_{2m+1}]$. In order to have the same orientation induced on the intersecting face, we reverse the orientation of every simplex $[F_k]$, for k even. Thus, $[F_i]$ and $-[F_{i+1}]$ induce the same orientation on $[F_i \cap F_{i+1}]$, for every $1 \leq i \leq 2m$. However, $[F_1]$ and $[F_{2m+1}]$ remain coherently oriented, which contradicts Theorem 1.3.7. The analysis for the case of $(n+1)$-non-orientable circuits is analogous.

(2) $\Rightarrow$ (1) proceeds by contradiction. We shall assume (2) and $\neg(1)$, i.e. $q+2 \notin \operatorname{spec}(L^q_{up}(\Sigma))$; by Theorem 1.3.7 the former is equivalent to non-existence of an orientation which induces incoherent orientations on neighbouring q-simplices. Namely, any attempt to assign incoherent orientations to neighbouring simplices would eventually result in two neighbouring simplices with coherent orientations.

More precisely, let F_{i_1} be an arbitrary $(q+1)$-face of Σ; and let $[F_{i_1}]$ be the *initial* positively oriented face. Let $[F_{i_1 i_2 \ldots i_m}]$ be a $(q+1)$-face of Σ which shares a q-face with $[F_{i_1 i_2 \ldots i_{m-1}}]$; assume both faces induce the same orientation on their intersecting face and are oriented incoherently. Then by $\neg(1)$ this construction will eventually lead us to a point where $F_{i_1 i_2 \ldots i_m} \equiv F_{i_1 i_2 \ldots i_k}$, but $[F_{i_1 i_2 \ldots i_m}] = -[F_{i_1 i_2 \ldots i_k}]$. Notice that by construction, $\{F_{i_1 i_2 \ldots i_m}, F_{i_1 i_2 \ldots i_{m+1}}, \ldots, F_{i_1 i_2 \ldots i_k}\}$ is a circuit. However, the only circuits which do not admit incoherent orientations are odd orientable or even non-orientable circuits, which contradicts (2).

1.3.4 Spectra of Some Special Classes of Simplicial Complexes

In this section we shall fully describe the spectrum of $L^q_{up}(\Sigma)$ for some basic classes of simplicial complexes:

- $(n-1)$-simplex
- n- path
- n-circuit
- n-star

Theorem 1.3.8 *Let Σ be a simplex on n vertices, i.e., an $(n-1)$-dimensional simplex. Then*

$$\operatorname{spec}(L^q_{up}(\Sigma)) = \left\{ 0^{(\binom{n-1}{q})}, \frac{n}{n-q-1}^{(\binom{n-1}{q+1})} \right\}.$$

Proof We shall proceed to construct an eigenfunction $f \in C^q(\Sigma, \mathbb{R})$, corresponding to the Eigenvalue $\frac{n}{n-q-1}$. In particular,

$$f = f_{[\bar{F}]}([F]) = \begin{cases} \operatorname{sgn}([F], \partial[\bar{F}]) & \text{if } F \text{ is facet of } (q+1)\text{-face } \bar{F} \\ 0 & \text{otherwise.} \end{cases}$$

There are exactly $\binom{n-1}{q+1}$ linearly independent functions of this form. In the following, we shall verify that the equality

$$(L^q_{up} f_{[\bar{F}]})[F] = \frac{n}{n-q-1} f([F])$$

holds for every q-dimensional face F of Σ. We shall distinguish three possibilities for F and $\bar{F}$:
(i) F is an arbitrary facet of $\bar{F}$. Then,

$$\begin{aligned}
(L^q_{up} f_{[\bar{F}]})([F]) &= \sum_{\substack{\bar{E} \in S_{q+1}: \\ F \in \partial \bar{E}}} \frac{w(\bar{E})}{w(F)} f_{[\bar{F}]}([F]) \\
&\quad + \sum_{\substack{F' \in S_q(L): \\ (\exists \bar{E} \in S_{q+1}(L)) F, F' \in \partial \bar{E}}} \frac{w(\bar{E})}{w(F)} \operatorname{sgn}([F], \partial[\bar{E}]) \operatorname{sgn}([F'], \partial[\bar{E}]) f_{[\bar{F}]}([F']) \\
&= \frac{1}{n-q-1} \sum_{\substack{\bar{E} \in S_{q+1}: \\ F \in \partial \bar{E}}} f_{[\bar{F}]}(F) \\
&\quad + \frac{1}{n-q-1} \sum_{\substack{F' \in S_q(L): \\ (\exists \bar{E} \in S_{q+1}(L)) F, F' \in \partial \bar{E}}} \operatorname{sgn}([F], \partial[\bar{E}]) \operatorname{sgn}([F'], \partial[\bar{E}]) f_{[\bar{F}]}([F']) \\
&= f_{[\bar{F}]}([F]) + \frac{q+1}{n-q-1} \operatorname{sgn}([F], \partial[\bar{F}]) \\
&= \frac{n}{n-q-1} f([F]).
\end{aligned}$$

(ii) F and $\bar{F}$ have q vertices in common, i.e., their intersection is a face of dimension $q-1$.
Then by definition $f_{\bar{F}}([F]) = 0$. Let $v_0, v_1, \ldots, v_{q+2} \in [n]$ be arbitrary vertices of Σ; then we shall assume without loss of generality that $\bar{F} = [v_0, \ldots, \hat{v_l}, \ldots, v_{q+2}]$ and $[F] = [v_0, \ldots, \hat{v_j}, \ldots, \hat{v_k}, \ldots, v_{q+2}]$ for $0 \leq j < k < l \leq q+2$. Therefore, there are exactly two q-faces, F_1 and F_2, in the boundary of $\bar{F}$, and two $(i+1)$-simplices, $\bar{F}_1$ and $\bar{F}_2$, of Σ, such that $F, F_1 \in \partial \bar{F}_1$ and $F, F_2 \in \partial \bar{F}_2$. In particular, $F_1 = [v_0, \ldots, \hat{v_k}, \ldots, \hat{v_l}, \ldots, v_{i+2}]$, $F_2 = [v_0, \ldots, \hat{v_j}, \ldots, \hat{v_l}, \ldots, v_{i+2}]$ and $\bar{F}_1 = [v_0, \ldots, \hat{v_k}, \ldots, v_{i+2}]$, $\bar{F}_2 = [v_0, \ldots, \hat{v_j}, \ldots, v_{i+2}]$. Now it is straightforward to calculate

$$\begin{aligned}
(L^q_{up} f_{[\bar{F}]})([F]) &= 0 + \operatorname{sgn}([F], \partial[\bar{F}_1])\operatorname{sgn}([F_1], \partial[\bar{F}_1])\, f_{[\bar{F}]})([F_1]) \\
&\quad + \operatorname{sgn}([F], \partial[\bar{F}_2])\operatorname{sgn}([F_2], \partial[\bar{F}_2])\, f_{[\bar{F}]})([F_2]) \\
&= \operatorname{sgn}([F], \partial[\bar{F}_1])\operatorname{sgn}([F_1], \partial[\bar{F}_1])\operatorname{sgn}([F_1], \partial[\bar{F}]) \\
&\quad + \operatorname{sgn}([F], \partial[\bar{F}_2])\operatorname{sgn}([F_2], \partial[\bar{F}_2])\operatorname{sgn}([F_2], \partial[\bar{F}]) \\
&= (-1)^j(-1)^{l-1}(-1)^k + (-1)^{k-1}(-1)^{l-1}(-1)^j \\
&= 0.
\end{aligned}$$

(iii) F and $\bar{F}$ have less than q vertices in common.
Obviously, there are no faces in the boundary of $\bar{F}$ which are $(q+1)$-up neighbours of F. This implies that $L^q_{up} f([F]) = 0$, which completes the proof.

Remark Let K be a $q+1$-combinatorial manifold, possibly with boundary. Then, by orienting it, that is by orienting its $(q+1)$-simplices coherently, we in fact choose a basis $B_{q+1}(\Sigma)$ of the vector space $C_{q+1}(\Sigma, \mathbb{R})$ consisting of elementary $(q+1)$-chains $[\bar{F}]$ that are oriented *coherently*.

For the subsequent calculations, the following elementary result will be useful.

Lemma 1.3.3 *When two matrices M and P commute, i.e., $MP = PM$, and when λ is a simple eigenvalue of P, then its corresponding eigenvector v is also an eigenvector of M.*

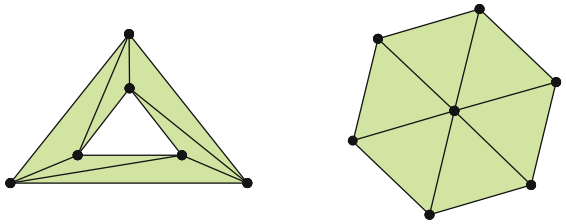
Proof Let $Pv = \lambda v$. Then $PMv = MPv = \lambda Mv$, and so, Mv is an eigenvector of P for λ. Since λ is simple, it must be a multiple of v.

Theorem 1.3.9 *Let Σ be an orientable n-circuit of length m. Then the eigenvalues of $L^n_{down}(\Sigma)$ are $n - \cos(2\pi i/m)$, $i = 0, 1, \ldots m-1$.*

Proof Let $F_1 < F_2 < \ldots < F_m$ be the ordering of n-simplices of Σ satisfying the conditions of Definition 1.3.13; and let $[F_1], [F_2], \ldots, [F_m]$ be a coherent orientation. Let $p : C^n(\Sigma, \mathbb{R}) \to C^n(\Sigma, \mathbb{R})$ be a map, such that $p([F_i]) = [F_{i+1}]$, for all $1 \leq i < m$, and $p([F_m]) = [F_1]$.

It is not difficult to check that

Fig. 1.1 Examples of 2-circuits



(a) 2-circuit of length 6 (b) 2-circuit of length 6

$$pL^n_{down} = L^n_{down}p \tag{1.99}$$

Let P be the matrix associated to the mapping p. P is a permutation matrix and its characteristic polynomial is $\lambda^m - 1 = 0$. Eigenvectors of P are $U_\theta = (1, \theta, \theta^2, \ldots \theta^{m-1})^T$, where θ is the m-th root of unity. Thus, the eigenfunctions of the map p are $u_\theta([F_i]) = \theta^{i-1}$.

With Lemma 1.3.3, we can now easily calculate the eigenvalues of L^n_{down}. Let $E_i := F_{i-1} \cap F_i$ for $2 \leq i \leq m-1$ and let $E_m := F_m \cap F_1$. We have

$$\begin{aligned}
L^n_{down}u_\theta([F_i]) = & \sum_{\substack{E \in S_{n-1}(L): \\ E \in \partial F_i}} \frac{w(F_i)}{w(E)}\theta^{i-1} \\
& + \frac{w(F_i)}{w(E_i)}\mathrm{sgn}([E_i], \partial[F_i])\mathrm{sgn}([E_i], \partial F_{i-1})\theta^{i-2} \\
& + \frac{w(F_i)}{w(E_{i+1})}\mathrm{sgn}([E_{i+1}], \partial[F_i])\mathrm{sgn}([E_{i+1}], \partial[F_{i+1}])\theta^{i} \\
= & (\frac{2}{2} + n - 1)\theta^{i-1} - \frac{1}{2}\theta^{i-2} - \frac{1}{2}\theta^{i} \\
= & \theta^{i-1}(i - \frac{\theta^{-1} + \theta}{2}) \\
= & \theta^{i-1}(n - \cos(\frac{2\pi i}{m})).
\end{aligned}$$

It is straightforward to check that a similar equality holds for $i = 1$ and $i = m$. Therefore the non zero spectrum of L^n_{down} is $\{n - \cos(2\pi i/m) \mid i = 0, 1, \ldots m-1\}$.

Remark The eigenvalues of an orientable n-circuit depend only on its length, thus there are different combinatorial structures which give the same eigenvalues of L^n_{down}. For example, 1, 1.5, 1.5, 2.5, 2.5, 3 are the eigenvalues of L^2_{down} of both simplicial complexes in Fig. 1.1.

An analysis similar to the one above can be carried out for a non-orientable n-circuit of length m. In that case we define p to be

$$p([F_k]) = \begin{cases} [F_{k+1}] \text{ , for } 1 \leq k < m \\ -[F_1] \text{ , for } k = m. \end{cases} \tag{1.100}$$

The remaining calculations are carried out as in Theorem 1.3.9. Thus,

Theorem 1.3.10 *Let Σ be a non-orientable n-circuit of length m. Then the eigenvalues of $L^n_{down}(\Sigma)$ are $n - \sin(2\pi i/m)$ for m even and $n + \cos(2\pi i/m)$ for m odd, where $i = 0, 1, \ldots m-1$.*

Proof The characteristic polynomial of the permutation matrix corresponding to the map p from (1.100) is

$$\begin{aligned} &\det \begin{vmatrix} -\lambda & 1 & 0 & \ldots & 0 \\ 0 & -\lambda & 1 & \ldots & 0 \\ & & & & \\ 0 & 0 & \ddots & \ddots & \vdots \\ -1 & 0 & \ldots & \ldots & -\lambda \end{vmatrix} = \\ &= (-1)^{m+1} \det \begin{vmatrix} 1 & 0 & \ldots & 0 \\ -\lambda & 1 & \ldots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & \ldots & -\lambda & 1 \end{vmatrix} + (-\lambda)^{m+m} \det \begin{vmatrix} -\lambda & 1 & 0 & \ldots & 0 \\ 0 & -\lambda & 1 & \ldots & 0 \\ & & & & \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \ldots & -\lambda \end{vmatrix} \\ &= (-1)^{m+1} + \lambda^m. \end{aligned}$$

Thus for m odd, the characteristic polynomal of the permutation matrix P is $\lambda^m + 1$. The Eigenvectors of P are $U_\theta = (-1, -\theta, (-\theta)^2, \ldots, (-\theta)^{m-1})^T$, where θ is the m-th root of unity. Using Lemma 1.3.3, we can now easily calculate the eigenvalues of L^n_{down}.
Let $E_k := F_{k-1} \cap F_k$ for $2 \leq k \leq m-1$ and let $E_m := F_m \cap F_1$. We have

$$\begin{aligned} L^n_{down} u_\theta([F_k]) = &\sum_{\substack{E \in S_{n-1}(L): \\ E \in \partial F_k}} \frac{w(F_k)}{w(E)} (-\theta)^{k-1} \\ &+ \frac{w(F_k)}{w(E_k)} \text{sgn}([E_k], \partial[F_k]) \text{sgn}([E_k], \partial F_{k-1}) (-\theta)^{k-2} \\ &+ \frac{w(F_k)}{w(E_{k+1})} \text{sgn}([E_{k+1}], \partial[F_k]) \text{sgn}([E_{k+1}], \partial[F_{k+1}]) (-\theta)^k \\ = &(\frac{2}{2} + n - 1)(-\theta)^{k-1} - \frac{1}{2}(-\theta)^{k-2} - \frac{1}{2}(-\theta)^k \\ = &(-\theta)^{k-1}(n - \frac{(-\theta)^{-1} - \theta}{2}) \\ = &(-\theta)^{k-1}(n - (-\cos(\frac{2\pi k}{m}) + i \sin(\frac{2\pi k}{m}) \\ &- \cos(\frac{2\pi k}{m}) - i \sin(\frac{2\pi k}{m}))/2) \\ = &(-\theta)^{k-1}(n + \cos(\frac{2\pi k}{m})) \end{aligned}$$

If $k = 1$, then

$$\begin{aligned} L^n_{down} u_\theta([F_1]) &= (\frac{2}{2} + n - 1)(-\theta)^m + \frac{1}{2}(-\theta)^{m-1} - \frac{1}{2}(-\theta)^1 \\ &= (-\theta)^m (n - \frac{(-\theta)^{-1} - \theta}{2}) \\ &= (-\theta)^m (n + \cos(2\pi)) \\ &= n + 1. \end{aligned}$$

A similar relation holds for $k = m$. For the case when m is even, the proof is analogous to the above proof for m odd. Therefore the non-zero spectrum of $L^n_{down}(\Sigma)$ is

$$\begin{cases} \{n - \sin(2\pi i/m) \mid i = 0, 1, \ldots m-1\} & \text{if } m \text{ is even} \\ \{n + \cos(2\pi i/m) \mid i = 0, 1, \ldots m-1\} & \text{otherwise.} \end{cases}$$

Corollary 1.3.3 *The eigenvalues of $L^n_{down}(\Sigma)$ of an n-path K of length m are $\lambda_k = n - \cos(\pi k/m)$, for $k = 0, \ldots, m-1$*

Proof Since there are no self-intersections of dimension $(n-1)$ in an n-path, every path is orientable. From Theorem 1.3.9, follows that in the spectrum of the n-th down Laplacian of an n-circuit of length $2m$, all eigenvalues appear twice, except $(n-1)$ and $(n+1)$. In particular, $\lambda_k = n - \cos(k\pi/m) = n - \cos((2m-k)\pi/m) = \lambda_{2m-k}$, for $k \neq 0$ and $k \neq m$. Let $\phi = \exp(ik\pi/m)$, then the eigenvector corresponding to λ_k, $0 \leq k \leq m$ is $u_k = (1, \exp(ik\pi/m, \ldots, \exp(i(2m-1)k\pi/m)^T$.

The function $v_k = u_k + u_{2m-k}$ is the eigenvector for the eigenvalue λ_k as well

$$v_k(m) = \mathrm{e}^{i\frac{\pi k}{m}} + \mathrm{e}^{i\frac{\pi(2m-k)}{m}} = \mathrm{e}^{i\frac{\pi k}{m}} + \mathrm{e}^{-i\frac{\pi k}{m}}.$$

It is now a straightforward calculation to see that the first m entries of v_k, for every $k = 0, 1, \ldots m-1$, constitute the Eigenvectors of Σ for the eigenvalue $n - \cos(\pi k/m)$.

This idea generalizes to paths with self-intersections of dimension $(n-1)$, but then it is necessary to distinguish among orientable and non-orientable paths. The eigenvalues of a star are described in the following theorem.

Theorem 1.3.11 *Let Σ be an n-star consisting of m, n-simplices. Then the non-zero eigenvalues of $L^n_{down}(\Sigma)$ are n with multiplicity $(m-1)$ and $(n+1)$ with multiplicity 1.*

Proof Let F_k, $k \in \{1, \ldots, m\}$, be an n-dimensional face of Σ and let $\bigcap_k F_k = E$. Let $p : B_n(\Sigma, \mathbb{R}) \to B_n(\Sigma, \mathbb{R})$ be a mapping, such that $p([F_k]) = [F_{k+1}]$. Since $F_k \cap F_i = E$, for any two n-faces of K, we can fix the orientations on the F_k such that they induce the same orientation on E. Now it is easy to check that

$$pL^n_{down} = L^n_{down} p.$$

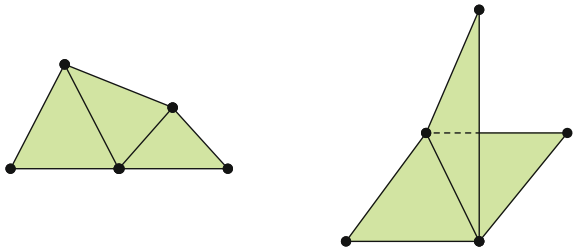


Fig. 1.2 Examples of a path and a star

(a) 2-path of length 3 (b) 2-star with 3 rays

Let θ denote an m-th root of unity different from 1 and u the eigenvector of p corresponding to it. Then

$$\begin{aligned} L^n_{down} u_\theta([F_k]) &= \sum_{E, E \in \partial F_k} \frac{w(F_k)}{w(E)} \theta^{k-1} + \sum_{F, F \neq F_k} \frac{w(F)}{w(E)} u_\theta([F]) \\ &= n\theta^{k-1} + \frac{1}{m}(1 + \theta + \ldots + \theta^{m-1}) \\ &= n\theta^{k-1}. \end{aligned}$$

Thus, u_θ is an eigenfunction of $L^n_{down}(\Sigma)$ corresponding to the eigenvalue n. The case when $\theta = 1$ results in the eigenvalue $n + 1$ (Fig. 1.2).

1.3.5 Cheeger-Type Inequalities

It is a natural question to ask for Cheeger-type inequalities for the higher Laplacians on simplicial complexes. There are some results for the highest order Laplacian in [41, 42], but not much seems to be known in general.

1.4 Hypergraphs

1.4.1 The Laplacians

We present the Laplace operators on oriented hypergraphs as natural generalizations of those on graphs, following [14, 43–46]. The basic idea underlying the definition of an oriented hypergraph is the following. Assigning an orientation to the edge of a graph means going from one of its vertices, considered as its tail or input, to the other, its head or output. Reversing the orientation means changing the roles of the

two vertices and going in the opposite direction. Thus, for an oriented hyperedge, we distinguish now two sets of vertices and move from the tail (input) set to the head (output) set. Here, more generally than for graphs, either of these sets could be empty.[1] Again, we can change the orientation by reversing roles and going in the opposite direction. And since Laplacians should be related to network flows, we treat all the members of the tail set as being parallel to each other, and the same for the head set.

In this section, we consider an oriented hypergraph Γ with vertex set V and hyperedge set H. For a vertex v, we let

$$\deg v := \left|\text{hyperedges containing } v\right| \tag{1.101}$$

and we assume that $\deg v > 0$ for all $v \in V$.

Definition 1.4.1 (Laplace operators) The Laplace operator for functions $f : V \to \mathbb{R}$ on the vertex set V of an oriented hypergraph is

$$\begin{aligned} L^0 f(v) := & \frac{\sum_{h_{\text{in}}: v \text{ input}} \left(\sum_{v' \text{ input of } h_{\text{in}}} f(v') - \sum_{w' \text{ output of } h_{\text{in}}} f(w') \right)}{\deg v} \\ & - \frac{\sum_{h_{\text{out}}: v \text{ output}} \left(\sum_{\hat{v} \text{input of } h_{\text{out}}} f(\hat{v}) - \sum_{\hat{w} \text{ output of } h_{\text{out}}} f(\hat{w}) \right)}{\deg v}. \end{aligned} \tag{1.102}$$

The Laplacian for functions $\gamma : H \to \mathbb{R}$ on the hyperedge set H, with $\gamma(h^+) = -\gamma(h^-)$ under a change of orientation, is

$$\begin{aligned} L^1 \gamma(h) := & \sum_{v_i \text{ input of } h} \frac{\sum_{h_{\text{in}}: v_i \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}: v_i \text{ output}} \gamma(h_{\text{out}})}{\deg v_i} \\ & - \sum_{v^j \text{ output of } h} \frac{\sum_{h'_{\text{in}}: v^j \text{ input}} \gamma(h'_{\text{in}}) - \sum_{h'_{\text{out}}: v^j \text{ output}} \gamma(h'_{\text{out}})}{\deg v^j}. \end{aligned} \tag{1.103}$$

When we have a graph, that is, when each hyperedge has a single input and a single output, these two operators reduce to those defined in 1.2 and 1.5,

$$L^0 f(v) = f(v) - \frac{1}{\deg v} \sum_{w \sim v} f(w)$$

$$L^1 \gamma(e) = \frac{1}{\deg v_0} \cdot \sum_{v_0 \in e' = [v_0, w]} \gamma(e') - \frac{1}{\deg v_1} \cdot \sum_{v_1 \in e'' = [v_1, w]} \gamma(e'') \quad \text{for } e = [v_0, v_1],$$

because the neighbors of v are outputs of the edges for which v is an input and conversely.

And we can generalize all the constructions underlying those operators to the case of oriented hypergraphs. This will then provide us again with powerful tools to analyze the spectra.

[1] But our convention here, which is different from that employed for simplicial complexes is that they should not both be empty.

We start with the *scalar products*.

Definition 1.4.2 For $f, g : V \to \mathbb{R}$, let

$$(f, g)_V := \sum_{v \in V} \deg v \cdot f(v) \cdot g(v). \tag{1.104}$$

For $\omega, \gamma : H \to \mathbb{R}$, let

$$(\omega, \gamma)_H := \sum_{h \in H} \omega(h) \cdot \gamma(h). \tag{1.105}$$

The *boundary* operator is next. It maps functions on vertices to functions on hyperedges that change their sign upon a change of orientation, as always.

Definition 1.4.3 For $f : V \to \mathbb{R}$ and $h \in H$, let

$$\delta f(h) := \sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j). \tag{1.106}$$

We then have

Lemma 1.4.1 *The adjoint of the operator δ w.r.t. the scalar products 1.104, 1.105 is*

$$\delta^*(\gamma)(v) = \frac{\sum_{h_{in}: v \text{ input}} \gamma(h_{in}) - \sum_{h_{out}: v \text{ output}} \gamma(h_{out})}{\deg v}. \tag{1.107}$$

And we then have the analogue of Lemma 1.2.2

Lemma 1.4.2

$$L^0 = \delta^* \delta \tag{1.108}$$

$$L^1 = \delta \delta^* \tag{1.109}$$

and the analogue of Corollary 1.2.1

Corollary 1.4.1 *We have*

$$(f, L^0 f)_V = (\delta f, \delta f)_H = (L^0 f, f)_V \tag{1.110}$$

and

$$(\gamma, L^1 \gamma)_H = (\delta^* \gamma, \delta^* \gamma)_V = (L^1 \gamma, \gamma)_H \tag{1.111}$$

for all f, γ.
In particular, the operators L^0 and L^1 are self-adjoint and nonnegative, and all their eigenvalues are real and nonnegative.
L^0 and L^1 have the same spectrum, except possibly for the multiplicity of the eigenvalue 0.

While the details require a more complicated track keeping of inputs and outputs, the basic ideas of the *proofs* are the same as before, and so, we do not spell out the details here. After all, the point is that we have identified *natural* constructions.

We now consider the matrix formulations of L^0 and L^1. We denote by $v_1, \dots, v_N$ the vertices of Γ and by $h_1, \dots, h_M$ its hyperedges. Given a hyperedge h, we say that two vertices v_i and v_j are *co-oriented in h* if they either are both inputs, or both outputs, for h. Conversely, we say that v_i and v_j are *anti-oriented in h* if they both belong to h but have opposite orientations.

Definition 1.4.4 The *degree matrix* of Γ is the $N \times N$ diagonal matrix

$$D := \operatorname{diag}(\deg v_1, \dots, \deg v_N). \tag{1.112}$$

The *incidence matrix* of Γ is the $N \times M$ matrix $\mathcal{I} := (\mathcal{I}_{ij})_{ij}$, where

$$\mathcal{I}_{ij} := \begin{cases} 1 & \text{if } v_i \text{ is an input of } h_j \\ -1 & \text{if } v_i \text{ is an output of } h_j \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, each row $\mathcal{I}_i$ of $\mathcal{I}$ represents a vertex v_i and each column $\mathcal{I}^j$ of $\mathcal{I}$ represents a hyperedge h_j.

The *adjacency matrix* of Γ is the $N \times N$ matrix $A := (A_{ij})_{ij}$, where $A_{ii} := 0$ for each $i = 1, \dots, N$ and, for $i \neq j$,

$$\begin{aligned} A_{ij} := &\left|\{\text{hyperedges in which } v_i \text{ and } v_j \text{ are anti-oriented}\}\right| \\ &- \left|\{\text{hyperedges in which } v_i \text{ and } v_j \text{ are co-oriented}\}\right|. \end{aligned}$$

The Laplacians can then be written, in matrix form, as

$$L^0 = \mathrm{Id} - D^{-1}A = D^{-1}\mathcal{I}\mathcal{I}^\top$$

and

$$L^1 = \mathcal{I}^\top D^{-1}\mathcal{I}.$$

In particular, this allows us to simplify (1.102) and write

$$L^0 f(v_i) = f(v_i) - \frac{1}{\deg v_i} \sum_{j \neq i} A_{ij} f(v_j),$$

for a given function $f : V \to \mathbb{R}$ and a vertex v_i.

Oriented hypergraphs were introduced by Shi in 1992 [47] as a generalization of signed graphs, that can be seen as oriented hypergraphs for which all hyperedges have size 2. Their corresponding spectral theory has been developed later. The adjacency matrix and the *algebraic Laplacian* $\mathcal{L} := \mathcal{I}\mathcal{I}^\top$ of an oriented hypergraph have been

first introduced in [48], whereas the Laplacians L^0 and L^1 have been introduced in [14].

As for graphs, when the hypergraph is *regular*, that is $\deg v \equiv const$, then the spectra of L^0, $\mathcal{L}$ and A differ only by an additive or multiplicative constant. For general hypergraphs, their spectral theory is different. We refer the reader to [48–60] for a vast literature on the adjacency and algebraic Laplacian matrices of oriented hypergraphs. We refer to [14, 17, 34, 43–46, 55, 56, 61, 62] for literature on the hypergraph Laplacians L^0 and L^1, on which we shall focus. We refer to [63, 64] for applications of the latter theory to dynamical systems on hypergraphs.

1.4.2 The Spectrum

Spectra of oriented hypergraphs can exhibit features not found for graphs. In particular, most of the results derived in Sect. 1.2.2 do not generalize to hypergraphs. Let the hypergraph have N vertices. Then L^0 has N real and nonnegative eigenvalues, possibly with multiplicities,

$$0 \leq \lambda_1 \leq \cdots \leq \lambda_N \leq N.$$

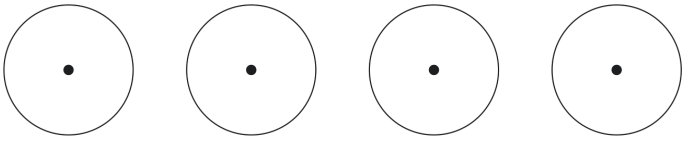
Likewise when we have M hyperedges, L^1 has M eigenvalues, the positive ones agreeing with those of L^0.

These eigenvalues always sum to N, since $L^0 = \mathrm{Id} - D^{-1}A$, therefore the trace of L^0 is N and this implies that also

$$\sum_{i=1}^{N} \lambda_i = N.$$

Examples:

1. Consider a hypergraph whose vertices $v_1, \ldots, v_N$ are only contained in *self-loops*, i.e. hyperedges of cardinality 1.

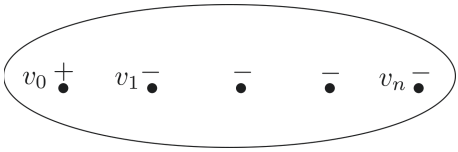


Then, for any function $f : V \to \mathbb{R}$, clearly

$$L^0 f(v_i) = f(v_i).$$

Hence, 1 is the only eigenvalue of L^0. In particular, 0 is not an eigenvalue in this case. This implies that Corollary 1.2.3 does no longer hold for the general case of hypergraphs.

2. Consider a hypergraph with vertices $v_0, v_1, \ldots, v_n$ and a single hyperedge h with v_0 as input and $v_1, \ldots, v_n$ as outputs.



Thus $\deg v = 1$ for all vertices, and

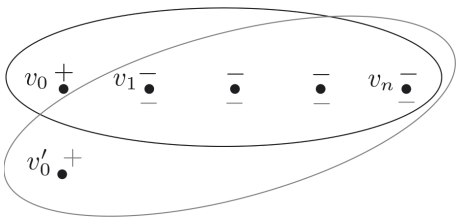
$$L^0 f(v_0) = f(v_0) - \sum_{i=1}^{n} f(v_i)$$

$$L^0 f(v_i) = \sum_{i=1}^{n} f(v_i) - f(v_0) \quad \text{for } i = 1, \ldots, n.$$

For $n = 1$, this is simply the line graph $K_{1,1}$. We let $i = 1, \ldots, n$. Eigenfunctions then are f_0 with $f_0(v_0) = n$, $f_0(v_i) = 1$ with the eigenvalue 0, $f_{i_0}(v_{i_0}) = 1$, $f_{i_0}(v_{j_0}) = -1$ for some pair $1 \le i_0 < j_0 \le n$ and $f_0(v_k) = 0$ for all other k, again with eigenvalue 0, and finally $f_{n+1}(v_0) = 1$, $f_{n+1}(v_i) = -1$ with the eigenvalue $n + 1$.

Thus, for $n > 1$, we have several eigenfunctions for the eigenvalue 0, but none of them is constant. This implies that the operator L^0 no longer obeys the maximum principle of Lemma 1.2.3.

3. We next double the vertex v_0, that is, introduce another vertex v_0' and a hyperedge h' with v_0' as input and $v_1, \ldots, v_n$ as outputs.



Then $\deg v_0 = \deg v_0' = 1$, but $\deg v_i = 2$ (as before, $i = 1, \ldots, n$). The previous eigenfunctions extend if we put $f(v_0') = f(v_0)$. We get another eigenfunction f' for the eigenvalue 1 with $f'(v_0) = 1$, $f'(v_0') = -1$, $f'(v_i) = 0$.

In the previous example, the duplication of a vertex produced the eigenvalue 1. This is always the case, as shown in Lemma 1.4.3 below, that generalizes the results on duplication for graphs that we systematically discussed in Sect. 1.2.3.

Definition 1.4.5 Two vertices v_i and v_j are *duplicates* of each other if the corresponding rows/columns of the adjacency matrix are the same, that is,

$$A_{il} = A_{jl} \quad \text{for each } l = 1, \ldots, N.$$

In particular, $A_{ij} = A_{jj} = 0$.

Remark In the case of graphs, Definition 1.4.5 coincides with the classical definition of duplicate vertices that we presented in Sect. 1.2.2.

Lemma 1.4.3 *Let v_i and v_j are duplicates of each other. Let $f : V \to \mathbb{R}$ be such that $f(v_i) = -f(v_j) \neq 0$ and $f = 0$ otherwise. Then, $L^0 f = f$, that is, 1 is an eigenvalue for L^0 and f is a corresponding eigenfunction.*

Proof It is easy to see that, by definition of f,

- $L^0 f(v_i) = f(v_i)$,
- $L^0 f(v_j) = f(v_j)$, and
- For each $l \neq i, j$,

$$\begin{aligned} L^0 f(v_l) &= -\frac{1}{\deg v_l}\big(A_{li} f(v_i) + A_{lj} f(v_j)\big) \\ &= -\frac{1}{\deg v_l}\big(A_{li} f(v_i) - A_{li} f(v_i)\big) \\ &= 0 = f(v_l). \end{aligned}$$

□

Corollary 1.4.2 *If there are n duplicate vertices, 1 is an eigenvalue with multiplicity at least $n - 1$.*

Proof Assume that $v_1, \ldots, v_n$ are duplicate vertices. For each $i = 1, \ldots, n-1$, let $f_i : V \to \mathbb{R}$ such that $f_i(v_i) = 1$, $f_i(v_{i+1}) = -1$ and $f_i = 0$ otherwise. Then, by Lemma 1.4.3 the f_i's are eigenfunctions corresponding to the eigenvalue 1. Also, $\dim(\text{span}(f_1, \ldots, f_{n-1})) = n - 1$, therefore the multiplicity of 1 is at least $n - 1$.

Similarly to duplicate vertices, we define and discuss *twin vertices*.

Definition 1.4.6 Two vertices v_i and v_j are *twins* of each other if they belong exactly to the same hyperedges, with the same orientations. In particular, $A_{ij} = -\deg v_i = -\deg v_j$ and $A_{il} = A_{jl}$ for all $l \neq i, j$.

Note that two vertices v_i and v_j cannot be both duplicates and twins of each other. In fact, if they are duplicates then $A_{ij} = 0$ while, if they are twins, then $A_{ij} < 0$. Also, while duplicate vertices exist for graphs, twin vertices cannot exist for graphs, since in this case each edge has one input and one output.

We now generalize the notions of duplicate vertices and twin vertices by defining *duplicate families of twin vertices*.

Definition 1.4.7 A family $V_1 \sqcup \ldots \sqcup V_l \subset V$ of vertices is an *l-duplicate family of t-twin vertices* if

- For each $a \in \{1, \ldots, l\}$, $|V_a| = t$ and the t vertices in V_a are twins of each other;
- For each $a, b \in \{1, \ldots, l\}$ with $a \neq b$, for each $v_i \in V_a$ and for each $v_j \in V_b$, v_i and v_j are duplicates of each other.

Proposition 1.4.1 *If Γ contains an l-duplicate family of t twins, then*

- *t is an eigenvalue with multiplicity at least $l - 1$, and*
- *0 is an eigenvalue with multiplicity at least $l(t - 1)$.*

Proof In order to show that t is an eigenvalue with multiplicity at least $l - 1$, consider the following $l - 1$ functions. For $a = 2, \ldots, l$, let $f_a : V \to \mathbb{R}$ such that $f_a := 1$ on V_1, $f_a := -1$ on V_a and $f_a := 0$ otherwise. Then,

- For each $v_j \in V_1$,

$$L^0 f_a(v_j) = 1 - \frac{1}{\deg v_j} \sum_{v_k \in V_1 \setminus \{v_j\}} - \deg v_j = 1 + t - 1 = t \cdot f_a(v_j);$$

- For each $v_i \in V_a$,

$$L^0 f_a(v_i) = -1 - \frac{1}{\deg v_i} \sum_{v_k \in V_a \setminus \{v_i\}} \deg v_i = -1 - (t - 1) = t \cdot f_a(v_i);$$

- For each $v_k \in V \setminus V_1 \sqcup V_a$,

$$L^0 f_a(v_k) = -\frac{1}{\deg v_k} \left(\sum_{v_j \in V_1} A_{jk} - \sum_{v_i \in V_a} A_{ik} \right) = 0 = t \cdot f_a(v_k).$$

Therefore, f_a is an eigenfunction for t. Furthermore, the functions $f_2, \ldots, f_l$ are linearly independent, hence t is an eigenvalue with multiplicity at least $l - 1$.

Similarly, in order to prove that 0 is eigenvalue with multiplicity at least $l(t - 1)$, let $V_a = \{v_1^a, \ldots, v_t^a\}$ and consider the $l(t - 1)$ functions $g_b^a : V \to \mathbb{R}$ defined as follows, for $a = 1, \ldots, l$ and $b = 2, \ldots, t$. Let $g_b^a(v_1^a) := 1$, $g_b^a(v_b^a) := -1$ and $g_b^a := 0$ otherwise. Then, each g_b^a is an eigenfunction for 0. Since, furthermore, these are $l(t - 1)$ linearly independent functions, 0 has multiplicity at least $l(t - 1)$.

1.4.3 Rayleigh Quotients and the Courant-Fischer-Weyl Scheme

The constructions of Sect. 1.2.3 naturally extend to oriented hypergraphs. In particular, all the eigenvalues of L^0 and L^1 can be characterized in terms of minmax values

of Rayleigh quotients. For instance, the largest eigenvalue λ_N of L^0 (which is also the largest eigenvalue of L^1 can be characterized in two different ways [14]

$$\lambda_N = \max_f \frac{(\delta f, \delta f)_H}{(f, f)_V} = \max_f \frac{\sum_{h\in H}\left(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j)\right)^2}{\sum_{v\in V} \deg v f(v)^2}$$

and

$$\lambda_N = \max_\gamma \frac{(\delta^*\gamma, \delta^*\gamma)_V}{(\gamma, \gamma)_H} = \max_\gamma \frac{\sum_{v\in V} \frac{1}{\deg v} \cdot \left(\sum_{h_{\text{in}}: v \text{ input}} \gamma(h_{\text{in}}) - \sum_{h_{\text{out}}: v \text{ output}} \gamma(h_{\text{out}})\right)^2}{\sum_{h\in H} \gamma(h)^2}.$$

As a consequence of these characterizations of λ_N one can prove, for instance, Theorem 1.4.1 below [45], showing a generalization of the following inequalities that hold for graphs,

$$\frac{N}{N-1} \leq \lambda_N \leq 2,$$

that we have seen in Sect. 1.2, to the general case of hypergraphs. Before stating it, we need to define *bipartite hypergraphs*, which naturally generalize bipartite graphs, as well as a few other preliminary definitions. As we shall see, analogously to the graph case, λ_N gives a measure of bipartitness for all oriented hypergraphs.

Definition 1.4.8 A hypergraph Γ is *bipartite* (Fig. 1.3) if one can decompose the vertex set as a disjoint union $V = V_1 \sqcup V_2$ such that, for every hyperedge h of Γ, either h has all its inputs in V_1 and all its outputs in V_2, or vice versa.

Definition 1.4.9 The *cardinality* of a hyperedge h, denoted $|h|$, is the number of vertices in h. A hypergraph is said to be *k-uniform* if all its hyperedges have cardinality k.

Clearly, graphs and signed graphs are 2-uniform hypergraphs.

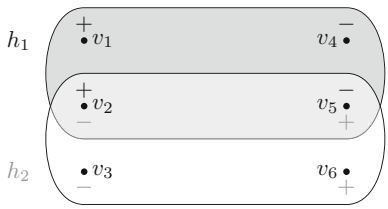


Fig. 1.3 A bipartite hypergraph with $V_1 = \{v_1, v_2, v_3\}$ and $V_2 = \{v_4, v_5, v_6\}$

Definition 1.4.10 A hypergraph $\hat{\Gamma} = (\hat{V}, \hat{H})$ is a *sub-hypergraph* of $\Gamma = (V, H)$, denoted $\hat{\Gamma} \subset \Gamma$, if $\hat{V} \subseteq V$ and

$$\hat{H} = \{h \cap \hat{V} : h \in H\}.$$

Given a sub-hypergraph $\hat{\Gamma} \subset \Gamma$, we let

$$\eta(\hat{\Gamma}) := \frac{\sum_{v \in \hat{V}} \frac{\deg_{\hat{\Gamma}}(v)^2}{\deg v}}{|\hat{H}|},$$

where $\deg_{\hat{\Gamma}}(v)$ denotes the degree of v in $\hat{\Gamma}$ and $|\hat{H}|$ is the number of hyperedges in $\hat{\Gamma}$.

Theorem 1.4.1 *For every connected, oriented hypergraph* Γ*,*

$$\lambda_N \leq \max_{h \in H} |h|, \tag{1.113}$$

with equality if and only if Γ *is bipartite and* $|h|$ *is constant for all* h*, and*

$$\lambda_N \geq \max_{\hat{\Gamma} \subset \Gamma \text{ bipartite}} \eta(\hat{\Gamma}). \tag{1.114}$$

Proof We first prove (1.113). Let $f : V \to \mathbb{R}$ be an eigenfunction for λ_N. Then,

$$\begin{aligned}\lambda_N &= \frac{\sum_{h \in H} \left(\sum_{v_i \text{ input of } h} f(v_i) - \sum_{v^j \text{ output of } h} f(v^j) \right)^2}{\sum_{v \in V} \deg v f(v)^2} \\ &\leq \frac{\sum_{h \in H} \left(\sum_{v \in h} |f(v)| \right)^2}{\sum_{i \in V} \deg v f(v)^2},\end{aligned}$$

with equality if and only if f has its nonzero values on a bipartite sub-hypergraph. Now, for each $h \in H$,

$$\begin{aligned}\left(\sum_{v \in h} |f(v)| \right)^2 &= \sum_{v \in h} f(v)^2 + \sum_{\{v,w\}:\, v \neq w \in h} 2 \cdot |f(v)| \cdot |f(w)| \\ &\leq \sum_{v \in h} f(v)^2 + \sum_{\{v,w\}:\, v \neq w \in h} \left(f(v)^2 + f(w)^2 \right) \\ &= \sum_{v \in h} f(v)^2 + \sum_{v \in h} (|h| - 1) f(v)^2 \\ &= |h| \cdot \sum_{v \in h} f(v)^2,\end{aligned}$$

with equality if and only if $|f|$ is constant on all $v \in h$. Therefore,

$$\begin{aligned}\frac{\sum_{h\in H}\left(\sum_{v\in h}|f(v)|\right)^2}{\sum_{i\in V}\deg v f(v)^2} &\le \frac{\sum_{h\in H}\sum_{v\in h}|h|\cdot\sum_{v\in h}f(v)^2}{\sum_{i\in V}\deg v f(v)^2}\\ &= \frac{\sum_{v\in V}\sum_{h\ni v}|h|\cdot f(v)^2}{\sum_{i\in V}\deg v f(v)^2}\\ &\le \left(\max_{h\in H}|h|\right)\cdot\frac{\sum_{v\in V}\deg v f(v)^2}{\sum_{i\in V}\deg v f(v)^2}\\ &= \max_{h\in H}|h|,\end{aligned}$$

where the first inequality is an equality if and only if $|f|$ is constant (since we assuming that Γ is connected), and the last inequality is an equality if and only if $|h|$ is constant for all h. Putting everything together, we have that

$$\lambda_N \le \max_{h\in H}|h|,$$

with equality if and only if $|h|$ is constant for all $|h|$ while $|f|$ is constant and it is defined on a bipartite sub-hypergraph (that is, $|f|$ is constant and Γ is bipartite). This proves the first claim.

It is left to prove (1.114). Given a bipartite sub-hypergraph $\hat{\Gamma}\subset\Gamma$, let $\gamma' : H\to\mathbb{R}$ be 1 on $\hat{H}$ and 0 otherwise. Then, up to changing (without loss of generality) the orientations of the hyperedges,

$$\begin{aligned}\lambda_N &= \max_{\gamma:H\to\mathbb{R}}\frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\left(\sum_{h_{\text{in}}:v\text{ input}}\gamma(h_{\text{in}})-\sum_{h_{\text{out}}:v\text{ output}}\gamma(h_{\text{out}})\right)^2}{\sum_{h\in H}\gamma(h)^2}\\ &\ge \frac{\sum_{v\in V}\frac{1}{\deg v}\cdot\left(\sum_{h_{\text{in}}:v\text{ input}}\gamma'(h_{\text{in}})-\sum_{h_{\text{out}}:v\text{ output}}\gamma'(h_{\text{out}})\right)^2}{\sum_{h\in H}\gamma'(h)^2}\\ &\ge \frac{\sum_{v\in\hat{V}}\frac{1}{\deg v}\cdot\left(\sum_{h_{\text{in}}:v\text{ input}}\gamma'(h_{\text{in}})-\sum_{h_{\text{out}}:v\text{ output}}\gamma'(h_{\text{out}})\right)^2}{\sum_{h\in H}\gamma'(h)^2}\\ &= \frac{\sum_{v\in\hat{V}}\frac{\deg_{\hat{\Gamma}}(v)^2}{\deg v}}{|\hat{H}|}.\end{aligned}$$

Since the above inequality is true for all $\hat{\Gamma}$, this proves (1.114).

Observe that, in the graph case, since $|h|=2$ for each (hyper)edge, (1.113) tells us that

$$\lambda_N \le 2,$$

with equality if and only if the graph is bipartite. (1.113) is therefore a generalization of the classical upper bound for λ_N of Corollary 1.2.4 to the case of hypergraphs.

Also, given a graph Γ, fix a vertex v and let $\hat{\Gamma}$ be the bipartite sub-graph of Γ given by the edges that have v as endpoint. Then, by (1.114),

$$\lambda_N \geq \eta(\hat{\Gamma}) = 1 + \sum_{w \sim v} \frac{1}{\deg w \cdot \deg v} \geq 1 + \sum_{w \sim v} \frac{1}{(N-1) \cdot \deg v} = 1 + \frac{1}{N-1} = \frac{N}{N-1}.$$

Hence, from (1.114), we can re-infer the fact that $\lambda_N \geq N/(N-1)$ for graphs.

In the following **examples**, we exhibit the sharpness of (1.114):

1. Let $\Gamma = K_N$ be the complete graph on N nodes. Fix a vertex v and let $\hat{\Gamma}$ be the bipartite sub-graph of Γ given by the edges that have v as endpoint. Then,

$$\eta(\hat{\Gamma}) = \frac{N}{N-1} = \lambda_N.$$

   Therefore, (1.114) is an equality for K_N.
2. Let $\Gamma = K_N \setminus \{(v_1, v_2)\}$ be the complete graph with an edge (v_1, v_2) removed. We know, from [19], that $\lambda_N = (N+1)/(N-1)$. Let $\hat{\Gamma}$ be the bipartite sub-graph of Γ given by the edges that have either v_1 or v_2 as endpoint. Then,

$$\eta(\hat{\Gamma}) = \frac{N+1}{N-1} = \lambda_N.$$

   Therefore, (1.114) is an equality for $\Gamma = K_N \setminus \{(v_1, v_2)\}$.
3. For a bipartite, k-uniform hypergraph Γ, by Theorem 1.4.1 $\lambda_N = k$. Also,

$$\eta(\Gamma) = \frac{\sum_{v \in V} \deg v}{M} = \frac{\sum_{h \in H} |h|}{M} = \frac{M \cdot k}{M} = k.$$

   Therefore, (1.114) is an equality also in this case.

Theorem 1.4.1 shows that, as in the graph case, the largest eigenvalue λ_N measures how different a hypergraph is from a bipartite one. Proposition 1.4.2 below shows that all bipartite hypergraphs that have the same vertices and hyperedges, independently of the input/output structure, are isospectral with each other. Before stating it, we define the *underlying hypergraph* and the *signless Laplacian* of Γ.

Definition 1.4.11 The *underlying hypergraph* of Γ is the oriented hypergraph obtained from Γ by letting each vertex be an input for all hyperedges in which it is contained. The *signless Laplacian* of Γ is the Laplacian L^0 of its underlying hypergraph.

Remark Assume that Γ is a graph and let Γ_+ be its underlying hypergraph. Then, the adjacency matrices of Γ and Γ_+ are such that $A(\Gamma_+) = -A(\Gamma)$, while the degree matrices of Γ and Γ_+ coincide. Therefore, the Laplacians of Γ and Γ_+ are

$$L^0(\Gamma) = \mathrm{Id} - D(\Gamma)^{-1}A(\Gamma) \quad \text{and} \quad L^0(\Gamma_+) = \mathrm{Id} + D(\Gamma)^{-1}A(\Gamma) = 2 \cdot \mathrm{Id} - L^0(\Gamma),$$

respectively. Hence, λ is an eigenvalue for $L^0(\Gamma)$ if and only if $2 - \lambda$ is an eigenvalue for $L^0(\Gamma_+)$.

Proposition 1.4.2 *If Γ is bipartite, it is isospectral to its underlying hypergraph, therefore, in particular, also to every other bipartite hypergraph that has the same underlying hypergraph as Γ.*

Proof Since Γ is bipartite, up to switching (without loss of generality) the orientations of some hyperedges we can assume that all the inputs are in V_1 and all the outputs are in V_2, with $V = V_1 \sqcup V_2$. Furthermore, by definition of L^1, which has the same nonzero spectrum as L^0, we can move a vertex from V_1 to V_2 or vice versa, by letting it be always an output or always an input, without affecting the spectrum. In particular, if we move all vertices to V_1, we obtain the underlying hypergraph of Γ.

We discuss two **examples,** namely the *hyperflowers*, that generalize star graphs, and the *c-complete hypergraphs*, that generalize complete graphs.

Example 4 Given $t \geq 1$, an oriented hypergraph $\Gamma = (V, H)$ on N nodes and M hyperedges is a *hyperflower with t twins* if (Fig. 1.4):

- The vertex set can be decomposed as $V = C \sqcup P$, where C is the *core* and P is given by tM *peripheral vertices* $v_{11}, \ldots, v_{t1}, \ldots, v_{1M}, \ldots, v_{tM}$;
- Forgetting about the input/output structure, the hyperedges are

$$h = C \sqcup \bigcup_{i=1}^{t} v_{ij} \ \text{ for } j = 1, \ldots, M.$$

If Γ is a bipartite hyperflower, by Proposition 1.4.2 we can assume, when computing the spectrum, that all its vertices are inputs for all the hyperedges in which they are contained. In this case, since all hyperedges have cardinality $N - tM + t$, by Theorem 1.4.1 $\lambda_N = N - tM + t$. Furthermore, by Proposition 1.4.1, t is an eigenvalue with multiplicity at least $M - 1$. We have therefore listed M eigenvalues whose sum is N. Since Γ has, in total, N eigenvalues whose sum is N, this implies that 0 has multiplicity $N - M$. Note that the star graph is a bipartite hyperflower with $t = N - M = 1$. Hence, the above computations show that the star graph has eigenvalues 0 with multiplicity 1, 1 with multiplicity $N - 2$ and 2 with multiplicity 1.

Example 5 Given $c \geq 2$, we say that Γ is a *c-complete hypergraph* if, forgetting about the input/output structure, its hyperedges are all possible $\binom{N}{c}$ hyperedges of

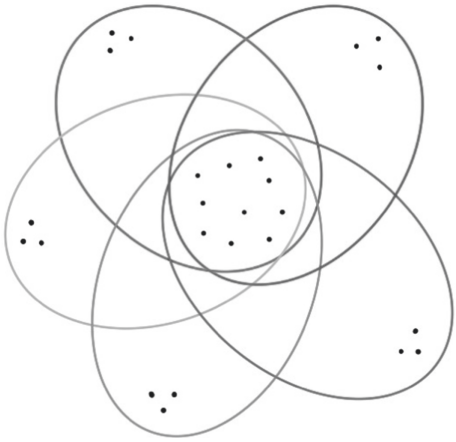


Fig. 1.4 A hyperflower with 3 twins

cardinality c. If Γ is a bipartite, c-complete hypergraph and we are interested in computing its spectrum, we can again assume, without loss of generality, that all vertices are always inputs for the hyperedges in which they are contained. In this case, by Theorem 1.4.1, $\lambda_N = c$. Moreover, observe that each vertex has degree $d := \binom{N-1}{c-1}$, while $a := A_{ij} = -\binom{N-2}{c-2}$ is constant for all $i \neq j$. Therefore, $\frac{a}{d} = -\frac{c-1}{N-1}$ and

$$L^0 f(v) = f(v) - \frac{a}{d}\left(\sum_{w \neq v} f(w)\right) = f(v) + \frac{c-1}{N-1}\left(\sum_{w \neq v} f(w)\right),$$

for all $v \in V$. Now, for each $i = 2, \ldots, N$, let $f_i(v_1) := 1$, $f_i(v_i) := -1$ and $f_i := 0$ otherwise. Then,

- $L^0 f_i(v_1) = 1 - \frac{c-1}{N-1} = \frac{N-c}{N-1} \cdot f_i(v_1)$,
- $L^0 f_i(v_i) = -1 + \frac{c-1}{N-1} = \frac{N-c}{N-1} \cdot f_i(v_i)$, and
- $L^0 f_i(v_j) = 0 = \frac{N-c}{N-1} \cdot f_i(v_j)$ for all $j \neq 1, i$.

Therefore, the f_i's are $N-1$ linearly independent eigenfunctions with eigenvalue $\frac{N-c}{N-1}$. This implies that the spectrum of Γ is given by c with multiplicity 1, and $\frac{N-c}{N-1}$ with multiplicity $N-1$.

In particular, if Γ is the complete graph K_N, we can apply the above computations to the underlying hypergraph of K_N (which is a signed graph) and say that this has eigenvalues 2 with multiplicity 1, and $\frac{N-2}{N-1}$ with multiplicity $N-1$. Since λ is an eigenvalue for the underlying hypergraph if and only if $2-\lambda$ is an eigenvalue for the original hypergraph, this implies that the eigenvalues of K_N are 0 with multiplicity 1 and $\frac{N}{N-1}$ with multiplicity $N-1$, as we already knew from Sect. 1.2.2.

1.4.4 Cheeger-Type Estimates

The Cheeger inequalities

$$\frac{1}{2}h^2 \leq \lambda_2 \leq 2h \tag{1.115}$$

for connected graphs (cf. (1.42)) have been generalized in [44] to the case of connected, k-uniform, bipartite hypergraphs, and it remains an open question whether they can be generalized for all connected oriented hypergraphs. The idea developed in [44] is the following. Given a connected graph Γ, its underlying hypergraph Γ_+ is a signed graph that has the same Cheeger constant as Γ. Moreover, λ is an eigenvalue for $L^0(\Gamma)$ if and only if $2-\lambda$ is an eigenvalue for $L^0(\Gamma_+)$. Therefore, the Cheeger inequalities in (1.115) can be equivalently reformulated in terms of the second largest eigenvalue of $L^0(\Gamma_+)$, as

$$\frac{1}{2}h(\Gamma_+)^2 \leq 2-\lambda_{N-1}(\Gamma_+) \leq 2h(\Gamma_+). \tag{1.116}$$

This equivalent formulation of the Cheeger inequalities in (1.115) can be used in order to prove a generalization for uniform, bipartite hypergraphs. We state the generalized inequalities, but we do not provide the proof here.

Let Γ be a connected, k-uniform, bipartite hypergraph. As for the case of graphs that we discussed in Sect. 1.2.4, given $S \subseteq V$, we let $\overline{S} := V \setminus S$ and

$$\mathrm{vol}(S) := \sum_{v \in S} \deg v.$$

Moreover, for $r \in \{1, \ldots, k\}$, we let

$$E_r(S) := \{e \in E : |e \cap S| = r\}.$$

Clearly, for each $r \in \{1, \ldots, k\}$, $E_r(S) = E_{k-r}(\overline{S})$. Also,

$$E_k(S) = \{e \in E : e \subseteq S\},$$

$$E_0(S) = \{e \in E : e \subseteq \overline{S}\}$$

and

$$\mathrm{vol}(S) = \sum_{r=1}^{k} r|E_r(S)|.$$

With the above notations, we can define the generalized Cheeger constant, as follows.

Definition 1.4.12 Given $\emptyset \neq S \subsetneq V$, let

$$h(S) := \frac{\sum_{r=1}^{k-1} |E_r(S)| r(k-r)}{\min\{\mathrm{vol}(S), \mathrm{vol}(\overline{S})\}}.$$

The *Cheeger constant* of Γ is

$$h := \min_{\emptyset \neq S \subsetneq V} h(S).$$

Observe that the quantity

$$\sum_{r=1}^{k-1} |E_r(S)| r(k-r)$$

appearing in the numerator of $h(S)$ counts the number of pairwise connections between S and $\overline{S}$. Furthermore, if Γ is a graph, then $k = 2$, $E_1(S)$ is the set of edges between S and $\overline{S}$, and the Cheeger constant defined above coincides with the one introduced by Pólya and Szegö, see (1.40).

The following theorem, proved in [44], generalizes (1.116) which is, on its turn, equivalent to the classical Cheeger inequalities in (1.115).

Theorem 1.4.2 *Let Γ be a connected, k-uniform, bipartite hypergraph. Then,*

$$\frac{1}{2(k-1)} h^2 \leq k - \lambda_{N-1} \leq 2(k-1)h.$$

Similarly, the Cheeger-like constant Q defined, for a graph, as

$$Q := \max_{e=(v,w)\in E} \left(\frac{1}{\deg v} + \frac{1}{\deg w} \right),$$

has been generalized in [45] for any oriented hypergraph, as

$$Q := \max_{h \in H} \left(\sum_{v \in h} \frac{1}{\deg v} \right).$$

The lower bound in Theorem 1.2.4, $Q \leq \lambda_N$, has been proved to hold for all oriented hypergraphs.

1.4.5 Generalizations

An oriented hypergraph can be seen as a classical hypergraph such that, in addition, for each vertex v and each hyperedge h, there exists a coefficient $C(v,h) \in \{-1, 0, +1\}$, where

$$C(v,h) = 0 \iff v \notin h, \tag{1.117}$$

while $C(v,h) = +1$ if and only if v is an input for h, and similarly $C(v,h) = -1$ if and only if v is an output for h. Various generalizations of this constructions have been studied in the context of spectral theory:

- *Chemical hypergraphs* are a generalization of oriented hypergraphs in which (1.117) does not necessarily hold, that is, one can have $C(v,h) = 0$ even if $v \in h$. The idea is that, in this case, v can be seen as *both* an input and an output for h, in which case v is said to be a *catalyst* for h. Chemical hypergraphs and their Laplacians have been introduced in [14], and the corresponding spectral results change based on the definition of generalized vertex degree that one considers. If

$$\deg v := |\{h \in H : C(v,h) \neq 0\}|,$$

then a chemical hypergraph Γ is isospectral to the oriented hypergraph obtained from Γ by removing catalysts from the hyperedges. If

$$\deg v := |\{h \in H : v \in h\}|,$$

by using the definition of Laplacian in (1.13), as in [14], the matrix formulation of L^0 becomes

$$L^0 = \hat{D} - D^{-1}A,$$

where $\hat{D}$ is the diagonal matrix with diagonal entries

$$\hat{D}_{ii} := \frac{\deg v_i - |\{h \in H : v_i \in h \text{ as catalyst}\}|}{\deg v_i}.$$

In this case, most of the results that we stated for oriented hypergraphs still hold. The main difference is that the eigenvalues are generally smaller, since now

$$\sum_{i=1}^{N} \lambda_i = N - \sum_{i=1}^{N} \frac{|\{h \in H : v_i \in h \text{ as catalyst}\}|}{\deg v_i}$$

can be smaller than N.

- *Complex unit hypergraphs* [56] are a generalization of oriented hypergraphs in which the coefficients $C(v,h)$ are from the complex unit circle, and (1.117) still holds. In this case, most of the spectral results for oriented hypergraphs can be generalized. The difference is that, instead of being symmetric operators, the Laplacians are now Hermitian operators and therefore the proofs require slightly different methods.
- *Hypergraphs with real coefficients* [17, 65] are a generalization of oriented hypergraphs in which (1.117) holds and the coefficients $C(v,h)$ are real numbers. In the case when, for all v and h, $C_{v,h} \geq 0$ and $\sum_{h \in H} C_{v,h} = 1$, we can see each coefficient $C_{v,h}$ as the *probability* of the vertex v to belong to the hyperedge h. In the case when the coefficients are integers, we can see each vertex as a chemical

element, each hyperedge as a chemical reaction and each coefficient $C_{v,h}$ as the chemical *stoichiometric coefficient* of the element v in the reaction h. Moreover, in the case when $C(v,h) =: w(h)$ only depends on h, for each hyperedge h and for each vertex v, we can see these hypergraphs as *weighted hypergraphs*.
If $\Gamma = (V, H, \{C(v,h) : v \in V \text{ and } h \in H\})$ is a hypergraph with real coefficients, its vertex degrees are defined by

$$\deg v := \sum_{h \in H} C(v,h)^2,$$

and it is easy to see that this generalizes (1.101). Moreover, the *degree matrix* of Γ is still defined by (1.112), the *incidence matrix* of Γ is defined by $\mathcal{I} := (\mathcal{I}_{ij})_{ij}$, where

$$\mathcal{I}_{ij} := C(v_i, h_j),$$

and the *adjacency matrix* of Γ is $A := (A_{ij})_{ij}$, where $A_{ii} := 0$ for all $i = 1, \ldots, N$ and, for $i \neq j$,

$$A_{ij} := -\sum_{h \in H} C(v_i,h) \cdot C(v_j,h).$$

The generalized *Laplacians* are then defined, in matrix form, as

$$L^0 := \mathrm{Id} - D^{-1}A = D^{-1}\mathcal{I}\mathcal{I}^\top \quad \text{and} \quad L^1 := \mathcal{I}^\top D^{-1}\mathcal{I},$$

and most of the properties that we discussed for oriented hypergraphs can be generalized to the more general setting of hypergraphs with real coefficients.

References

1. E. Llanos, W. Leal, A. Bernal, G. Restrepo, J. Jost, P. Stadler, A network model of the chemical space provides similarity structure to the system of chemical elements, in *Complex networks 2019: the 8th international conference on complex networks and their applications; December 10–12, 2019 Lisbon, Portugal; book of abstract* ed. by H. Cherifi. International Conference on Complex Networks & Their Applications (2019), pp. 308–310
2. E. Llanos, W. Leal, H.D. Luu, J. Jost, P. Stadler, G. Restrepo, Exploration of the chemical space and its three historical regimes. Proc. Nat. Acad. Sci. U.S.A. **116**(26), 12660–12665 (2019)
3. U. Schöning, Graph isomorphism is in the low hierarchy. J. Comput. Syst. Sci. **37**(3), 312–323 (1988)
4. A. Banerjee, J. Jost, Spectral plots and the representation and interpretation of biological data. Theor. Biosci. **126**(1), 15–21 (2007)
5. A. Banerjee, J. Jost, Spectral plot properties: towards a qualitative classification of networks. Netw. Heterog. Media **3**(2), 395–411 (2008)
6. A. Banerjee, J. Jost, Graph spectra as a systematic tool in computational biology. Discrete Appl. Math. **157**(10), 2425–2431 (2009)
7. M. Eidi, A. Farzam, W. Leal, A. Samal, J. Jost, Edge-based analysis of networks: curvatures of graphs and hypergraphs. Theor. Biosci. **139**(4), 337–348 (2020)

8. F. Bauer, B.B. Hua, J. Jost, S.P. Liu, G.F. Wang, The geometric meaning of curvature. Local and nonlocal aspects of Ricci curvature, in *Modern Approaches to Discrete Curvature*, number 2184 in LNM, ed. by L. Najman, P. Romon (Springer, 2017)
9. P. Joharinad, J. Jost, *Riemannian Geometry and Geometric Analysis*, 7th edn. (Springer, 2017)
10. F. Chung, *Spectral Graph Theory* (American Mathematical Society, 1997)
11. J. Jost, *Mathematical methods in biology and neurobiology* (Springer, London, 2014)
12. D. Horak, J. Jost, Interlacing inequalities for eigenvalues of discrete Laplace operators. Ann. Global Anal. Geom. **43**(2), 177–207 (2013)
13. D. Horak, J. Jost, Spectra of combinatorial Laplace operators on simplicial complexes. Adv. Math. **244**, 303–336 (2013)
14. J. Jost, R. Mulas, Hypergraph Laplace operators for chemical reaction networks. Adv. Math. **351**, 870–896 (2019)
15. A. Banerjee, J. Jost, On the spectrum of the normalized graph Laplacian. Linear Algebra Appl. **428**(11–12), 3015–22 (2008)
16. A. Banerjee, J. Jost, Laplacian spectrum and protein-protein interaction networks. arXiv:0705.3373
17. J. Jost, R. Mulas, Normalized Laplace operators for hypergraphs with real coefficients. J. Complex Netw. (2021). https://doi.org/10.1093/comnet/cnab009
18. J. Jost, *Postmodern analysis*, 3rd edn. (Springer, 2005)
19. K. Das, S. Sun, Extremal graph on normalized Laplacian spectral radius and energy. Electron. J. Linear Algebra **29**(1), 237–253 (2016)
20. J. Jost, R. Mulas, F. Münch, Spectral gap of the largest Eigenvalue of the normalized graph Laplacian. Commun. Math. Stat. (2021). https://doi.org/10.1007/s40304-020-00222-7
21. J. Dodziuk, Difference equations, isoperimetric inequality and transience of certain random walks. Trans. Amer. Math. Soc. **284**(2), 787–794 (1984)
22. N. Alon, V. Milman, λ_1, isoperimetric inequalities for graphs, and superconcentrators. J. Combin. Theory Ser. B **38**(1), 73–88 (1985)
23. G. Pólya, S. Szegö, Isoperimetric inequalities in mathematical physics. Ann. Math. Stud. **27** (1951)
24. J. Jost, Mathematical principles of topological and geometric data analysis. To appear
25. F. Bauer, B.B. Hua, J. Jost, The dual Cheeger constant and spectra of infinite graphs. Adv. Math. **251**, 147–194 (2014)
26. F. Bauer, J. Jost, Bipartite and neighborhood graphs and the spectrum of the normalized graph Laplacian. Comm. Anal. Geom. **21**, 787–845 (2013)
27. L. Trevisan, Max cut and the smallest eigenvalue. SIAM J. Comput. **41**(6), 1769–1786 (2012)
28. K.C. Chang, The spectrum of the 1-Laplace operator. Commun. Contemp. Math. **11**, 865–894 (2009)
29. K.C. Chang, Spectrum of the 1-Laplacian and Cheeger's Constant on Graphs. J. Graph Theor. **81**, 167–207 (2016)
30. K.C. Chang, S. Shao, D. Zhang, Spectrum of the signless 1-Laplacian and the dual Cheeger constant on graphs. arXiv:1607.00489
31. K.C. Chang, S. Shao, D. Zhang, The 1-Laplacian Cheeger cut: theory and algorithms. J. Comput. Math. 443–467 (2015)
32. M. Hein, T. Bühler, An inverse power method for nonlinear Eigenproblems with applications in 1–spectral clustering and sparse PCA, in *NIPS*, pp. 847–855 (2010)
33. M. Hein, S. Setzer, Beyond spectral clustering—tight relaxations of balanced graph cuts. Adv. Neural Inf. Process. Syst. **24**, 2366–2374 (2011)
34. J. Jost, R. Mulas, D. Zhang, p-Laplace operators for oriented hypergraphs. Vietnam J. Math. Spec Issue 60th Birthday Bernd Sturmfels, (2021). https://doi.org/10.1007/s10013-021-00525-4
35. J. Jost, D. Zhang, Interactions between discrete and continuous optimization and critical point theory via multi-way Lovász extensions. arXiv:2003.06021
36. U. von Luxburg, A tutorial on spectral clustering. Stat. Comput. **17**, 395–416 (2007)

37. J. Jost, R. Mulas, Cheeger-like inequalities for the largest eigenvalue of the graph Laplace operator. J. Graph Theory (2021). https://doi.org/10.1002/jgt.22664
38. F. Bauer, Normalized graph Laplacians for directed graphs. Linear Algebra Appl. **436**, 4193–4222 (2012)
39. F.M. Atay, S. Liu, Cheeger constants, structural balance, and spectral clustering analysis for signed graphs. Discrete Math. **343**(1), 111616 (2020)
40. B. Eckmann, Harmonische Funktionen und Randwertaufgaben in einem Komplex. Comment. Math. Helv. **17**(1), 240–255 (1944)
41. J. Steenbergen, C. Klivans, S. Mukherjee, A Cheeger-type inequality on simplicial complexes. Adv. Appl. Math. **56**, 56–77 (2014)
42. A. Gundert, M. Szedlak, Higher dimensional discrete Cheeger inequalities, in *Proceedings of the Annual Symposium on Computational Geometry* (2014)
43. E. Andreotti, R. Mulas, Signless Normalized Laplacian for Hypergraphs. arXiv:2005.14484
44. R. Mulas, A Cheeger cut for uniform hypergraphs. Graphs and Combinatorics **37**, 2265–2286 (2021). https://doi.org/10.1007/s00373-021-02348-z
45. R. Mulas, Sharp bounds for the largest Eigenvalue. Math. Notes **109**, 102–109 (2021)
46. R. Mulas, D. Zhang, Spectral theory of Laplace Operators on oriented hypergraphs. *Discrete Math.* (2021)
47. C.-J. Shi, A signed hypergraph model of the constrained via minimization problem. Microelectron. J. **23**(7), 533–542 (1992)
48. N. Reff, L.J. Rusnak, An oriented hypergraphic approach to algebraic graph theory. Linear Algebra Appl. **437**, 2262–2270 (2012)
49. G. Chen, V. Liu, E. Robinson, L.J. Rusnak, K. Wang, A characterization of oriented hypergraphic Laplacian and adjacency matrix coefficients. Linear Algebra Appl. **556**, 323–341 (2018)
50. V. Chen, A. Rao, L.J. Rusnak, A. Yang, A characterization of oriented hypergraphic balance via signed weak walks. Linear Algebra Appl. **485**, 442–453 (2015)
51. L. Duttweiler, N. Reff, Spectra of cycle and path families of oriented hypergraphs. Linear Algebra Appl. **578**, 251–271 (2019)
52. W. Grilliette, J. Reynes, L.J. Rusnak, Incidence hypergraphs: injectivity, uniformity, and matrix-tree theorems. Linear Algebra Appl. **634**, 77–105 (2022)
53. W. Grilliette, L.J. Rusnak, Incidence Hypergraphs: Box Products and the Laplacian. arXiv:2007.01842
54. O. Kitouni, N. Reff, Lower bounds for the Laplacian spectral radius of an oriented hypergraph. Australas. J. Combin. **74**(3), 408–422 (2019)
55. R. Mulas, Spectral classes of hypergraphs. Australas. J. Combin. **79**, 495–514 (2021)
56. R. Mulas, N. Reff, Spectra of Complex Unit Hypergraphs. arXiv:2011.10458
57. N. Reff, Spectral properties of oriented hypergraphs. Electron. J. Linear Algebra **27** (2014)
58. N. Reff, Intersection graphs of oriented hypergraphs and their matrices. Australas. J. Combin. **65**(1), 108–123 (2016)
59. E. Robinson, L.J. Rusnak, M. Schmidt, P. Shroff, Oriented hypergraphic matrix-tree type theorems and bidirected minors via Boolean order ideals. J. Algebraic Combin. (2017)
60. L.J. Rusnak, Oriented hypergraphs: Introduction and balance. Electron. J. Combin. **20** (2013)
61. A. Abiad, R. Mulas, D. Zhang, Coloring the normalized Laplacian for oriented hypergraphs. Linear Algebra Appl. **629**, 192–207 (2021). https://doi.org/10.1016/j.laa.2021.07.018
62. R. Mulas, R.J. Sánchez-García, B.D. MacArthur, Geometry and symmetry in biochemical reaction systems. Theory Biosci. **140**, 265–277 (2021). https://doi.org/10.1007/s12064-021-00353-7
63. T. Böhle, C. Kuehn, R. Mulas, J. Jost, Coupled hypergraph maps and chaotic cluster synchronization. EPL (2021). https://doi.org/10.1209/0295-5075/ac1a26
64. R. Mulas, C. Kuehn, J. Jost, Coupled dynamics on hypergraphs: master stability of steady states and synchronization. Phys. Rev. E **101**, 062313 (2020)
65. Y. Hirono, T. Okada, H. Miyazaki, Y. Hidaka, Structural reduction of chemical reaction networks based on topology. arXiv:2102.07687

Chapter 2
Random Simplicial Complexes: Models and Phenomena

Omer Bobrowski and Dmitri Krioukov

Abstract We review a collection of models of random simplicial complexes together with some of the most exciting phenomena related to them. We do not attempt to cover all existing models, but try to focus on those for which many important results have been recently established rigorously in mathematics, especially in the context of algebraic topology. In application to real-world systems, the reviewed models are typically used as null models, so that we take a statistical stance, emphasizing, where applicable, the entropic properties of the reviewed models. We also review a collection of phenomena and features observed in these models, and split the presented results into two classes: phase transitions and distributional limits. We conclude with an outline of interesting future research directions.

Keywords Random simplicial complexes · Geometric complexes · Maximum entropy principle · Canonical ensembles · Phase transitions · Percolation · Persistent homology · Betti numbers · Limit theorems

2.1 Introduction

Simplicial complexes serve as a powerful tool in algebraic topology, a field of mathematics fathered by Poincaré at the end of the 19th century. This tool was built—or discovered, depending on one's philosophical view—by topologists in order to study topology. Much more recently, over the past decade or so, this tool was also discovered by network and data scientists who study complex real-world systems in which interactions are not necessarily diadic. The complexity of interactions in these sys-

O. Bobrowski
Viterbi Faculty of Electrical and Computer Engineering, Technion - Israel Institute of Technology, 32000 Haifa, Israel

D. Krioukov (✉)
Network Science Institute, Departments of Physics, Mathematics, and Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA
e-mail: dima@northeastern.edu

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_2

tems is amplified by their stochasticity, making them difficult or impossible to predict, and suggesting that these intricate systems should be modeled by random objects. In other words, the combination of stochasticity and high-order interactions in real-world complex systems suggests that models of *random* simplicial complexes may be useful models of these systems.

From the mathematical perspective, the study of random simplicial complexes combines combinatorics and probability with geometry and topology. As a consequence, the history of random simplicial complexes is quite dramatic. The drama is that this history is super-short, compared to the histories of probability and topology taken separately. For example, the first and simplest model of random simplicial complexes, the Linial-Meshulam model [1], appeared only in 2006. The main reason behind this dramatic delay is that probability and topology had been historically at nearly opposite extremes in terms of methods, skills, intuition, and esthetic (dis)likes among mathematicians. Fortunately, the wall is now dismantled, and over the last 15 years or so, the field of random topology has been growing explosively, as our review attempts to convey through the lens of random simplicial complexes.

We do not attempt to review *all* existing models and results related to random simplicial complexes, which is a mission impossible. Instead, we try to focus on those models for which some exciting phenomena—which we review as well—have been rigorously established in mathematics, especially in topology. It is not entirely coincidental that a majority of these models are particulary attractive not only from the topological and probabilistic perspectives, but also from the statistical and information-theoretic perspectives. This is because these models tend to be statistically unbiased, in the sense that they are canonical models satisfying the maximum entropy principle. As a consequence, they can be used as the correct null models of real-world complexes exhibiting certain structural features.

We take this statistical stance in our review of models in Sect. 2.2. This review is then followed by the review of some of the most exciting phenomena related to these models, which we split between phase transitions (Sect. 2.3) and distributional limit theorems (Sect. 2.4). Many of the presented results are higher-dimensional analogues of the well-known phenomena in random graphs related to connectivity, giant component, cycles, etc., therefore we preamble, where possible, the higher-dimensional statements with brief recollections of their one-dimensional counterparts. The focus on mathematics taken in this review precludes us unfortunately from covering many exciting subjects, such as models of growing complexes [2–7], their applications, and phenomena in them including percolation [8–13]. Yet in the concluding Sect. 2.6 that outlines our view on interesting future directions, we also comment briefly on some applications and their implications for different models of random simplicial complexes.

2.2 Review of Models

In application to real-world systems, the models of random simplicial complexes reviewed here are typically used as null models reproducing a particular property of interest. In other words, these models are not intended to be "correct models" of real-world systems, but they are intended to be *correct null models* of these systems. By "correct" we mean here a model that is statistically unbiased, and by "statistically unbiased" we mean a model that maximizes entropy across all models that have a desired property. Therefore, we begin this section with a brief recollection of basic facts behind the maximum entropy principle and canonical ensembles, which we call canonical models in this review. We then observe that, with a few exceptions, the reviewed models are higher-dimensional generalizations of 1-dimensional random simplicial complexes, which are random graphs. Therefore, we preamble, where possible, the definition of a higher-dimensional model with its 1-dimensional counterpart. We finish the section with a short summary of the maximum entropy properties of the reviewed models.

2.2.1 Maximum Entropy Principle and Canonical Models

The maximum entropy principle [14] formalizes the concept of statistical unbiasedness of a null model. Indeed, Shannon entropy is the unique measure of information satisfying the basic axioms of continuity, monotonicity, and system and subset independence [15], and the maximum entropy principle follows directly from these axioms coupled with the additional consistency axioms of uniqueness and representation invariance [16–18]. For these reasons, a maximum-entropy model is the unique model that contains all the information that the model is asked to model, and, more importantly, that does *not* contain any other junk information.

Formally, let $\mathcal{X}$ be a space of graphs or simplicial complexes, and $\mathbb{P}$ a probability distribution on $\mathcal{X}$ corresponding to a *model* of $\mathcal{X}$: $\mathbb{P}(X)$ is the probability with which the model generates $X \in \mathcal{X}$. The *Shannon entropy* of $\mathbb{P}$ is $S(\mathbb{P}) = -\sum_{X\in\mathcal{X}} \mathbb{P}(X) \log \mathbb{P}(X)$. Let $x_q : \mathcal{X} \to \mathbb{R}, q = 1, 2, \ldots, Q$, be a finite collection of functions that we call *properties* of X, and let $y_q \in \mathbb{R}$ be a set of numbers that we will associate with *values* of properties x_q. In general, properties x_q can take values in spaces that are more sophisticated than $\mathbb{R}$, but $\mathbb{R}$ suffices for us here. Denote $\mathbf{x} = \{x_q\}_{q=1}^{Q}$, $\mathbf{y} = \{y_q\}_{q=1}^{Q}$, and let ρ be a probability distribution on $\mathbb{R}^Q$.

Given a pair of properties $\mathbf{x}$ and their values $\mathbf{y}$, the *microcanonical model* is then the one that maximizes entropy under the sharp constraints that the values of the properties $\mathbf{x}$ of X must be equal to $\mathbf{y}$ exactly: $\mathbb{P}_{\text{mic}}(\mathbf{x}, \mathbf{y}) = \text{argmax}_{\mathbb{P}} \{S(\mathbb{P}) : \mathbf{x}(X) = \mathbf{y}\}$. This means that $\mathbb{P}_{\text{mic}}(\mathbf{x}, \mathbf{y})$ is the uniform distribution over all such Xs:

$$\mathbb{P}_{\text{mic}}[\mathbf{x}, \mathbf{y}](X) = \frac{1}{\mathcal{N}(\mathbf{x}, \mathbf{y})},$$

where $\mathcal{N}(\mathbf{x}, \mathbf{y}) = |\{X \in \mathcal{X} : \mathbf{x}(X) = \mathbf{y}\}|$.

The *canonical model* is the one that maximizes entropy under the soft constraints that the values of the properties $\mathbf{x}$ are equal to $\mathbf{y}$ in expectation: $\mathbb{P}_{\text{can}}(\mathbf{x}, \mathbf{y}) = \text{argmax}_{\mathbb{P}} \{S(\mathbb{P}) : \mathbb{E}\,\mathbf{x} = \mathbf{y}\}$. If the properties $\mathbf{x}$ are sufficiently nice (e.g., satisfy certain convexity assumptions [19, 20]), then the measure $\mathbb{P}_{\text{can}}(\mathbf{x}, \mathbf{y})$ is a Gibbs measure

$$\mathbb{P}_{\text{can}}[\mathbf{x}, \mathbf{y}](X) = \exp[-\boldsymbol{\lambda}(\mathbf{y}) \cdot \mathbf{x}(X)]/Z,$$

where $Z = \sum_{X \in \mathcal{X}} \exp[-\boldsymbol{\lambda}(\mathbf{y}) \cdot \mathbf{x}(X)]$ and the parameters $\boldsymbol{\lambda}(\mathbf{y})$ solve the system of equations $\mathbb{E}\,\mathbf{x} = -\partial \log Z/\partial \boldsymbol{\lambda} = \mathbf{y}$. The solution exists and is unique under the same niceness assumptions [19, 20]. Since Gibbs measures are known as *exponential families* in statistics, canonical models of random graphs are called *exponential random graph models* there. Consequently, canonical models of simplicial complexes are *exponential random simplicial complexes* [21].

Finally, the *hypercanonical model* $\mathbb{P}_{\text{hyp}}(\mathbf{x}, \rho)$ is the canonical model with random $\mathbf{y} \sim \rho$. The measure $\mathbb{P}_{\text{hyp}}(\mathbf{x}, \rho)$ is thus the ρ-mixture of the Gibbs measures,

$$\mathbb{P}_{\text{hyp}}[\mathbf{x}, \rho](X) = \int \mathbb{P}_{\text{can}}[\mathbf{x}, \mathbf{y}](X)\, d\rho(\mathbf{y}),$$

reproducing a desired distribution ρ of the values $\mathbf{y}$ of the properties $\mathbf{x}$ of X.

It is important to notice that $\mathbb{P}_{\text{can}}[\mathbf{x}, \mathbf{y}](X)$ depends on X only via $\mathbf{x}(X)$. For this reason, the properties $\mathbf{x}$ are called *sufficient statistics* [22]: it suffices to know $\mathbf{x}(X)$ to know $\mathbb{P}_{\text{can}}[\mathbf{x}, \mathbf{y}](X)$; no further details about X are needed. It follows that all X's with the same value of $\mathbf{x}(X)$ are equally likely in the canonical model, so that it is a probabilistic mixture of the corresponding microcanonical models, while the hypercanonical model is a probabilistic mixture of the canonical models.

Canonical models of random graphs and simplicial complexes tend to be more tractable than their microcanonical counterparts, explaining why the best-studied models of random simplicial complexes are based on canonical models of random graphs, versus microcanonical ones. We begin with the simplest models based on the Erdős-Rényi graphs.

2.2.2 Complexes Based on Homogeneous Random Graphs

The Erdős-Rényi graphs $G(n, p)$ and $G(n, M)$. Perhaps the best-studied random graph model is the $G(n, M)$ model of random graphs with n nodes and M edges. All such graphs are equiprobable in the model, so that the model is microcanonical. It was introduced in 1951 by Solomonoff and Rapoport [23], and then studied by Erdős and Rényi in 1959 [24]. Its canonical counterpart, introduced by Gilbert in 1959 [25], is the $G(n, p)$ ensemble of random graphs in which every possible edge between the $\binom{n}{2}$ pairs of nodes exists independently with probability p. The sufficient statistic in this canonical model is the number of edges M. If $\binom{n}{2} p/M \to 1$ in the large graph

limit $n \to \infty$, then the microcanonical $G(n, M)$ and canonical $G(n, p)$ models are asymptotically equivalent according to all definitions of such equivalence [26–28].

The Linial-Meshulam complex $Y_2(n, p)$. The constructive definition of the $G(n, p)$ model can be rephrased as follows: take a complete 0-complex, which is a set of n vertices, and then add 1-simplexes (edges) to it at all possible $\binom{n}{2}$ locations independently with probability p. The result is a random 1-dimensional complex $Y_1(n, p) = G(n, p)$.

The Linial-Meshulam model [1] is a straightforward 2-dimensional analogy of $G(n, p)$: take a complete 1-complex, which is the complete graph of size n, and then add 2-simplexes (filled triangles) to it at all possible $\binom{n}{3}$ locations independently with probability p. The result is a random 2-dimensional complex $Y_2(n, p)$.

The d-dimensional Linial-Meshulam complexes $Y_d(n, p)$ and $Y_d(n, M)$. The Linial-Meshulam complex admits the natural generalization to any dimension $d = 1, 2, \ldots, n-1$ [29] in the following way. Take a complete $(d-1)$-complex, and then add d-simplexes to it at all possible $\binom{n}{d+1}$ locations independently with probability p. We denote this random complex by $Y_d(n, p)$.

The microcanonical version $Y_d(n, M)$ of canonical $Y_d(n, p)$ is also well defined. Here, we can take a complete $(d-1)$-complex, add exactly M d-simplexes to it chosen uniformly at random out of all the $\binom{\binom{n}{d+1}}{M}$ possibilities. Note that $Y_1(n, M) = G(n, M)$. The microcanonical $Y_d(n, M)$ model gained less consideration than the canonical $Y_d(n, p)$ one, but some of its aspects were studied in [30, 31]. It should be easy to show that $Y_d(n, M)$ is asymptotically equivalent to $Y_d(n, p)$, but this has not been done.

The random flag complexes $X(n, p)$ and $X(n, M)$. The Linial-Meshulam complexes are just one way to generalize the Erdős-Rényi graphs to higher dimensions. Another straightforward generalization is to extend $G(n, p)$ into a random flag complex $X(n, p)$ [32, 33]. Recall that the *flag complex* of a graph G is the simplicial complex X obtained by filling all the $(k+1)$-cliques in G with k-simplexes, for all $k = 1, 2, \ldots, n-1$. We can similarly define $X(n, M)$ as the flag complex of $G(n, M)$. To the best of our knowledge, the $X(n, M)$ model has not been considered in the past.

The multi-parameter complexes $X(n, \mathbf{p})$ and $X(n, \mathbf{M})$. A further generalization of $G(n, p)$, subsuming both the Linial-Meshulam and flag complexes, is the multi-parameter complex $X(n, \mathbf{p})$ considered in [34–37]. Let $\mathbf{p} = (p_1, p_2, \ldots, p_{n-1})$ where $p_k \in [0, 1]$, $k = 1, 2, \ldots, n-1$, is the probability of the existence of simplexes of dimension k in $X(n, \mathbf{p})$. Given this vector of probabilities, the complex $X(n, \mathbf{p})$ is then defined as follows. First, take n vertices and add edges to them independently with probability p_1, i.e., generate a $G(n, p_1)$ graph. Second, considering this graph as a 1-skeleton of a 2-complex, go over all the triangles (i.e. 3-cliques) in it, and fill each of them with a 2-simplex independently with probability p_2. Do not stop here, but continue in this fashion—given the $(k-1)$-skeleton, go over all the $(k+1)$-cliques in it, and fill each of them with a k-simplex independently with probability p_k—until $k = n-1$.

Among the $G(n, p)$-based complexes discussed thus far, the $X(n, \mathbf{p})$ complex is the most general since it subsumes both the flag complex,

$$X(n, p) = X(n, (p, 1, 1, \ldots, 1)),$$

and the Linial-Meshulam complex,

$$Y_d(n, p) = X(n, (\underbrace{1, 1, \ldots, 1}_{d-1}, p, \underbrace{0, 0, \ldots, 0}_{n-d-1})).$$

The $X(n, \mathbf{p})$ model is canonical, and to specify its sufficient statistics, we define a *k-shell* $\partial\sigma$ to be the complete $(k-1)$-dimensional boundary of a *(potential)* k-simplex σ in a complex X. That is, X may or may not contain σ (σ can be filled or empty), but if $\partial\sigma$ is a k-shell, then all its simplexes are filled, so that $\partial\sigma$ can be thought of as a "pre-k-simplex", in the sense that it is ready to be filled with σ, but might eventually be left empty.

As shown in [21] (see also Sect. 2.2.6), the sufficient statistics in $X(n, \mathbf{p})$ are not only the numbers $M_{s,k}$ of k-simplexes of each dimension k, but also the numbers $M_{c,k}$ of k-shells. Therefore, the microcanonical counterpart of $X(n, \mathbf{p})$ is $X(n, \mathbf{M})$, where $\mathbf{M} = ((M_{c,1}, M_{s,1}), (M_{c,2}, M_{s,2}), \ldots, (M_{c,n-1}, M_{s,n-1}))$, $M_{c,1} = \binom{n}{2}$, and $M_{s,1} = M$. This model has not been considered in the past, but it is another natural higher-dimensional generalization of $G(n, M)$, while $X(n, M)$ and $Y_d(n, M)$ are special cases of $X(n, \mathbf{M})$.

All other relationships between the discussed complexes are shown in Fig. 2.1.

2.2.3 Complexes Based on Inhomogeneous Random Graphs

2.2.3.1 General Complexes

The inhomogeneous random graph $G(n, \hat{p})$. The edge existence probability in the random Erdős-Rényi graph $G(n, p)$ does definitely not have to be the same p for all edges. Each edge i, j can have a different existence probability p_{ij}, $i, j = 1, 2, \ldots, n$. Such random graphs are known as inhomogeneous or generalized random graphs [38–41]. We denote them by $G(n, \hat{p})$, where $\hat{p}$ is the matrix of edge existence probabilities, $\hat{p} = \{p_{ij}\}$.

The multi-parameter complex $X(n, \hat{\mathbf{p}})$. A straightforward generalization of $G(n, \hat{p})$ to higher dimensions was introduced in [21]. It is also a generalization of $X(n, \mathbf{p})$, and it is defined as follows. Let $\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_{n-1})$ be a vector of simplex existence probabilities for all possible simplexes of all possible dimensions $k = 1, 2, \ldots, n-1$: $\hat{p}_1 = \{p_{ij}\}$, $\hat{p}_2 = \{p_{ijl}\}$, and $\hat{p}_k = \{p_{\sigma_k}\}$ collects the $\binom{n}{k+1}$ existence probabilities for all possible $\binom{n}{k+1}$ simplexes σ_k of dimension k. Given such a vector of tensors $\hat{\mathbf{p}}$, the random complex $X(n, \hat{\mathbf{p}})$ is then generated similarly to $X(n, \mathbf{p})$:

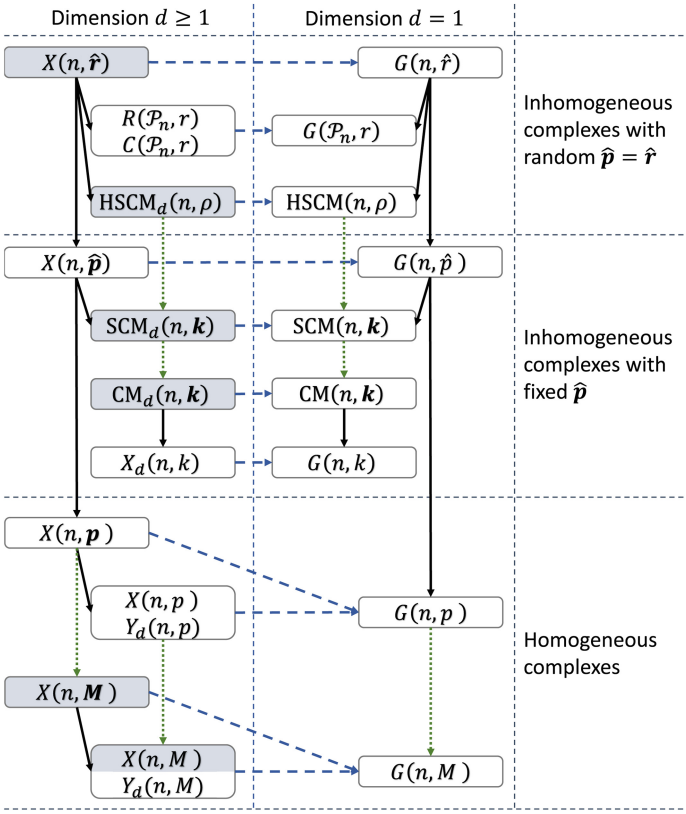


Fig. 2.1 The relations between the considered models of random simplicial complexes. The vertical solid lines connect more general complexes to their special cases. The horizontal dashed lines connect higher-dimensional complexes to their 1-dimensional cases or 1-skeletons. The vertical dotted lines connect probabilistic mixtures (e.g., canonical models) to their constituents (e.g., microcanonical models). The shaded models have not been considered before. The review of models in Sect. 2.2 proceeds roughly against the arrow directions, from the least general $G(n, M)$ to the most general $X(n, \hat{\mathbf{r}})$

starting with an inhomogeneous random graph $G(n, \hat{p}_1)$, go over all the 3-cliques (2-shells) in this graph, and fill them with 2-simplexes σ_2 independently with probability p_{σ_2}. Proceed to higher dimensions $k = 3, 4, \ldots, n-1$ in a similar fashion, filling k-shells with k-simplexes independently with probabilities p_{σ_k}. The complex $X(n, \mathbf{p})$ is a special case of $X(n, \hat{\mathbf{p}})$—the one with $\hat{p}_k \equiv p_k$.

2.2.3.2 Configuration Models

The configuration models $\mathrm{SCM}(n, \mathbf{k})$ and $\mathrm{CM}(n, \mathbf{k})$. The inhomogeneous random graphs $G(n, \hat{p})$ are quite general, subsuming many other important models of random

graphs. In particular, they encompass the soft configuration model SCM$(n, \mathbf{k})$ [38–44], which is the canonical model of random graphs with a given sequence of expected degrees $\mathbf{k} = (k_1, k_2, \ldots, k_n)$, $k_i \geq 0$, $i = 1, 2, \ldots, n$. The SCM$(n, \mathbf{k})$ model is the $G(n, \hat{p})$ with $\hat{p}$ given by

$$p_{ij} = \frac{1}{e^{\lambda_i + \lambda_j} + 1}, \tag{2.1}$$

where the parameters λ_i's (known as Lagrange multipliers) solve the system of n equations given by

$$\sum_j p_{ij} = k_i. \tag{2.2}$$

This equation guarantees the expected degree of node i to be k_i.

The sufficient statistics in SCM$(n, \mathbf{k})$ are the degrees of all nodes in a graph. Therefore, SCM$(n, \mathbf{k})$'s microcanonical counterpart, the configuration model CM$(n, \mathbf{k})$ [45, 46], is the uniform distribution over all the graphs with the degree sequence $\mathbf{k}$. Note that $\mathbf{k}$ can be any sequence of nonnegative real numbers in SCM$(n, \mathbf{k})$, but in CM$(n, \mathbf{k})$, $\mathbf{k}$ is a graphical sequence of nonnegative integers. A sequence $\mathbf{k}$ is called *graphical*, if there exists a graph whose degree sequence is $\mathbf{k}$. The necessary and sufficient conditions for $\mathbf{k}$ to be graphical are the Erdős-Gallai conditions [47].

An important special case of CM$(n, \mathbf{k})$ is the random k-regular graph $G(n, k) =$ CM$(n, (k, k, \ldots, k))$.

The d-dimensional configuration models SCM$_d(n, \mathbf{k})$ and CM$_d(n, \mathbf{k})$. Recall that the degree k_i of a vertex (0-simplex) i is the number of edges (1-simplexes) that contain i. In a similar vein, the *degree* k_σ of d-simplex σ is defined to be the number of $(d+1)$-simplices that contain σ.

The d-dimensional soft configuration model SCM$_d(n, \mathbf{k})$ is a generalization of SCM$(n, \mathbf{k})$, where now $\mathbf{k} = \{k_\tau\}$ is a sequence of $\binom{n}{d}$ expected degrees $k_\tau \geq 0$ of all $(d-1)$-simplexes τ. To construct SCM$_d(n, \mathbf{k})$ we take the complete $(d-1)$-skeleton on n vertices, and add all possible d-simplexes σ independently with probability

$$p_\sigma = \frac{1}{e^{\sum_\tau \lambda_\tau \mathbb{1}\{\tau < \sigma\}} + 1}, \tag{2.3}$$

where the summation is over all $(d-1)$-faces τ of σ (as implied by the standard $\tau < \sigma$ notation), and the parameters λ_τ solve the system of $\binom{n}{d}$ equations

$$\sum_\sigma p_\sigma \mathbb{1}\{\tau < \sigma\} = k_\tau. \tag{2.4}$$

Note that SCM$_d(n, \mathbf{k})$ is a special case of $X(n, \hat{\mathbf{p}})$:

$$\text{SCM}_d(n, \mathbf{k}) = X(n, (\underbrace{1, 1, \ldots, 1}_{d-1}, \hat{p}_d, \underbrace{0, 0, \ldots, 0}_{n-d-1})),$$

where $\hat{p}_d = \{p_\sigma\}$ is given by (2.3). If $d = 1$, then Eqs. (2.3 and 2.4) reduce to Eqs. (2.1 and 2.2), respectively, and we have $\mathrm{SCM}_1(n, \mathbf{k}) \equiv \mathrm{SCM}(n, \mathbf{k})$.

The $\mathrm{SCM}_d(n, \mathbf{k})$ is the canonical model of random complexes whose sufficient statistics are the degrees of all the $(d-1)$-simplexes in a complex. Therefore, $\mathrm{SCM}_d(n, \mathbf{k})$'s microcanonical counterpart, the configuration model $\mathrm{CM}_d(n, \mathbf{k})$, is the uniform distribution over complexes with the complete $(d-1)$-skeleton and the degree sequence of $(d-1)$-simplexes equal to $\mathbf{k}$. In $\mathrm{CM}_d(n, \mathbf{k})$, $\mathbf{k}$ is a realizable sequence of $\binom{n}{d}$ nonnegative integers, versus any sequence of $\binom{n}{d}$ nonnegative real numbers in $\mathrm{SCM}_d(n, \mathbf{k})$. The conditions for $\mathbf{k}$ to be realizable, analogous to the Erdős-Gallai graphicality conditions in the $d = 1$ case, are at present unknown.

An important special case of the configuration model is the random k-regular complex $X_d(n, k) = \mathrm{CM}_d(n, (k, k, \ldots, k))$ in which all $(d-1)$-simplexes have degree k [48]. For $k = 1$, such a complex is also known as an (n, d)-Steiner system, whose randomized construction is due to [49].

With the exception of $X_d(n, k)$, the d-dimensional configuration models $\mathrm{CM}_d(n, \mathbf{k})$ and $\mathrm{SCM}_d(n, \mathbf{k})$ have not been considered in the past. The configuration models defined and studied in [50, 51] are very different and unrelated to any other model considered above. We review them in Sect. 2.2.5.2.

2.2.4 Complexes Based on Inhomogeneous Random Graphs with Random Connection Probabilities

2.2.4.1 General Complexes

The inhomogeneous random graph with random connection probabilities $G(n, \hat{r})$. Note that the edge probabilities p_{ij} in inhomogeneous random graphs $G(n, \hat{p})$ can themselves be random [40, 41, 52, 53]. In that case, we replace $\hat{p}$ with $\hat{r} = \{r_{ij}\}$ which is a random matrix with entries in $[0, 1]$, and the corresponding inhomogeneous random graph is denoted $G(n, \hat{r})$. This class of random graphs is extremely general and subsumes a great variety of many important random graph models, for example $G(n, \hat{p})$.

Of a particular interest is the case where

$$r_{ij} = p(X_i, X_j),$$

where $\{X_1, \ldots, X_n\}$ is a set of random variables in some measure space S, and $p : S \times S \to [0, 1]$ is a fixed function called the *connection probability function*. This class of models is quite general and subsumes, for instance, the stochastic block models [54], hidden variable models [52, 53], latent space models [55–57], random geometric graphs [58, 59], and many other popular models. If the X_i's are independent random variables uniformly distributed in $[0, 1]$, and p is a symmetric integrable function, then p is also known as a *graphon* in the theory of graph limits [60, 61].

The multi-parameter complex $X(n, \hat{\mathbf{r}})$. The most general class of models considered in this chapter is a version of the multi-parameter complex $X(n, \hat{\mathbf{p}})$ with random simplex existence probabilities $\hat{\mathbf{p}}$. To emphasize their randomness, we denote this complex by $X(n, \hat{\mathbf{r}})$, where $\hat{\mathbf{r}}$ is random $\hat{\mathbf{p}}$. For maximum generality, the probabilities of existence of different simplexes, including simplexes of different dimensions, do not have to be independent random variables, so that the space of $X(n, \hat{\mathbf{r}})$ is parameterized by joint probability distributions over the vector of tensors $\hat{\mathbf{r}}$, equivalent to all nonempty subsets of $\{1, \ldots, n\}$. As evident from Fig. 2.1, all other models of random complexes and graphs are special cases of $X(n, \hat{\mathbf{r}})$. In particular, $X(n, \hat{\mathbf{p}})$ is a special degenerate case of $X(n, \hat{\mathbf{r}})$. At this level of generality, the complex $X(n, \hat{\mathbf{r}})$ has not been considered or defined before, yet it is simply a higher-dimensional version of the well-studied $G(n, \hat{r})$.

2.2.4.2 Hypersoft Configuration Model

The hypersoft configuration model $\mathrm{HSCM}(n, \rho)$. The hypersoft configuration model is the hypercanonical model of random graphs with a given expected degree distribution ρ [40, 41, 44, 52, 53, 62]. It is defined as the $\mathrm{SCM}(n, \mathbf{k})$ in which the expected degree sequence $\mathbf{k}$ is not fixed but random: the expected degree k_i of every vertex i is an independent random variable with distribution ρ, $k_i \sim \rho$. In other words, to generate an $\mathrm{HSCM}(n, \rho)$ graph, one first samples k_is independently from the distribution ρ, then solves the system of equations (2.2) to find all λ_is, and finally creates edges with probabilities p_{ij} (2.1). The $\mathrm{HSCM}(n, \rho)$ is thus a probabilistic mixture of $\mathrm{SCM}(n, \mathbf{k})$'s with random $\mathbf{k} \sim \rho^n$. The degree distribution in $\mathrm{HSCM}(n, \rho)$ is the mixed Poisson distribution with mixing ρ [40, 41].

An equivalent definition of $\mathrm{HSCM}(n, \rho)$ is based on the observation that the distribution ρ defines the joint distribution Ψ of $\Lambda = \{\lambda_i\}$ via (2.2). This means that an equivalent procedure to generate an $\mathrm{HSCM}(n, \rho)$ graph is to sample Λ directly from Ψ first, and then create edges with probabilities p_{ij} (2.1). This definition makes it explicit that the $\mathrm{HSCM}(n, \rho)$ is a special case of $G(n, \hat{r})$.

For a given ρ, it may difficult to find the explicit form of the distribution Ψ. However, in many cases with important applications—including the Pareto ρ with a finite mean, for instance—it has been shown that the canonical connection probability (2.1) and its classical limit approximation

$$p_{ij}^{\mathrm{CL}} = \min\left(1, e^{-\lambda_i} e^{-\lambda_i}\right) = \min\left(1, \frac{k_i k_j}{\bar{k} n}\right), \tag{2.5}$$

where $\bar{k}$ is the mean of ρ, lead to asymptotically equivalent random graphs [28]. In such cases, the random Lagrange multipliers λ_i's are asymptotically independent, $\Psi = \psi^n$, $\lambda_i \sim \psi$, with the distribution ψ defined by the distribution ρ via the change of variables $\lambda_i = \ln\left(\sqrt{\bar{k} n}/k_i\right)$ where $k_i \sim \rho$. Generating such a graph is extremely simple: first sample either the k_i's or λ_i's independently from the distribution ρ or ψ, respectively, and then generate edges with probabilities p_{ij}^{CL} (2.5).

The d-dimensional hypersoft configuration model $\mathrm{HSCM}_d(n, \rho)$. A straightforward generalization of $\mathrm{HSCM}(n, \rho)$ to dimension d is achieved by defining the $\mathrm{HSCM}_d(n, \rho)$ as the probabilistic mixture of $\mathrm{SCM}_d(n, \mathbf{k})$ with random $\mathbf{k} \sim \rho^{\binom{n}{d}}$: the expected degree k_τ of every $(d-1)$-simplex τ is an independent random variable with distribution ρ. In other words, to generate a random $\mathrm{HSCM}_d(n, \rho)$ complex, one first prepares the complete $(d-1)$-complex of size n, then samples the $\binom{n}{d}$ expected degrees k_τ of all the $(d-1)$-simplexes τ independently from the distribution ρ, then solves the system of equations (2.4) to find all the λ_τ's, and finally creates the d-simplexes σ independently with probability p_σ (2.3). For the same reasons as in the $d = 1$ case, the $\mathrm{HSCM}_d(n, \rho)$ model is a special case of $X(n, \hat{\mathbf{r}})$ for any $d \geq 1$. The model has not been previously considered, except the $d = 1$ case.

2.2.4.3 Geometric Complexes

An important class of random complexes are geometric complexes, which are a special case of $X(n, \hat{\mathbf{r}})$ whose 1-skeletons are random geometric graphs. Their randomness comes from random locations of vertices in a space, and this randomness is often modeled by either binomial or Poisson point processes whose definitions we recall next.

Let S be a metric space with a probability distribution P on it. The definitions below apply to any sufficiently nice pair of S and P. However, for the sake of simplicity, in the subsequent sections of this chapter, the space S will always be a flat d-dimensional torus $S = \mathbb{T}^d = [0, 1]^d / \{0 \sim 1\}$, while P will be its Lebesgue measure λ, i.e., $P(A) = \lambda(A)/\lambda(\mathbb{T}^d) = \lambda(A)$, where $\lambda(A)$ is the Euclidean volume of $A \subseteq \mathbb{T}^d$.

The binomial point process is simply a set of n random points $\mathcal{X}_n = \{X_1, \ldots, X_n\} \subset S$ sampled independently from P. In the simplest case $S = \mathbb{T}^1$, we simply sample X_i's from the uniform distribution, $X_i \sim \mathcal{U}(0, 1)$, viewed as a circle. The distribution of the number of points in any measurable subset $A \subseteq S$ is binomial with mean $nP(A)$, giving the process its name.

The Poisson point process is the binomial process with a random number of points N sampled from the Poisson distribution with mean n, i.e. $\mathcal{P}_n = \mathcal{X}_N$, where $N \sim \mathrm{Pois}(n)$. The distribution of the number of points in any measurable $A \subseteq S$ is Poisson with mean $nP(A)$, where n is the *rate* of the process. Another key property of the Poisson process is *spatial independence*: the numbers of points in any pair of disjoint subsets of S are independent random variables. This property, absent in the binomial process, makes the Poisson process favorable in probabilistic analysis.

In the $n \to \infty$ limit, the two process are asymptotically identical—the binomial process converges to the Poisson process in a proper sense [63]. For this reason, and given that the Poisson process is easier to deal with, we will restrict ourselves in this chapter to the Poisson process $\mathcal{P}_n$ on the torus $\mathbb{T}^d$. These setting turn out to provide the most elegant results, while not losing much in terms of generality.

The random geometric graph $G(\mathcal{P}_n, r)$. We start with $G(\mathcal{X}_n, r)$, which is a special case of $G(n, \hat{r})$, in which

$$r_{ij} = p(X_i, X_j) = \mathbb{1}\left\{d(X_i, X_j) \leq r\right\},$$

where $d(X_i, X_j)$ is the distance between $X_i, X_j \in \mathcal{X}_n$ in $\mathbb{T}^d$, and $r > 0$ is the connectivity radius parameter [58, 59]. In other words, the vertices of this random graph are a realization of the binomial process, and two vertices are connected if the distance between them in $\mathbb{T}^d$ is at most r. To generate the random geometric graph $G(\mathcal{P}_n, r)$ we first sample $N \sim \text{Pois}(n)$, and then generate $G(\mathcal{X}_N, r)$ as described above.

The random Vietoris-Rips complex $R(\mathcal{P}_n, r)$. This is the flag complex over the random geometric graph $G(\mathcal{P}_n, r)$. It is thus a straightforward higher-dimensional generalization of $G(\mathcal{P}_n, r)$. Note that we can consider $R(\mathcal{X}_n, r)$ for the binomial process as well, which is a special case of $X(n, \hat{\mathbf{r}})$. Then $R(\mathcal{P}_n, r) = R(\mathcal{X}_N, r)$ with $N \sim \text{Pois}(n)$.

The random Čech complex $C(\mathcal{P}_n, r)$. This is a different higher-dimensional generalization of $G(\mathcal{P}_n, r)$. The $C(\mathcal{P}_n, r)$ rule is: draw balls $B_{r/2}(X_i)$ of radius $r/2$ around each of the points $X_i \in \mathcal{P}_n$, then look for all the intersections of these balls, and for every $(k+1)$-fold intersection of the balls, add the corresponding k-simplex to the complex. Note that the binomial version $C(\mathcal{X}_n, r)$ can still be considered as a special case of $X(n, \hat{\mathbf{r}})$ with a different rule for the creation of higher-dimensional simplexes, compared to $R(\mathcal{X}_n, r)$.

While the 1-skeleton of both $R(\mathcal{P}_n, r)$ and $C(\mathcal{P}_n, r)$ is the same random geometric graph $G(\mathcal{P}_n, r)$, higher-dimensional simplexes in these complexes are in general different, as illustrated in Fig. 2.2. The following useful relation was proved in [64] for any $\alpha \leq \sqrt{\frac{d+1}{2d}}$:

$$R(\mathcal{P}_n, \alpha r) \subset C(\mathcal{P}_n, r) \subset R(\mathcal{P}_n, r).$$

A powerful property of the Čech complex is due to the Nerve Lemma [65] stating that under mild conditions we have the following isomorphism between the homology groups H_k:

$$H_k(C(\mathcal{P}_n, r)) \cong H_k(B(\mathcal{P}_n, r)),$$

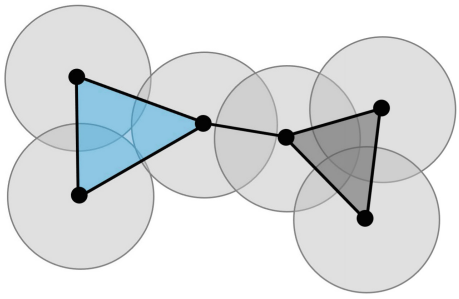


Fig. 2.2 The Čech and Rips complexes. We draw balls of a fixed radius r around the points, and consider their intersection. The edges in both cases correspond to the geometric graph. While both triangles belong to the Rips complex, the left triangle is not included in the Čech complex

where $B(\mathcal{P}_n, r) = \bigcup_{X_i \in \mathcal{P}_n} B_{r/2}(X_i)$. This property, absent in the Rips complex, is very useful in the probabilistic analysis of the Čech complex, allowing one to switch back-and-forth between combinatorics and stochastic geometry.

2.2.5 *Complexes Based on Random Hypergraphs*

2.2.5.1 General Complexes

The main reason why the Linial-Meshulam model $Y_d(n, p)$ is defined on top of the complete $(d-1)$-skeleton, is simplicity. The complete $(d-1)$-complex underlying the d-complex $Y_d(n, p)$ simplifies its topological analysis drastically. Much less simple, but also much more general, is the $X(n, \hat{\mathbf{p}})$ complex in which the probability of existence of simplexes of different dimensions are all different, but this complex is still relatively simple from the probabilistic perspective, since simplexes of higher dimensions are created in a conditionally independent manner, with the conditions being the existence of required simplexes of lower dimensions. Very recently, this simplicity was sacrificed even further for the sake of even greater generality in a class of models based on random hypergraphs [66, 67], which we briefly review next.

The inhomogeneous random hypergraph $H(n, \hat{\mathbf{p}})$. This is a straightforward generalization of the inhomogeneous random graph $G(n, \hat{p})$ to hypergraphs: every possible hyperedge σ (which is a nonempty subset of $\{1, \ldots, n\}$) in $H \sim H(n, \hat{\mathbf{p}})$ exists independently with probability $p_\sigma \in \hat{\mathbf{p}}$. Note that H is not a simplicial complex with high probability, unless $\hat{\mathbf{p}}$ is specially designed for H to be a complex.

The lower and upper complexes $\underline{Z}(n, \hat{\mathbf{p}})$ and $\overline{Z}(n, \hat{\mathbf{p}})$. Introduced in [66, 68], these are both based on the $H(n, \hat{\mathbf{p}})$. The lower complex $\underline{Z} \sim \underline{Z}(n, \hat{\mathbf{p}})$ is the largest simplicial complex that the random $H \sim H(n, \hat{\mathbf{p}})$ contains, while the upper complex $\overline{Z} \sim \overline{Z}(n, \hat{\mathbf{p}})$ is the smallest simplicial complex that contains the random $H \sim H(n, \hat{\mathbf{p}})$, so that $\underline{Z} \leq H \leq \overline{Z}$. In other words, a simplex σ is included in the lower complex $\underline{Z}$ if and only if $\sigma \in H$ and also all its faces are hyperedges of H. The other way around, every hyperedge $\sigma \in H$ is included in the upper complex $\overline{Z}$ as a simplex together with all its faces, even if some of these faces do not happen to be in H.

Compared to the graph-based models reviewed in the previous sections, the hypergraph-based models are much more difficult to analyze, primarily because the distributions of the skeletons are highly nontrivial and carry intricate dependency structure. However, an exciting fact about these two models is that they are dual in some strict topological sense (Alexander duality), and therefore statements about one of the models can be translated with ease to corresponding statements about the other [66, 67].

At this point, it may be tempting to proceed similarity to the previous X-sections, and let the hyperedge existence probabilities $\hat{\mathbf{p}}$ be random $\hat{\mathbf{r}}$, leading to $H(n, \hat{\mathbf{r}})$, $\underline{Z}(n, \hat{\mathbf{r}})$, and $\overline{Z}(n, \hat{\mathbf{r}})$. However, nothing new is gained by doing so, since the $X(n, \hat{\mathbf{r}})$

model is already general enough as it includes any possible joint distribution of simplex existence probabilities, including the distributions describing $\underline{Z}(n, \hat{\mathbf{p}})$ and $\overline{Z}(n, \hat{\mathbf{p}})$. That is, the lower and upper models $\underline{Z}(n, \hat{\mathbf{p}})$ and $\overline{Z}(n, \hat{\mathbf{p}})$ with nonrandom $\hat{\mathbf{p}}$ are special cases of the more general $X(n, \hat{\mathbf{r}})$ model with the joint distribution of $\hat{\mathbf{r}}$ set equal to the joint distribution of simplex existence probabilities in $\underline{Z}(n, \hat{\mathbf{p}})$ and $\overline{Z}(n, \hat{\mathbf{p}})$, while the lower and upper models $\underline{Z}(n, \hat{\mathbf{r}})$ and $\overline{Z}(n, \hat{\mathbf{r}})$ with random $\hat{\mathbf{r}}$ are equivalent to $X(n, \hat{\mathbf{r}}_*)$ with matching joint distributions of $\hat{\mathbf{r}}_*$.

2.2.5.2 *Z*-configuration Models

One interesting special case of the Z-complexes from the previous section is the Z-version of the d-dimensional (soft) configuration model considered in [50].

The Z-configuration models Z-CM$_d(n, \mathbf{k})$ and Z-SCM$_d(n, \mathbf{k})$. The *d-degree* $k_\sigma(d)$ of a d'-simplex σ is defined to be the number of simplexes of dimension $d > d'$ that contain σ [2], and the Z-configuration models are defined by a vector $\mathbf{k} = \{k_i(d)\}$ of a given sequence of (expected) d-degrees of vertices $i = 1, 2, \ldots, n$, as opposed to the vector of (expected) d-degrees of $(d-1)$-simplexes in the (S)CM$_d(n, \mathbf{k})$ in Sect. 2.2.3.2. No low-dimensional skeleton is formed in the Z-SCM$_d(n, \mathbf{k})$, while the probability of d-simplex σ is

$$p_\sigma = \frac{1}{e^{\sum_i \lambda_i \mathbb{1}\{i<\sigma\}} + 1}, \tag{2.6}$$

where the n parameters λ_i of vertices i solve the system of n equations

$$\sum_\sigma p_\sigma \mathbb{1}\{i < \sigma\} = k_i(d). \tag{2.7}$$

If simplex σ is added to the complex, then so are all its lower-dimensional faces. The Z-SCM$_d(n, \mathbf{k})$ is a special case of the $\bar{Z}(n, \hat{\mathbf{p}})$ with $\hat{\mathbf{p}} = (0, \ldots, 0, \hat{p}_d, 0, \ldots, 0)$ and $\hat{p}_d = \{p_\sigma\}$ given by (2.6).

The model is canonical with the d-degree of all vertices being its sufficient statistics, so that the corresponding microcanonical model Z-CM$_d(n, \mathbf{k})$ is the uniform distribution over all the d-complexes with the d-degree sequence of the vertices equal to $\mathbf{k}$. Note that these complexes satisfy the additional constraint: they do not contain any simplex of dimension $d' < d$ that is not a face of a d-simplex. The hypercanonical model Z-HSCM$_d(n, \rho)$ is the Z-SCM$_d(n, \mathbf{k})$ with random $k_i(d) \sim \rho$.

2.2.6 Random Simplicial Complexes as Canonical Models

Table 2.1 summarizes what is known concerning the reviewed complexes as canonical models. To convey this knowledge succinctly, it suffices to focus only on the canonical

Table 2.1 Canonical models of random simplicial complexes

Models	Spaces $\mathcal{S}$	Properties $\mathbf{x}$	References	Comments
$G(n,p)$, $G(n,M)$	$\mathcal{G}_n$	M	[39]	(1)
$X(n,p)$, $X(n,M)$	$\mathcal{F}_n$	M	[21]	(1)
$Y_d(n,p)$, $Y_d(n,M)$	$\mathcal{C}_{n,d}$	$M_{s,d}$	[21]	(1)
$X(n,\mathbf{p})$, $X(n,\mathbf{M})$	$\mathcal{C}_n$	$\mathbf{M}=\{M_{c,d},M_{s,d}\}_{d=1}^{n-1}$	[21]	(1, 2)
$G(n,\hat{p})$, $G(n,\hat{r})$	$\mathcal{G}_n$	$\hat{\sigma}_1=\{\sigma_1\}$	[39]	(1)
$X(n,\hat{\mathbf{p}})$, $X(n,\hat{\mathbf{r}})$	$\mathcal{C}_n$	$\hat{\boldsymbol{\sigma}}=\{\{\partial\sigma_d\},\{\sigma_d\}\}_{d=1}^{n-1}$	[21]	(1, 2)
((H)S)CM$(n,\cdot)$	$\mathcal{G}_n$	$\mathbf{k}=\{k_i\}$	[39]	(3)
((H)S)CM$_d(n,\cdot)$	$\mathcal{C}_{n,d}$	$\mathbf{k}=\{k_{\sigma_{d-1}}\}$	here	(4)
Z-((H)S)CM$_d(n,\cdot)$	$\mathcal{Z}_{n,d}$	$\mathbf{k}=\{k_i(d)\}$	[50]	(5)

The first column lists the canonical models from Sect. 2.2, along with their micro- and hypercanonical counterparts (if any), whose entropy maximization properties have been established. The second and third columns document the constraints (space of complexes and their properties) under which the model entropy is maximized. The fourth column contains references to where this maximization has been established and to further details. The last column refers to pertinent comments in the text

models *per se*, since their microcanonical constituents and hypercanonical mixtures are coupled to them as discussed in Sect. 2.2.1.

To define a canonical model, one needs to specify not only the sufficient statistics $\mathbf{x}$ (whose expectations can take any admissible values $\mathbf{y}$ in a particular canonical model), but also the space $\mathcal{S}$ of allowed complexes over which the entropy of a canonical model is maximized. Table 2.1 uses the following notations for such spaces $\mathcal{S}$ of *labeled* graphs and complexes:

- $\mathcal{G}_n$: all graphs of size n;
- $\mathcal{F}_n$: flag complexes of size n;
- $\mathcal{C}_n$: all complexes of size n;
- $\mathcal{C}_{n,d}$: d-complexes of size n with a complete $(d-1)$-skeleton;
- $\mathcal{Z}_{n,d}$: d-complexes of size n whose all d'-simplexes, for all $d'<d$, are faces of d-simplexes.

For the sufficient statistics, Table 2.1 uses the following notations that rely on the definition of a d-shell, which can be found at the end of Sect. 2.2.2:

- $M_{c,d}$ and $M_{s,d}$: numbers of d-shells and d-simplexes ($M_{s,1}=M$, the number of edges);
- $\partial\sigma_d$ and σ_d: d-shell and d-simplex ($\sigma_0=i$, a vertex; $\sigma_1=\{i,j\}$, an edge);
- k_{σ_d} and $k_{\sigma_d}(d')$: $(d+1)$-degree and generalized d'-degree of σ_d $(d<d'<n)$.

Finally, the comments referred to in Table 2.1 are as follows:

1. The values $\mathbf{y}$ of properties $\mathbf{x}$ are not shown, but they are assumed to match properly the model parameters. That is, in a general canonical model we have that $\mathbb{E}\,\mathbf{x}=\mathbf{y}$, so that the $G(n,p)$ model, for example, is the canonical model with $\mathbf{x}=M$ and $\mathbf{y}=\mathbb{E}\,M=\binom{n}{2}p$. Similarly, the $X(n,\hat{\mathbf{p}})$ model is the canonical

model with $\mathbb{E}\,\sigma_d = p_{\sigma_d}\mathbb{E}\,\partial\sigma_d$ and $\mathbb{E}\,\partial\sigma_d = \prod_{\tau_{d-1}<\sigma_d}\mathbb{E}\,\tau_{d-1}$, giving the expected values of the sufficient statistics as functions of the model parameters p_{σ_d}.

2. The presence of the constraints on the existence of the boundaries of d-simplexes, their d-shells, *in addition* to the constraints on the existence of d-simplexes themselves, may appear surprising at first. These stem from the conditional nature of the definition of these complexes. The models with these additional constraints removed, e.g. the canonical model over $\mathcal{C}_n$ with $\mathbf{x} = \left\{M_{s,d}\right\}_{d=1}^{n-1}$, are still entirely unknown. This is not surprising, since such models appear to be combinatorially intractable [21].
3. Note that the constraints under which entropy is maximized in a hypercanonical model are generally intractable. The proofs that these constraints are the degree distributions in the dense and sparse HSCM graphs appear in [44] and [62], respectively. That is, the HSCM is the unbiased maximum-entropy model of random graphs with a given degree *distribution*, versus an (expected) degree *sequence* in the (S)CM.
4. The d-dimensional ((H)S)CM models have not been considered before. The proof that $\mathrm{SCM}_d(n, \mathbf{k})$ maximizes entropy subject to the expected $(d-1, d)$-degree sequence constraints is a straightforward adjustment of notations in the corresponding proof for a general canonical model [14], so that it is omitted here for brevity.
5. The hypercanonical version of the model ($\mathrm{Z\text{-}HSCM}_d(n, \rho)$) has not been considered before. Efficient algorithms to sample from a generalized version of $\mathrm{Z\text{-}CM}_d(n, \mathbf{k})$ are considered in [51].

2.3 Phase Transitions

Phase transitions are very interesting and important phenomena occurring in random structures in statistical physics. In this section we briefly recall the most fundamental types of phase transitions that have been studied in random graphs, and then describe their higher dimensional analogues in simplicial complexes.

2.3.1 Homological Connectivity

The very first theorem proved for random graphs in [24] was the phase transition when the $G(n, M)$ random graph becomes connected. The following is the equivalent result for the $G(n, p)$ random graph.

Theorem 2.1 *Let $G \sim G(n, p)$. Then*

$$\lim_{n\to\infty}\mathbb{P}\left(G\,is connected\right) = \begin{cases} 1 & np = \log n + w(n), \\ 0 & np = \log n - w(n), \end{cases}$$

where $w(n)$ is any function satisfying $w(n) \to \infty$. In addition, if $np = \log n + c$ for $c \in (0, \infty)$, and denoting by N_{comp} the number of connected components in G, then

$$(N_{\text{comp}} - 1) \xrightarrow{D} \text{Pois}\left(e^{-c}\right),$$

where $\xrightarrow{D}$ denotes convergence in distribution. This convergence implies that

$$\lim_{n\to\infty} \mathbb{P}\left(G \text{ is connected}\right) = e^{-e^{-c}}.$$

The main idea behind the proof of Theorem 2.1 is to show that around $np = \log n$, the random graph consists only of a giant connected component and isolated vertices. In that case, the phase transition for connectivity can be achieved by analysing the number of isolated vertices.

Note that in the language of homology, the event 'G is connected' can be phrased as 'the 0th homology group $H_0(G)$ is trivial', and that N_{comp} is equal to the 0th Betti number $\beta_0(G)$. Thus, it is tempting to try to generalize this phase transition for higher degrees of homology, and search for the point (value of p or np) where the kth homology group H_k becomes trivial.

2.3.1.1 The Random d-complex

We start with the Linial-Meshulam random d-complex $Y_d(n, p)$. Recall that $G(n, p) = Y_1(n, p)$, and that Theorem 2.1 studies H_0. Similarly, the following results studies H_{d-1} in $Y_d(n, p)$.

Theorem 2.2 *Let $Y \sim Y_d(n, p)$. Then*

$$\lim_{n\to\infty} \mathbb{P}\left(H_{d-1}(Y) = 0\right) = \begin{cases} 1 & np = d\log n + w(n), \\ 0 & np = d\log n - w(n), \end{cases}$$

where $w(n) \to \infty$.

In addition, if $np = d\log n + c$, $c \in (0, \infty)$, then

$$\beta_{d-1}(Y) \xrightarrow{D} \text{Pois}\left(\frac{e^{-c}}{d!}\right),$$

which implies that

$$\lim_{n\to\infty} \mathbb{P}\left(H_{d-1}(Y) = 0\right) = e^{-\frac{e^{-c}}{d!}}.$$

Note that taking $d = 1$ in Theorem 2.2, nicely recovers the graph result in Theorem 2.1. Thus, Linial and Meshulam named this phase transition 'homological connectivity'. The phase transition itself was proved first for $d = 2$ and for homol-

ogy in $\mathbb{Z}_2$ coefficients in [1], and then for any d and homology in $\mathbb{Z}_m$ coefficients in [29]. The Poisson limit was proved a decade later [30].

Aside from the analogy between the statements of Theorems 2.2 and 2.1, the general idea behind the proof also shares some similarity. To prove Theorem 2.2 one can show that around $np = d \log n$, the only possible cycles in H_{d-1}[1] are those generated by isolated $(d-1)$-simplexes. An 'isolated' $(d-1)$-simplex is such that it is not included in any d-dimensional simplex. The fact that isolated $(d-1)$-simplexes yield nontrivial H_{d-1} is relatively easy to prove. The much harder part here is to show that these are the *only* possible cycles.

2.3.1.2 The Random Flag Complex

The random flag complex $X(n, p)$ differs from the random d-complex $Y_d(n, p)$ in two important ways. Firstly, $X(n, p)$ has random homology in all possible degrees k, rather than just $k = d$ and $k = d-1$ as in $Y_d(n, p)$. Secondly, note that in $Y_d(n, p)$ both H_{d-1} and H_d are *monotone*—H_{d-1} is decreasing in p, while H_d is increasing. This is not the case for the flag complex, where except for H_0, none of the homology groups is monotone. The following result was proved by Kahle [33].

Theorem 2.3 *Let $X \sim X(n, p)$. Then,*

$$\lim_{n\to\infty} \mathbb{P}\left(H_k(X) = 0\right) = \begin{cases} 1 & np^{k+1} = \left(\frac{k}{2} + 1 + \epsilon\right) \log n, \\ 0 & np^{k+1} = \left(\frac{k}{2} + 1 - \epsilon\right) \log n. \end{cases}$$

Note that here as well, taking $k = 0$ agrees with the phase transition in Theorem 2.1. The phase transition here is also a consequence of the vanishing of the isolated k-simplexes. However, as oppose to the proofs in [1, 29] which mainly consist of combinatorial arguments, the proof in [33] goes in a different way, employing *Garland's method*. Briefly, Garland's method [69] is a powerful tool that allows "breaking" the computation of homology into local pieces (the *links* of the faces), and to invoke spectral graph arguments.

2.3.1.3 The Multi-parameter Random Complex

Phase transitions are usually described as a rapid change of system behavior, in response to a small change in a system parameter value (e.g. p or r above). Since the multi-parameter complex $X(n, \mathbf{p})$ has an infinite sequence of parameters, we cannot describe a suitable phase transition *per se*. Nevertheless, interesting results were presented in [70], showing that there are different regions within this high-dimensional parameter space, where in each region there is a single dominant homology group H_k.

[1] More accurately, the proof actually looks at cocycles in H^{d-1}.

To describe the result we need a few definitions. Let $d > 0$ and let $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_+$. Define

$$\psi_k(\alpha) = \sum_{i=1}^{d} \binom{k}{i} \alpha_i, \quad k = 1, \ldots, d.$$

These functions are then used to define different domains

$$\mathcal{D}_k = \left\{ \alpha \in \mathbb{R}^d_+ : \psi_k(\alpha) < 1 < \psi_{k+1}(\alpha) \right\},$$

for $k = 0, \ldots, d-1$, and where $\psi_0 \equiv 0$. We also define $\mathcal{D}_d = \{\alpha : \psi_d(\alpha) < 1\}$. Finally, define

$$\tau_k(\alpha) = \sum_{i=1}^{k} (1 - \psi_i(\alpha)), \qquad e(\alpha) = \min_{0 \le i \le d} (1 - \psi_i(\alpha)).$$

Theorem 2.4 *([70]) Let $\mathbf{p} = (p_1, p_2, \ldots, p_d, 0, 0, \ldots)$, be such that $p_i = n^{-\alpha_i}$, where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ may be a function of n, and $\lim_{n\to\infty} \alpha(n) = \alpha_*$. Let $X \sim X(n, \mathbf{p})$, and suppose that $\alpha_* \in \mathcal{D}_k$. Then, with high probability*

$$\beta_k(X) \approx n^{\tau_k(\alpha)}(k+1)!,$$

and for all $j \neq k$ we have

$$\beta_j(X) = O(n^{-e(\alpha)} \beta_k(X)).$$

In other words, in each $\mathcal{D}_k$ the k-th Betti numbers is quite large, and all other Betti numbers are negligible compared to it. One can think of this results as a "multi-parameter phase transition" so that the system behavior changes as one moves from one region $\mathcal{D}_k$ to another. The proofs in [37] relies mainly on counting faces and using Morse inequalities.

2.3.1.4 Random Geometric Complexes

With respect to connectivity, the random geometric graph $G(\mathcal{P}_n, r)$ behaves quite similarly to the $G(\mathcal{P}_n, p)$ random graph. Denoting by $\Lambda = \omega_d n r^d$ the expected degree equal to the expected number of points in a ball of radius r, with ω_d standing for the volume of the unit-ball in $\mathbb{R}^d$, the following was shown in [71].

Theorem 2.5 *Let $G \sim G(\mathcal{P}_n, r)$. Then*

$$\lim_{n\to\infty} \mathbb{P}(G \, is connected) = \begin{cases} 1 & \Lambda = \log n + w(n), \\ 0 & \Lambda = \log n - w(n), \end{cases}$$

where $w(n)$ is any function satisfying $w(n) \to \infty$.

In other words, in both models connectivity is achieved once the expected degree is larger than $\log n$.

Moving to higher dimensions, the geometric models start to exhibit different behavior than the combinatorial ones. The main difference is the following. In $Y_d(n, p)$ and $X(n, p)$, when $p \to 1$, homological connectivity describes the stage where homology becomes trivial. This is due to the fact that there is no structure underlying the complex. In the geometric complexes, the vertices are sampled over a metric space S, which might has its own intrinsic homology. Thus, for r large enough, one should expect the homology of the random complex ($C(\mathcal{P}_n, r)$ or $R(\mathcal{P}_n, r)$) to "converge" to the homology of S, rather than to vanish. To account for this, the event we will refer to as the k-th homological connectivity here, is defined in [72] as

$$\mathcal{H}_{k,r} := \{\forall s \geq r : H_k(C(\mathcal{P}_n, s)) \cong H_k(S)\}$$

for the Čech complex, and similarly for the Rips complex, with 'C' replaced by 'R'. Note that $\mathcal{H}_{k,r}$ is a monotone event (i.e. $\mathcal{H}_{k,r_1} \subset \mathcal{H}_{k,r_2}$ for all $r_1 \leq r_2$).

In the Čech complex, the phase transition in $\mathcal{H}_{k,r}$ is by now fully understood. The following was proved in [72].

Theorem 2.6 *Let $S = \mathbb{T}^d$, then for $1 \leq k \leq d-2$,*

$$\lim_{n\to\infty} \mathbb{P}\left(\mathcal{H}_{k,r}\right) = \begin{cases} 1 & \Lambda = 2^d(\log n + (k-1)\log\log n + w(n)), \\ 0 & \Lambda = 2^d(\log n + (k-1)\log\log n - w(n)), \end{cases}$$

for any $w(n) \to \infty$. In addition,

$$\lim_{n\to\infty} \mathbb{P}\left(\mathcal{H}_{d-1,r}\right) = \lim_{n\to\infty} \mathbb{P}\left(\mathcal{H}_{d,r}\right) = \begin{cases} 1 & \Lambda = 2^d(\log n + (d-1)\log\log n + w(n)), \\ 0 & \Lambda = 2^d(\log n + (d-1)\log\log n - w(n)). \end{cases}$$

While the proof there considers only the flat torus case, the results in [73] indicate that similar results could be achieved for any compact Riemannian manifold. The key idea behind the proof of Theorem 2.6 was to consider the evolution of complex $C(\mathcal{P}_n, r)$ as r is increased, and to look for *critical faces*. By that we mean simplexes that facilitate changes in $H_k(C(\mathcal{P}_n, r))$ when they first enter the complex. The analysis of critical faces employs Morse theory [74, 75]. The proof then consists of the following arguments:

1. If Λ is large enough ($\Lambda \gg \log n$), we have coverage, i.e. $B(\mathcal{P}_n, r) = \mathbb{T}^d$. Then, by the Nerve Lemma [65] we have that $H_k(C(\mathcal{P}_n, r)) \cong H_k(\mathbb{T}^d)$.
2. If $\Lambda = 2^d(\log n + (k-1)\log\log n + w(n))$ we can show that w.h.p. no more critical faces that modify H_k enter the complex. In other words, H_k has reached its limit. From the previous argument the limit is $H_k(\mathbb{T}^d)$, and thus we conclude that $\mathcal{H}_{k,r}$ holds.

3. If $\Lambda = 2^d(\log n + (k-1)\log\log n - w(n))$, critical faces still enter the complex and make chances in H_k. Thus $\mathcal{H}_{k,r}$ does not hold (w.h.p.).

Similarly to Theorems 2.1 and 2.2, the results in [72] also include a Poisson limit for the counting of the critical faces, when $\Lambda = 2^d(\log n + (k-1)\log\log n + c)$. This result is the analogue of the isolated vertices in the random graph, since: (a) Critical faces are the 'obstructions' to homological connectivity, in the same way that isolated vertices are for connectivity. (b) It can be shown that critical faces are indeed isolated when they first enter the complex. Since the exact definition of the critical faces require more details, we will not present the exact theorem here, but refer the reader to [72, 76].

Homological connectivity for the Rips complex $R(\mathcal{P}_n, r)$ is still an open problem to date. The most recent result in this direction appeared in [77]. By using Discrete Morse Theory [78], Kahle was able to prove the following.

Theorem 2.7 *Let S be a compact and convex subset of $\mathbb{R}^d$. Then*

$$\mathbb{E}\left\{\beta_k(R(\mathcal{P}_n, r))\right\} = O\left(n\Lambda^k e^{-c\Lambda}\right).$$

In particular, if $\Lambda = \frac{1+\epsilon}{c}\log n$ then $\beta_k = 0$ (w.h.p.).

Note that this result is for compact and convex sets S, for which $H_k(S) = 0$ for all $k \geq 1$. While this result is a significant step, it is still incomplete since: (a) It does not provide a sharp phase transition. In particular, there is no reason to believe that the constant c provided by the proof, is the optimal one. (b) It does not apply to general manifolds.

2.3.2 Emergence of Homology

Another fundamental result proved from random graph concerns the appearance of cycles. The following was proved in [47] (for the $G(n, M)$ model, but the statement for $G(n, p)$ is equivalent).

Theorem 2.8 *Let $G \sim G(n, p)$, then*

$$\lim_{n\to\infty} \mathbb{P}\left(G\, contains cycles\right) = \begin{cases} 1 & np = c \geq 1, \\ \gamma(c) & np = c \in (0,1), \\ 0 & np = o(1), \end{cases}$$

where $\gamma(c) = 1 - \sqrt{1-c}\exp\left(\frac{c}{2} + \frac{c^2}{4}\right)$.

Note that the event 'G contains cycles' can be phrased as $H_1(G) \neq 0$. In other words, Theorem 2.8 describes a phase transition for the emergence of H_1 in $G(n, p) = X_1(n, p)$.

In the following, we will review the results known to date about the emergence of k-cycles for the various random simplicial complex models.

2.3.2.1 The Random d-complex

For the random d-complex, the following is an aggregate of a collection of works.

Theorem 2.9 *Let $Y \sim Y_d(n, p)$. There exists $c_d^* > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left(H_d(Y) \neq 0\right) = \begin{cases} 1 & np = c > c_d^*, \\ \gamma_d(c) & np = c \in (0, c_d^*), \\ 0 & np = o(1), \end{cases}$$

where $\gamma_d(c) = 1 - \exp\left(-\frac{c^{d+2}}{(d+2)!}\right)$. In addition, the only d-cycles in $c \in (0, c_d^*)$ are generated by empty $(d+1)$-shells.

The first case was proved in [79], the middle case was proved in [80], and the last case was proved in [81]. An equation defining the critical values can be found in [79], while numerical approximations [80] yield $c_2^* \approx 2.754$, $c_3^* \approx 3.907$, $c_4^* \approx 4.962, \ldots, c_{1000}^* \approx 1001$. So roughly, it looks like $c_d^* \approx d + 1$.

Another closely related phase transition is for collapsibility. Briefly, a d-complex is collapsible, if we can iteratively erase pairs of simplexes in dimension d and $d-1$, without changing the topology of the complex. In [82] a phase transition for collapsibility was shown at critical values $c_d^{\text{col}} < c_d^*$.

2.3.2.2 The Random Flag Complex

Recall that the homology $H_k(X(n, p))$ is non-monotone (in p). In Sect. 2.3.1.2 we saw that the largest value of p such that $H_k(X(n, p))$ is non-trivial is when $np^{k+1} = \Theta(\log n)$. In this section we look for the smallest possible value of p.

Theorem 2.10 *Let $X \sim X(n, p)$, and suppose that $\lim_{n\to\infty} np^k < \infty$. Then,*

$$\lim_{n\to\infty} \mathbb{P}\left(H_k(X) \neq 0\right) = \begin{cases} 1 & np^k \geq k + 1 + \epsilon, \\ 0 & np^k \leq n^{-\epsilon}, \end{cases}$$

for any $\epsilon > 0$.

The upper case was proved in [33], while the lower case was proved in [32]. While this result is not as sharp as the previous ones, it still implies a phase transition when $np^k = \Theta(1)$. Together with Theorem 2.3, we have that $H_k(X) \neq 0$ for p in the interval

$$\left[\left(\frac{k+1+\epsilon}{n}\right)^{\frac{1}{k}}, \left(\frac{\left(\frac{k}{2}+1-\epsilon\right)\log n}{n}\right)^{\frac{1}{k+1}}\right].$$

2.3.2.3 Random Geometric Complexes

The behavior of random geometric complexes is similar in many ways to that of the random flag complex. In particular, for each k we have two phase transitions – one for the emergence of the first k-cycles, and another for homological connectivity (Sect. 2.3.1.4). Here we present the former phase transition.

For the Čech complex, combining the results from [77, 83], we have the following.

Theorem 2.11 *For $1 \leq k \leq d-1$,*

$$\lim_{n\to\infty} \mathbb{P}\left(H_k(C_r(\mathcal{P}_n)) \neq 0\right) = \begin{cases} 1 & n\Lambda^{k+1} = \omega(1), \\ \gamma_k(c) & n\Lambda^{k+1} = c \in (0,\infty), \\ 0 & n\Lambda^{k+1} = o(1), \end{cases}$$

where $\gamma_k(c) = 1 - \exp(-cA_k)$, for some constant $A_k > 0$.

Note that the transition occurs when $\Lambda = \Theta(n^{-\frac{1}{k+1}}) \to 0$. In other words, the first cycles appear in a sparse regime where the expected degree is small. Consequently, it was shown in [77] that in this regime the only k-cycles that appear are the smallest possible ones – i.e., $(k+1)$-shells, which consist of $(k+2)$ vertices. The phase transition in Theorem 2.11 then essentially describes the appearance of these empty $(k+1)$-shells.

A similar result was proved for the Rips complex.

Theorem 2.12 *For any $k \geq 1$,*

$$\lim_{n\to\infty} \mathbb{P}\left(H_k(R_r(\mathcal{P}_n)) \neq 0\right) = \begin{cases} 1 & n\Lambda^{2k+1} = \omega(1), \\ \tilde{\gamma}_k(c) & n\Lambda^{2k+1} = c \in (0,\infty), \\ 0 & n\Lambda^{2k+1} = o(1), \end{cases}$$

where $\tilde{\gamma}_k(c) = 1 - \exp(-c\tilde{A}_k)$, for some constant $\tilde{A}_k > 0$.

There are two main differences between Theorems 2.11 and 2.12. Firstly, due to the Nerve Lemma [65], the Čech complex on $\mathbb{T}^d$ (as well as any subset of $\mathbb{R}^d$) cannot have any k-cycles for $k > d$. The Rips complex, on the other hand, may have k-cycles for any k. Secondly, the transition here occurs when $\Lambda = \Theta\left(n^{-\frac{1}{2k+1}}\right)$, as opposed to $\Lambda = \Theta\left(n^{-\frac{1}{k+1}}\right)$ in Theorem 2.11. This difference stems from the fact that empty shells cannot appear in the Rips complex. Thus, it can be shown that the smallest

possible k-cycle is achieved by generating an empty cross-polytope, which consists of $2k + 2$ vertices.

2.3.3 Percolation-Related Phenomena

Percolation theory studies the emergence of infinite or "giant" connected component in various random media. The study of higher-dimensional analogues is at a preliminary stage, and we describe here the most recent results.

In the case of random graphs, giant components are commonly defined in terms of the number of vertices. Specifically, for both the $G(n, p)$ and $G(\mathcal{P}_n, r)$ models, by a "giant component" we refer to a component that consists of $\Theta(n)$ many vertices. As opposed to connectivity which we discussed earlier, this notion does not have a simple analogue in higher dimensions. In this section we present two different notions – one of the random d-complex and another for the random Čech complex.

2.3.3.1 Random d-complex

Starting with the $G(n, p)$ graph, the following was proved in [47].

Theorem 2.13 *Let $G \sim G(n, p)$, and $np = c \in (0, \infty)$. Denote by L_n the largest connected component in G. Then with high probability*

$$L_n = \begin{cases} \Theta(n) & c > 1, \\ \Theta(n^{2/3}) & c = 1, \\ \Theta(\log n) & c < 1. \end{cases}$$

In other words, a giant component appears for $c > 1$. Further, it can be shown that in this case, the giant component is unique.

In order to generalize the emergence of the giant component to higher dimension, the notion of *shadow* was introduced in [84]. Suppose that X is a d-dimensional simplicial complex with a complete $(d-1)$-skeleton. The shadow of X, denoted $\mathrm{SH}(X)$, is the set of all d-dimensional faces σ such that (a) $\sigma \notin X$, and (b) adding σ to X generates a new d-cycle.

In the case of the graph (i.e., $d = 1$) the shadow are all the edges not in the graph, such that adding them generates a new cycle. In that case, it is easy to show that when $c > 1$ we have $|\mathrm{SH}(X)| = \Theta\left(\binom{n}{2}\right)$, so that the shadow has a *positive density*, and we say that the graph has a *giant shadow*. In [80], this phenomenon was generalized for higher dimensions.

Theorem 2.14 *Let $Y \sim Y_d(n, p)$, and $np = c \in (0, \infty)$. Then with high probability,*

$$|\mathrm{SH(X)}| = \begin{cases} \Theta\left(\binom{n}{d+1}\right) & c > c_d^*, \\ \Theta(n) & c < c_d^*, \end{cases}$$

where c_d^* is the critical value defined in Sect. 2.3.2.

In other words, when $c > c_d^*$ a giant shadow emerges in the random d-complex, similarly to what happens in the graph case.

2.3.3.2 Random Geometric Complexes

Similarly to the $G(n, p)$ model, the random geometric graph $G(\mathcal{P}_n, r)$ exhibits a phase transition for the emergence of a giant component. This phase transition occurs when the expected degree Λ is constant, and the corresponding critical value is denoted as λ_c.

In the context of geometric complexes, a different generalization for percolation was studied recently [85, 86]. Consider the random Čech complex $C(\mathcal{P}_n, r)$ defined in Sect. 2.2, where $\mathcal{P}_n$ is a homogeneous Poisson process on the d-torus $\mathbb{T}^d$. One of the nice facts about the torus, is that it has non-trivial homology for all $0 \leq k \leq d$, and in particular $\beta_k(\mathbb{T}^d) = \binom{d}{k}$. In [85] the notion of a "giant k-cycle" was introduced, by which we mean any k-cycle in $C(\mathcal{P}_n, r)$ that corresponds to any of the nontrivial cycles in $H_k(\mathbb{T}^d)$.

To define this more rigorously, we consider the union of balls $B(\mathcal{P}_n, r)$, and the fact that $H_k(B(\mathcal{P}_n, r)) \cong H_k(C(\mathcal{P}_n, r))$. We then take the inclusion map $i : B(\mathcal{P}_n, r) \hookrightarrow \mathbb{T}^d$, and its induced map $i_* : H_k(B(\mathcal{P}_n, r)) \to H_k(\mathbb{T}^d)$. By a *giant k-cycle* we refer to all nontrivial elements in the image $\mathrm{Im}(i_*)$. Finally, we define the events

$$\mathcal{E}_k := \{\mathrm{Im}(i_*) \neq 0\}, \quad \mathcal{A}_k = \left\{\mathrm{Im}(i_*) = H_k(\mathbb{T}^d)\right\}.$$

In other words, $\mathcal{E}_k$ is the event that some giant cycles exist, while $\mathcal{A}_k$ is the event that all of them are present in $B(\mathcal{P}_n, r)$ (and correspondingly in $C(\mathcal{P}_n, r)$). The following was proved in [85].

Theorem 2.15 *Suppose that $\Lambda = \omega_d n r^d = \lambda \in (0, \infty)$.*

1. *There exist $\lambda_1^0 \leq \lambda_2^0 \ldots \leq \lambda_{d-1}^0$, such that if $\lambda < \lambda_k^0$ then,*

$$\mathbb{P}(\mathcal{A}_k) \leq \mathbb{P}(\mathcal{E}_k) \leq e^{-C_k^0 n^{1/d}},$$

   for some $C_k^0 > 0$.
2. *There exist $\lambda_1^1 \leq \lambda_2^1 \ldots \leq \lambda_{d-1}^1$, such that if $\lambda > \lambda_k^1$ then,*

$$\mathbb{P}(\mathcal{E}_k) \geq \mathbb{P}(\mathcal{A}_k) \geq 1 - e^{-C_k^1 n^{1/d}},$$

   for some $C_k^1 > 0$.

In addition, we have $\lambda_k^0 \leq \lambda_k^1$ *for all* k, *and* $\lambda_1^0 = \lambda_1^1 = \lambda_c$.

In other words, Theorem 2.15 suggests that there is a phase transition describing the emergence of the giant k-cycles. However in order to complete the proof, one needs to show that $\lambda_k^0 = \lambda_k^1$ for all k.

At this stage the giant shadow and the giant cycle phenomena are incomparable. However, it should be noted that both occur in the regime where the expected degree (np or Λ) is finite, similarly to the giant component. It is an interesting question whether these are merely two ways to view the same phenomenon, or they describe completely different structures. Finally, we should note that other ideas for extending percolation-type phenomena to higher dimension exist in the literature.

2.3.4 The Fundamental Group

Given a space X, the fundamental group $\pi_1(X)$ represents the space of equivalence classes of loops in X, where the equivalence is based on homotopy (a continuous transformation of one loop into the other). The space X is simply connected if and only if $\pi_1(X)$ is trivial.

Note that the first homology group $H_1(X)$ also represents loops, but under a different equivalence relation (being a boundary). In fact, it can be shown that $H_1(X)$ is an abelianization of $\pi_1(X)$. Thus, it is possible that $H_1(X)$ is trivial while $\pi_1(X)$ is not. Consequently, the vanishing threshold for π_1 is larger than the threshold for H_1 as the following statements show.

We start with the random d-complex. Here the only relevant dimension is $d = 2$, since otherwise the fundamental group is trivial. The following was proved in [87, 88].

Theorem 2.16 *Let* $Y \sim Y_2(n, p)$. *Then,*

$$\lim_{n\to\infty} \mathbb{P}\left(\pi_1(Y) = 0\right) = \begin{cases} 1 & np = cn^{\frac{1}{2}}, \\ 0 & np = n^{\frac{1}{2}-\epsilon}, \end{cases}$$

where c *is any constant bigger than* $\sqrt{3^3/4^4}$.

Further, when $np = n^{1/2-\epsilon}$ it is shown in [87] that the fundamental group is hyperbolic. It is also conjectured in [88] that in fact $c = \sqrt{3^3/4^4}$ is the exact sharp threshold for simple connectivity.

For the random flag complex, the existing results [89] are even less sharp.

Theorem 2.17 *Let* $X \sim X(n, p)$, *then for any*

$$n^{\frac{1}{2}+\epsilon} \leq np \leq n^{\frac{2}{3}-\epsilon},$$

with high probability $\pi_1(X)$ *is nontrivial and hyperbolic.*

Note that the algebraic structure of the fundamental group is more intricate than $H_1(X)$, and thus several other properties have been studied in the literature. For example: its geometric and cohomological dimension [90], torsion [91], property T [33], freeness [92], and more.

2.4 Limit Theorems

The results in the previous section focused on a rather qualitative analysis of phase transitions. Another direction of study is to consider some numerical marginals of homology (e.g., the Betti numbers) and find their limiting distribution. In this section we will briefly review some of the results in this direction.

We will not discuss the proofs here, however we note that in most cases they rely on *Stein's method* [93, 94]. Briefly, this method allows one to prove central limit theorems and convergence to the normal and Poisson distributions in cases that involve sums of random variables that are *not* independent.

Notation. In this section we will use $a_n \approx b_n$ to denote that $a_n/b_n \to 1$ as $n \to \infty$, and $a_n \ll b_n$ to denote that $a_n/b_n \to 0$. In addition, $\xrightarrow{D}$ stands for convergence in distribution, and $N(\mu, \sigma^2)$ for the normal distribution with mean μ and variance σ^2.

2.4.1 Betti Numbers

As we have seen in Sect. 2.3.2, the k-th homology of $X(n, p)$ is nontrivial roughly when p is between $n^{-1/k}$ and $n^{-1/(k+1)}$ (up to constants). It was shown in [32] that when $n^{-1/k} \ll p \ll n^{-1/(k+1)}$ then

$$\mathbb{E}\{\beta_k(X(n,p))\} \approx \frac{n^k}{(k+1)!} p^{\binom{k+1}{2}}.$$

In addition, the following central limit theorem was proved in [83].

Theorem 2.18 *Let $X \in X(n, p)$, and let $n^{-1/k} \ll p \ll n^{-1/(k+1)}$. Then*

$$\frac{\beta_k(X) - \mathbb{E}\{\beta_k(X)\}}{\sqrt{\mathrm{Var}\,(\beta_k(X))}} \xrightarrow{D} N(0,1).$$

In the geometric complexes case, the results usually split according to the expected degree Λ. In the case where $\Lambda \to 0$ (sometimes called the sparse regime), the homology is dominated by empty shells, and thus it is relatively easy to compute the Betti numbers. For example, it was proved in [77] that

$$\mathbb{E}\{\beta_k(C(\mathcal{P}_n, r))\} \approx A_k n \Lambda^{k+1},$$

where $A_k > 0$ is a known constant [77]. Note that the limit of the right hand side can be zero, finite, or infinite. Consequently we have three possible limits. The following was proved in [83].

Theorem 2.19 *Let $C \sim C(\mathcal{P}_n, r)$, and suppose that $\Lambda \to 0$.*

1. *If $n\Lambda^{k+1} \to 0$ then $\beta_k(C) = 0$ with high probability.*
2. *If $n\Lambda^{k+1} = \alpha \in (0, \infty)$ then $\beta_k(C) \xrightarrow{D} \mathrm{Pois}(A_k\alpha)$.*
3. *If $n\Lambda^{k+1} \to \infty$ then $\frac{\beta_k(C) - \mathbb{E}\{\beta_k(C)\}}{\sqrt{\mathrm{Var}(\beta_k(C))}} \xrightarrow{D} N(0,1)$.*

Similar results hold for the Rips complex as well, just with $n\Lambda^{2k+1}$ replacing $n\Lambda^{k+1}$ [83].

The regime where Λ is finite and non-vanishing is sometimes referred to as the *thermodynamic limit*. It can be shown that most of the cycles in either the Čech or Rips complexes are generated in this regime. However, exact counting here is much more difficult. It was shown in [77] that $\mathbb{E}\{\beta_k(C(\mathcal{P}_n, r))\} = \Theta(n)$ (and the same for the Rips). However the exact expectation is not known.

Nevertheless, one can prove laws of large numbers as well as central limit theorems for the Betti number functionals. This type of results was first introduced in [95], and has been improved over the years [96–98]. The main tools used are stabilization methods for random point processes (e.g. [99]). The result is as follows.

Theorem 2.20 *Let $C \sim C(\mathcal{P}_n, r)$, and suppose that $\Lambda = \lambda \in (0, \infty)$. Then*

$$\frac{\beta_k(C) - \mathbb{E}\{\beta_k(C)\}}{\sqrt{n}} \xrightarrow{D} N(0, \sigma^2(\lambda)),$$

for some $\sigma^2(\lambda) > 0$.

A similar result applies to the Rips complex as well. In addition in [98, 100] the binomial process $\mathcal{X}_n$ was studied as well, and similar limiting theorems were proved.

In the dense regime, i.e., $\Lambda \to \infty$ there are very few results. In particular, even the scale of the expected Betti numbers is not known. The only case analyzed is the Poisson limit when $\Lambda \approx \log n$ described earlier, as part of the homological connectivity phenomenon.

2.4.2 Topological Types

In [101] a completely different type of limit theorem was proved. There, the goal was to study the distribution of the different topological types that may appear in a random Čech complex. Let X be a simplicial complex, and let $\mathcal{C}(X)$ be the set of all connected components of X (so that $\beta_0(X) = |\mathcal{C}(X)|$). Define the empirical measure

$$\mu_X := \frac{1}{\beta_0(X)} \sum_{C \in \mathcal{C}} \delta_{[C]},$$

where δ stands for the Dirac delta measure, and $[C]$ stands for the equivalence class of all components homotopy equivalent to C (i.e. so that one component can be "continuously transformed" into the other, see [102]).

Theorem 2.21 *Let $C \sim C(\mathcal{P}_n, r)$, and denote $\hat{\mu}_n := \mu_C$. Then the random measure $\hat{\mu}_n$ converges in probability to a universal probability measure. The support of the distribution is the set of all connected (finite) Čech complexes in $\mathbb{R}^d$.*

Note that universality here means that the limit is the same regardless of the underlying manifold S.

2.4.3 Persistent Homology

In the context of Topological Data Analysis (TDA), which is discussed in detail in Chap. 3, one is often more interested in studying *persistent* homology rather than just the (static) homology. In this section we consider the persistent homology over filtrations of either Čech or Rips complexes, namely $\{C(\mathcal{P}_n, r)\}_{r=0}^{\infty}$ or $\{R(\mathcal{P}_n, r)\}_{r=0}^{\infty}$, respectively.

2.4.3.1 Limit Theorems

One useful quantity that can be extracted from persistent homology are the *persistent Betti numbers*. Briefly, for any $s \leq t$, we denote by $\beta_k^{(s,t)}$ the number of cycles born before radius s and die after radius r (see [96] for a formal definition). In [96] the following central limit theorem was proved.

Theorem 2.22 *Let $C \sim C(\mathcal{P}_n, r)$, and suppose that $ns^d = \alpha$ and $nt^d = \beta$, for some $\alpha \leq \beta \in (0, \infty)$. Then*

$$\frac{\beta_k^{(s,t)}(C) - \mathbb{E}\left\{\beta_k^{(s,t)}(C)\right\}}{\sqrt{n}} \xrightarrow{D} N(0, \sigma^2(\alpha, \beta)),$$

for some $\sigma^2(\alpha, \beta) > 0$.

Note that this theorem describes the persistent Betti numbers in the thermodynamic limit, where most cycles are born and die. A similar theorem applies for $R(\mathcal{P}_n, r)$, and in fact [96] present this central limit theorem for an even larger general class of geometric complexes. In [97] Theorem 2.22 was further extended to a multidimensional central limit theorem (i.e. for a sequence $(\alpha_1, \beta_1), \ldots, (\alpha_m, \beta_m)$). Finally, note that Theorem 2.20 is in fact a special case of Theorem 2.22 in the case where $\alpha = \beta$.

Recall that a common representation for persistent homology is via a *persistence diagram*, which is essentially a finite subset of $\Delta = \{(x, y) : y \geq x\} \subset \mathbb{R}^2$, where the coordinates of each point correspond to the birth and death times of a persistence

interval. Thus, persistence diagrams attached to random finite simplicial complexes, can be thought of as a random subset of Δ, or equivalently a random discrete measure on Δ. One can then employ the theory of random sets and random measures to study the limit of such diagrams.

Consider the random Čech filtration $\{C(\mathcal{P}_n, r)\}_{r=0}^{\infty}$, and let $\mu_k(n)$ be the discrete measure representing the persistence diagram in degree k. The main result in [96] shows that there exists a limiting Radon measure on Δ,

$$\mu_k = \lim_{n\to\infty} \frac{\mu_k(n)}{n}.$$

The same is true for the Rips filtration as well. This statement is essentially a law of large numbers for the persistence diagram as a Radon measure in $\mathbb{R}^2$.

A different approach was taken in [103], where persistence diagrams were studied from the point of view of set-topology. In that case it was shown that each persistence diagram can be roughly split into three different areas. The limit of the persistence diagram in the *sparse* region is empty. In the *intermediate* region it has a limiting Poisson process distribution. Finally, in the dense region, the persistence diagram converges to a solid (nonrandom) 2d region.

2.4.3.2 Maximal Cycles

Another interesting and significant type of limit for persistence diagrams was presented in [104]. There, it was assumed that S is a unit box (though the results should not depend on that). The quantity considered was the maximal death/birth ratio among all persistent cycles of degree k, denoted Π_k.

Theorem 2.23 *Let Π_k be the maximally persistent cycle for either the Rips or Čech complex. Then there exist $A_k < B_k$ such that with high probability*

$$A_k \left(\frac{\log n}{\log\log n}\right)^{1/k} \leq \Pi_k \leq B_k \left(\frac{\log n}{\log\log n}\right)^{1/k}.$$

Theorem 2.23 provides the scaling for the multiplicative persistence value in a random geometric complex. Since the unit box has no intrinsic homology, this result can be thought of as describing the homology of *noisy* cycles, i.e., cycles that are not part of the underlying topology. It can be shown that the same result applies for other compact spaces S, as long as we do not count cycle that belong to $H_k(S)$. Recalling Sect. 2.3.3.2, such "signal" cycles (those in $H_k(S)$) are born when Λ is constant, implying that the birth time is $r \propto n^{-1/d}$. On the other hand, the signal cycles die when r is a constant (depends on S and not on n). Therefore, for the signal cycles we have (death/birth) $\propto n^{1/d}$. In other words, taking death/birth as a measure of persistence, Theorem 2.23 shows that asymptotically noisy cycles are significantly less persistent than the signal ones, and one should be able to differentiate between them, assuming a large enough sample.

2.5 Other Directions

In Sects. 2.3 and 2.4 we presented some of the most fundamental results proved in the mathematical literature on random simplicial complexes. There are numerous important studies that we omit for space reasons, but would like to mention them briefly here.

- **Spectrum and expansion:** Similarly to graphs, one can define an adjacency and Laplacian operators for simplicial complexes, as well as different types of expansion properties. These were studied mainly in the the context of the random d-complex $Y_d(n, p)$ [48, 105–107].
- **Functional limit theorems:** One can examine functionals such as the Betti numbers and the Euler characteristic in a dynamic setting, and seek a limit in the form of a stochastic (Gaussian) process. In [103, 108], geometric complexes were studied, and the dynamics was the growing connectivity radius in the complex. The results show that the limiting process is indeed Gaussian. In [109], the random flag complex was studied. Here, the set of edges used to construct the complex turns on and off, following stationary Markov dynamics. The limiting processes are shown to be Ornstein-Uhlenbeck (Gaussian) processes. These results were further extended to the multiparameter complex [110].
- **Spanning acycles:** Similarly to the study of spanning trees in graphs, one can consider *spanning acycles* in simplicial complex. Suppose that K is a complete $(d-1)$-skeleton on n vertices. A set of d-faces S is considered a spanning acycle if $\beta_{d-1}(K \cup S) = \beta_d(K \cup S) = 0$. In other words, adding S to K kills all the existing $(d-1)$-cycles (hence "spanning"), while not generating any new d-cycles (hence "acycle"). This definition invites the study of spanning acycles in the random d-complex. In [111, 112] the weights of faces in the minimal spanning acycle and their connection to the persistence diagram were studied.

2.6 Future Directions

We hope our review illustrates well that the recent progress in the fundamental study of random simplicial complexes in mathematics has been extremely rapid. As far as applications to real-world systems in network/data science are concerned, the progress has been as rapid, but the field is definitely less mature that its 1-dimensional counterpart dealing with applications of random graph models. There, a collection of important and ubiquitous properties of real-world networks abstracted as graphs have been long identified and well studied. Such properties include sparsity, scale-free degree distributions, degree correlations, clustering, small-worldness, community structure, self-similarity, and so on [113, 114]. Consequently, there have been an impressive body of work on null models that reproduce these properties in the statistically unbiased manner, and on growing network models that attempt to shed

light on possible mechanisms that might lead to the formation of these properties in real-world systems [113, 114].

When the same systems are modeled as simplicial complexes or hypergraphs, the list of important properties one should be looking at, is much less understood. Consequently, the world of models dealing with such properties is much less explored, as it is not entirely clear what properties one should be most concerned with in the first place.

In addition to its relative infancy, the other reason why the world of random complex models has not yet been navigated as much, is that models of even simplest higher-dimensional properties tend to be much more complicated, not necessarily at the technical level as much as at the conceptual level.

Consider the null models of the degree distribution, i.e. the configuration models, for example. Since, as opposed to graphs, there is not one but $\binom{n}{2}$ notions of degree in a simplicial complex of size n (d-degrees of d'-simplexes for any pair of dimensions d, d', $0 \leq d' < d \leq n-1$), it may not be immediately clear what degrees to focus on. The degrees that the (S)CM$_d(n, \mathbf{k})$ and Z-(S)CM$_d(n, \mathbf{k})$ models reproduce in Sects. 2.2.3.2 and 2.2.5.2, are the d-degrees of $(d-1)$- and 0-dimensional simplexes, respectively, with very different constraints imposed on the underlying simplicial structure.

In that regard, these two models are the two simplest extremal versions of a great variety of configuration models we can think of. Such models can reproduce sequences of d-degrees of d'-simplexes for any pair of dimensions d, d', in which case the defining Eqs. (2.3, 2.4) remain the same, except that simplexes τ in those equations are now understood to be of dimension d'. We may also want to jointly reproduce different combinations of degree sequences of different dimensions d, d' satisfying structural constraints. Such models, based on generalized degree sequences observed in real data, could serve as more adequate and more realistic null models of the data. The first step in this direction has been made in [51].

Another example is homology, which we discussed at length in this chapter. Homology is known to be a powerful tool to capture information about the qualitative structure of simplicial complexes, and its theory in the random setting is constantly growing. However, as a tool to analyze high-dimensional networks, it is not yet as clear what part of the information provided by homology is useful and for what purposes. For example, while H_0 (connected components) and H_1 (cycles) are quite intuitive, the role of H_2 as a network descriptor is not as clear. Neither is it entirely clear what useful information is contained in the number of k-cycles, their physical size, their persistent lifetime, etc. So while the theoretical homological analysis is fruitful and highly interesting, it tends to be a challenge to see how and when to use it in application to real-world networks.

The degree distribution and homology are certainly just two out of many properties that the future work may find interesting and important to model in more realistic models of random simplicial complexes. For now, the list of such models is quite rudimentary for the reasons above, so that it is not surprising that rigorous mathematical results are available only for very basic models. Yet we have seen that even

for these simplest models, the spectrum of available results is extremely rich and complex.

We have also seen that these results mostly address direct problems, whereas in dealing with real-world data one often faces inverse problems. For example, the results presented in Sect. 2.4.3 tell us facts about the persistence diagrams in "laboratory-controlled" settings, i.e. for given complexes in given spaces with known topology—a direct problem. An inverse problem would be to make any statistical inferences about the space topology from a persistence diagram coming from real-world data. Some progress in this direction has been made over the past decade [115–119], but this is still an active and important area of research to pursue.

Finally, we comment on one potentially very interesting direction of future research in mathematics of random complexes. One of the most fundamentally important recent achievements in mathematics of random graphs is the development and essential completion of the theory of limits of dense graphs known as graphons [60, 61]. One of the main results in that theory is that, even though there are many very different notions of graph convergence, they all are equivalent for dense graphs, and if a family of dense graphs converge, they converge to a (random) graphon [120]. As mentioned in Sect. 2.2.4.1, a graphon is an integrable connection probability function $W : [0, 1]^2 \to [0, 1]$ modulo a certain equivalence relation, and a graphon-based random graph of size n—known as a W-random graph in the graphon theory—is $G \sim G(n, \hat{r})$, where $r_{ij} = W(X_i, X_j)$ and $X_i \sim U(0, 1)$.

Can an analogous theory be constructed—or discovered, depending on one's philosophical view—for the limits of dense complexes, e.g. *complexons*? Complexons may probably be related to *hypergraphons* [121–124], the limits of dense hypergraphs, via the lower and upper complexes $\underline{Z}(n, \hat{\mathbf{p}})$ and $\overline{Z}(n, \hat{\mathbf{p}})$ discussed in Sect. 2.2.5.1.

References

1. N. Linial, R. Meshulam, Homological connectivity of random 2-complexes. Combinatorica **26**, 475 (2006). https://doi.org/10.1007/s00493-006-0027-9
2. G. Bianconi, C. Rahmede, Complex quantum network manifolds in dimension $d>2$ are scale-free. Sci. Rep. **5**, 13979 (2015). https://doi.org/10.1038/srep13979
3. G. Bianconi, C. Rahmede, Z. Wu, Complex quantum network geometries: evolution and phase transitions. Phys. Rev. E. **92**, 022815 (2015). https://doi.org/10.1103/PhysRevE.92.022815
4. G. Bianconi, C. Rahmede, Network geometry with flavor: from complexity to quantum geometry. Phys. Rev. E. **93**, 53 (2016). https://doi.org/10.1103/PhysRevE.93.032315
5. O.T. Courtney, G. Bianconi, Weighted growing simplicial complexes. Phys. Rev. E. **95**, 062301 (2017). https://doi.org/10.1103/PhysRevE.95.062301
6. N. Fountoulakis, T. Iyer, C. Mailler, H. Sulzbach, *Dynamical Models for Random Simplicial Complexes* (2019). arXiv:1910.12715
7. K. Kovalenko, I. Sendiña-Nadal, N. Khalil, A. Dainiak, D. Musatov, A.M. Raigorodskii, K. Alfaro-Bittner, B. Barzel, S. Boccaletti, Growing scale-free simplices. Commun. Phys. **4**, 43 (2021). https://doi.org/10.1038/s42005-021-00538-y

8. G. Bianconi, I. Kryven, R.M. Ziff, Percolation on branching simplicial and cell complexes and its relation to interdependent percolation. Phys. Rev. E. **100**, 062311 (2019). https://doi.org/10.1103/PhysRevE.100.062311
9. I. Iacopini, G. Petri, A. Barrat, V. Latora, Simplicial models of social contagion. Nat. Commun. **10**, 2485 (2019). https://doi.org/10.1038/s41467-019-10431-6
10. M.T. Schaub, A.R. Benson, P. Horn, G. Lippner, A. Jadbabaie, Random walks on simplicial complexes and the normalized hodge 1-laplacian. SIAM Rev. **62**, 353 (2020). https://doi.org/10.1137/18M1201019
11. L.V. Gambuzza, F. Di Patti, L. Gallo, S. Lepri, M. Romance, R. Criado, M. Frasca, V. Latora, S. Boccaletti, Stability of synchronization in simplicial complexes. Nat. Commun. **12**, 1255 (2021). https://doi.org/10.1038/s41467-021-21486-9
12. Y. Lee, J. Lee, S.M. Oh, D. Lee, B. Kahng, Homological percolation transitions in growing simplicial complexes. Chaos. An Interdiscip. J. Nonlinear Sci. **31**(2021). https://doi.org/10.1063/5.0047608
13. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, G. Petri, Networks beyond pairwise interactions: structure and dynamics. Phys. Rep. **874**, 1 (2020). https://doi.org/10.1016/j.physrep.2020.05.004
14. E.T. Jaynes, Information theory and statistical mechanics. Phys. Rev. **106**, 620 (1957). https://doi.org/10.1103/PhysRev.106.620
15. C.E. Shannon, A mathematical theory of communication. Bell. Syst. Tech. J. **27**, 379 (1948). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x
16. J. Shore, R. Johnson, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. IEEE Trans. Inf. Theo. **26**, 26 (1980). https://doi.org/10.1109/TIT.1980.1056144
17. Y. Tikochinsky, N.Z. Tishby, R.D. Levine, Consistent inference of probabilities for reproducible experiments. Phys. Rev. Lett. **52**, 1357 (1984). https://doi.org/10.1103/PhysRevLett.52.1357
18. J. Skilling, The axioms of maximum entropy, in *booktitle Maximum-Entropy and Bayesian Methods in Science and Engineering* (Springer Netherlands, Dordrecht, 1988) pp. 173–187. https://doi.org/10.1007/978-94-009-3049-0_8
19. F. Rassoul-Agha, T. Seppäläinen, *A Course on Large Deviations with an Introduction to Gibbs Measures*. Series Graduate Studies in Mathematics, vol. 162 (American Mathematical Society, Providence, Rhode Island, 2015). https://doi.org/10.1090/gsm/162
20. S. Chatterjee, *Large Deviations for Random Graphs*. Series Lecture Notes in Mathematics, vol. 2197 (Springer International Publishing, Cham, 2017). https://doi.org/10.1007/978-3-319-65816-2
21. K. Zuev, O. Eisenberg, D. Krioukov, Exponential random simplicial complexes. J. Phys. A: Math. Theoret. **48**, 465002 (2015) http://iopscience.iop.org/article/10.1088/1751-8113/48/46/465002/meta
22. T.M. Cover, J.A. Thomas, *Elements of Information Theory* (Wiley, Hoboken, NJ, 2005). https://doi.org/10.1002/047174882X
23. R. Solomonoff, A. Rapoport, Connectivity of random nets. B Math. Biophys. **13**, 107 (1951). https://doi.org/10.1007/BF02478357
24. P. Erdős, A. Rényi, On random graphs. Publ. Math. **6**, 290 (1959)
25. E.N. Gilbert, Random graphs. Ann. Math. Stat **30**, 1141 (1959). https://doi.org/10.1214/aoms/1177706098
26. K. Anand, G. Bianconi, Entropy measures for networks: toward an information theory of complex topologies. Phys. Rev. E. **80**, 045102 (2009). https://doi.org/10.1103/PhysRevE.80.045102
27. T. Squartini, J. de Mol, F. den Hollander, D. Garlaschelli, Breaking of ensemble equivalence in networks. Phys. Rev. Lett. **115**, 268701 (2015). https://doi.org/10.1103/PhysRevLett.115.268701
28. S. Janson, Asymptotic equivalence and contiguity of some random graphs. Random Struct. Algor. **36**, 26 (2010). https://doi.org/10.1002/rsa.20297

29. R. Meshulam N. Wallach, Homological connectivity of random k-dimensional complexes. Random Struct. Algo. **34**, 408 (2009) . http://onlinelibrary.wiley.com/doi/10.1002/rsa.20238/abstract
30. M. Kahle, B. Pittel *Inside the critical window for cohomology of random k-complexes*. Random Struct. Alg. **48**, 102 (2016). http://onlinelibrary.wiley.com/doi/10.1002/rsa.20577/full
31. T. Luczak, Y. Peled, Integral homology of random simplicial complexes. Discrete and Comput. Geom. **59**, 131 (2018)
32. M. Kahle, Topology of random clique complexes. Discrete Mathe. **309**, 1658 (2009)
33. M. Kahle, Sharp vanishing thresholds for cohomology of random flag complexes. Ann. Mathe. **179**, 1085 (2014). https://doi.org/10.4007/annals.2014.179.3.5
34. M. Kahle, Topology of random simplicial complexes: a survey. AMS Contemp. Math. **620**, 201–221 (2014). https://doi.org/10.1090/conm/620/12367
35. A. Costa, M. Farber, in *Proceedings of the IMA Conference on Mathematics of Robotics*, Series and number number September (Institute of Mathematics and its Applications, 2015) pp. 1–8. https://doi.org/10.19124/ima.2015.001.01
36. A. Costa, M. Farber, Random Simplicial Complexes, in *Configuration Spaces Geometric Topology Represent Theory* (2016) pp. 129–153. https://doi.org/10.1007/978-3-319-31580-5_6
37. A. Costa, M. Farber, Large random simplicial complexes. I, J. Topol. Anal. **08**, 399 (2016). https://doi.org/10.1142/S179352531650014X
38. B. Söderberg, General formalism for inhomogeneous random graphs. Phys. Rev. E **66**, 066121 (2002). https://doi.org/10.1103/PhysRevE.66.066121
39. J. Park, M.E.J. Newman, Statistical mechanics of networks. Phys. Rev. E. **70**, 066117 (2004). https://doi.org/10.1103/PhysRevE.70.066117
40. T. Britton, M. Deijfen, A. Martin-Löf, Generating simple random graphs with prescribed degree distribution. J. Stat. Phys. **124**, 1377 (2006). https://doi.org/10.1007/s10955-006-9168-x
41. B. Bollobás, S. Janson, O. Riordan, The phase transition in inhomogeneous random graphs. Random Struct. Algor. **31**, 3 (2007). https://doi.org/10.1002/rsa.20168
42. G. Bianconi, The entropy of randomized network ensembles. EPL **81**, 28005 (2008). https://doi.org/10.1209/0295-5075/81/28005
43. D. Garlaschelli, M. Loffredo, Maximum likelihood: extracting unbiased information from complex networks. Phys. Rev. E. **78**(2008). https://doi.org/10.1103/PhysRevE.78.015101
44. S. Chatterjee, P. Diaconis, A. Sly, Random graphs with a given degree sequence. Ann. Appl. Probab. **21**, 1400 (2011). https://doi.org/10.1214/10-AAP728
45. E.A. Bender, E.R. Canfield, The asymptotic number of labeled graphs with given degree sequences. J. Comb. Theo. Ser. A **24**, 296 (1978). https://doi.org/10.1016/0097-3165(78)90059-6
46. M. Molloy, B. Reed, A critical point for random graphs with a given degree sequence. Random Struct. Algor. **6**, 161 (1995). https://doi.org/10.1002/rsa.3240060204
47. P. Erdős, A. Rényi, On the evolution of random graphs. Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960)
48. A. Lubotzky, Z. Luria, R. Rosenthal, Random Steiner systems and bounded degree coboundary expanders of every dimension. Dis. Comput. Geom. **62**, 813 (2019)
49. P. Keevash, The existence of designs. arXiv preprint arXiv:1401.3665 (2014)
50. O.T. Courtney, G. Bianconi, Generalized network structures: the configuration model and the canonical ensemble of simplicial complexes. Phys. Rev. E. **93**, 062311 (2016). https://doi.org/10.1103/PhysRevE.93.062311
51. J.-G. Young, G. Petri, F. Vaccarino, A. Patania, Construction of and efficient sampling from the simplicial configuration model. Phys. Rev. E. **96**, 032312 (2017). https://doi.org/10.1103/PhysRevE.96.032312
52. G. Caldarelli, A. Capocci, P. De Los Rios, M.A. Muñoz, Scale-free networks from varying vertex intrinsic fitness. Phys. Rev. Lett. **89**, 258702 (2002). https://doi.org/10.1103/PhysRevLett.89.258702

53. M. Boguñá, R. Pastor-Satorras, Class of correlated random networks with hidden variables. Phys. Rev. E. **68**, 036112 (2003). https://doi.org/10.1103/PhysRevE.68.036112
54. P.W. Holland, K.B. Laskey, S. Leinhardt, Stochastic blockmodels: First steps. Soc Networks **5**, 109 (1983). https://doi.org/10.1016/0378-8733(83)90021-7
55. A.P. Sorokin, *Social Mobility* (Harper, New York, 1927)
56. D.D. McFarland, D.J. Brown, Social distance as a metric: a systematic introduction to smallest space analysis, in *Bonds of Pluralism: The Form and Substance of Urban Social Networks* (John Wiley, New York, 1973), pp. 213–252
57. P.D. Hoff, A.E. Raftery, M.S. Handcock, Latent space approaches to social network analysis. J. Am. Stat. Assoc. **97**, 1090 (2002). https://doi.org/10.1198/016214502388618906
58. E.N. Gilbert, Random plane networks. J. Soc. Ind. Appl. Math. **9**, 533 (1961). https://doi.org/10.1137/0109045
59. M. Penrose, *Random Geometric Graphs* (Oxford University Press, Oxford, 2003) https://doi.org/10.1093/acprof:oso/9780198506263.001.0001
60. L. Lovász, *Large Networks and Graph Limits* (American Mathematical Society, Providence, RI, 2012)
61. S. Janson, Graphons, cut norm and distance, couplings and rearrangements. NYJM Monogr. **4** (2013)
62. P. van der Hoorn, G. Lippner, D. Krioukov, Sparse maximum-entropy random graphs with a given power-law degree distribution. J. Stat. Phys. **173**, 806 (2018). https://doi.org/10.1007/s10955-017-1887-7
63. G. Last M. Penrose, *Lectures on the Poisson Process* (Cambridge University Press, Cambridge, 2017). https://doi.org/10.1017/9781316104477
64. V. De Silva, R. Ghrist, Coverage in sensor networks via persistent homology. Algebraic Geometric Topol. **7**, 339 (2007)
65. K. Borsuk, On the imbedding of systems of compacta in simplicial complexes. Fundamenta Math. **35**, 217 (1948)
66. M. Farber, L. Mead, T. Nowik, Random simplicial complexes, duality and the critical dimension. J. Topol. Anal. **1** (2019). https://doi.org/10.1142/S1793525320500387
67. M. Farber, L. Mead, Random simplicial complexes in the medial regime. Topol. Appl. **272**, 107065 (2020). https://doi.org/10.1016/j.topol.2020.107065
68. O. Cooley, N.D. Giudice, M. Kang, P. Sprüssel, Vanishing of cohomology groups of random simplicial complexes. Random Struct. Algor. **56**, 461 (2020)
69. H. Garland, p-adic curvature and the cohomology of discrete subgroups of p-adic groups. Ann. Math. **375** (1973). http://www.jstor.org/stable/1970829
70. A. Costa, M. Farber, Large random simplicial complexes. III; the critical dimension. J. Knot Theory Its Ramifications **26**, 1740010 (2017). https://doi.org/10.1142/S0218216517400107
71. M.D. Penrose, The longest edge of the random minimal spanning tree. Ann. Appl. Probab. **340** (1997). http://www.jstor.org/stable/2245234
72. O. Bobrowski, Homological Connectivity in Random Čech Complexes. arXiv:1906.04861 [math] (2019)
73. O. Bobrowski, G. Oliveira, Random Čech complexes on riemannian manifolds. Random Struct. Alg. **54**, 373 (2019)
74. J.W. Milnor, *Morse Theory* (Princeton University Press, 1963)
75. V. Gershkovich, H. Rubinstein, Morse theory for Min-type functions. Asian J. Math. **1**, 696 (1997)
76. O. Bobrowski, M. Schulte, D. Yogeshwaran, Poisson process approximation under stabilization and Palm coupling. arXiv preprint arXiv:2104.13261 (2021)
77. M. Kahle, Random geometric complexes. Discrete Comput. Geom. **45**, 553 (2011)
78. R. Forman, A user's guide to discrete Morse theory. Sém. Lothar. Combin. **48**, 35pp (2002). http://www.emis.ams.org/journals/SLC/wpapers/s48forman.pdf
79. L. Aronshtam, N. Linial, T. Luczak, R. Meshulam, Collapsibility and vanishing of top homology in random simplicial complexes. Discrete Comput. Geom. **49**, 317 (2013). http://link.springer.com/article/10.1007/s00454-012-9483-8

80. N. Linial, Y. Peled, On the phase transition in random simplicial complexes. Ann. Math. Sec. Ser. **184**, 745 (2016). https://doi.org/10.4007/annals.2016.184.3.3
81. D. Kozlov, The threshold function for vanishing of the top homology group of random d-complexes. Proc. Am. Math. Soc. **138**, 4517 (2010) http://www.ams.org/proc/2010-138-12/S0002-9939-2010-10596-8/
82. L. Aronshtam, N. Linial, The threshold for d-collapsibility in random complexes. Random Struct. Algor. (2015) http://onlinelibrary.wiley.com/doi/10.1002/rsa.20585/full
83. M. Kahle, E. Meckes, Limit the theorems for Betti numbers of random simplicial complexes. Homology, Homotopy Appl. **15**, 343 (2013) http://projecteuclid.org/euclid.hha/1383943681
84. N. Linial, I. Newman, Y. Peled, Y. Rabinovich, Extremal hypercuts and shadows of simplicial complexes. Israel J. Math. **229**, 133 (2019)
85. O. Bobrowski, P. Skraba, Homological percolation: the formation of giant k-cycles. Int. Math. Res. Notices (2020). https://doi.org/10.1093/imrn/rnaa305
86. O. Bobrowski, P. Skraba, Homological percolation and the Euler characteristic. Phys. Rev. E **101**, 032304 (2020)
87. E. Babson, C. Hoffman, M. Kahle, The fundamental group of random 2-complexes. J. Am. Math. Soc. **24**, 1 (2011). http://www.ams.org/jams/2011-24-01/S0894-0347-2010-00677-7/
88. Z. Luria, Y. Peled, On simple connectivity of random 2-complexes. ArXiv preprint arXiv:1806.03351 (2018)
89. E. Babson, Fundamental groups of random clique complexes. ArXiv preprint arXiv:1207.5028 (2012)
90. A. Costa, M. Farber, D. Horak, Fundamental groups of clique complexes of random graphs. Trans. London Math. Soc. **2**, 1 (2015). https://doi.org/10.1112/tlms/tlv001
91. A.E. Costa, M. Farber, Geometry and topology of random 2-complexes. Isr. J. Math. **209**, 883 (2015). https://doi.org/10.1007/s11856-015-1240-2
92. A. Newman, Freeness of the random fundamental group. J. Topol. Anal. **12**, 29 (2020)
93. C. Stein (Organization IMS, 1986)
94. N. Ross et al., Fundamentals of Stein's method. Prob. Surv. **8**, 210 (2011)
95. D. Yogeshwaran, E. Subag, R.J. Adler, Random geometric complexes in the thermodynamic regime. Prob. The. Related Fields, **1** (2016) http://link.springer.com/article/10.1007/s00440-015-0678-9
96. Y. Hiraoka, T. Shirai, K.D. Trinh, Limit theorems for persistence diagrams. Ann. Appl. Prob. **28**, 2740 (2018)
97. J.T. Krebs, W. Polonik, On the asymptotic normality of persistent Betti numbers. arXiv preprint arXiv:1903.03280 (2019)
98. K.D. Trinh et al. On central limit theorems in stochastic geometry for add-one cost stabilizing functionals. Electron. Commun. Prob. **24** (2019)
99. M.D. Penrose, J.E. Yukich, Central limit theorems for some graphs in computational geometry. Ann. Appl. Probab. 1005 (2001)
100. A. Goel, K.D. Trinh, K. Tsunoda, Strong law of large numbers for Betti numbers in the thermodynamic regime. J. Statist. Phys. **174**, 865 (2019)
101. A. Auffinger, A. Lerario, E. Lundberg, Topologies of random geometric complexes on riemannian manifolds in the thermodynamic limit. Int. Math. Res. Not. (2020). https://doi.org/10.1093/imrn/rnaa050
102. A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, 2002)
103. T. Owada, O. Bobrowski et al., Convergence of persistence diagrams for topological crackle. Bernoulli **26**, 2275 (2020)
104. O. Bobrowski, M. Kahle, P. Skraba, Maximally persistent cycles in random geometric complexes. Ann. Appl. Prob. **27**, 2032 (2017a). https://projecteuclid.org/euclid.aoap/1504080025
105. A. Gundert, U. Wagner, On eigenvalues of random complexes. Israel J. Math. **216**, 545 (2016). https://doi.org/10.1007/s11856-016-1419-1
106. D. Dotterrer M. Kahle, Coboundary expanders. J. Topol. Analys. **4**, 499 (2012). http://www.worldscientific.com/doi/pdf/10.1142/S1793525312500197

107. A. Knowles, R. Rosenthal, Eigenvalue confinement and spectral gap for random simplicial complexes. Random Struct. Algor. **51**, 506 (2017)
108. A.M. Thomas, T. Owada, Functional limit theorems for the Euler characteristic process in the critical regime. Adv. Appl. Probab. **53**, 57 (2021)
109. G.C. Thoppe, D. Yogeshwaran, R.J. Adler et al., On the evolution of topology in dynamic clique complexes. Adv. Appl. Prob. **48**, 989 (2016)
110. T. Owada, G. Samorodnitsky, G. Thoppe, *Limit Theorems for Topological Invariants of the Dynamic Multi-parameter Simplicial Complex* (Stochastic Proc, Their Appl, 2021)
111. P. Skraba, G. Thoppe, D. Yogeshwaran, Randomly Weighted d-complexes: minimal spanning acycles and persistence diagrams, [math] (2017), arXiv: 1701.00239
112. Y. Hiraoka, T. Shirai, Minimum spanning acycle and lifetime of persistent homology in the Linial-Meshulam process. Random Struct. Algor. **51**, 315 (2017)
113. A.-L. Barabási, *Network Science* (Cambridge University Press, Cambridge, UK, 2016)
114. M.E.J. Newman, *Networks* (Oxford University Press, Oxford, 2018)
115. P. Niyogi, S. Smale, S. Weinberger, A topological view of unsupervised learning from noisy data. SIAM J. Comput. **40**, 646 (2011)
116. B.T. Fasy, F. Lecci, A. Rinaldo, L. Wasserman, S. Balakrishnan, A. Singh, Confidence sets for persistence diagrams. Ann. Statis. **42**, 2301 (2014)
117. F. Chazal, B. Fasy, F. Lecci, B. Michel, A. Rinaldo, L. Wasserman, Robust topological inference: distance to a measure and kernel distance. J. Mach. Learn. Res. **18**, 5845 (2017)
118. O. Bobrowski, S. Mukherjee, J.E. Taylor, Topological consistency via kernel estimation. Bernoulli **23**, 288 (2017)
119. Y. Reani O. Bobrowski, *Cycle Registration in Persistent Homology with Applications in Topological Bootstrap*. arXiv preprint arXiv:2101.00698 (2021)
120. P. Diaconis, S. Janson, Graph limits and exchangeable random graphs. Rend di Matemtica **28**, 33 (2008)
121. W. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem. Ann. Math. **166**(3), 897–946 (2007). https://doi.org/10.4007/annals.2007.166.897
122. G. Elek, B. Szegedy, A measure-theoretic approach to the theory of dense hypergraphs. Adv. Math. **231**(3–4), 1731–1772 (2012). https://doi.org/10.1016/j.aim.2012.06.022
123. Y. Zhao, Hypergraph limits: a regularity approach. Random Struct. Algor. **47**(2), 205–226 (2015). https://doi.org/10.1002/rsa.20537
124. K. Balasubramanian, D. Gitelman, H. Liu, Nonparametric modeling of higher-order interactions via hypergraphons. J. Mach. Learn. Res. **22**(146), 1–35 (2021)

Chapter 3
Persistent Homology: A Topological Tool for Higher-Interaction Systems

Francesco Vaccarino, Ulderico Fugacci, and Sara Scaramuccia

Abstract The aim of this chapter is to give a handy but thorough introduction to persistent homology and its applications. The chapter's path is made by the following steps. First, we deal with the constructions from data to simplicial complexes according to the kind of data: filtrations of data, point clouds, networks, and topological spaces. For each construction, we underline the possible dependence on a fixed scale parameter. Secondly, we introduce the necessary algebraic structures capturing topological informations out of a simplicial complex at a fixed scale, namely the simplicial homology groups and the Hodge Laplacian operator. The so-obtained linear structures are then integrated into the multiscale framework of persistent homology where the entire persistence information is encoded in algebraic terms and the most advantageous persistence summaries available in the literature are discussed. Finally, we introduce the necessary metrics in order to state properties of stability of the introduced multiscale summaries under perturbations of input data. At the end, we give an overview of applications of persistent homology as well as a review of the existing tools in the broader area of Topological Data Analysis (TDA).

3.1 Introduction

Persistent homology is an emerging tool to identify robust topological features underlying the higher-order structure of (high-dimensional) data and complex dynamical systems (such as brain dynamics, molecular folding, distributed sensing). In the

F. Vaccarino (✉) · S. Scaramuccia
SmartData@PoliTO, Dipartimento di Scienze Matematiche, Politecnico di Torino, C.so Duca degli Abruzzi no 24, Torino 10129, Italy
e-mail: francesco.vaccarino@polito.it

S. Scaramuccia
e-mail: sara.scaramuccia@polito.it

U. Fugacci
IMATI, CNR, Via de Marini 6, Genova 16149, Italy
e-mail: ulderico.fugacci@cnr.it

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_3

broader sense, persistence aims to associate meaningful mathematical objects to filtrations of data, point clouds, networks, topological spaces, simplicial complexes etc. Here, a filtration is a sequence of spaces, point clouds, networks, etc. together with maps connecting them. Very often these maps, if not always, will be inclusions and the sequence will results from filtering on similarities, or weights, in a way akin to hierarchical clustering. More specifically, we will be interested in considering filtrations of simplicial complexes, higher dimensional tetrahedral foams, combinatorial topological spaces naturally associated to the input data. Attached to simplicial complexes there are the homology groups that represent the lack of k-connectivity: the zeroth group measures usual connectivity, the first one the lack of tiles, e.g. three cliques, the second one the lack of four cliques and so on. The objects counted by the homology groups are *voids* as dis-junctions, cycles, empty boxes and so on. The homology groups corresponding to the stages of the filtration at the relevant scales, are connected via linear mappings induced by the filtration's maps (by so called functoriality). Persistence studies the entanglement of these information into a single object called the persistence module, which is a linear space endowed with a structure of graded module encoding births and deaths of the $voids$ along the filtration.

The main feature characterizing the persistence approach is to associate a data filtration with a unique well defined algebraic summary, computable by using linear algebra tools, whose behaviour represent in a principled way the topological properties of the data at the various scales in a single summary. This differentiate in an unequivocal way persistence from its germane, clustering. Indeed, while clustering analyze the behaviour of data with respect to e.g. some similarity by representing the process of aggregating data at the various scales but implicitly forces to choose a significative scale where to stop the process, persistence formalize and compress the information of the aggregation process itself. Furthermore, when persistence is applied to the higher degrees homologies it deals with information that no longer can be seen as a mere representation of the connectivity properties of the data along their characteristic scales.

In this chapter, we review the main notions in persistent homology according to the persistence pipeline. In Sect. 3.2, we address the main constructions to obtain a simplicial complex, or a filtration of simplicial complexes, according to the kind of input data. In Sect. 3.3, we introduce the notion of simplicial homology and its relation to the Hodge Laplacian as a way to deal with topological features of a simplicial complex, into linear algebraic terms. In Sect. 3.4, the multiscale approach of persistent homology is introduced and its main summaries are reviewed. Stability properties of metrics to compare changes in the data filtration to changes in the persistence summaries are reviewed in Sect. 3.5. In Sect. 3.6, we review some applications of persistent homology and provide the necessary references for tools implementing the persistence pipeline in practice.

Notations

- Σ Simplicial Complex (finite and $\emptyset \notin \Sigma$)
- σ, τ, ρ simplices
- $\mathbb{F}, \mathbb{R}, \mathbb{Z}, \mathbb{Z}_2$ an arbitrary field, real numbers, integers, $\mathbb{Z}/2\mathbb{Z}$
- $(v_0, \ldots, v_q)$ q-simplex (simplex of dimension q)
- $[v_0, \ldots, v_q]$ oriented q-simplex
- Σ_q the set of the q-simplices of Σ
- $\Sigma_{(q)}$ the q-skeleton of Σ (i.e. the set of the p-simplices of Σ with $p \leq q$)
- $C_q(\Sigma; \mathbb{F})$ (short C_q or $C_q(\Sigma)$) group of q-chains of Σ with coefficients in $\mathbb{F}$
- $\mathscr{C}_q$ canonical basis of $C_q(\Sigma; \mathbb{F})$
- $\partial_q : C_q(\Sigma; \mathbb{F}) \to C_{q-1}(\Sigma; \mathbb{F})$ boundary map
- $\delta^q : C^q(\Sigma; \mathbb{F}) \to C^{q+1}(\Sigma; \mathbb{F})$ co-boundary map
- D_q matrix representing the map $\partial_q : C_q(\Sigma; \mathbb{F}) \to C_{q-1}(\Sigma; \mathbb{F})$ with respect to the basis $\mathscr{C}_q, \mathscr{C}_{q-1}$
- $B_q(\Sigma; \mathbb{F})$ (short B_q or $B_q(\Sigma)$) q-boundaries of Σ with coefficients in $\mathbb{F}$
- $Z_q(\Sigma; \mathbb{F})$ (short Z_q or $Z_q(\Sigma)$) q-cycles of Σ with coefficients in $\mathbb{F}$
- $H_q(X; \mathbb{F})$ (short H_q or $H_q(\Sigma)$) q^{th} homology group of Σ with coefficients in $\mathbb{F}$
- $\beta_q(\Sigma; \mathbb{F})$ (short β_q or $\beta_q(\Sigma)$) q^{th} Betti number of Σ with coefficients in $\mathbb{F}$
- $\ker(f)$ kernel of f
- $\mathrm{im}(f)$ image of f
- $\dim(\Sigma)$ dimension of Σ
- $\mathcal{H}_q$ Laplacian kernel
- $\varphi : \emptyset \subseteq \Sigma^0 \subseteq \Sigma^1 \subseteq \ldots \Sigma^m := \Sigma$ filtration of Σ
- C_q^i short for $C_q(\Sigma^i; \mathbb{F})$
- H_q^i short for $H_q(\Sigma^i; \mathbb{F})$
- $H_q(\varphi) := \bigoplus_i H_q(\Sigma^i; \mathbb{F})$ the q^{th} persistence module of φ
- (a, b), $(a, +\infty)$ persistence pairs

3.2 From Data to Simplicial Structures

In this section, we review and discuss the main constructions to obtain the main higher-order representation in the persistence pipeline, namely a *simplicial complex* associated to the original input datum [49]. We begin with Sect. 3.2.1 introducing simplicial complexes. We proceed further in Sect. 3.2.2 by introducing general constructions for simplicial complexes. In particular, we focus on the Nerve and the Flag complexes. Then, we review specific constructions according to the kind of the given data. We discuss point cloud data, graph and complex networks, partially ordered sets, and functions. A summary of the discussed methods according to the kind of input datum is depicted in Table 3.1.

Table 3.1 A summary of the possible simplicial complex construction (columns) introduced in this section according to the input data (rows)

Input Data	Simplicial complex			
	Nerve	Flag	Order complex	Sublevel set
Point cloud	Delaunay, Čech, α-shapes	VR, Witness		
Weighted graph		✓		✓
Poset			✓	
Function				Filtering function, Vertex-induced

3.2.1 Simplicial Complexes

Simplicial complexes are a classical mathematical tool for representing discrete shapes [78]. The elementary building blocks which form a simplicial are called simplices.

A *simplex* σ of dimension q (also addressed as q-simplex) is the convex hull of $q+1$ affinely independent points in the Euclidean space. Practically speaking, a 0-simplex is just a point, a 1-simplex an edge, a 2-simplex a triangle, a 3-simplex a tetrahedron, and so on. Given a q-simplex σ, any simplex τ which is the convex hull of a non-empty subset of the points generating σ is called a face of σ. Conversely, σ is called a coface of τ.

A (finite) *simplicial complex* Σ is a finite set of simplices such that:

- each face of a simplex in Σ belongs to Σ;
- each non-empty intersection of any two simplices in Σ is a face of both.

Worth to be noticed that formally there is no obstruction in considering the empty simplex as included in any simplicial complex Σ. Its inclusion is typically allowed in theoretical frameworks and it will lead to the definitions of augmented chain complexes and of reduced homology. For the sake of simplicity, in this chapter we will only consider simplicial complexes not including the empty simplex.

The two above claimed conditions are visually depicted in Fig. 3.1. In particular, we highlight the fact that once a simplex (such as the tetrahedron $ABCD$) belongs to a simplicial complex Σ, then all its faces (vertices, edges, and triangle in the depicted case) also belong to Σ. Moreover, in accordance with the second condition, notice that given any two simplices of the depicted simplicial complex Σ their intersection is empty or another simplex of Σ which is face of both.

By definition, a simplicial complex is a collection of simplices lying in an Euclidean space $\mathbb{R}^d$ of a certain dimension d. In order to adopt the notion of simplicial complex to a larger variety of datasets, it is useful to introduce a more general

definition of simplicial complexes which do not necessarily need for a geometric realization. An *abstract simplicial complex* Σ on a finite set V is a collection of non-empty subsets of V, called *simplices*, with the property of being closed under inclusion (see Fig. 3.2 for an example). More explicitly, if σ is a simplex which belongs to Σ and τ is a non-empty subset of V such that $\tau \subset \sigma$, then τ belongs to Σ. Given a simplicial complex Σ, the elements of V are called *vertices* of Σ and a simplex $\sigma \in \Sigma$ is called a *q-simplex* (equivalently, a simplex of dimension q) if it consists of $q + 1$ vertices.

Given a simplicial complex or an abstract simplicial complex Σ, the maximum of the dimensions of its simplices is called the dimension of Σ and denoted as $\dim(\Sigma)$. Moreover, we will denoted by Σ_q the set of the q-simplices of Σ and by $\Sigma_{(q)}$ the q-skeleton of Σ (i.e. the set of the p-simplices of Σ with $p \leq q$).

In spite of the different definitions, the two presented notions of simplicial complexes are strictly related. In fact, given a simplicial complex Σ, one can associate to it an abstract simplicial complex Σ' defined as the collection of the simplices of Σ. Vice versa, given an abstract simplicial complex Σ, it always possible to retrieve a simplicial complex Σ' whose associated abstract simplicial complex coincides with Σ [78]. The simplicial complexes Σ and Σ' depicted in Figs. 3.1 and 3.2 , respectively, represent an example of this correspondence. Σ' is the abstract simplicial complex associated to Σ, while Σ is a geometric realization of Σ'.

In the following, in the few cases in which we would like to distinguish between simplicial complexes and abstract simplicial complexes, we will adopt the term "geometric" simplicial complex to refer to the former. Differently, we will simply use the term "simplicial complex" to address both the structures.

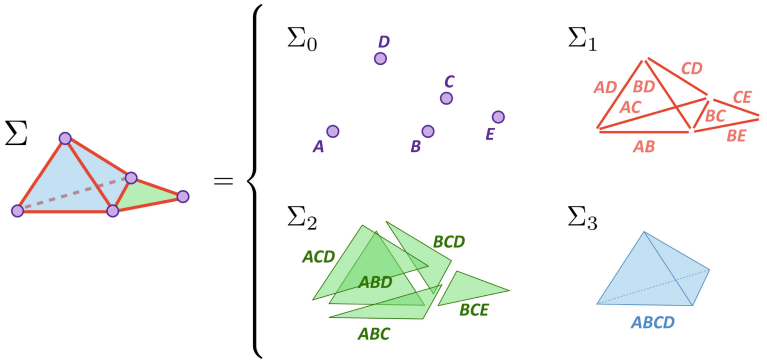


Fig. 3.1 A (geometric) simplicial complex Σ (on the left). The collection of simplices of the various dimensions Σ consists of (on the right)

$$\Sigma' = \left\{ \begin{array}{l} \{A\},\{B\},\{C\},\{D\},\{E\}, \\ \{A,B\},\{A,C\},\{A,D\},\{B,C\}, \\ \{B,D\},\{B,E\},\{C,D\},\{C,E\}, \\ \{A,B,C\},\{A,B,D\},\{A,C,D\}, \\ \{B,C,D\},\{B,C,E\}, \\ \{A,B,C,D\} \end{array} \right\} \begin{array}{l} \Sigma'_0 \\ \cup \\ \Sigma'_1 \\ \cup \\ \Sigma'_2 \\ \cup \\ \Sigma'_3 \end{array}$$

Fig. 3.2 An abstract simplicial complex Σ' on $V = \{A, B, C, D, E\}$. Σ' is the abstract simplicial complex associated to the geometric simplicial complex Σ depicted in Fig. 3.1

3.2.2 Nerve Complexes, Flag Complexes, and Other Constructions of Simplicial Complexes

As already mentioned, Topological Data Analysis aims at describing and characterizing data in terms of their shape. In order to satisfy this assumption, the first required step in the persistence pipeline concerns associating to the input dataset a simplicial complex that will provide the data with a suitable topological structure needed for the next steps of the topological-based analysis.

The kinds of datasets eligible for this process of "translation" into a simplicial complex are various and heterogeneous and cover a vast majority of the datasets which a researcher could face in its work.

For the sake of simplicity, in this document we will focus mainly on point clouds embedded in an Euclidean space, (weighted) graphs or complex networks, functions, and sets endowed with a relation of partial order among its elements. In spite of this, most of the presented constructions can be generalized/adapted to the case of a finite collection of elements endowed with a notion of proximity enabling to cover a wide plethora of datasets. More properly, with the term "proximity" we mean a semi-metric, i.e. a distance not necessarily satisfying the triangle inequality.

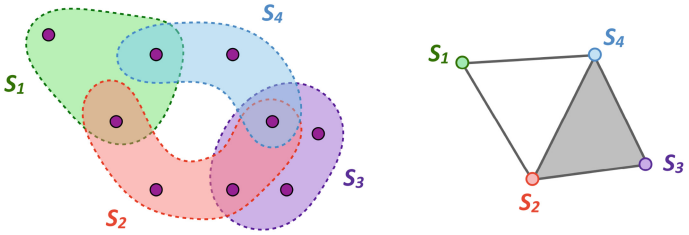
Two common tools adopted in this step assigning a simplicial complex to an arbitrary dataset are the Nerve complex and the Flag complex.

Nerve complex

Let us consider a finite collection S of sets in $\mathbb{R}^n$. The *Nerve complex* $Nrv(S)$ of S is the abstract simplicial complex generated by the non-empty common intersections among the sets of S [49]. More precisely,

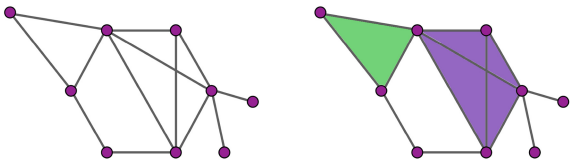
$$Nrv(S) := \left\{\sigma \subseteq S \mid \bigcap_{s\in\sigma} s \neq \emptyset\right\}. \tag{3.1}$$

As depicted in Fig. 3.3, intuitively the Nerve complex produces a simplicial complex whose vertices are in correspondence with the sets of S, whose edges represents the non-empty intersections among two sets in S, whose triangles the non-empty intersections among three sets in S, and so on. One of the most desirable properties of this construction relies on its capability of "preserving the shape" of the input



(a) (b)

Fig. 3.3 A cover S (**a**) and its Nerve complex $Nrv(S)$ (**b**)



(a) (b)

Fig. 3.4 A (undirected) graph G (**a**) and its Flag complex $Flag(G)$ (**b**)

collection of sets. In fact, (a simplified version of the) Nerve theorem ensures that, if all the sets in S are convex sets of $\mathbb{R}^n$, then the Nerve complex of S and the union of the sets in S are homotopy equivalent. Intuitively, this represents a theoretical guarantee ensuring us that the proposed construction respect the shape of the data and, specifically, its homology.

Flag complex

Given a (undirected) graph $G = (V, E)$, its *Flag complex* $Flag(G)$ is the abstract simplicial complex having as simplices all the subsets σ of vertices of G such that, for any two vertices u, v in σ, (u, v) is an edge of the graph G (see Fig. 3.4 for an example).

In other words, $Flag(G)$ consists of all the cliques of the graph G. For this reason, it is also called *clique complex*. Intuitively, the Flag complex of a graph G captures the connectivity among the vertices of G: two connected vertices are turned into an edge of the simplicial complex, three fully-connected vertices into a triangle, and so on.

Each of the next subsections will focus on a specific kind of dataset showing how the introduced tools can be applied in order to transform the input data into a simplicial complex. Let us notice how, in most of the cases, the described constructions will not produce a single simplicial complex but a collection of simplicial complexes Σ^a depending on a parameter a such that, if $a \leq b$, then $\Sigma^a \subseteq \Sigma^b$.

Table 3.2 A summary of the properties of simplicial complexes obtained from point cloud data embedded in $\mathbb{R}^n$

		Properties		
		Output complex	Dimension	Parameter dependence
Methods	Delaunay	Geometric	n	-
	Čech	Abstract	Arbitrary	✓
	Vietoris-Rips	Abstract	Arbitrary	✓
	α-shapes	Geometric	n	✓
	Witness	Abstract	Arbitrary	✓

3.2.2.1 Point Clouds

Let us consider a finite set of points V in $\mathbb{R}^n$ and let us investigate several different techniques to associate to it a simplicial complex. Table 3.2 highlights some properties of the methods to be introduced in this section: the simplicial complex might be a purely combinatorial object (abstract) or embedded in some space (geometric); the maximum dimension reached by the simplices in the complex might be dependent on the original point cloud dimension or not; and the construction might be dependent on a fixed parameter or not. The last property is particularly relevant when the parameter is left free to vary. This multi-scale framework will be treated in Sect. 3.4.

Delaunay triangulations

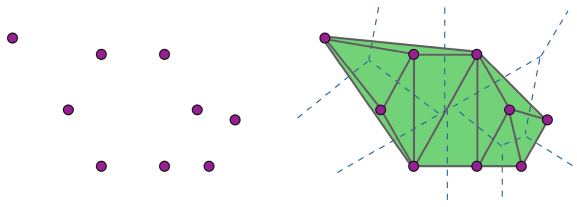
One of the most traditional ways to build a simplicial complex from a point cloud is the so called Delaunay triangulation [37]. This construction, originally described for sets of points in $\mathbb{R}^2$ but generalizable to arbitrary dimensions, aims at producing a triangulation of the convex hull of V (i.e. the smallest convex set containing V) free of long and skinny triangles. More properly, a *Delaunay triangulation* $Del(V)$ of V is a triangulation of the convex hull of V such that the circumcircle of any triangle does not contain any point of V in its interior. A Delaunay triangulation of V can be achieved by computing the Nerve complex of the Voronoi regions of the points in V, where the *Voronoi region* $R_V(u)$ of a vertex u of V is the set of all points in $\mathbb{R}^n$ whose Euclidean distance to u is not greater than their distance to the any other vertex v of V. More formally,

$$R_V(u) := \{x \in \mathbb{R}^n \mid \forall\, v \in V, d(x,u) \leq d(x,v)\}, \tag{3.2}$$

where d is the Euclidean distance.

Figure 3.5 depicts an example of Delaunay triangulation. Specifically, Fig. 3.5b shows a Delaunay triangulation of the point cloud V represented in Fig. 3.5a.

The points in $V \subseteq \mathbb{R}^n$ will be said in general position if no $n+2$ points lie on a common $(n-1)$-dimensional sphere. E.g., for $n = 2$, the points of V are in general



(a) (b)

Fig. 3.5 A point cloud V in $\mathbb{R}^2$ (**a**) and its Delaunay triangulation $Del(V)$ (**b**). In (**b**), the Voronoi regions in which $\mathbb{R}^2$ is subdivided are delimited by dotted lines

position if and only if no four or more points are co-circular. The hypothesis that the points in V are in general position enables to ensure that the Delaunay triangulation of V is unique and its realization as the Nerve complex of the Voronoi regions of the vertices of V produces not simply an abstract simplicial complex but a geometric simplicial complex embedded in $\mathbb{R}^n$.

Čech complexes

Given a point cloud V in $\mathbb{R}^n$ and fixed a scalar parameter $\epsilon > 0$, the *Čech complex* $\check{C}ech_\epsilon(V)$ associated with V and ϵ is the abstract simplicial complex defined as:

$$\check{C}ech_\epsilon(V) := \left\{\sigma \subseteq V \mid \sigma \neq \emptyset \text{ and } \bigcap_{v\in\sigma} B(v, \epsilon/2) \neq \emptyset\right\}, \tag{3.3}$$

where $B(v, \epsilon/2)$ is the closed ball of radius $\epsilon/2$ centered in v (see Fig. 3.6a for an example) [36, 49].

One can notice that the above definition arises from the idea of replacing each vertex of V with a ball of radius $\epsilon/2$ and then taking the Nerve complex associated to this collection of balls. While considering the Čech complex of a point cloud V is one of the most natural way in order to associate a simplicial complex to V, its computation is definitely a time consuming procedure. For this reason, several "approximated" versions of the Čech complex have been proposed in the literature.

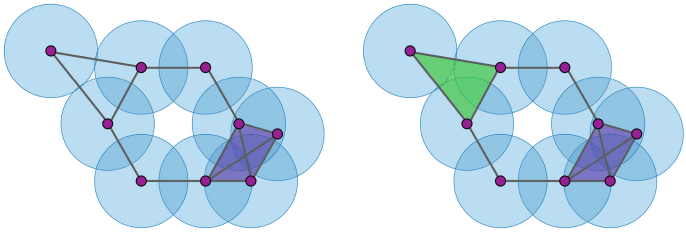
Vietoris-Rips complexes

A first approximated version of the Čech complex is represented by the Vietoris-Rips complex [49, 102]. Fixed a scalar parameter $\epsilon > 0$, the *Vietoris-Rips (VR) complex* $VR_\epsilon(V)$ associated with V and ϵ is the abstract simplicial complex defined as:

$$VR_\epsilon(V) := \{\sigma \subseteq V \mid \sigma \neq \emptyset, \text{ and } \forall u, v \in \sigma, d(u, v) \leq \epsilon\}, \tag{3.4}$$

where d is the Euclidean distance (see Fig. 3.6b for an example).

By definition, it easy to notice that the Vietoris-Rips complex $VR_\epsilon(V)$ coincides with the Flag complex of the 1-skeleton of the Čech complex $\check{C}ech_\epsilon(V)$. Moreover, the fact that the Vietoris-Rips complex represents an approximation of the Čech complex is ensured by the property claiming that for any $\epsilon > 0$,



(a) (b)

Fig. 3.6 The Čech complex $\check{C}ech_\epsilon(V)$ (**a**) and the Vietoris-Rips complex $VR_\epsilon(V)$ (**b**) associated with the point cloud V depicted in Fig. 3.5a and ϵ. The depicted closed balls are of radius $\epsilon/2$

$$VR_{\frac{\sqrt{2}}{2}\epsilon}(V) \subseteq \check{C}ech_\epsilon(V) \subseteq VR_\epsilon(V). \tag{3.5}$$

Alpha-shapes

The previously described constructions present several pros and cons. Delaunay triangulations produce geometrical simplicial complexes embedded in the same Euclidean space in which the input point cloud lies but, as a drawback, the obtained complex, being a triangulation of a convex hull, is quite trivial from a topological point of view (in fact, it has the same homology of a space consisting of a single point). On the contrary, Čech and Vietoris-Rips complexes better preserve the shape of the input dataset but they produce abstract simplicial complexes of arbitrarily high dimension. The idea behind alpha-shapes is to define a new strategy which could satisfies the positive aspects achieved by the previously described constructions and, at the same time, avoid to have their drawbacks [35, 40]. Given a point cloud V in $\mathbb{R}^n$ and fixed a scalar parameter $\epsilon > 0$, this can simply achieved by defining the *alpha-shape* $Alpha_\epsilon(V)$ associated with V and ϵ as the Nerve complex generated by the regions obtained as the intersections between the closed balls of radius $\epsilon/2$ centered in the points of V and the Voronoi regions of the points in V. More precisely,

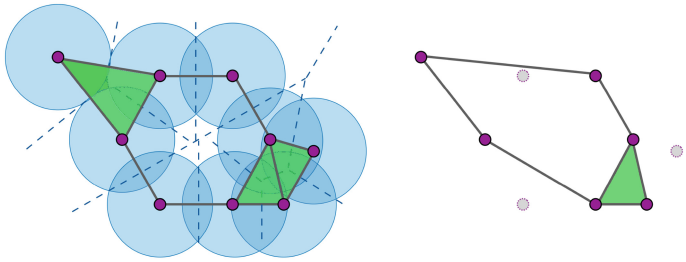
$$Alpha_\epsilon(V) := \left\{\sigma \subseteq V \mid \sigma \neq \emptyset \text{ and } \bigcap_{v\in\sigma} A(v, \epsilon/2) \neq \emptyset\right\}, \tag{3.6}$$

where $A(v, \epsilon/2)$ is defined as $B(v, \epsilon/2) \cap R_V(v)$ (see Fig. 3.7a for an example). So, since by definition $A(v, \epsilon/2) \subseteq B(v, \epsilon/2)$, we have that for any $\epsilon > 0$,

$$Alpha_\epsilon(V) \subseteq \check{C}ech_\epsilon(V). \tag{3.7}$$

Witness complexes

When dealing with point clouds consisting of a huge number of elements, the previously introduced constructions (even if sometimes still tractable) require a considerable amount of computational resources. *Witness complexes* aim at fixing this poten-



(a) (b)

Fig. 3.7 The alpha-shape $Alpha_\epsilon(V)$ (**a**) and the witness complex (**b**) associated with the point cloud V depicted in Fig. 3.5a and ϵ. The depicted closed balls in (**a**) are of radius $\epsilon/2$

tial issue by constructing a simplicial complex still well representing the "shape" of the point cloud V in input but having a vertex sets of cardinality equal to just a fraction of V [30]. While a detailed definition of this class of complexes is out of the scope of this chapter, we can intuitively describe the witness complex of V as follows. In a nutshell, a witness complex is built by choosing a set of "landmark" points from V and then constructing a simplicial complex by using the remaining (non-landmark) data points as witnesses to the existence of edges or simplices spanned by combinations of landmark points (see Fig. 3.7b for an example). Analogously to the previous constructions, the definition of the witness complex associated with a point cloud V depends on a scalar parameter $\epsilon > 0$.

3.2.2.2 Graphs and Complex Networks

Given a complex network represented as a graph $G = (V, E)$ a natural way to associate with it a simplicial complex is by considering its Flag complex $Flag(G)$ [102]. If (as it is common is several application domains) the edges of graph G are weighted by a function $w : E \to \mathbb{R}$, fixed a scalar parameter $\epsilon > 0$, let us denote as $G_\epsilon = (V_\epsilon, E_\epsilon, w_\epsilon : E_\epsilon \to \mathbb{R})$, the subgraph of G such that:

- its vertex set V_ϵ coincides with V;
- its set of edges E_ϵ consists of the edges of G having weight lower or equal than ϵ;
- its weight function w_ϵ is the restriction of w to the edge set E_ϵ.

Combining this definition with the notion of Flag complex, one can define the Flag complex associated with G and ϵ as the Flag complex $Flag(G_\epsilon)$ of G_ϵ.

3.2.2.3 Functions

Given a simplicial complex Σ a function $f : \Sigma \to \mathbb{R}$ is called a *filtering function* if, whenever σ is a face of τ with $\sigma, \tau \in \Sigma$, then $f(\sigma) \leq f(\tau)$. Given a filtering

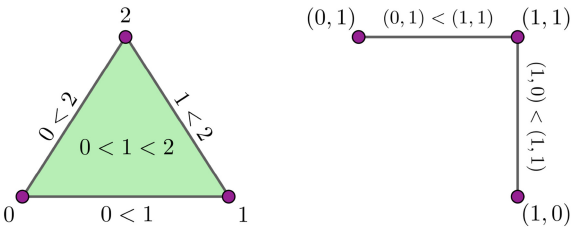


Fig. 3.8 The order complexes associated to $[2] = \{0, 1, 2\}$ (**a**) and to $\{(0, 1), (1, 0), (1, 1)\}$ (**b**). The considered partial orders $\leq$ are the standard ordering of natural numbers in the first case, while its component-wise extension in the second one

function, it is possible to associate with f a collection of simplicial subcomplexes of Σ depending on a parameter $r \in \mathbb{R}$. Specifically, choosing a scalar value $r \in \mathbb{R}$, the sublevel set of f with respect to r is defined as

$$\Sigma^r := f^{-1}\big((-\infty, r]\big). \tag{3.8}$$

In several applicative scenarios, the function f could be defined on a domain which is not a simplicial complex. In spite of this, in most of the cases this does not represents a serious obstruction since one can often reconduct the situation to the previously described case. For instance, if the domain of f is a point cloud V, one can:

- construct from V a simplicial complex Σ by adopting one of the techniques described in Subsect. 3.2.2.1;
- extend f to Σ by defining, for any $\sigma \in \Sigma$,

$$f(\sigma) := \max\{f(v) \mid v \text{ is a vertex of } \sigma\}. \tag{3.9}$$

3.2.2.4 Partially-Ordered Sets

Order complexes

A *partially-ordered set*, usually called a *poset*, $(S, \leq)$ is a set S endowed with a partial order $\leq$ defined among its elements. Given a (finite) poset $(S, \leq)$, the *order complex* $\Sigma(S, \leq)$ associated to it is the abstract simplicial complex on S whose vertex set coincides with S and whose simplices are the finite chains of $(S, \leq)$, i.e. the finite totally-ordered subsets of S [73, 82].

For example, let us consider the subset $[n] \subseteq \mathbb{N}$ consisting of the first $n + 1$ natural numbers, i.e. $[n] = \{0, 1, 2, \ldots, n\}$. By considering the standard ordering $\leq$ on natural numbers, $([n], \leq)$ is a totally-ordered set (in particular, a poset). As

depicted in Fig. 3.8a for the case $n = 2$, the order complex associated to $([n], \leq)$ consists of a n-simplex and all its faces.

Differently, let us consider the subset $S := \{(0, 1), (1, 0), (1, 1)\} \subseteq \mathbb{N}^2$ and endow it with the partial order $\leq$ defined by: $(x, y) \leq (x', y')$ if and only if $x \leq x'$ and $y \leq y'$. Then, as depicted in Fig. 3.8b, the order complex associated to $(S, \leq)$ consists of two 1-simplices connected by a vertex (corresponding to the point $(1, 1)$). Please notice that the edge connecting $(0, 1)$ and $(1, 0)$ is missing since the two points are not comparable with respect to the considered partial order.

3.3 From Simplicial Structure to Linear Algebra

In the previous section, we described how to endow data with a shape through the combinatorial and topological notion of a simplicial complex. In this section, we introduce the notion of simplicial homology and its relation to the Hodge Laplacian as a way to capture topological features coming from a simplicial complex into linear algebraic terms. The notions introduced in this section have to be seen as preliminary ones to persistent homology, that is the core multiscale notion of this chapter introduced in Sect. 3.4. Differently from persistent homology, a simplicial homology group and a Hodge Laplacian refer to a single scale parameter characterizing the shape we have associated to our data, and provides the building blocks for the multiscale approach of persistent homology.

3.3.1 Simplicial Homology

Homology is a mathematical tool able to describe a shape in terms of its holes [78]. In order to define the homology of a simplicial complex Σ, we need to move from the combinatorial notion of a simplicial complex to the algebraic structures of chain groups.

Let σ be a q-simplex spanned by the vertices $v_0, v_1, \ldots, v_q$. Two orderings of the vertices of σ are defined equivalent if they differ by an even permutation. If $q > 0$, the orderings of the vertices of σ fall into two equivalence classes called orientations of σ. A pair consisting of a simplex and an orientation of it will be called *oriented simplex*. We will adopt the symbol $[v_0, v_1, \ldots, v_q]$ to denote the oriented simplex generated by $v_0, v_1, \ldots, v_q$ and the equivalence class of the specific ordering $(v_0, v_1, \ldots, v_q)$. When clear from the context, we will simply use σ to denote either a simplex or an oriented simplex.

Given a simplicial complex Σ and a field $\mathbb{F}$, we define the q^{th} *chain group* $C_q(\Sigma; \mathbb{F})$ of Σ as the vector space over $\mathbb{F}$ generated by the oriented q-simplices of Σ and where $-\sigma$ coincides with the simplex σ endowed with the opposite orientation. An element of $C_q(\Sigma; \mathbb{F})$ is called *q-chain* and it is a finite linear combination $\sum_i \lambda_i \sigma_i$ with coefficients in $\mathbb{F}$ of oriented q-simplices. Once an orientation for each

simplex of Σ_q is fixed, a canonical basis $\mathscr{C}_q$ of $C_q(\Sigma;\mathbb{F})$ is determined. Properly, $\mathscr{C}_q$ consists of the q-chains corresponding to the oriented q-simplices of Σ. In the following when it will not cause any ambiguity, we will often write $C_q(\Sigma)$ or simply C_q for in place of $C_q(\Sigma;\mathbb{F})$. For the sake of readability, an analogous simplification will be adopted to all the other notations presented in the following.

Chain groups of a simplicial complex Σ are connected by linear maps $\partial_q : C_q(\Sigma;\mathbb{F}) \to C_{q-1}(\Sigma;\mathbb{F})$ called *boundary operators*. Given an element of the basis $\mathscr{C}_q$ corresponding to the oriented q-simplex $\sigma = [v_0, v_1, \ldots, v_q]$ of Σ, $\partial_q(\sigma)$ is defined as:

$$\partial_q(\sigma) := \sum_{i=0}^{q} (-1)^i [v_0, \ldots, \hat{v}_i, \ldots, v_q], \tag{3.10}$$

where $\hat{v}_i$ means that vertex v_i is not present (see Fig. 3.9 for an example).

The boundary ∂_q is extended to each q-chain of Σ by linearity. In the following, we will denote by D_q the matrix representing the map $\partial_q : C_q(\Sigma;\mathbb{F}) \to C_{q-1}(\Sigma;\mathbb{F})$ with respect to the basis $\mathscr{C}_q, \mathscr{C}_{q-1}$.

We denote as $Z_q(\Sigma;\mathbb{F}) := \ker \partial_q$ the $\mathbb{F}$-vector space of the *q-cycles* of Σ and as $B_q(\Sigma;\mathbb{F}) := \operatorname{im} \partial_{q+1}$ the $\mathbb{F}$-vector space of the *k-boundaries* of Σ. Examples of boundaries and cycles are depicted in Fig. 3.9.

It is immediate to check that, for any q, $\partial_q \partial_{q+1} = 0$ or, equivalently, that $B_q(\Sigma) \subseteq Z_q(\Sigma)$. This fact ensures that the quotient

$$H_q(\Sigma;\mathbb{F}) := \frac{Z_q(\Sigma;\mathbb{F})}{B_q(\Sigma;\mathbb{F})} \tag{3.11}$$

is a well-defined vector space over $\mathbb{F}$. The space $H_q(\Sigma;\mathbb{F})$ will be called the k^{th} *homology group* of Σ with coefficients in $\mathbb{F}$.

By definition of $H_q(\Sigma;\mathbb{F})$, q-cycles of Σ are partitioned in equivalence classes called *homology classes*. Two q-cycles are said *homologous* if they belong to the

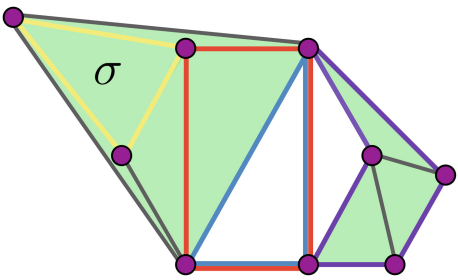


Fig. 3.9 Examples of boundaries and cycles of a simplicial complex Σ. The yellow edges represent the boundary of the triangle σ. All the highlighted collections of edges represent 1-cycles. The yellow and the purple cycles are also 1-boundaries while the red and the light blue ones are not. Moreover, the red and the light blue cycles are homologous, i.e. they are two representatives of the same homology class (the triangular hole they both include)

same homology class (see Fig. 3.9 for an example). Given a q-cycle c its homology class (i.e. the collection of all the q-cycles homologous to c) will be denoted by $[c]$.

The dimension of $H_q(\Sigma;\mathbb{F})$ as $\mathbb{F}$-vector space is called the k^{th} *Betti number* of Σ and denoted by $\beta_q(\Sigma;\mathbb{F})$.

Intuitively, homology spots the "holes" of a simplicial complex Σ. In fact, a non-null element of $H_q(\Sigma)$ is an equivalence class of homologous cycles that are not the boundary of any $(q+1)$-chain of Σ. The number β_q of such classes represents, in dimension 0, the number of connected components of Σ, in dimension 1, the number of its tunnels and its loops, in dimension 2, the number of voids or cavities, and so on.

3.3.2 Hodge Decomposition

In this section, the ground ring of coefficients will be the field of real numbers $\mathbb{R}$ unless otherwise stated. We fix an arbitrary finite simplicial complex Σ, and we set $C_q := C_q(\Sigma;\mathbb{R})$ so that $\partial_q : C_q \to C_{q-1}$ will stand for the boundary operator for any $q = 0,\ldots,\dim(\Sigma)$. Furthermore, we set $H_q := H_q(\Sigma;\mathbb{R})$. We will work in a simplified setting, adopting for the entire section the canonical basis $\mathscr{C}_q$ of $C_q(\Sigma;\mathbb{F})$ for C_q i.e. the one in bijection with (equivalence classes of oriented) $q-$dimensional simplices and this will allow us to identify ∂_q with its matrix D_q w.r.t. the bases $\mathscr{C}_q$ and $\mathscr{C}_{q-1}$,

$$D_q := M_{\partial_q}^{\mathscr{C}_q,\mathscr{C}_{q-1}}. \tag{3.12}$$

The *Hodge q-Laplacian*, that is also called *q-th Combinatorial Laplacian*, can be defined both in terms of cohomology, coboundaries and adjoint, as it has been proposed in the previous chapter, than as follows in terms of homology, boundary maps and their matrices

$$\mathcal{L}_q := \partial_q^*\partial_q + \partial_{q+1}\partial_{q+1}^*, \tag{3.13}$$

where ∂_q^* denotes the adjoint of ∂_q w.r.t some inner product defined on the chains groups. We fix the standard inner (scalar) product on C_q which makes $\mathscr{C}_q$ an orthonormal basis so that the matrix of ∂_q^* with respect to $\mathscr{C}_{q-1}$, $\mathscr{C}_q$ is D_q^T the transposed of D_q. We write

$$L_q := D_q^T D_q + D_{q+1}D_{q+1}^T. \tag{3.14}$$

We will now study some properties of the Laplacians by taking advantage of this formulation and following [68].

The fundamental theorem of topology states that $D_{q+1}D_q = 0$ for all $q \geq 0$. Therefore, it makes sense to study the Laplacians as follows.

Definition 3.1 Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$ be two real matrices such that $AB = \mathbf{0} \in \mathbb{R}^{n\times n}$, then

$$\mathrm{im}(B) \subseteq \ker(A) \tag{3.15}$$

and we set

$$H_{A,B} := \ker(A)/\operatorname{im}(B), \tag{3.16}$$

for the homology group with respect to A and B.
Accordingly, their *Hodge Laplacian* is

$$L_{A,B} := A^T A + BB^T. \tag{3.17}$$

Note that $L_{A,B}^T = L_{A,B}$.

Theorem 3.1 *Let A and B be as above. Then, the following hold*

1. $H_{A,B} \cong \ker(L_{A,B})$;
2. $\ker(L_{A,B}) = \ker(A) \cap \ker(B^T)$;
3. $\operatorname{im}(L_{A,B}) = \operatorname{im}(A^T) \oplus \operatorname{im}(B)$;
4. *(Hodge Decomposition) There is an orthogonal direct sum decomposition*

$$\mathbb{R}^n \cong \operatorname{im}(A^T) \oplus \ker(L_{A,B}) \oplus \operatorname{im}(B); \tag{3.18}$$

5. *Hodge Decomposition versus Fredholm Alternative*

$$\mathbb{R}^n = \overbrace{\operatorname{im}(A^T) \oplus \underbrace{\ker(L_{A,B})}_{\ker(A)}}^{\ker(B^T)} \oplus \operatorname{im}(B). \tag{3.19}$$

Proof See [68].

Now, by substituting $A = \partial_q$ and $B = \partial_{q+1}$ in Theorem 3.1 we obtain the classical result on Hodge decomposition.

Theorem 3.2 *Let ∂_q, ∂_{q+1} and L_q be as above. Then, the following hold*

1. $H_q \cong \ker(L_q)$;
2. $\ker(L_q) = \ker(D_q) \cap \ker(D_{q+1}^T)$;
3. $\operatorname{im}(L_q) = \operatorname{im}(D_q^T) \oplus \operatorname{im}(D_{q+1})$;
4. *(Hodge Decomposition) There is an orthogonal direct sum decomposition*

$$C_q = \operatorname{im}(D_q^T) \oplus \ker(L_q) \oplus \operatorname{im}(D_{q+1}); \tag{3.20}$$

5. *Hodge Decomposition versus Fredholm Alternative*

$$C_q = \overbrace{\operatorname{im}(D_q^T) \oplus \underbrace{\ker(L_q)}_{\ker(D_q)}}^{\ker(D_{q+1}^T)} \oplus \operatorname{im}(D_{q+1}). \tag{3.21}$$

We will present some interesting applications of the Hodge decomposition in Sect. 3.6.

3.4 Multiscale Topology a.k.a. Persistent Homology

As described in Sect. 3.3, shapes can be associated with homology groups capturing topological features. Instead, the idea of persistent homology [16, 39] is to deal with, not simply the homology group at a specific scale, but rather the homology groups of filtrations of shapes, that is shapes at multiple scales. This section aims at introducing persistent homology along with the most known persistent homology summaries. Theoretical comparisons among such summaries are provided. The same notions will be extensively used in Sect. 3.5 to discuss stability results for persistence.

3.4.1 Persistent Homology

As described in Sect. 3.2.1, given a dataset there are multiple ways in order to assign to it the structure of a simplicial complex and, therefore, to study the "shape" of the original dataset by computing the homology of the associated complex. Certainly, in most of cases, the presented construction strategies depend on the choice of a scalar parameter ϵ while the considered dataset gives us no clue as to which ϵ is preferable. In spite of this, fixing a construction strategy and varying the value ϵ produces a collection φ of encapsulated simplicial complexes to be studied in a multiscale framework. Since our simplicial complexes are always finite, we can restrict to consider finite collections of encapsulated simplicial complexes were indexes take integer values only.

A *filtration* of Σ is a finite sequence of subcomplexes $\varphi := \{\Sigma^a \,|\, 0 \leq a \leq m\}$ of Σ such that $\emptyset = \Sigma^0 \subseteq \Sigma^1 \subseteq \cdots \subseteq \Sigma^m = \Sigma$. We refer to Σ^a as the *step* a in the filtration φ. In Fig. 3.11, we see a filtration obtained from a set V of 7 points by constructing the VR-complex $VR_\epsilon(\mathrm{V})$ for 7 selected values of parameter ϵ (see Sect. 3.2.2.1).

Let C_q^a be the short notation for $C_q(\Sigma^a; \mathbb{F})$ introduced in Sect. 3.3.1. Consider an inclusion $\Sigma^a \subseteq \Sigma^b$. Then, we get induced a map $f_q^{a,b} : C_q^a \longrightarrow C_q^b$ defined over each q-simplex σ of the basis of C_q^a by $f_q(\sigma) = \sigma$ and then extended linearly. The linear map $f_q^{a,b}$ is a *q-chain map* since it satisfies

$$f_q(\partial_q \sigma) = \partial_q f_q(\sigma). \tag{3.22}$$

This implies obviously the *homology functoriality*. Indeed, Equation (3.22) ensures that f_q preserves q-cycles and q-boundaries. Hence, the homology construction can be applied not simply to all filtration steps in φ to obtain a collection

of vector spaces $H_q(\Sigma^a; \mathbb{F})$, shortly indicated by H_q^a, but also to all inclusion maps to get maps $i_q^{a,b} : H_q^a \longrightarrow H_q^b$ at homology level defined by

$$i_q^{a,b}([c]) = [f_q^{a,b}(c)], \quad \text{for any } k\text{-cycle } c. \tag{3.23}$$

More generally, the homology functoriality holds due to the inclusion $\Sigma^a \subseteq \Sigma^b$ being a *simplicial map*. That is, a map $s : \Sigma \longrightarrow \Sigma'$ satisfying

$$\text{for all } \sigma, \tau \in \Sigma, \quad \sigma \subset \tau \quad \Rightarrow \quad s(\sigma) \subset s(\tau).$$

Thus, given a filtration $\varphi = \{\Sigma^a \mid 0 \leq a \leq m\}$, we can define the following. For $a, b \in \{0, \ldots, m\}$ such that $a \leq b$, the (a, b)-*persistent* q-*homology group* $H_q^{a,b}(\Sigma)$ is

$$H_q^{a,b}(\varphi) := \text{im}(i_q^{a,b}). \tag{3.24}$$

Since homology captures cycles in a shape by factoring out the boundary cycles, *persistent homology* [16, 39] captures whether cycles that are non-boundary elements in a certain step of the filtration and will turn into boundaries in some subsequent step. The persistence, along a filtration, of a cycle gives quantitative information about the relevance of the cycle itself for the shape.

3.4.2 Representing: from Persistence to Topological Summaries

The purpose of this section is that of introducing the most relevant ways of representing the persistent homology information. Our main focus is on the information encoded by each descriptor. Thus, we begin by exposing the algebraic correspondence that provides a complete and standard way of encoding the whole persistent homology information. We review some equivalent descriptors such as persistence diagrams [38], barcodes, and persistent Betti numbers [44, 94]. We will go through the diverse summaries by following the guiding example in Fig. 3.11. Afterwards, we move to other summaries encoding the persistent homology information in a more structured way. Statistical approaches require desciptor spaces being suitable not simply for measurements (covered in Sect. 3.5) but where collections of descriptors can be summed up, weighted and averaged. We close this part by reviewing some proposals in statistical and learning direction to encode the persistent homology information, namely persistence landscapes [13], persistence images [1] and other kernel representations. A comparison of persistent homology summaries over the same example is contained Fig. 3.10.

In order to deal with the totality of the persistence information is common to consider the following definition.

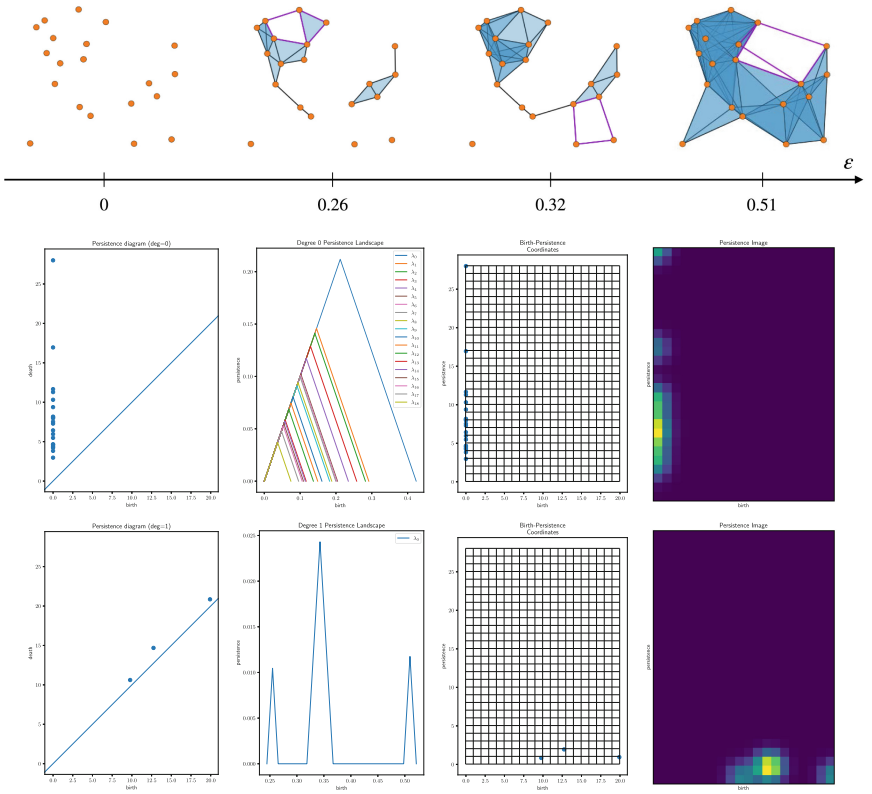


Fig. 3.10 A Vietoris-Rips filtration for increasing values of the parameter ε (top). From left to right the persistence diagram, persistence landscapes, persistence surface, and persistence image, for homology degree 0 (middle) and 1 (bottom)

Given a filtration $\varphi = \{\Sigma^a \mid 0 \leq a \leq m\}$, the q^{th}*-persistence module* $H_q(\varphi)$ is the collection of the groups H_q^a, with $0 \leq a \leq m$, connected by the linear maps $i_q^{a,b}$, with $0 \leq a \leq b \leq m$.

As explained in [16], a q^{th}-persistence module $H_q(\varphi)$ can be thought of as a finitely generated graded $\mathbb{F}[x]$-module $M = \bigoplus_{a \in \mathbb{N}} H_q^a$ where, each H_q^a is taken as the set of homogeneous elements of grade a, and the action $x^{b-a}h$ over an element h of grade a is defined by $i_q^{a,b}(h)$. Hence, $x^{b-a}h$ belongs to H_q^b for all $h \in H_q^a$, that is the graded module definition is met.

Theorem 3.3 (Structure Theorem) *Any finitely generated graded* $\mathbb{F}[x]$*-module can be decomposed as a finite direct sum of finitely generated* $\mathbb{F}[x]$*-modules as follows*

$$M \cong \bigoplus_{i=1}^{\alpha} x^{a_i}\mathbb{F}[x] \oplus \bigoplus_{j=1}^{\beta} \left(\left(x^{b_j}\mathbb{F}[x]\right) / \left(x^{c_j - b_j}\mathbb{F}[x]\right)\right), \tag{3.25}$$

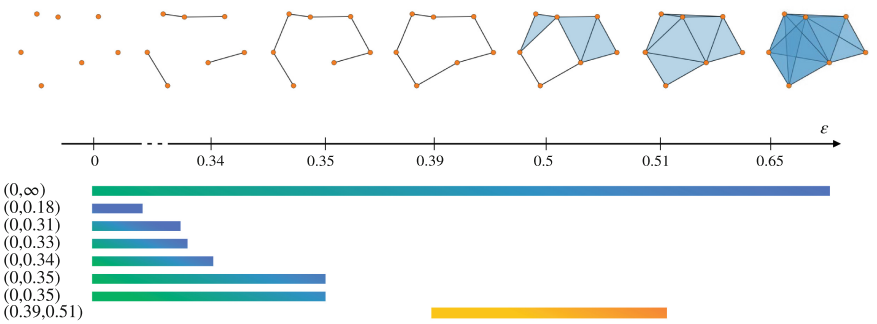


Fig. 3.11 A filtration with the corresponding persistence pairs and barcode. All bars starting at 0 represent homological classes of degree 0. At the bottom, only one bar corresponds to the life-span of a cycle

where $a_i, b_i, c_i \in \mathbb{N}$ *with* $b_i < c_i$*, and, for any* $n \in \mathbb{N}$, $x^n\mathbb{F}[x]$ *is the ideal generated by* x^n *inside* $\mathbb{F}[x]$*. Furthermore, the decomposition is unique up to reordering of the summands.*

The summands entailing only one parameter a_i' capture the generators of M appearing at grade a_i', whereas the other summands capture generators appearing at grade a_j and disappearing at grade b_j. The former kind of summands form what is called the free part of the module decomposition, the latter kind form the torsion part.

By applying the Structure Theorem to the graded $\mathbb{F}[x]$-module $\bigoplus_{a\in\mathbb{N}} H_q^a$ associated with the q^{th}-persistence module $H_q(\varphi)$, we can encode, up to isomorphisms, the entire information of persistent homology by means of pairs of two kinds. Each summand in the free part provides us with a pair (a, ∞), representing the persistence of a single homology class appearing at step a and never vanishing along the filtration. Each summand in the torsion part provides us with a pair (a, b), for a single class appearing at step a and vanishing at step b. The example in Fig. 3.11 has only the pair $(0, \infty)$ for the homology degree 0 which never vanishes. All the other persistence pairs in the example vanishes.

Tame persistence module

In order to obtain persistence modules from a filtering function as introduced in Sect. 3.2.2.3 and in view of Sect. 3.5 on the stability of persistence summaries, we generalize the persistence module definition to indexes in the set of real numbers $\mathbb{R}$ totally ordered under the standard order relation $\leq$.

An $\mathbb{R}$-*persistence module* [23] $\mathscr{F}$ with coefficients in $\mathbb{F}$ and real indexes consists of a collection $\{F^a\}_{a\in\mathbb{R}}$ of $\mathbb{F}$-vector spaces along with a collection $\{f^{a,b}\}_{a\leq b\in\mathbb{R}}$ of linear maps such that

$$\begin{aligned} \forall a \leq b \leq c \in A, \quad & f^{a,c} = f^{b,c} \circ f^{a,b} \\ \forall a \in A, \quad & f^{a,a} = \mathrm{id}_{F^a}. \end{aligned} \tag{3.26}$$

An $\mathbb{R}$-persistence module $\mathscr{F}$ is said to be *tame* if and only if,

$$\begin{aligned} \forall a \in A, \quad & \dim_{\mathbb{F}} F^a < +\infty \\ f^{a,b} \quad & \text{is an isomorphism,} \\ & \text{except for a finite number of index pairs} (a, b). \end{aligned} \tag{3.27}$$

The notion of an $\mathbb{R}$-persistence module generalizes the one with finite natural indexes introduced at the beginning of this section. Indeed as an example, the q^{th}-persistence module $H_q(\varphi)$ of a filtration $\varphi = \{\Sigma^a \mid 0 \leq i \leq m, a \in \mathbb{N}\}$ with $m \in \mathbb{N}$ can be made into an $\mathbb{R}$-persistence module. First, we extend the indexes to the integers $\mathbb{Z}$. The collection of groups $\{H^i\}$ is extended by setting $H^i = 0$ if $i < 0$, and $H^i = H^m$ if $i > m$. The collection of linear maps $\{f^{i,j}\}$ is extended by setting $f^{i,i+1} = 0$ if $i + 1 < 0$, and $f^{i,i+1}$ the identity map if $i > m$. Then, all possible compositions define $f^{i,j}$ for general indexes $i \leq j$. The extension to indexes in $\mathbb{R}$ is obtained by setting, for all real numbers $a \leq b$, $H^a = H^i$ with the integer i such that $i \leq a < i + 1$, $H^b = H^j$ with the integer j such that $j \leq b < j + 1$, and $f^{a,b} = f^{i,j}$. Moreover, it is obvious to check that the obtained $\mathbb{R}$-persistence module is tame.

This extension allows us to consider relevant examples of tame $\mathbb{R}$-persistence modules such as the one obtained via sublevel sets of a filtering function introduced in Sect. 3.2.2.3, or the ones obtained via filtrations from point clouds, as the one in Fig. 3.11, introduced in Sect. 3.2.2.1, or the ones from graph data introduced in Sect. 3.2.2.2.

Persistence diagrams

Fix the notation $\bar{\mathbb{R}} = (\mathbb{R} \cup \{\infty\})$ and $\bar{\mathbb{N}} = (\mathbb{N} \cup \{\infty\})$. Consider a tame $\mathbb{R}$-persistence module $\mathscr{F}$. The finite multi-set of points obtained from the Structure Theorem applied to $\mathscr{F}$ defines the *persistence diagram* [38] as

$$\begin{aligned} & \mathbb{R} \times (\mathbb{R} \cup \{\infty\}) \xrightarrow{\mathrm{PD}(\mathscr{F})} \mathbb{N} \cup \{\infty\} \\ \text{if } a \neq b \quad & (a, b) \longmapsto \#\{\text{summands in Equation (13) corresponding to } (a, b)\} \\ \text{if } a = b \quad & (a, a) \longmapsto \infty. \end{aligned} \tag{3.28}$$

The elements belonging to the persistence diagram are called *persistence pairs*. By the Structure Theorem, the persistence diagram is a complete invariant for a persistence module. The persistence diagram of the filtration in Fig. 3.11 consists of the pairs depicted in column on the left. Notice that the pair $(0, 0.35)$ appears with multiplicity 2.

An invariant for a persistence module equivalent to the persistence diagram is the *barcode*. The barcode $\mathrm{Bar}(\mathscr{F})$ is the finite collection $\{(b_i, d_i)\}_{i \in I}$ of points (b_i, d_i) varying in the support of $\mathrm{PD}(\mathscr{F})$ counted $\mathrm{PD}(\mathscr{F})(b, d)$ times. The barcode of the fil-

tration in Fig. 3.11 consists of the bars of different length according to the persistence of the corresponding homology class as the value of the parameter ϵ varies.

As treated in [23], the tameness condition ensures that $\mathscr{F}$ admits a representation as a finitely generated module by a suitable reindexing. Hence, the definition in Equation (3.28) applies. The tameness condition ensures that a sequence of nested discretizations of $\mathscr{F}$ converges to a unique persistence diagram with respect to a similarity notion between persistence modules called the interleaving distance which will be introduced in Sect. 3.5.

Persistent Betti numbers

Given a tame persistence module $\mathscr{F} = (\{F^a\}, \{f^{a,b}\})$ with coefficients in $\mathbb{F}$ and indexes in $(\mathbb{R}, \leq)$, we define the *persistent Betti numbers* [44] of $\mathscr{F}$ as the function $\beta : \mathbb{R} \times \bar{\mathbb{R}} \to \mathbb{N}$ defined by

$$\beta(a, b) := \begin{cases} \dim f^{a,b}(F^a) & \text{if } a \leq b \\ 0 & \text{if } a > b. \end{cases} \tag{3.29}$$

Equivalently, given the barcode $\text{Bar}(\mathscr{F}) = \{(b_i, d_i)\}_{0 \leq i \leq N}$, we obtain the same function by defining

$$\beta(a, b) = \#\{0 \leq i \leq N \mid \text{s.t. } b_i \leq a \leq b \leq d_i\}. \tag{3.30}$$

The Representation Theorem in [21] , or the Triangle Lemma in [27], guarantees that knowing β is equivalent to knowing the persistence diagram. For a given homology degree, the value at (a, b) of the persistent Betti numbers of the filtration can be easily read off from Fig. 3.11 by counting the number of bars of full length in the ϵ range from a to b.

Persistence landscapes

A limitation of a persistence diagram in encoding the persistence information is that the mean of multiple persistence diagrams as functions might be ambiguous. To overcome this, *persistence landscapes* are introduced in [13].

The *persistence landscape* of $\mathscr{F}$ is the function $\lambda : \mathbb{N} \times \mathbb{R} \to \bar{\mathbb{R}}$ such that

$$\lambda(k, t) := \sup\{m \geq 0 \mid \beta(t - m, t + m) \geq k\}. \tag{3.31}$$

Notice that the transformation $t = \frac{a+b}{2}, m = \frac{b-a}{2}$ sends the diagonal of the persistence diagram domain to the horizontal axis as we can see by comparing persistence diagrams and landscapes in Fig. 3.12.

The persistence landscapes are invertible in the sense that the persistent diagram can be recovered from the persistence landscapes. Indeed, persistent Betti numbers $\beta(a, b)$ at a point (a, b) can be recovered. Consider the set of integers $I_{(a,b)} = \{k \mid \text{s.t. supp}\lambda(k, \cdot) \supseteq [a, b]\}$. It follows that $\beta(a, b) = \sum_{k \in I_{(a,b)}} \lambda(k, \frac{a+b}{2})$. The definition of the function λ can be extended to the domain $\mathbb{R}^2$ by setting $\lambda(\lceil x \rceil, t)$.

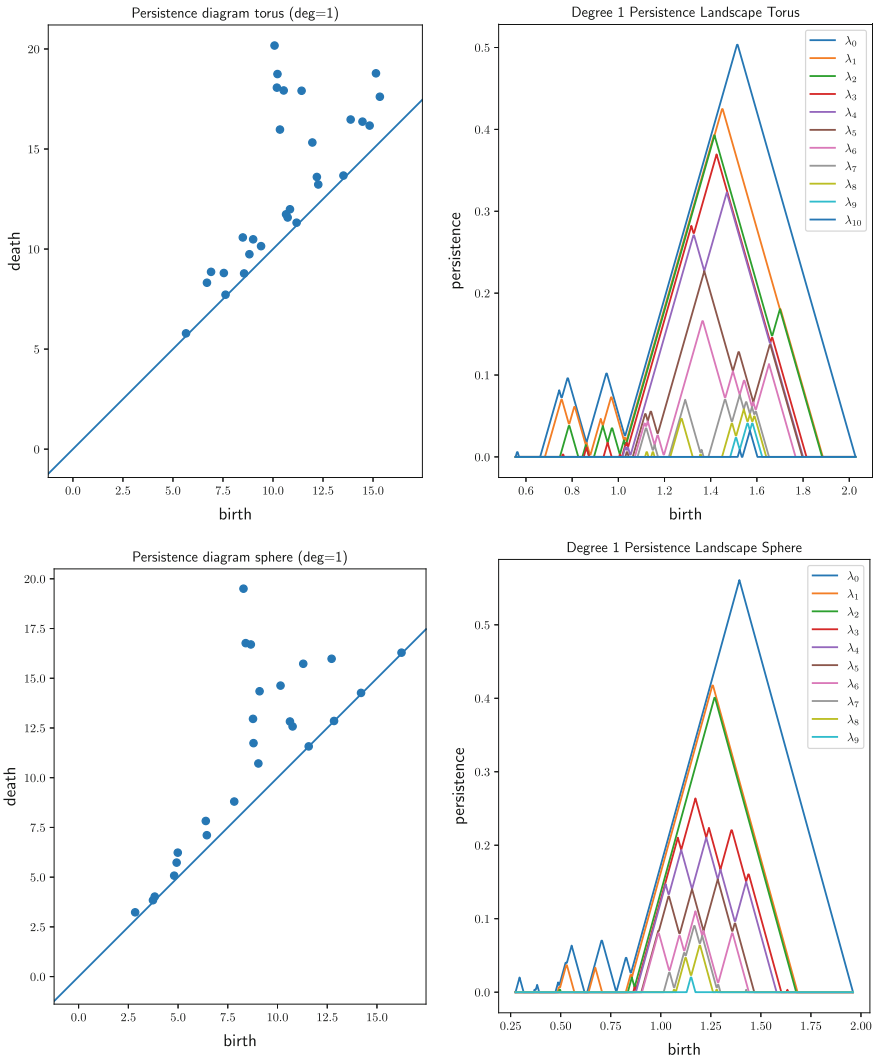


Fig. 3.12 Persistence diagrams and landscapes for the first homology degree of a sampling of 100 points over a torus (top) and a sphere (bottom)

The *mean* $\bar{\lambda}$ of $\lambda^1, \ldots, \lambda^n$ persistence landscapes is defined by

$$\bar{\lambda}(x,t) := \frac{1}{n}\sum_{i=1}^{n} \lambda^i(x,t). \tag{3.32}$$

Persistence images

Recently, a very powerful summary for persistence has been introduced, called *persistence images* [1]. Actually, their representation approximizes the information in a persistence diagram and a persistence image is not invertible in the sense of persistence landscapes.

Nonetheless, the key strength point is that persistence images represent persistence as a finite dimensional vector in $\mathbb{R}^n$. The finite dimensional representation makes persistence images a favorable way for integrating persistence into deep neural architectures thus opening the way for an interplay between deep learning and TDA (see Sect. 3.6).

A persistence diagram D is transformed into a *persistence surface* $\rho_D : \mathbb{R}^2 \to \mathbb{R}$

$$\rho_D(x) := \sum_{y \in T(D)} \omega(y) k_G(x, y), \tag{3.33}$$

where $T(D)$ is the set of birth-persistence pairs obtained from D by applying $(a, b) \mapsto (a, b - a)$, the weight function $\omega : \mathbb{R}^2 \to \mathbb{R}$ is zero on the horizontal axis and step-wise differentiable, the function k_G is the Gaussian kernel defined by $k_G(x, y) = \exp -\frac{\|x-y\|^2}{2\sigma^2}$, for some parameter σ.

The *persistence image* $I(\rho_D)$ of D is obtained from ρ_D by subdividing its domain into n pixels P by means of a regular grid and by computing the integral of ρ_D over each P. In Fig. 3.13, we see an example with the birth-persistence diagram $T(D)$ and the corresponding persistence image obtained by discretization.

Homological scaffolds

Homological scaffolds are effective and compact summaries of the homological features of weighted networks capable of simultaneously make their homological properties amenable to networks theoretical methods [87].

Given a complex network represented as a weighted graph $G = (V, E, w : E \to \mathbb{R})$, let B be a collection of representative cycles of the 1-dimensional homological classes occurring during the filtration φ of Flag complexes associated to G. Given an edge e of G, it is possible to associate to it a weight $\omega(e)$ defined as it follows:

$$\omega(e) := \sum_{g \in B, e \in g} \omega_g, \tag{3.34}$$

where ω_g is a weight value associated to the homological class of g. Based on the weight ω, the *homological scaffold* of the network G is defined as the weighted graph $G' = (V', E', \omega : E' \to \mathbb{R}_{>0})$ where:

- $V' := V$;
- E' consists of the edges e of E for which $\omega(e) > 0$.

The scaffold is called *persistence homological scaffold* by defining ω_g as the lifespan π_g of g, while it is called *frequency homological scaffold* by setting, for any cycle g, $\omega_g := 1$.

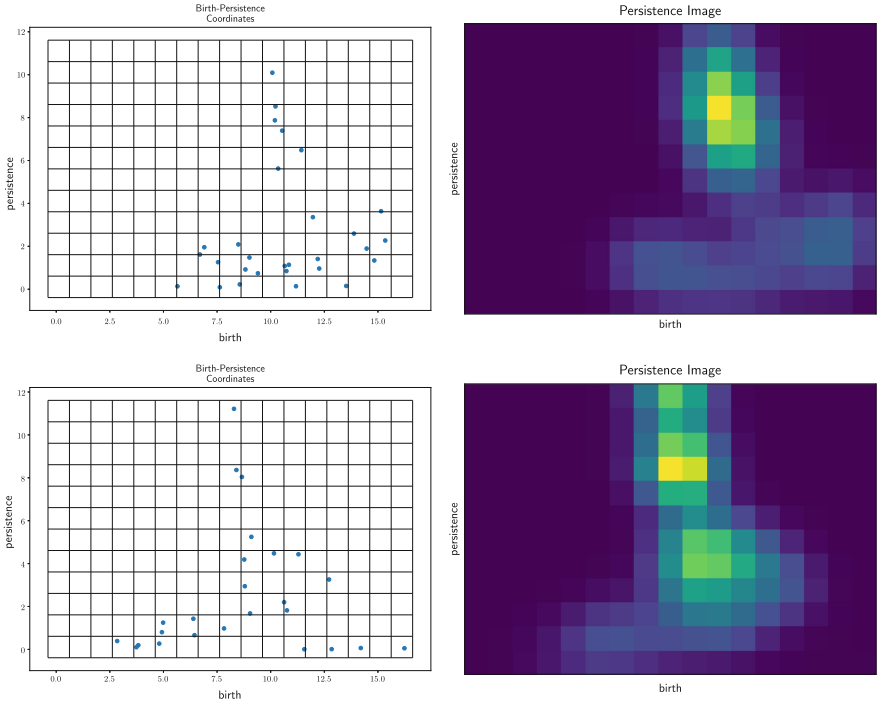


Fig. 3.13 Birth-persistence diagrams and persistence images for the first homology degree of a sampling of 100 points over a torus (top) and a sphere (bottom)

Despite their proven applicative capabilities, (persistence and frequency) homological scaffolds theoretically depend on the choice of representative cycles. In order to address this potential issue and thanks to the recent advances in the computation of minimal homology bases [32], a quasi-canonical version of the scaffold, called minimal, has been introduced in [48].

Formally, given a complex network $G = (V, E, w : E \to \mathbb{R})$, let us consider the minimal homology basis B^ϵ of $H_1(Flag(G_\epsilon))$ (i.e. a collection of representative cycles of minimal total length whose classes form a basis of $H_1(Flag(G_\epsilon))$). Moreover, let us define B^* as $\bigsqcup_\epsilon B^\epsilon$. Similarly to the frequency homological scaffold, the *minimal homological scaffold* of G is defined as the weighted graph $G' = (V', E', \omega : E' \to \mathbb{R}_{>0})$ where:

- $V' := V$;
- $\omega(e) := \sum_{g \in B^*, e \in g} 1$ (i.e. the number of cycles in B^* containing e);
- E' consists of the edges e of E for which $\omega(e) > 0$.

In Fig. 3.14, an example of the minimal scaffold consisting of minimal 1-homology generators appeared along the filtration is depicted.

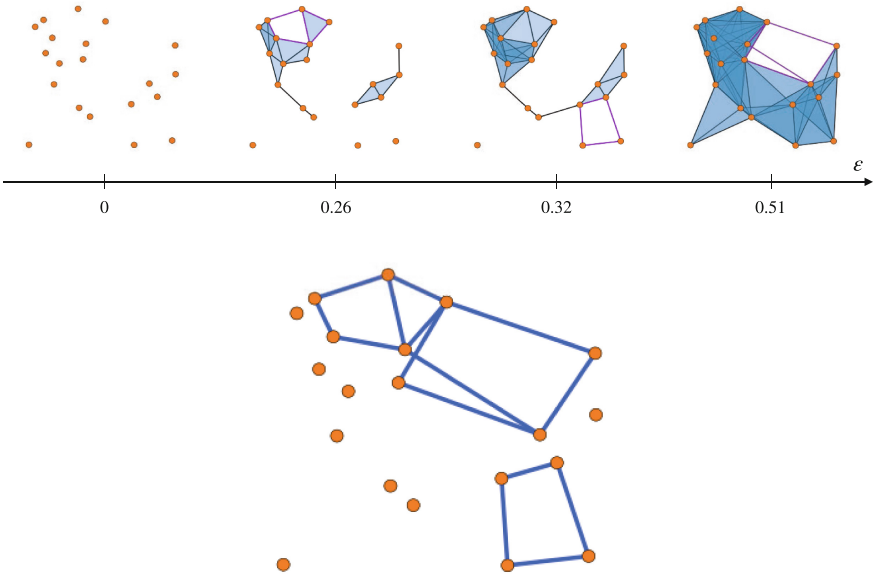


Fig. 3.14 At the top, filtration with minimal 1-homology generators (loops) highlighted. Below, the associated minimal scaffold. Image available under courtesy of the authors of [48]

3.4.2.1 Representing the Topological Information Through Kernels

Topological information and, specifically, persistence diagrams are effective data discriminants. On the other hand, interfacing persistence diagrams directly with statistics and Machine Learning poses technical difficulties, because the space of persistence diagrams is not endowed with a structure of an inner product or, more properly, of a Hilbert space structure. This lack prevents that lengths, angles, and means of persistence diagrams can be defined as well as that kernel-based learning methods such as kernel PCA and support vector machines for classification [92] can be adopted.

Given an input space X, a *kernel* for X is a map $k : X \times X \to \mathbb{R}$ such that there exist a Hilbert space H and a map $\phi : X \to H$, called *feature map* for which, for any $x, y \in X$, $k(x, y) = \langle \phi(x), \phi(y) \rangle$. Equivalently, k is a kernel if the following diagram commutes:

$$\begin{array}{ccc} X \times X & \xrightarrow{k} & \mathbb{R} \\ {\scriptstyle \phi \times \phi}\downarrow & \nearrow_{\langle \cdot, \cdot \rangle} & \\ H \times H & & \end{array} \tag{3.35}$$

Recall that a Hilbert space H is a complete metric space with respect to the distance induced by the inner product $\langle \cdot, \cdot \rangle$. A widely adopted assumption for an

inner product $\langle \cdot, \cdot \rangle$ is the request of being positive semi-definite. In such a case, a kernel will be called reproducing kernel.

Worth to be mentioned that the definition a kernel k does not require the explicit knowledge of its feature map ϕ but just a theoretical guarantee of its existence. Results as the theorems of Moore-Aronszajn and of Mercer characterize the conditions under which a function $k : X \times X \to \mathbb{R}$ can serve as a kernel [59]. For instance, Moore-Aronszajn's theorem claims that any finite function $k : X \times X \to \mathbb{R}$ is a reproducing kernel as long as it is finite, symmetric, and positive semi-definite.

The idea of defining a kernel for the space X of persistence diagrams has been introduced in the late nineties in [33, 43] but it has been widely adopted just in the last few years. In the framework of persistence-based kernels, the feature map is usually explicitly provided and the chosen Hilbert space is typically the $L^2(\mathbb{R}^2)$ space of the square-integrable functions. In the following, we will list and briefly discuss the main features of all (to the best of our knowledge) persistence-based kernels introduced in the literature.

Roughly, we can classify them into three groups: persistence landscapes [13]; Gaussian kernels [1, 61, 92]; sliced Wasserstein kernels [19].

Persistence landscapes as kernels

Persistence landscapes, introduced in [13] and previously described in Sect. 3.4.2, allow for representing any persistence diagram as a square-integrable function. Even if not previously mentioned, the procedure assigning a L^2 function $\lambda(D)$ to a persistence diagram D can be considered as a feature map from the space of the persistence diagrams to $L^2(\mathbb{R}^2)$. According with this perspective, a kernel k based on persistence landscapes can be plainly defined as

$$k(D, D') = \|\lambda(D) - \lambda(D')\|_{L^2}, \tag{3.36}$$

where D and D' are two persistence diagrams.

Gaussian kernels

The second class of persistence-based kernels is defined thanks to the explicit introduction of a feature map ϕ. In [92] and in [61], a similar approach is adopted. Focusing on [92], let denote as Ω the region $\{x = (x_1, x_2) \in \mathbb{R}^2 \,|\, x_2 \geq x_1\} \subseteq \mathbb{R}^2$ above the diagonal and as δ_x the Dirac delta centered at the point x. The feature map ϕ will be defined on a given persistence diagram D, as the solution $u : \Omega \times \mathbb{R}_{\geq 0} \to \mathbb{R}$ of the partial differential equation

$$\begin{cases} \Delta_x u = \frac{\partial u}{\partial t} & \text{in } \Omega \times \mathbb{R}_{\geq 0} \\ u = 0 & \text{on } fr(\Omega) \times \mathbb{R}_{\geq 0} \\ u = \sum_{y \in D} \delta_y & \text{in } \Omega \times \{0\}. \end{cases} \tag{3.37}$$

A solution of the above equation is achieved by posing

$$\phi_\sigma^D(x) := \frac{1}{4\pi\sigma} \sum_{y \in D} \left(e^{\left(-\frac{\|x-y\|^2}{4\sigma}\right)} - e^{\left(-\frac{\|x-y'\|^2}{4\sigma}\right)} \right), \tag{3.38}$$

with $y' = (b, a)$ if $y = (a, b)$.

So, we obtain the following explicit expression for the associated kernel

$$k(D, D') = \frac{1}{8\pi\sigma} \sum_{y \in D, z \in D'} \left(e^{\left(-\frac{\|y-z\|^2}{8\sigma}\right)} - e^{\left(-\frac{\|y-z'\|^2}{8\sigma}\right)} \right). \tag{3.39}$$

Intuitively, given a persistence diagram D, the above described procedure returns an L^2 function $\phi(D)$ having a Gaussian peak centered at each point of the considered diagram D. Moreover, in order to obtain a stable kernel, the height of each peak is set as proportional to the distance of the corresponding point (a, b) from the diagonal of the first quadrant of $\mathbb{R}^2$.

Even if still based on Gaussian peaks, a different approach has been adopted in [1]. As previously mentioned, persistence images enable to convey the persistent homology information through a vector. This is achieved by: (i) transforming any persistence diagram D into a surface ρ_D obtained by centering at each point of D a suitably weighted Gaussian peak; (ii) discretizing ρ_D by decomposing the domain into a regular grid of pixels; (iii) representing the obtained result through a heat map $I(\rho_D)$. Among other advantages, such a vectorization process promptly provides a notion of kernel. In fact, given two persistence diagrams D and D', one can define their kernel simply as the inner product of their associated vectors $I(\rho_D)$ and $I(\rho_{D'})$:

$$k(D, D') := \langle I(\rho_D), I(\rho_{D'}) \rangle. \tag{3.40}$$

Sliced Wasserstein kernel

In various contexts, a standard way to construct a kernel is to exponentiate the negative of an Euclidean distance. In [19], the authors aim at adopting a similar approach in order to define a kernel between persistence diagrams. According with this idea, given two persistence diagrams D and D' a kernel can be defined as:

$$k(D, D') := e^{\left(-\frac{f(D,D')}{2\sigma^2}\right)}. \tag{3.41}$$

An important theorem in [8] (Theorem 3.2.2, page 74) ensures that the above formula defines a valid positive definite kernel for any $\sigma > 0$ if and only if f is negative semi-definite. Unfortunately, none of the already introduced distances between persistence diagrams can serve as f since does not satisfy the property of negative semi-definiteness. In order to overcome this limitation, a new distance d_{SW}, called *sliced Wasserstein distance*, has been introduced in [19]. Sliced Wasserstein distance

represents an approximation of 1^{st}-Wasserstein distance and, thanks to the fact of being provably negative semi-definite, it can be chosen as f in Eq. (3.41) in order to define a valid positive definite kernel of persistence diagrams.

3.5 Comparing Multiscale Summaries: Distances and Stability

Stability is a concept rooted in celestial mechanics and, thereafter, in the study of dynamical systems. Loosely speaking, a quantity related to the system is stable if small perturbations will result in relatively small changes of it. In this spirit stability has been studied also in the area of persistence. Namely, several results have been obtained about the behaviour of the topological summaries presented in the previous sections with respect to perturbations of the input data. In particular, this has been addressed by comparing suitable set distances variations, as for example the Gromov-Hausdorff distance, with the corresponding modifications of their e.g. persistence diagrams and the other summaries. This section is devoted to introduce and elucidate this topic and the main results on it.

3.5.1 Distances on Data

In this section, we collect the definitions of the dissimilarity measures applying to input data within the persistence pipeline (see Sect. 3.2). The first one is the *natural pseudo-distance* [62] and it applies to continuous functions used to filter a domain by sublevel sets. In the case of continuous functions over the same domain, the natural pseudo-distance is usually replaced by the actual distance defined by the L_∞-norm.

The second distance to be introduced is the *Gromov-Hausdorff* [24] distance which applies to finite metric spaces such as the point clouds discussed in Sect. 3.2.2.1.

Natural pseudo-distance

A topological space X is called *triangulable* [78] if there exists a (finite) simplicial complex homemorphic to X. A continuous function $f : X \longrightarrow \mathbb{R}$ is called *tame* if and only if it induces a filtration on X via sublevel sets of the form $X^a = f^{-1}((-\infty, a])$ (see Sect. 3.2.2.3) such that each X^a has finite dimensional homology. Notice that the tameness condition on functions implies the tameness condition for the $\mathbb{R}$-persistence module obtained through sublevel sets discussed in Sect. 3.4.2. A pair (X, f) consisting of a triangulable space X and a tame function is called a *size pair*.

Let (X, f) be a size pair. The *natural pseudo-distance* [62] between two pairs (X, f), (Y, g) is defined by

$$\mathrm{d_N}\left((X, f), (Y, g)\right) := \begin{cases} \inf_{h \in \mathrm{Hom}(X,Y)} \|f - g \circ h\|_\infty, & \text{if } \mathrm{Hom}(X, Y) \neq \emptyset \\ +\infty & \text{if } \mathrm{Hom}(X, Y) = \emptyset, \end{cases} \tag{3.42}$$

where $\mathrm{Hom}(X, Y)$ is the set of all homeomorphisms from X to Y.

In the case $Y = X$, it is preferable to consider the L_∞-*norm* inducing an actual distance function between f and g instead of their natural pseudo-distance. Indeed, if there exists an homeomorphism h such that $f = g \circ h$, then $\mathrm{d_N}((X, f), (X, g)) = 0$ even though their L_∞-norm is in general non-trivial.

Gromov-Hausdorff distance

A *finite metric space* is a pair (X, d_X) consisting of a finite set X and a metric function $d_X : X \times X \longrightarrow \mathbb{R}$. For example, we can think of X endowed with the metric function inherited from some embedding into a finite dimensional Euclidean space. A *correspondence* r *between* X and Y is a subset of $X \times Y$ such that the canonical projection from the Cartesian product $\pi_X : X \times Y \longrightarrow X$ and $\pi_Y : X \times Y \longrightarrow Y$ are both surjective. The set of all correspondences between X and Y will be denoted by $X \rightleftharpoons Y$. A correspondence generalizes a bijection between two sets with different cardinalities.

Once we have a correspondence $r \in X \rightleftharpoons Y$ between the underlying sets, we can define its *distorsion* as

$$\mathrm{Dis}(r) := \sup\{|d_X(x, y) - d_Y(x', y')| \mid (x, y), (x', y') \in r\}. \tag{3.43}$$

The *Gromov-Hausdorff distance* [47] between two finite metric spaces (X, d_X), (Y, d_Y) is defined as

$$\mathrm{d_{GH}}\left((X, d_X), (Y, d_Y)\right) := \frac{1}{2} \inf\{\mathrm{Dis}(r) \mid r \in X \rightleftharpoons Y\}. \tag{3.44}$$

3.5.2 Distances on Persistence Summaries

In this section, we collect the definitions of the dissimilarity measures applying to persistent homology summaries (see Sect. 3.4.2).

We first review the *interleaving distance* [23] applying to $\mathbb{R}$-persistence modules. Afterwards, we review the *p-landscape distance* [13] applying in the same way to both $\mathbb{R}$-persistence modules and persistence diagrams. Later, we review the distances on persistence diagrams: the *matching/bottleneck distance* [27, 62], the p^{th}*-Wasserstein distance* [28], and the *Hausdorff distance* [27].

We close this section by stating the Isometry Theorem [14, 65] for relating the interleaving distance and the introduced distances on diagrams.

Interleaving distance

Let $\mathscr{F}, \mathscr{G}$ be $\mathbb{R}$-persistence modules as in Eq. (3.26) and $\epsilon > 0$. We say that $\mathscr{F}$ and $\mathscr{G}$ are ϵ-*interleaved* if and only if there exists a family $\{\Phi^a : F^a \longrightarrow G^{a+\epsilon}\}_{a\in\mathbb{R}}$ and a family $\{\Psi^a : G^a \longrightarrow F^{a+\epsilon}\}_{a\in\mathbb{R}}$ such that the following diagrams commute

$$\begin{array}{ccc} F^{a-\epsilon} & \xrightarrow{f^{a-\epsilon,a+\epsilon}} & F^{a+\epsilon} \\ & {\scriptstyle \Phi^{a-\epsilon}} \searrow \quad \nearrow {\scriptstyle \Psi^a} & \\ & G^a & \end{array} \qquad \begin{array}{ccc} & F^{a+\epsilon} & \xrightarrow{f^{a+\epsilon,b+\epsilon}} F^{b+\epsilon} \\ {\scriptstyle \Psi^a} \nearrow & & \nearrow {\scriptstyle \Psi^b} \\ G^a & \xrightarrow{g^{a,b}} & G^b \end{array}$$

$$\begin{array}{ccc} & F^a & \\ {\scriptstyle \Psi^{a-\epsilon}} \nearrow & & \searrow {\scriptstyle \Phi^a} \\ G^{a-\epsilon} & \xrightarrow{g^{a-\epsilon,a+\epsilon}} & G^{a+\epsilon} \end{array} \qquad \begin{array}{ccc} F^a & \xrightarrow{f^{a,b}} & F^b \\ {\scriptstyle \Phi^a} \searrow & & \nearrow {\scriptstyle \Phi^b} \\ & G^{a+\epsilon} & \xrightarrow{g^{a+\epsilon,b+\epsilon}} G^{b+\epsilon}. \end{array} \tag{3.45}$$

The *interleaving distance* between real-indexed persistence modules is defined by

$$\mathrm{d_I}(\mathscr{F}, \mathscr{G}) := \inf\{\epsilon > 0 \mid \mathscr{F}, \mathscr{G} \text{ are } \epsilon\text{-interleaved}\}. \tag{3.46}$$

Distance on persistence landscapes

Let λ, λ' be the persistence landscapes associated to two persistence modules $\mathscr{F}$, $\mathscr{F}'$, respectively. Fix p a real number $p \geq 1$ or $p = \infty$. Then, the *p-landscape distance* is defined by means of the L_p-norm as follows

$$\Lambda_p(\mathscr{F}, \mathscr{F}') := \|\lambda - \lambda'\|_p. \tag{3.47}$$

In the same way, if λ, λ' are the persistence landscapes associated to two persistence diagrams D, D' we define $\Lambda_p(D, D') := \|\lambda - \lambda'\|_p$.

Distances on persistence diagrams

Let D_1 and D_2 be persistence diagrams as defined in Eq. (3.28). Some pseudo-metrics require the notion of a matching between persistence diagrams. A *matching* between D_1 and D_2 is simply a bijection between the underlying sets. Notice that the persistence diagrams are not finite metric spaces like those in Eq. (3.44). Despite there might be points with finite multiplicity in different number within two persistence diagrams, such a cardinality discrepancy is overcome by matching points to the diagonal when needed. The *bottleneck distance* [27] (also called *matching distance* [62]) is defined as

$$\mathrm{d_B}(D_1, D_2) := \inf_{\substack{\pi: D_1 \to D_2 \\ bijection}} \sup\{\|x - \pi(x)\|_\infty \mid x \in D_1\}, \tag{3.48}$$

where the $\|\|_\infty$-norm is always taken for points as part of $\mathbb{R}^2$.

The bottleneck distance can be seen as part of a family of pseudo-metrics. For any integer $1 \leq p \leq +\infty$, the p^{th}-*Wasserstein* distance [28] is defined as

$$\mathrm{d_W^p}(D_1, D_2) := \inf_{\substack{\pi: D_1 \to D_2 \\ bijection}} \left(\sum_{x \in D_1} \|x - \pi(x)\|_\infty^p \right)^{\frac{1}{p}}. \tag{3.49}$$

The *Hausdorff* distance [27] does not require a bijection between the underlying sets and it is defined as

$$\mathrm{d_H}(D_1, D_2) := \max \left\{ \sup_{x \in D_1} \{ \inf_{y \in D_2} \{ \|x - y\|_\infty \} \}, \sup_{y \in D_2} \{ \inf_{x \in D_1} \{ \|x - y\|_\infty \} \} \right\}. \quad (3.50)$$

Let (X, f), (Y, g) be two pairs consisting of a triangulable topological space and a continuous real-valued function on it. Suppose f and g are both tame functions, meaning that the respective $\mathbb{R}$-persistence modules $\mathscr{F}$ and $\mathscr{G}$ obtained by sublevel sets are tame in the sense of Equation (3.27). Let D_f, D_g be the persistence diagrams associated to the sublevel set filtrations induced by f and g, respectively. It follows that the following relations among persistence pseudo-metrics hold:

$$\mathrm{d_H}(D_f, D_g) \leq \mathrm{d_B}(D_f, D_g), \quad (3.51)$$

$$\mathrm{d_B}(D_f, D_g) = \mathrm{d_I}(\mathscr{F}, \mathscr{G}). \quad (3.52)$$

By Eq. (3.51), we get that, on persistence diagrams, a stability result with respect to the bottleneck distance implies the corresponding stability with respect to the Hausdorff distance. Equation (3.52) is usually referred to as the Isometry Theorem [14, 65].

3.5.3 Stability Results

In this section, we collect the main results concerning the stability with respect to the dissimilarity measures introduced in Sect. 3.5.1 and Sects. 3.5.2.

We begin by stating Theorem 3.4 and Theorem 3.5 for the stability in terms of bottleneck distance of persistence diagrams and natural pseudo-distance of size pairs [29]. We proceed further with Theorem 3.6 and Theorem 3.7 for extending to persistence landscapes the stability results just stated [13]. Afterwards, we state Theorem 3.8 for the stability of the bottleneck distance of persistence diagrams with respect to the Gromov-Hausdorff distance of finite metric spaces [28]. We conclude this part with a focus on the stability of the *persistence-based kernels* which have been introduced in Sect. 3.4.2.1.

Theorem 3.4 *Let (X, f), (Y, g) be two pairs consisting of a triangulable topological space and a tame function on it. Let D_f, D_g be the persistence diagrams associated to the sublevel set filtrations induced by f and g, respectively. Then, it holds that*

$$\mathrm{d_B}(D_f, D_g) \leq \mathrm{d_N}((X, f), (Y, g)).$$

In the case of $X = Y$, the stronger result with respect to the L_∞-norm applies [29].

Theorem 3.5 *Let (X, f), (X, g) be two pairs consisting of a triangulable topological space and a tame function on it. Let D_f, D_g be the persistence diagrams associated to the sublevel set filtrations induced by f and g, respectively. Then, it holds that*

$$\mathrm{d}_B(D_f, D_g) \leq \|f - g\|_\infty.$$

Notice that, by Eq. (3.52), the two stability results apply also to the interleaving distance between the corresponding persistence modules. That is usually referred to as *algebraic stability theorem* [23]. Moreover, recently the stability result has been sharpened and generalized to interval decomposable persistence modules and block decomposable $\mathbb{R}^n$-persistence modules [12]. The former ones apply to modified persistence frameworks: zig-zag persistence [76] and Reeb graphs (a survey in [10]). The latter ones apply to the the case of persistence generalized to filtrations determined by multiparameters [17] instead of a single one.

In the case of persistence landscapes, the stability is stated by the following and proved in [13].

Theorem 3.6 *Let (X, f), (X, g) be two pairs consisting of a triangulable topological space and a function on it. Let $\mathscr{F}$, $\mathscr{G}$ be the two persistence modules associated to the sublevel set filtrations induced by f and g, respectively. Then, it holds that*

$$\Lambda_\infty(\mathscr{F}, \mathscr{G}) \leq \|f - g\|_\infty.$$

In the same way, the stability of persistence landscapes with respect to the p^{th}-Wasserstein distance of persistence diagrams is also proven [13].

Theorem 3.7 *Let D_1, D_2 be two persistence diagrams. Fix p a real number $p \geq 1$ or $p = \infty$. Then, it holds that*

$$\Lambda_\infty(D_1, D_2) \leq d_W^p(D_1, D_2).$$

In the case of filtrations built on top of a finite metric space such as those introduced in Sect. 3.2.1, the following result holds [24].

Theorem 3.8 *Let (X, d_X), (Y, d_Y) be two finite metric spaces. Let D_X^{VR}, D_Y^{VR} be the persistence diagrams associated to the Vietoris-Rips complex constructions over (X, d_X), (Y, d_Y), respectively. Then, it holds that*

$$\mathrm{d}_B(D_X^{VR}, D_Y^{VR}) \leq \mathrm{d}_{GH}((X, d_X), (Y, d_Y)).$$

By applying Formula (3.5), the result can be extended to the Čech construction.

Stability of persistence-based kernels

Analogously to the the previously discussed topological summaries, it is crucial that also for the information conveyed by a persistence-based kernel to be stable. Given a (pseudo) distance $d : X \times X \to \mathbb{R}$ of a space X, a kernel k for X is called *stable* with respect to d if there exists a constant $C > 0$ such that, for any $x, y \in X$,

$$\|\phi(x) - \phi(y)\|_H \leq C \cdot d(x, y), \tag{3.53}$$

where $\| \cdot \|_H$ is the norm induced by the inner product of H.

All the described persistence-based kernels satisfy some stability result. We refer to the papers in which each kernel is introduced for the details about the specific stability properties they satisfy. As an example, we report here the stability results satisfied by the scale-space kernel. Stability of scale-space kernel with respect to 1^{st}-Wasserstein distance of persistence diagrams

$$\|\phi_\sigma(D) - \phi_\sigma(E)\|_{L^2} \leq \frac{1}{2\sigma\sqrt{\pi}} \mathrm{d}_{\mathrm{W}}^1(D, E). \tag{3.54}$$

3.6 Applications

In this section, we review applications of Topological Data Analysis (TDA) exploiting the persistent homology pipeline. Our purpose is not to provide a comprehensive list. Surveys on the topic may be found in [25, 84]. Instead, we aim at showing how TDA general properties are exploited and adapted to specific application fields. Generally speaking, TDA is well-appreciated since it:

- summarizes **multiscale** information according to the relevance of homology classes along the scale range on data;
- captures **higher-order relations** among data entities;
- captures **mesoscale** information on complex data since homology mediates between local properties (holes) inserted in global contexts (homologous holes are not necessarily close each other);
- captures **coarse** information;
- captures **robust** information.

The reader interested in computational aspects of the persistence pipeline is referred to the standard algorithm for persistence introduced in [16]. A list of softwares implementing several persistent homology algorithms comprises: PHAT [6] which is written in C++ and Python along with its distributed version DIPHA [5], the tool Ripser [4], and the more recent tool GIOTTO [98]. A persistence tool specifically designed for complex networks is jHoles [11]. The reader interested in a comprehensive comparison of algorithms and tools for persistent homology is referred to the roadmap work in [81].

Libraries for performing general TDA computations are available in C++ and Python through the Gudhi library [72] along with its R software interface [42]. A previous C++ library for persistent homology is Dionysus [77]. A comprehensive tool-kit for TDA not limited to the persistence pipeline is available at [85] in C++ and Python.

In the following part, we focus on the fields of *complex networks*, *neuroscience*, and *biology* where the contribution of TDA is well recognized. We proceed by presenting two specific scopes where the q-Hodge Laplacian (see Sect. 3.3.2) finds applications. In the final part, we present a collection of other applications to be easily accessible by domain experts.

Complex networks and social science

Most of applications in complex networks exploits the ability of TDA of capturing higher-order relations among data entities. Graphs and weighted graphs naturally encode the pairwise relation of complex networks. The simplicial complex constructions introduced in Sect. 3.2.2.2 allow to endow a (weighted) graph with a (filtered) higher-order structure.

Some applications are focused on translations of graph concepts into simplicial complex terms. In [9], the percolation is extended from graph to simplicial complexes. In [54], the clique complex is introduced. This allows for a persistence multiscale summary of higher-order structures [93] to track community cliques. This leads the authors to introducing a new non-local measure which is proven to be complementary to standard centrality and comparison measures. A classification of network models which integrates topological information is proposed in [97].

Some applications in the field of complex networks exploit the ability of persistence in capturing mesoscale information. For instance, in [89] authors apply persistence to detect new non-local structures. In [88], a new and robust filtration is introduced which is proven to be richer than both metrical and clique filtrations.

Due to the formalization into (weighted) graph terms, applications to Social Science exploit a framework similar to that of complex networks. In [2], topology is used to detect a correlation between socio-economic indicators and spatial structures in the city environment. In [99], topology is combined to non-linear dimensional reduction to provide new contagion maps. In [20], authors apply the degree 0 and 1 persistence to analyze collaboration networks and differentiate them from the random case. In [95], persistence is applied for selecting features out of the high-dimensional point clouds of clients connected by weighted interaction links.

Neuroscience

The multiscale summaries provided by TDA have shown relevant impact on the study of structural and functional brain connectivity. The brain connectivity is often represented by a connectivity matrix based on physical adjacency or correlation measures between brain regions. This can be easily turned into weighted graph terms thus allowing for the application of TDA and, specifically, the persistence pipeline.

TDA supports the exploration and visualization purposes especially by means of the Mapper graph [96]. The Mapper graph is a topological signature alternative to persistence. The signature captures the adjacency of connected components with respect to scalar function level sets. The Mapper signature is exploited in [83] to show coherence of topological information on structural and functional brain aspects.

Coming back to the persistence pipeline, we first focus on structural aspects. In [75], authors investigate the dynamics of the Kuramoto model in the case of higher-

order relations. In [7], graph filtrations are exploited to study tree-like structures of brain arteries to find correlations between age and sex higher than previously obtained without higher-order structures.

As for the functional aspects, in [46] the clique complex construction is used to derive the critical role of the hippocampal circuits in shaping the geometrical structure of correlations in the brain of rats. In [87], the scaffold persistence summary is exploited to discriminate between healthy and psycoactive patients. The function brain connectivity is also studied under fMRI data (Functional Magnetic Resonance Images) in [70] to introduce a new topology-based measure based on the scaffold summary. In [101], vineyards, that is continuous families of persistence diagrams, are used to classify between resting and gaming subjects to obtain more robust results compared to Principal Component Analysis.

Brain functionalities other than the connectivity are also investigated under data represented by functions. Image data of cortical surfaces are classified in [26] according to persistence summaries. Most applications are focused on EEG signals treated in several works. In [100] denoising of signals is performed by means of persistence landscape summaries. The original signal and the transformed one are found to share topological properties. The independence from scales and translations of topological features is exploited over the frequent space after Fourier transform to isolate the signal. In [52, 53], EEG signals are treated for finding evidences of correlations between head positioning and hypnotizability of patients. In [15], authors propose feature selection to separate signals of able-bodied from amputee subjects.

Biology

The field of biology is experimenting a growing interest towards topological data analysis. A first motivation relies on the high number of features known to be involved in many biological studies. This is the case of the work in [90] where TDA is integrated into feature selection to detect subgroups of features based on similarity measures.

A second case is that of gene sequences. The gene sequence evolutions can be represented by treelike structures where points are gene sequences. As stated in [22], treelike structures accurately reflect vertical evolution (mutations over generations) but not horizontal evolution (genomic mixture between individuals). Tree structures capture vertical evolution. They consist of a single connected component without loops. Hence, homology in degree 0 might be associated to vertical evolution, whereas higher-order homological features might be associated with horizontal evolution. Higher-order homological features are captured by starting from the same initial gene sequences and considering the Vietoris-Rips complex with respect to some distance measuring genetic dissimilarity. These assumptions are applied in [41] to characterize and study antibiotic resistance. Another gene sequence application is in the scope of pattern recognition. TDA is applied to finding patterns within gene expressions [31] with no periodicity assumption, unlike, for instance, in the case of those retrieved by the Fourier transform.

A third case of relevant interest is the study of proteins. Protein data are particularly suited for geometrical representations, that is graphs or point clouds embedded

in some metric space. Most persistent loops are found to appear in correspondence of active protein sites [56]. For instance, DNA can be modeled by weighted graph structures where vertexes are atoms, edges are bonds and weights represent biological properties. In [74], the attitude of persistence of capturing mesoscale features is used to reveal local functional properties, where local is not a priori determined but retrieved. Weights are locally associated to persistence features within the generalized framework of weighted persistent homology. Moreover, in [45], protein compressibiity is related to the protein geometric structure. The simplicial model involved is that of α-shapes. A correlation between the topological measure and the experimentally-determined compressibility of most proteins is found.

HodgeRank

In [68], the Hodge decomposition has been successfully applied in the setting of statistical ranking. The main problem discussed in [68] is that of determining a global ranking from a dataset comprising a number of alternatives ranked by a number of voters. This is a relevant problem which sprung form several fields and it is of paramount relevance for, e.g., online global platform as Netflix, Amazon, Google, Ebay, etc.

Besides deducing a global ranking from the data, whenever possible, the authors also showed a way to analyze the obstructions not permitting a statistically meaningful global ranking. Their methods, collectively called HodgeRank, analyze pairwise rankings represented as edge flows on a graph using combinatorial Hodge theory also to provide a mean to determine a global ranking that also comes with a *certificate of reliability* for its validity. Alternatives are $V = \{1, \ldots, m\}$ and voters $\Lambda = \{1, \ldots, n\}$ For each voter $\alpha \in \Lambda$ ranking matrix

$$Y^{\alpha} \in \mathbb{R}^{n \times n}, \quad Y^{\alpha}_{ij} = -Y^{\alpha}_{ji}$$

Weight function $w : \Lambda \times V \times V \to [0, \infty)$

$$w^{\alpha}_{ij} = \begin{cases} 1 \text{ if } \alpha \text{ compared } (i, j) \\ 0 \text{ otherwise} \end{cases}$$

are provided. Let Σ be the clique complex whose one skeleton is the graph $G := (V, E)$ with vertices the alternatives and adjacency matrix $(w_{i,j})$ i.e. edge are between compared alternatives. Hodge theory then provides an orthogonal decomposition of the assigned pairwise comparison flow into three components: a gradient flow that is globally acyclic, a harmonic flow that is locally acyclic but globally cyclic, and a curl flow that is locally cyclic. Namely

$$Y = Y^{(g)} + Y^{(h)} + Y^{(c)}$$

from

$$C^1(\Sigma;\mathbb{R}) = \overbrace{\mathrm{im}(D_1^T) \oplus \underbrace{\ker(L_1)}_{\ker(D_1)}}^{\ker(D_2^T)} \oplus \mathrm{im}(D_2).$$

Usually, in this two dimensional case the components of the decomposition have traditional name caming for physics, that is: the elements of $\mathrm{im}(D_1^T)$ are the *gradient flows*; the elements of $\ker(L_1)$ are the *harmonic flows*; the elements of $\mathrm{im}(D_2)$ are the *curl flows*. The gradient flows component induces a global ranking of the alternatives and this can be computed via a linear least squares problem. Furthermore, the $l-2$-norm of the least squares residual, provides a measure of the validity of the global ranking induced by the gradient flow component. If the residual is small, then the gradient flow accounts for most of the variation in the given data and, therefore, the global ranking obtained from it is expected to be a majority consensus. Viceversa, if the residual is large, then the data manifest cyclic inconsistencies and no reasonable global ranking is expected to exist. The curl flow and harmonic flow components of an edge flow quantify respectively the local and global cyclic inconsistencies. Inconsistencies of a local nature will cause a dominant curl flow component, while a dominant harmonic flow component reflects a global nature of the inconsistencies. If in addition the harmonic flow component is small, then most of the inconsistencies happen locally and this could be read as the global ranking is sound on a coarse scale (infra-cluster ranking) but not on a finer scale (intra-cluster ranking).

Communication networks

In [3] the authors used the Hodge decomposition to analyze signals residing over a simplicial complex. Although they focused on signals defined over the edges of a complex of order two, i.e. including triangles, their findings and tools can be directly translated to analyze signals defined over higher order structures. What would be missing in the higher order cases would be concepts as solenoidal or irrotational behaviors. A signal over over a simplicial complex Σ is a family of functions $s^k : \Sigma_k \to \mathbb{R}$, with $k = 0, \ldots, \dim \Sigma$. By using the decomposition (4) and (5) in Theorem 3.2 one can orthogonally decompose s^k as

$$s^k = D_k^T s^{k-1} + s_H^k + D_{k+1} s^{k+1}$$

where $L_k(s_H^k) = 0$ i.e. it is an harmonic. Taking advantage of this representation the authors were able to prove that the Lovàsz extension of the triple-wise coupling function gives rise to a measure of the curl of edge signals along triangles. They also proposed a method to infer the structure of a second order simplicial complex from flow data and we showed that in applications over real wireless traffic data, the proposed approach can significantly outperform methods based only on graph representations. Furthermore, they analyze discrete vector fields and showed an application to the recovery of the RNA velocity field to predict the evolution of living cells. In such a case, using the eigenvectors of the first combinatorial Laplacian

have been able to highlight irrational and solenoidal behaviors that would have been difficult to highlight using only the eigenvectors of the zeroth one.

TDA applied to other fields

The number fields where TDA has found or is gaining success is experimenting an impressive grow.

In particular, thanks to the translation into vector representations and the statistically significant improvements in representing persistent homology (Sect. 3.4.2.1), the last few years have seen the flourishing of the interplay between TDA and Machine Learning. To this scope, the survey [51] on TDA and deep learning provides an interesting discussion on contributions of TDA in Machine learning as well as the converse.

Finally, we close this section by providing a not exhaustive list of applications which can possibly be of interest to field experts. The list entails: physics [34, 57, 58, 86], medicine [7, 63, 67, 80], chemistry [64], image analysis [18], shape study [60], object recognition [66], fractal geometry [71], quantum computing [69], machine monitoring [55], automated productivity [50], data analysis in biomechanics [91], and material science [79].

Acknowledgements SS and FV acknowledge the support from the Italian MIUR Award "Dipartimento di Eccellenza 2018–2022" - CUP: E11G18000350001 and the SmartData@PoliTO center for Big Data and Machine Learning.

References

1. H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, L. Ziegelmeier, Persistence images: a stable vector representation of persistent homology. J. Mach. Learn. Res. **18**, 1–35 (2017)
2. P. Bajardi, M. Delfino, A. Panisson, G. Petri, M. Tizzoni, Unveiling patterns of international communities in a global city using mobile phone data. EPJ Data Sci. **4**, 1–17 (2015)
3. S. Barbarossa, S. Sardellitti, Topological signal processing over simplicial complexes. IEEE Trans. Signal Process. **68**, 2992–3007 (2020)
4. U. Bauer, *Ripser: Efficient Computation of Vietoris-rips Persistence Barcodes*. https://github.com/Ripser/ripser. February 2021. Preprint
5. U. Bauer, M. Kerber, J. Reininghaus, Distributed computation of persistent homology. In *2014 Proceedings of the Sixteenth Workshop on Algorithm Engineering and Experiments (ALENEX)* (SIAM, 2014), pp. 31–38
6. U. Bauer, M. Kerber, J. Reininghaus, H. Wagner, *Phat: Persistent Homology Algorithm Toolbox*. http://bitbucket.org/phat-code/phat (2013)
7. P. Bendich, J.S. Marron, E. Miller, A. Pieloch, S. Skwerer, Persistent homology analysis of brain artery trees. Ann. Appl. Statistics **10**(1), 198 (2016)
8. C. Berg, J.P.R. Christensen, P. Ressel, *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*, vol. 100 (Springer, Heidelberg, 1984)
9. G. Bianconi, R.M. Ziff, Topological percolation on hyperbolic simplicial complexes. Phys. Rev. E **98**, 052308 (2018)
10. S. Biasotti, D. Giorgi, M. Spagnuolo, B. Falcidieno, Reeb graphs for shape analysis and applications. Theor. Comput. Sci. **392**(1–3), 5–22 (2008)

11. J. Binchi, E. Merelli, M. Rucco, G. Petri, F. Vaccarino, jholes: a tool for understanding biological complex networks via clique weight rank persistent homology. Electronic Notes Theor Comput Sci **306**, 5–18 (2014)
12. H.B. Bjerkevik, On the stability of interval decomposable persistence modules, in *Discrete & Computational Geometry* (2021), pp. 1–30
13. P. Bubenik, Statistical topological data analysis using persistence landscapes. J. Mach. Learn. Res. **16**(1), 77–102 (2015)
14. P. Bubenik, J.A. Scott, Categorification of persistent homology. Discrete & Computational Geometry **51**(3), 600–627 (2014)
15. E. Campbell, A. Phinyomark, A.H. Al-Timemy, R.N. Khushaba, G. Petri, E. Scheme, Differences in emg feature space between able-bodied and amputee subjects for myoelectric control, in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)* (IEEE, 2019), pp. 33–36
16. G. Carlsson, A. Zomorodian, Computing persistent homology. Discrete Comput. Geometry **33**(2), 249–274 (2005)
17. G. Carlsson, A. Zomorodian, The theory of multidimensional persistence. Discrete Comput. Geometry **42**(1), 71–93 (2009)
18. G. Carlsson, A. Zomorodian, A. Collins, L.J. Guibas, Persistence barcodes for shapes. Int. J. Shape Modeling **11**(02), 149–187 (2005)
19. M. Carrière, M. Cuturi, S. Oudot, Sliced Wasserstein Kernel for persistence diagrams, in *Proceedings of the 34th International Conference on Machine Learning* (PMLR, 2017), pp. 664–673
20. C.J. Carstens, K.J. Horadam, Persistent homology of collaboration networks, in *Mathematical problems in engineering* (2013)
21. A. Cerri, B. Di Fabio, M. Ferri, P. Frosini, C. Landi, Betti numbers in multidimensional persistent homology are stable functions. Math. Methods Appl. Sci. **36**, 1543–1557 (2013)
22. J.M. Chan, G. Carlsson, R. Rabadan, Topology of viral evolution. Proc. Natl. Acad. Sci. **110**(46), 18566–18571 (2013)
23. F. Chazal, D. Cohen-Steiner, M. Glisse, L.J. Guibas, S.Y. Oudot, Proximity of persistence modules and their diagrams, in *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry*, Aarhus, Denmark (ACM Press, 2009), pp. 237–246
24. F. Chazal, D. Cohen-Steiner, L.J. Guibas, F. Mémoli, S.Y. Oudot, Gromov-hausdorff stable signatures for shapes using persistence, in *Computer Graphics Forum*, vol. 28 (Wiley Online Library, 2009), pp. 1393–1403
25. F. Chazal, B. Michel, *An Introduction to Topological Data Analysis: Fundamental and Practical Aspects for Data Scientists*. *arXiv preprint* arXiv:1710.04019 (2017)
26. M.K. Chung, P. Bubenik, P.T. Kim, Persistence diagrams of cortical surface data, in *Information Processing in Medical Imaging* (Springer, Heidelberg, 2009), pp. 386–397
27. D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Stability of persistence diagrams. Discrete Comput. Geometry **37**(1), 103–120 (2007)
28. D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Y. Mileyko, Lipschitz functions have l p-stable persistence. Foundations Comput. Math. **10**(2), 127–139 (2010)
29. M. D'Amico, P. Frosini, C. Landi, Natural pseudo-distance and optimal matching between reduced size functions. Acta Applicandae Mathematicae **109**(2), 527–554 (2010)
30. V. De Silva, G.E. Carlsson, Topological estimation using witness complexes. SPBG **4**, 157–166 (2004)
31. M.-L. Dequeant, S. Ahnert, H. Edelsbrunner, T.M.A. Fink, E.F. Glynn, G. Hattem, A. Kudlicki, Y. Mileyko, J. Morton, A.R. Mushegian, et al, Comparison of pattern detection methods in microarray time series of the segmentation clock. PLoS One **3**(8), e2856 (2008)
32. T.K Dey, T. Li, Y. Wang, Efficient algorithms for computing a minimal homology basis, in *Latin American Symposium on Theoretical Informatics* (Springer, Heidelberg, 2018), pp. 376–398
33. P. Donatini, P. Frosini, A. Lovato, Size functions for signature recognition, in *Vision Geometry VII*, vol. 3454 (International Society for Optics and Photonics, 1998), pp. 178–183

34. I. Donato, M. Gori, M. Pettini, G. Petri, S. De Nigris, R. Franzosi, F. Vaccarino, Persistent homology analysis of phase transitions. Phys. Rev. E **93**(5), 052138 (2016)
35. H. Edelsbrunner, Alpha shapes-a survey. Tessellations Sci. **27**, 1–25 (2010)
36. H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, vol. 10 (Springer Science & Business Media, 2012)
37. H. Edelsbrunner, et al., *Geometry and Topology for Mesh Generation* (Cambridge University Press, 2001)
38. H. Edelsbrunner, J. Harer, Persistent homology—a survey. Contemporary Math. **453**, 257–282 (2008)
39. H. Edelsbrunner, D. Letscher, A. Zomorodian, Topological persistence and simplification. Discrete Comput. Geometry **28**(4), 511–533 (2002)
40. H. Edelsbrunner, E.P. Mücke, Three-dimensional alpha shapes. ACM Trans. Graphics (TOG) **13**(1), 43–72 (1994)
41. K.J. Emmett, R. Rabadan, Characterizing scales of genetic recombination and antibiotic resistance in pathogenic bacteria using topological data analysis, in *International Conference on Brain Informatics and Health* (Springer, Heidelberg, 2014), pp. 540–551
42. B.T. Fasy, J. Kim, F. Lecci, C. Maria, V. Rouvreau, *Tda: Statistical Tools for Topological Data Analysis*. Software available at https://cran.r-project.org/package=TDA (2014)
43. M. Ferri, P. Frosini, A. Lovato, C. Zambelli, Point selection: a new comparison scheme for size functions (with an application to monogram recognition), in *Asian Conference on Computer Vision* (Springer, Heidelberg, 1998), pp. 329–337
44. P. Frosini, C. Landi, Size functions and formal series. Applicable Algebra Eng. Commun. Comput. **12**(4), 327–349 (2001)
45. M. Gameiro, Y. Hiraoka, S. Izumi, M. Kramar, K. Mischaikow, V. Nanda, A topological measurement of protein compressibility. Jpn. J. Industrial Appl. Math. **32**(1), 1–17 (2015)
46. C. Giusti, E. Pastalkova, C. Curto, V. Itskov, Clique topology reveals intrinsic geometric structure in neural correlations. Proc. Natl. Acad. Sci. U. S. A. **112**(44), 13455–13460 (2015)
47. M. Gromov, *Metric Structures for Riemannian and Non-Riemannian Spaces* (Springer Science & Business Media, 2007)
48. M. Guerra, A. De Gregorio, U. Fugacci, G. Petri, F. Vaccarino, Homological scaffold via minimal homology bases. Sci. Rep. **11**(1), 1–17 (2021)
49. C. Gunnar, Topology and data. Bull. Am. Math. Soc. **46**(2), 255–308 (2009)
50. W. Guo, A.G. Banerjee, Toward automated prediction of manufacturing productivity based on feature selection using topological data analysis, in *IEEE International Symposium on Assembly and Manufacturing* (2016), pp. 31–36
51. C. Hofer, R. Kwitt, M. Niethammer, A. Uhl, Deep learning with topological signatures, in *Proceedings of the 31st International Conference on Neural Information Processing Systems* (2017), pp. 1633–1643
52. E. Ibáñez-Marcelo, L. Campioni, D. Manzoni, E.L. Santarcangelo, G. Petri, Spectral and topological analyses of the cortical representation of the head position: does hypnotizability matter? Brain and Behavior **9**(6), e01277 (2019)
53. E. Ibáñez-Marcelo, L. Campioni, A. Phinyomark, G. Petri, E.L Santarcangelo, The case of hypnotizability, Topology highlights mesoscopic functional equivalence between imagery and perception. NeuroImage **200**, 437–449 (2019)
54. H. Kannan, E. Saucan, I. Roy, A. Samal, Persistent homology of unweighted complex networks via discrete morse theory. Sci. Rep. 9 (2019)
55. F.A. Khasawneh, E. Munch, Chatter detection in turning using persistent homology. Mech. Syst. Signal Process. **70**, 527–541 (2016)
56. V. Kovacev-Nikolic, P. Bubenik, D. Nikolić, G. Heo, Using persistent homology and dynamical distances to analyze protein binding. Statistical Appl. Genetics Molecular Biol. **15**(1), 19–38 (2016)
57. M. Kramár, A. Goullet, L. Kondic, K. Mischaikow, Persistence of force networks in compressed granular media. Phys. Rev. E **87**(4), 042207 (2013)

58. M. Kramár, R. Levanger, J. Tithof, B. Suri, M. Xu, M. Paul, M.F. Schatz, K. Mischaikow, Analysis of kolmogorov flow and Rayleigh–bénard convection using persistent homology. Physica D: Nonlinear Phenomena **334**, 82–98 (2016)
59. D.P. Kroese, Z. Botev, T. Taimre, R. Vaisman, *Data Science and Machine Learning: Mathematical and Statistical Methods* (CRC Press, 2019)
60. V. Kurlin, A fast and robust algorithm to count topologically persistent holes in noisy clouds, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 1458–1463
61. G. Kusano, Y. Hiraoka, K. Fukumizu, Persistence weighted gaussian kernel for topological data analysis, in *International Conference on Machine Learning* (PMLR, 2016), pp. 2004–2013
62. C. Landi, P. Frosini, New pseudodistances for the size function space, in *SPIE Proceedings of Vision Geometry VI*, vol. 3168 (1997), pp. 52–60
63. P. Lawson, A.B. Sholl, J.Q.Brown, B.T. Fasy, C. Wenk, Persistent homology for the quantitative evaluation of architectural features in prostate cancer histology. Sci. Rep. **9**(1), 1–15 (2019)
64. Y. Lee, S.D. Barthel, P. Dłotko, S. Mohamad Moosavi, K. Hess, B. Smit,Quantifying similarity of pore-geometry in nanoporous materials. Nature Commun. **8**(1), 1–8 (2017)
65. M. Lesnick, The theory of the interleaving distance on multidimensional persistence modules. Foundations Comput. Math. **15**(3), 613–650 (2015)
66. C. Li, M. Ovsjanikov, F. Chazal, Persistence-based structural recognition, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 1995–2002
67. L. Li, W.-Y. Cheng, B.S. Glicksberg, O. Gottesman, R. Tamler, R. Chen, E.P. Bottinger, J.T. Dudley, Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Sci. Translational Med. **7**(311), 311ra174 (2015)
68. L..-H.. Lim, Hodge laplacians on graphs. Siam Rev. **62**(3), 685–715 (2020)
69. S. Lloyd, S. Garnerone, P. Zanardi, Quantum algorithms for topological and geometric analysis of data. Nature Commun. **7**(1), 1–7 (2016)
70. L.-D. Lord, P. Expert, H.M. Fernandes, G. Petri, T.J. Van Hartevelt, F. Vaccarino, G. Deco, F. Turkheimer, M.L. Kringelbach, Insights into brain architectures from the homological scaffolds of functional connectivity networks. Frontiers Syst. Neurosci. **10**, 85 (2016)
71. R. MacPherson, B. Schweinhart, Measuring shape with topology. J. Math. Phys. **53**(7) (2012)
72. C. Maria, J.-D. Boissonnat, M. Glisse, M. Yvinec, The gudhi library: simplicial complexes and persistent homology, in *International Congress on Mathematical Software* (Springer, Heidelberg, 2014), pp. 167–174
73. J. May Peter, Finite topological spaces, in *Notes for REU* (2003)
74. Z. Meng, D. Vijay Anand, Y. Lu, J. Wu, K. Xia, Weighted persistent homology for biomolecular data analysis. Sci. Rep. **10** (2020)
75. A.P. Millán, J.J. Torres, G. Bianconi, Explosive higher-order kuramoto dynamics on simplicial complexes. Phys. Rev. Lett. page to appear 2019
76. N. Milosavljević, D. Morozov, P. Skraba, Zigzag persistent homology in matrix multiplication time, in *Proceedings of the Twenty-Seventh Annual Symposium on Computational Geometry* (ACM, 2011), pp. 216–225
77. D. Morozov, *Dionysus*. Software available at http://www.mrzv.org/software/dionysus (2012)
78. J.R. Munkres, *Elements of Algebraic Topology* (Perseus Books, Advanced book classics, 1984)
79. T. Nakamura, Y. Hiraoka, A. Hirata, E.G. Escolar, Y. Nishiura, Persistent homology and many-body atomic structure for medium-range order in the glass. Nanotechnology **26**(30), 304001 (2015)
80. M. Nicolau, A.J. Levine, G. Carlsson, Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. Proc. Natl. Acad. Sci. **108**(17), 7265–7270 (2011)
81. N. Otter, M.A. Porter, U. Tillmann, P. Grindrod, H.A .Harrington, A roadmap for the computation of persistent homology. EPJ Data Sci. **6**(1), 17 (2017)

82. A. Patania, G. Petri, F. Vaccarino, P-persistent homology of finite topological spaces. Rendiconti del Seminario Matematico **75**(2), 27–45 (2017)
83. A. Patania, P. Selvaggi, M. Veronese, O. Dipasquale, P. Expert, G. Petri, Topological gene expression networks recapitulate brain anatomy and function. Netw. Neurosci. **3**(3), 744–762 (2019)
84. A. Patania, F. Vaccarino, G. Petri, Topological analysis of data. EPJ Data Sci. **6**(1), 7 (2017)
85. P. Pearson, D. Muellner, G. Singh, *Tdamapper: Analyze High-dimensional Data Using Discrete Morse Theory* (CRAN, Vienna, Austria, 2015)
86. J.A. Perea, J. Harer, Sliding windows and persistence: an application of topological methods to signal analysis. Foundations Comput. Math. **15**(3), 799–838 (2015)
87. G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer, F. Vaccarino, Homological scaffolds of brain functional networks. J. Roy. Soc. Interface **11**(101), 20140873 (2014)
88. G. Petri, M. Scolamiero, I. Donato, F. Vaccarino, Networks and cycles: a persistent homology approach to complex networks, in *Proceedings of the European Conference on Complex Systems 2012* (Springer, 2013), pp. 93–99
89. G. Petri, M. Scolamiero, I. Donato, F. Vaccarino, Topological strata of weighted complex networks. PloS one **8**(6) (2013)
90. A. Phinyomark, R.N. Khushaba, E. Ibáñez-Marcelo, A. Patania, E. Scheme, G. Petri, Navigating features: a topologically informed chart of electromyographic features space. J. Roy. Soc. Interface **14**(137), 20170734 (2017)
91. A. Phinyomark, G. Petri, E. Ibáñez-Marcelo, S.T. Osis, R. Ferber, Analysis of big data in gait biomechanics: current trends and future directions. J. Med. Biol. Eng. **38**(2), 244–260 (2018)
92. J. Reininghaus, S. Huber, U. Bauer, R. Kwitt, A stable multi-scale kernel for topological machine learning, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 07-12-June (IEEE Computer Society, 2015), pp. 4741–4748
93. B. Rieck, U. Fugacci, J. Lukasczyk, H. Leitte, Clique community persistence: a topological visual analysis approach for complex networks. IEEE Trans. Visualization Comput. Graphics **24**(1), 822–831 (2018)
94. V. Robins, Towards computing homology from finite approximations. Topol. Proc. **24**(1), 503–532 (1999)
95. K.B. Schebesch, R.W. Stecking, Topological data analysis for extracting hidden features of client data, in *Operations Research Proceedings 2015* (Springer, Heidelberg, 2017), pp. 483–489
96. G. Singh, F. Mémoli, G.E. Carlsson, Topological methods for the analysis of high dimensional data sets and 3d object recognition. SPBG **91**, 100 (2007)
97. A.E. Sizemore, C. Giusti, A. Kahn, J.M. Vettel, R.F. Betzel, D.S. Bassett, Cliques and cavities in the human connectome. J. Comput. Neurosci. **44**(1), 115–145 (2018)
98. G. Tauzin, U. Lupo, L. Tunstall, J.B. Pérez, M. Caorsi, A.M. Medina-Mardones, A. Dassatti,K. Hess, giotto-tda: a topological data analysis toolkit for machine learning and data exploration. J. Mach. Learn. Res. **22**(39), 1–6 (2021)
99. D. Taylor, F. Klimm, H.A. Harrington, M. Kramár, K. Mischaikow, M.A. Porter, P.J. Mucha, Topological data analysis of contagion maps for examining spreading processes on networks. Nature Commun. **6**(1), 1–11 (2015)
100. Y. Wang, H. Ombao, M.K. Chung, Topological data analysis of single-trial electroencephalographic signals. Ann. Appl. Statistics **12**(3), 1506–1534 (2017)
101. J. Yoo, E.Y. Kim, Y.M. Ahn, J.C. Ye, Topological persistence vineyard for dynamic functional brain connectivity during resting and gaming stages. J. Neurosci. Methods **267**(15), 1–13 (2016)
102. A. Zomorodian, Fast construction of the vietoris-rips complex. Comput. Graphics **34**(3), 263–271 (2010)

Chapter 4
Flow-Based Community Detection in Hypergraphs

Anton Eriksson, Timoteo Carletti, Renaud Lambiotte, Alexis Rojas, and Martin Rosvall

Abstract To connect structure, dynamics and function in systems with multibody interactions, network scientists model random walks on hypergraphs and identify communities that confine the walks for a long time. The two flow-based community-detection methods Markov stability and the map equation identify such communities based on different principles and search algorithms. But how similar are the resulting communities? We explain both methods' machinery applied to hypergraphs and compare them on synthetic and real-world hypergraphs using various hyperedge-size biased random walks and time scales. We find that the map equation is more sensitive to time-scale changes and that Markov stability is more sensitive to hyperedge-size biases.

With new introduction, experiments, and conclusion sections, this text reuses some text from Carletti, T., Fanelli, D., Lambiotte, R.: Random walks and community detection in hypergraphs. J. Phys. Complex. 2 015011 (2021) and Eriksson, A., Edler, D., Rojas, A., Rosvall, M.: Mapping flows on hypergraphs: How choosing random-walk model and network representation matters for community detection. Comm. Phys. (2021).

A. Eriksson · A. Rojas · M. Rosvall (✉)
Umeå University, Umeå, Sweden
e-mail: martin.rosvall@umu.se

A. Eriksson
e-mail: anton.eriksson@umu.se

A. Rojas
e-mail: alexis.rojas-briceno@umu.se

T. Carletti
University of Namur, Namur, Belgium
e-mail: timoteo.carletti@unamur.be

R. Lambiotte
University of Oxford, Oxford, UK
e-mail: renaud.lambiotte@maths.ox.ac.uk

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_4

4.1 Introduction

Researchers model and map flows on networks to identify important nodes and detect significant communities [1–6]. From small to large system scales, random walk-based methods help to uncover the inner workings of networks [7, 8]. When standard network models with dyadic relations between pairs of nodes fail to adequately represent a system's interactions, researchers turn to higher-order models of complex systems [9, 10], including multilayer networks [11–14] for multitype interactions, memory networks [15–17] for multistep interactions, and simplicial complexes [18–21] and hypergraphs [22–25] for multibody interactions.

While several methods can identify flow-based communities in multilayer [14, 26, 27] and memory [15–17] networks with higher-order Markov dynamics, researchers have focused on combinatorial methods to identify communities in hypergraphs [28–33] and only recently begun to unravel flow-based community structures associated with random walks guided by hyperedge sizes [23]. Two such methods are Markov stability and the map equation. Both algorithms exploit random walkers' tendency to stay unexpectedly long times in assortative communities, albeit in different ways. Markov stability measures the fraction of random walks that reside in the community where they started after time t compared with stationarity. Instead, the map equation measures the shortest possible modular codelength required to describe the random walker on the network with given communities. Also their optimisation algorithms differ and it remains unclear how their detected communities in hypergraphs compare.

For Markov stability, we have previously analysed random walks on hypergraphs with a parameter controlling the bias of the dynamics towards small or large hyperedges and identified widely differing communities [23]. For the map equation and its optimisation algorithm Infomap [34], we have derived and clustered unipartite, bipartite, and multilayer network representations of hypergraph flows with different advantages [35]. Both papers highlight that the network and the research question should decide which model to use, but did not compare the two methods. For example, when modelling the flow of ideas with random walks in a co-authorship hypergraph, how does the organisation of authors in communities, at scales from research groups to research areas, change with community-detection method?

We compare Markov stability and the map equation applied to random walks on hypergraphs with hyperedge-size bias. After explaining the random-walk model and the flow-based community-detection methods using a schematic hypergraph for illustration, we consider experiments on three real-world hypergraphs: a zoo hypergraph with 101 nodes, a collaboration hypergraph with 361 nodes, and a fossil-record hypergraph with 13,276 nodes. We find that the map equation responds faster to time-scale changes and that Markov stability is more sensitive to hyperedge-size biases.

4.2 Random Walks on Hypergraphs

We consider hypergraphs $\mathcal{H}(V, E)$ with n nodes $V = \{v_1, \ldots, v_n\}$ and m hyperedges $E = \{E_1, \ldots, E_m\}$. Each hyperedge $E_\alpha \subset V$ is an unordered collection of nodes. A hypergraph can be encoded by its incidence matrix $e_{i\alpha}$, where we use Roman indexes for nodes and Greek indexes for hyperedges:

$$e_{i\alpha} = \begin{cases} 1 & \mathrm{v}_i \in \mathrm{E}_\alpha \\ 0 & \text{otherwise}\,. \end{cases} \tag{4.1}$$

The $n \times n$ adjacency matrix of the hypergraph is $\mathbf{A} = \mathbf{e}\mathbf{e}^\top$, whose entry A_{ij} represents the number of hyperedges containing both nodes i and j with the diagonal elements set to zero. We also define the $m \times m$ hyperedge matrix $\mathbf{B} = \mathbf{e}^\top\mathbf{e}$, whose entry $B_{\alpha\beta}$ counts the number of nodes the two hyperedges share in the original hypergraph, $E_\alpha \cap E_\beta$. $\mathbf{B}$ corresponds to the weighted adjacency matrix of the dual hypergraph, where hyperedges of the original hypergraph become nodes of the new structure.

A random walk process on a hypergraph can be defined by the transition probability, T_{ij}, allowing the walker to move across any pair of nodes (i, j). The resulting continuous-time Markov process can be defined by

$$\dot{p}_i(t) = \sum_j p_j(t)T_{ji} - \sum_j p_i(t)T_{ij}\,, \tag{4.2}$$

where $p_i(t)$ is the probability of finding the walker on node i at time t. As often assumed when dealing with Markov processes, $\mathbf{p} = (p_1, \ldots, p_n)$ is a row vector. With normalised transition probabilities such that $\sum_j T_{ij} = 1$, we rewrite the continuous-time Markov process as

$$\dot{p}_i = \sum_j p_j(T_{ji} - \delta_{ij}) = -\sum_j p_j L_{ji}\,, \tag{4.3}$$

where $L_{ij} := \delta_{ij} - T_{ij}$ is the random-walk Laplace operator.

Assigning weights to nodes and hyperedges biases the transition probabilities and leads to different probability flows between nodes [22, 23, 25].

Fixed weights and independent moves [25].
Setting a weight $\omega(E_\alpha)$ to each hyperedge E_α and a weight $\gamma_{E_\alpha}(i)$ to each node $i \in E_\alpha$ allows us to decompose the random walk into three steps:

1. A random walker on node i chooses an incident hyperedge E_α proportional to its weight,

$$Q_{\alpha i} = \frac{\omega(E_\alpha)e_{i\alpha}}{\sum_\beta \omega(E_\beta)e_{i\beta}}\,. \tag{4.4}$$

2. The random walker picks a node $j \neq i$ of hyperedge E_α proportional to its weight,
$$R_{j\alpha} = \frac{\gamma_{E_\alpha}(j) e_{j\alpha}}{\sum_{\ell \neq i} \gamma_{E_\alpha}(\ell) e_{\ell\alpha}} . \tag{4.5}$$
3. The random walker moves to node j with probability
$$T_{ij} = \sum_\alpha R_{j\alpha} Q_{\alpha i} . \tag{4.6}$$

Several random-walk processes can be defined with various node and hyperedge weights. With hyperedge size as a proxy for the higher-order interactions' nature, we consider the transition probability

$$T_{ij}^{(\sigma)} = \frac{K_{ij}^{(\sigma)}}{\sum_{\ell \neq i} K_{i\ell}^{(\sigma)}} \quad \forall i \neq j \quad \text{and} \quad T_{ii}^{(\sigma)} = 0 , \tag{4.7}$$

with

$$K_{ij}^{(\sigma)} = \sum_\alpha (B_{\alpha\alpha} - 1)^\sigma e_{i\alpha} e_{j\alpha} \quad \forall i \neq j \quad \text{and} \quad K_{ii}^{(\sigma)} = 0 , \tag{4.8}$$

for some real σ, which biases random walks to hyperedges depending on their size [22, 23]. By setting $\omega(E_\alpha) = (B_{\alpha\alpha} - 1)^{\sigma+1}$ and $\gamma_{E_\alpha}(j) = (B_{\alpha\alpha} - 1)$ for all $j \in E_\alpha$, Eq. (4.7) takes the form of Eq. (4.6). From

$$\sum_{\ell \neq i} K_{i\ell}^{(\sigma)} = \sum_{\ell \neq i} \sum_\beta (B_{\beta\beta} - 1)^\sigma e_{i\beta} e_{\ell\beta} = \sum_\beta (B_{\beta\beta} - 1)^{\sigma+1} e_{i\beta} ,$$

where we used $\sum_{\ell \neq i} e_{\ell\beta} = B_{\beta\beta} - 1$, we can conclude

$$\begin{aligned} T_{ij}^{(\sigma)} &= \frac{1}{\sum_\beta (B_{\beta\beta} - 1)^{\sigma+1} e_{i\beta}} \sum_\alpha (B_{\alpha\alpha} - 1)^\sigma e_{i\alpha} e_{j\alpha} \\ &= \frac{1}{\sum_\beta (B_{\beta\beta} - 1)^{\sigma+1} e_{i\beta}} \sum_\alpha (B_{\alpha\alpha} - 1)^{\sigma+1} e_{i\alpha} \frac{e_{j\alpha}}{B_{\alpha\alpha} - 1} \\ &= \sum_\alpha Q_{\alpha i} R_{j\alpha} . \end{aligned}$$

For large positive values of σ, hyperedges with many nodes contribute more to $T_{ij}^{(\sigma)}$ and guide the random process. For large negative values of σ, the large hyperedges' contributions are negligible and only the small hyperedges drive the random walk process. In this way, σ is a hyperedge-size bias parameter.

When $\sigma = 0$, the random walk moves on the so-called clique-reduced multigraph, where each pair of nodes is connected by a number of edges equal to the number of hyperedges containing that pair in the hypergraph. The transition matrix takes the form

$$T_{ij}^{(0)} = \frac{K_{ij}^{(0)}}{\sum_{\ell \neq i} K_{i\ell}^{(0)}} = \frac{A_{ij}}{\sum_{\ell \neq i} A_{i\ell}} ,$$

where we used the hyperadjacency matrix $A_{ij} = \sum_{\alpha} e_{i\alpha} e_{j\alpha}$. The clique reduced multigraph is different from the projected network obtained by associating to each hyperedge a clique of the same size. The latter can be interpreted as the unweighted version of the clique-reduced multigraph.

The matrix $K_{ij}^{(\sigma)}$ given by Eq. (4.8) can be considered as the weighted adjacency matrix of an undirected network. With the associated Laplace operator

$$L_{ij}^{(\sigma)} = \delta_{ij} - T_{ij}^{(\sigma)} = \delta_{ij} - \frac{K_{ij}^{(\sigma)}}{\sum_{\ell \neq i} K_{i\ell}^{(\sigma)}} , \tag{4.9}$$

the continuous-time random walk is

$$\dot{p}_i(t) = -\sum_j p_j(t) L_{ji}^{(\sigma)} , \tag{4.10}$$

and the continuous-time transition matrix, which forwards the continuous-time random-walk process by time t, is

$$\boldsymbol{T}^{(\sigma)}(t) = e^{-t\boldsymbol{L}^{(\sigma)}} . \tag{4.11}$$

Based on this projection, we can invoke standard results about random walks and in particular prove that the stationary state [1]

$$\pi_j = \frac{d_j^{(\sigma)}}{\sum_{\ell} d_{\ell}^{(\sigma)}} , \tag{4.12}$$

where $d_j^{(\sigma)} = \sum_{\ell \neq j} K_{j\ell}^{(\sigma)}$ is the strength of node j in the weighted graph. This quantity is an immediate generalisation of the standard node degree that accounts for different hyperedge sizes.

[1] To lighten the notation, we do not show explicitly the dependence on σ on the asymptotic distribution $\boldsymbol{\pi}$.

The standard formulation of the map equation uses a discrete-time random-walk process. For discrete time steps k, the discrete-time version of Eq. (4.3) takes the form

$$p_i(k+1) = \sum_j p_j(k) T_{ji} \,, \tag{4.13}$$

with the same stationary distribution as for the continuous-time process.

4.3 Flow-Based Community Detection in Hypergraphs

To illustrate how the flow-based community-detection methods known as Markov stability and the map equation identify communities in hypergraphs, we explain their disparate machinery and illustrate with a schematic hypergraph.

4.3.1 Markov Stability

The Markov stability is a quality function for partitioning a network into communities based on the persistence of random-walk flows inside a group of nodes.

Consider an ergodic random walk process in its stationary state. The Markov stability [36, 37] of a partition at time t is defined as the difference between the probability of a random walker to be in the same community at time 0 and at time t and the analogous quantity computed once the system settles in the stationary state. At a given time t, Markov stability is large when random walkers are unlikely to have escaped their initial community. Because the process is assumed to be ergodic and knowledge of the initial conditions is lost asymptotically, the second probability is equal to the probability of two randomly chosen walkers residing in the same community.

Markov stability and autocovariance of the walker signal
The Markov stability is equivalent to the autocovariance of a signal encoding the sequence of communities visited by a random walker in the stationary state.

Starting from the continuous-time random walk in Eq. (4.3) or Eq. (4.10) and with a partition encoded by an $n \times \mathcal{C}$ indicator matrix $\boldsymbol{C}$, we assign the values X_α ($\alpha = 1, ..., \mathcal{C}$) to the vertices of each community. The autocovariance of the sequence of values $X(t)$ is

$$\mathrm{cov}\,[X(0)X(t)] = \langle X(0)X(t)\rangle - \langle X(0)\rangle\langle X(t)\rangle = \boldsymbol{X}^\top \boldsymbol{R}(t, \boldsymbol{C})\boldsymbol{X} \,, \tag{4.14}$$

where $\langle X(t)\rangle$ is the expectation of the random variable $X(t)$, $\boldsymbol{X}$ is the $1 \times \mathcal{C}$ column vector of labels assigned to the $\mathcal{C}$ communities and $\boldsymbol{R}(t, \boldsymbol{C})$ is the $\mathcal{C} \times \mathcal{C}$ clustered covariance matrix,

$$\boldsymbol{R}(t, \boldsymbol{C}) = \boldsymbol{C}^{\top} \left[\boldsymbol{\Pi} \exp(-t\boldsymbol{L}) - \boldsymbol{\pi}\boldsymbol{\pi}^{\top} \right] \boldsymbol{C} \,, \tag{4.15}$$

where $\boldsymbol{\Pi} = \mathrm{diag}(\boldsymbol{\pi})$ is a diagonal matrix encoding the stationary distribution. This $\mathcal{C} \times \mathcal{C}$ matrix does not depend on our arbitrary choice for the values X_α.

Since $(\boldsymbol{\Pi} \exp(-t\boldsymbol{L})_{ij}$ measures the flow of probability from node i to node j over time t, the first term of Eq. (4.15) measures the probability flow between two communities α and β over a time t. In a partition with a strong assortative community structure, the probability flows should be captured for long times within communities, and large entries in $\boldsymbol{R}(t, \boldsymbol{C})$ should be concentrated on its diagonal. Following this reasoning, the Markov stability $r(t, \boldsymbol{C})$ is defined as the trace of the clustered autocovariance matrix [36, 38]:

$$r(t, \boldsymbol{C}) = \mathrm{Tr}\left[\boldsymbol{R}(t, \boldsymbol{C})\right] \,. \tag{4.16}$$

The Markov stability quantifies the quality of a partition at different time scales. For every t, maximising $r(t, \boldsymbol{C})$ gives the best partition of the network into communities, resulting in a sequence of optimal partitions at different times. Time acts as a resolution parameter that enables us to tune the typical size of the communities in the optimal partition, as longer times typically lead to fewer and larger communities. Markov stability shows connections with several concepts related to community detection. For small t, the first order of its Taylor expansion recovers, up to a constant, a parametric generalisation [39] of Newman-Girvan modularity.

4.3.1.1 Markov Stability for Hypergraphs

The random walk process above defined in Eq. (4.10) corresponds to a random walk on a weighted undirected network. The link weights are self-consistently defined starting from the process on the hypergraph and thus taking into account the multibody interactions encoded by the size of the hyperedges. Exploiting this relation, we can optimise Markov Stability on the weighted network to partition the hypergraph nodes into communities. Consider again a partition of the nodes of a hypergraph into $\mathcal{C}$ non-overlapping communities, encoded by the $n \times \mathcal{C}$ indicator matrix $\mathbf{C}$. Expanding Eq. (4.16), the corresponding Markov stability is explicitly given by

$$r(t, \mathbf{C}) = \mathrm{Tr}\left[\mathbf{C}^{\top} \left(\Pi e^{-t\mathbf{L}^{(\sigma)}} - \boldsymbol{\pi}^{\top}\boldsymbol{\pi} \right) \mathbf{C} \right] \,, \tag{4.17}$$

where $\boldsymbol{\pi}$ is the asymptotic solution of Eq. (4.12), Π is the diagonal matrix containing $\boldsymbol{\pi}$ on the diagonal, $\mathbf{L}^{(\sigma)}$ is the random-walk Laplace matrix and $\boldsymbol{\pi}^{\top}\boldsymbol{\pi}$ is the matrix whose (i, j) entry is $\pi_i \pi_j$.

As an illustration, we compute the Markov stability for a schematic hypergraph where nodes have two features represented by letters (A, B or C) and numbers (1 or 2) (Fig. 4.1). The hypergraph has six nodes, $\{A_1, A_2, B_1, B_2, C_1, C_2\}$, and five

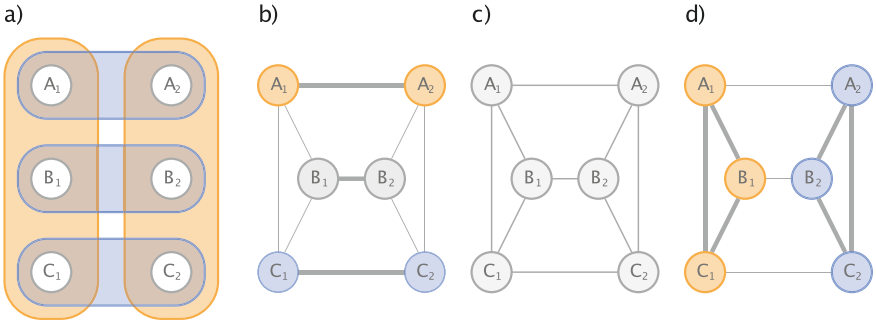


Fig. 4.1 A schematic hypergraph in (**a**) projected into three different weighted networks with hyperedge-size bias $\sigma = -1$ in (**b**), 0 in (**c**), and 1 in (**d**)

hyperedges, three of size two connecting nodes with the same letter, and two of size three connecting nodes with the same number. With $\mathbf{K}^{(\sigma)}$ defined in Eq. (4.8), the hyperedges of size two have a weight of 1^σ to the associated link, whereas the hyperedges of size three have a weight of 2^σ. The corresponding transition probabilities in Eq. (4.7) are

$$\begin{aligned} T_{X_1,X_2} &= \frac{1}{2 \times 2^\sigma + 1}\,, \\ T_{X_1,Y_1} &= \frac{2^\sigma}{2 \times 2^\sigma + 1}\,, \\ T_{X_1,Y_2} &= 0\,, \\ T_{X_1,X_1} &= 0\,, \\ T_{X_2,X_2} &= 0 \end{aligned}$$

for all X, Y in $\{A, B, C\}$.

When $\sigma = 0$, the transition probabilities are $T_{X_1,X_2} = T_{X_1,Y_1} = 1/3$. Thus, a walker remains twice as likely in the same 3-hyperedge than in a 2-hyperedge. The Markov stability thus returns two modules for sufficiently large Markov times ($t > 1$). For $\sigma = 1$, $T_{X_1,X_2} = 1/5$ and $T_{X_1,Y_1} = 2/5$. Then, a walker is four times more likely to remain in the same 3-hyperedge than in a 2-hyperedge. For $\sigma = -1$, $T_{X_1,X_2} = 1/2$ and $T_{X_1,Y_1} = 1/4$, and the probability to stay in the 3-hyperedge is the same as leaving it. In the limit as $\sigma \to \infty$,

$$\lim_{\sigma\to\infty} T_{X_1,X_2} = 0\,, \lim_{\sigma\to\infty} T_{X_1,Y_1} = 1/2 \quad \forall X, Y \in \{A, B, C\}\,,$$

hops among nodes with the same number are strongly favoured and a walker remains for a long time in the same 3-hyperedge. In the other limit,

$$\lim_{\sigma\to-\infty} T_{X_1,X_2} = 1\,, \lim_{\sigma\to-\infty} T_{X_1,Y_1} = 0 \quad \forall X, Y \in \{A, B, C\}$$

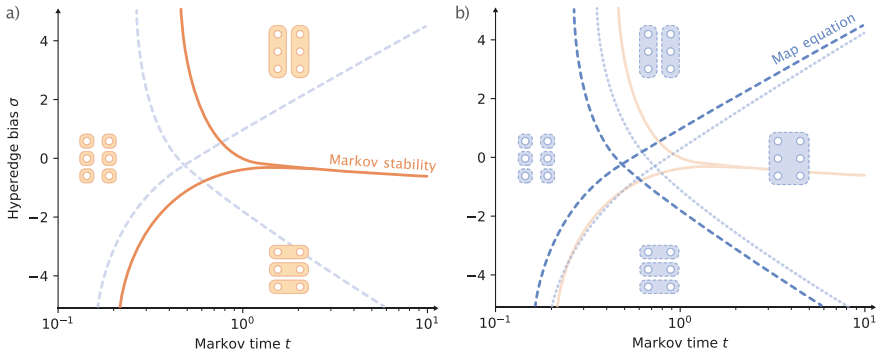


Fig. 4.2 Optimal solutions of the schematic hyper graph in Fig. 4.1 for different hyper edge biases and Markov times. Solutions from Fig. 4.1b–d and the singleton solution with six communities, demarcated with a) Markov stability in solid orange lines, and b) the map equation in dashed blue lines. In b), the blue dotted line is the map equation applied to a continuous time process

a walker spends longer time in the 2-hyperedges. In the first case, Markov stability favours two communities for sufficiently large Markov times, while it favours three communities in the second case (Fig. 4.2a).

4.3.2 The Map Equation

Like Markov stability, the map equation is a quality function for partitioning a network into communities based on the persistence of random-walk flows inside groups of nodes. But instead of using autocovariance, the map equations casts the problem of finding flow-based communities in networks into a minimum-description-length problem [5].

The map equation measures, in bits, the optimal codelength L per step of a discrete random walk on a network for a given node partition M with C modules. When all nodes are in the same module, the map equation is simply the Shannon entropy H of the node-visit rates $\mathcal{P} = \{\pi_i\}$.

In partitions with more than one module, the map equation combines within and between module codelengths for describing flows within and between modules. For modules $\alpha = 1, \ldots, C$ with

entry flow rates	$q_{\alpha\curvearrowleft} = \sum_{i\notin\alpha, j\in\alpha} \pi_i T_{ij}$,
exit flow rates	$q_{\alpha\curvearrowright} = \sum_{i\in\alpha, j\notin\alpha} \pi_i T_{ij}$,
entry flow rate random variable	$\mathcal{Q} = \{q_{\alpha\curvearrowleft}\}$
with total flow rate	$q_{\curvearrowleft} = \sum_\alpha q_{\alpha\curvearrowleft}$,
exit and node-visit rate random variables	$\mathcal{P}_\alpha = \{q_{\alpha\curvearrowright}, \pi_{i\in\alpha}\}$
with total flow rate	$p_{\alpha\circlearrowright} = q_{\alpha\curvearrowright} + \sum_{i\in\alpha} \pi_i$,

the map equation takes its general two-level form

$$L(\mathsf{M}) = q_{\curvearrowleft} H(\mathcal{Q}) + \sum_{\alpha} p_{\alpha\circlearrowleft} H(\mathcal{P}_{\alpha}). \tag{4.18}$$

The first term is the codelength for between-module movements, followed by the sum of codelengths for within-module movements over all modules.

This standard formulation of the map equation encodes every transition of a random walker. For Markov times other than 1, the map equation uses a linearised continuous-time transition matrix,

$$\tilde{\boldsymbol{T}}(t) = \begin{cases} (1-t)\boldsymbol{I} + t\boldsymbol{T} & t < 1 \\ t\boldsymbol{T} & t \geq 1, \end{cases} \tag{4.19}$$

which captures Markov times below 1 with self-links and Markov times above 1 with transition rates proportional to the average rate of the underlying Poisson process [40]. At Markov time 1, it recovers the discrete-time process in Eq. (4.13). Unlike the exponential of the Laplacian, the linearisation keeps the transition matrix sparse also for Markov times larger than zero.[2]

The linearisation is appealing for the map equation because the node-visit rates π_i remain the same for all Markov times t—since the relative visit rates at steady state do not depend on how often the visits are sampled—and the module exit and entry rates change linearly—since the number of random walkers moving along any link between nodes during time t is directly proportional to t. Therefore, we can define

$$q_{\alpha\curvearrowleft}(t) = t q_{\alpha\curvearrowleft} \tag{4.20}$$

$$q_{\alpha\curvearrowright}(t) = t q_{\alpha\curvearrowright}. \tag{4.21}$$

With time dependence on all variables in Eq. (4.18) that depend on the module entry or exit flow rates, the map equation for Markov time t takes the form

$$L(\mathsf{M}, t) = q_{\curvearrowleft}(t) H(\mathcal{Q}(t)) + \sum_{\alpha} p_{\alpha\circlearrowleft}(t) H(\mathcal{P}_{\alpha}(t)). \tag{4.22}$$

While the standard formulation of the map equation can be applied directly to the continuous-time transition matrix for various Markov times in Eq. (4.11), deriving the matrix is computationally expensive for large networks. For long Markov times, it also generates dense networks with longer clustering times. In contrast, the map equation for Markov time t in Eq. (4.22) has no overhead compared with the standard map equation. From a flow modelling perspective, the map equation for Markov time

[2] Markov stability can also apply the linearisation to speed up calculations [37], but here we use the exponential of the Laplacian for Markov stability in all examples.

t counts each time a random walker moves across a modular boundary during time t. Instead, the standard map equation applied to the continuous-time transition matrix counts at most one boundary-crossing since it only considers the final position after time t. For short Markov times, the two approaches are similar. Bur for longer Markov times, the map equation for Markov time t generally favours solutions with fewer modules [40].

To identify the optimal partition—the one that compresses the modular description the most—we use the community-detection algorithm Infomap [41] available on www.mapequation.org. Infomap is to the map equation what the Louvain [42] and the Leiden [43] methods are to the objective function modularity [44], which favours partitions with a high internal density of links compared with a statistical null model. Infomap uses a similar search algorithm as the Leiden method but tries to find the node assignment that minimises the map equation's codelength. Infomap also finds deeper hierarchical partitions—from top-level supermodules with multiple levels of submodules down to leaf-level modules containing the nodes—if such multilevel solutions give higher modular compression [45]. The search algorithm works on many network types, including weighted, directed, bipartite [46], and multilayer [26] networks. Infomap can also identify overlapping multilevel communities in higher-order network flows modelled with so-called memory networks [41]. Recent work introduces a Bayesian estimate of the map equation for identifying flows on sparse networks with missing links [47], but here we focus on the standard map equation and identify hard two-level partitions in weighted unipartite networks obtained with the hypergraph projection described in Sect. 4.2.

4.3.2.1 The Map Equation for Hypergraphs

We have previously derived unipartite, bipartite, and multilayer network representations of hypergraph flows and analysed their different advantages when mapping flows on hypergraphs with Infomap [35]. We found that when the research question does not require hyperedge assignments from using bipartite networks or overlapping modules from using multilayer networks, a unipartite network representation like the one we use here provides a good trade-off between speed and ability to reveal modular regularities.

We illustrate the map equation for hypergraphs with the schematic example in Fig. 4.1 for which the stationary distribution of the random walk is independent of σ such that $\pi_i^{(\sigma)} = \pi_i$ for all nodes i. The one-module codelength is

$$\begin{aligned} L(\mathsf{M}_1, t) &= H(\mathcal{P}) \\ &= H(\pi_{A_1}, \pi_{A_2}, \pi_{B_1}, \pi_{B_2}, \pi_{C_1}, \pi_{C_2}) \\ &= 2.58 \text{ bits for all values of } \sigma \text{ and } t. \end{aligned} \tag{4.23}$$

When a network has modular regularities, a partition captures the modular flows when the random walker spends long times within the modules with few transitions

between them. The codelength is shorter than in the one-module solution because the information required to specify a random walker's position in a module decreases with its size. But for partitions with too many modules, the information required for describing between-module movements exceeds the gain from using small modules. The optimal partition has the shortest codelength. Its node assignment best captures the modular regularities of flows on the network.

Using the three-module solution in Fig. 4.1b, the codelengths for the different hyperedge-size biased random walks parametrised with σ are

$$
\begin{aligned}
L(\mathsf{M}_3, t) = {} & q_{\curvearrowleft} H(q_{1\curvearrowleft}, q_{2\curvearrowleft}, q_{3\curvearrowleft}) \\
& + (q_{1\curvearrowleft} + \pi_{A_1} + \pi_{A_2}) H(q_{1\curvearrowleft}, \pi_{A_1}, \pi_{A_2}) \\
& + (q_{2\curvearrowleft} + \pi_{B_1} + \pi_{B_2}) H(q_{2\curvearrowleft}, \pi_{B_1}, \pi_{B_2}) \\
& + (q_{3\curvearrowleft} + \pi_{C_1} + \pi_{C_2}) H(q_{3\curvearrowleft}, \pi_{C_1}, \pi_{C_2}) \\
= {} & \begin{cases}
\mathbf{1.00}\ \mathbf{1.00}\ \mathbf{1.00} & \text{bits for } \sigma \to -\infty, \\
\mathbf{2.30}\ 3.17\ 4.58 & \text{bits for } \sigma = -1, \\
2.61\ 3.67\ 5.41 & \text{bits for } \sigma = 0, \\
2.84\ 4.05\ 6.04 & \text{bits for } \sigma = 1, \text{ and} \\
\underbrace{3.17}_{t=0.5}\ \underbrace{4.58}_{t=1}\ \underbrace{6.92}_{t=2} & \text{bits for } \sigma \to \infty.
\end{cases}
\end{aligned}
\tag{4.24}
$$

This three-module letter-based solution gives shorter codelengths than the one-module solution for small σ-values and short Markov times, indicated by bold numbers in Eq. (4.24). For these values, the random walker spends long times within the three letter-based modules with few transitions between.

Using the two-module solution in Fig. 4.1d, the codelengths for the different hyperedge-size biased random walks parametrised with σ are

$$
\begin{aligned}
L(\mathsf{M}_2, t) = {} & q_{\curvearrowleft} H(q_{1\curvearrowleft}, q_{2\curvearrowleft}) \\
& + (q_{1\curvearrowleft} + \pi_{A_1} + \pi_{B_1} + \pi_{C_1}) H(q_{1\curvearrowleft}, \pi_{A_1}, \pi_{B_1}, \pi_{C_1}) \\
& + (q_{2\curvearrowleft} + \pi_{A_2} + \pi_{B_2} + \pi_{C_2}) H(q_{2\curvearrowleft}, \pi_{A_2}, \pi_{B_2}, \pi_{C_2}) \\
= {} & \begin{cases}
3.46\ 4.58\ 6.34 & \text{bits for } \sigma \to -\infty, \\
2.74\ 3.46\ 4.58 & \text{bits for } \sigma = -1, \\
\mathbf{2.44}\ 3.00\ 3.87 & \text{bits for } \sigma = 0, \\
\mathbf{2.17}\ \mathbf{2.56}\ 3.19 & \text{bits for } \sigma = 1, \text{ and} \\
\underbrace{\mathbf{1.58}}_{t=0.5}\ \underbrace{\mathbf{1.58}}_{t=1}\ \underbrace{\mathbf{1.58}}_{t=2} & \text{bits for } \sigma \to \infty.
\end{cases}
\end{aligned}
\tag{4.25}
$$

This two-module number-based solution gives shorter codelengths than the one-module solution for large σ-values and short Markov times, again indicated by bold numbers in Eq. (4.25). For these values, the random walker instead spends long times within the two number-based modules so the codelength savings from using smaller modules exceed the extra cost from switching modules.

4.3.3 Comparing Markov Stability with the Map Equation

While both Markov stability and the map equation identify flow-based communities in hypergraphs, they highlight different structures. Both methods prefer the solution with one node in each community for minimal Markov times when nearly all flow remains in the nodes. But the Markov time at which they highlight non-trivial solutions differs. For Markov stability this happens when more flows stay in larger communities than in singleton communities compared with respective stationary expectation. For the map equation, it happens when describing transitions between communities costs more than gained from describing within-community movements in singleton communities. Which one comes first depends on the network structure. For the schematic network in Fig. 4.1, the map equation first prefers non-trivial solutions (Fig. 4.2b). But in the larger networks analysed in Sect. 4.4, Markov stability highlights non-trivial solutions for shorter Markov times than the map equation.

The map equation strictly prefers the one-community solution over a non-trivial solution for long Markov times. It prefers a two-community solution only when describing transitions between the communities costs no more than gained from describing within-community movements in smaller communities. The map equation's modular code structure with costs in bits for both community exits and entries has a regularising effect as long as the network is connected. In contrast, Markov stability prefers any solution that traps the slightest more flows within the communities for Markov time t than expected after infinite time. For example, only the map equation prefers the one-community solution for the longer Markov times included in Fig. 4.2. As a result, Markov stability typically has a longer range of non-trivial solution and responds slower to changes in Markov time.

For the schematic hypergraph, both methods prefer two communities with flows biased to larger hyperedges and three communities when biased to smaller hyperedges (Fig. 4.2). The map equation's continuous-time approximation slightly shifts the optimal solutions to longer Markov times in this case.

4.4 Experiments

To illustrate how Markov stability and the map equation differ applied to real hypergraphs, we analysed a zoo hypergraph, a collaboration hypergraph, and a hypergraph from fossil records.

4.4.1 A Zoo Hypergraph

The zoo hypergraph has been built using data from the UCI Machine Learning Depository [48]; it consists of 101 animals nodes and 20 features hyperedges such as fur, having a tail, and the number of legs. We identified communities for random-

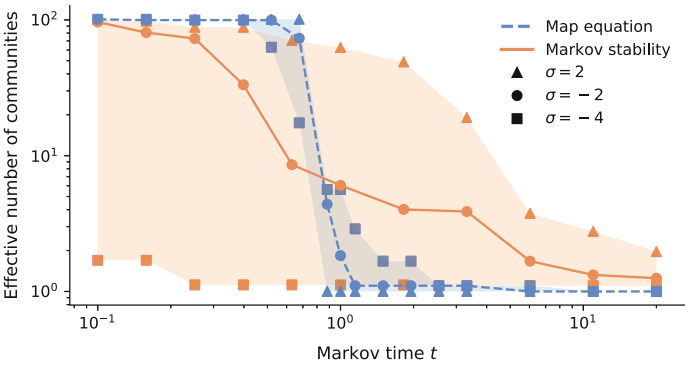


Fig. 4.3 Effective number of communities in the zoo hypergraph for different hyperedge-size bias σ and Markov time t. Filled areas between the highest and lowest number of effective communities for Markov stability (orange) and the map equation (blue)

walk bias $\sigma = -4$, -2, and 2 as a function of Markov time for Markov stability using public code [49] and the map equation using Infomap [34].

The effective number of communities, measured as the perplexity of the relative community sizes $\mathcal{S}$, $2^{H(\mathcal{S})}$, decreases gradually for Markov stability and more abruptly for the map equation (Fig. 4.3). Already at Markov time 0.1, Markov stability applied to the zoo hypergraph with hyperedge-size bias $\sigma = -4$ identifies less than two effective communities. The bias toward small hyperedges fragments the network into weakly connected components. The penalty term $-\boldsymbol{\pi}^\top \boldsymbol{\pi}$ in Eq. (4.17) with global flows makes Markov stability sense this large-scale network structure even if most flows—more than 90 per cent in this case—remain at the node for the short Markov time.

In contrast, the map equation prefers the singleton solution for the three tested hyperedge-size biases for Markov time up to 0.5. A short Markov time makes encoding each node in its singleton community cheap since most flows remain there in each step. The extra cost from describing movements in larger communities is higher than the gain from describing fewer transitions between communities until the Markov time approaches 1.

Markov stability identifies the largest communities with flows biased to small hyperedges for all Markov times. The map equation applied to a continuous-time process shows similar behaviour. But for the approximation shown in Fig. 4.3, it favours the one-community solution with flows biased to large hyperedges for Markov time about 0.9 or higher. Encoding all nodes in the same community best compresses the interconnected flows biased to large hyperedges.

The various community structures at Markov time 1 illustrates that Markov stability and the map equation work differently. They identify different communities for similar flows and similar communities for different flows. Between $\sigma = -2$ and -4, Markov stability and the map equation find closest solutions for differing σ. Markov stability for $\sigma = -2$ and the map equation for $\sigma = -4$ find communities that best capture animal classes (Fig. 4.4).

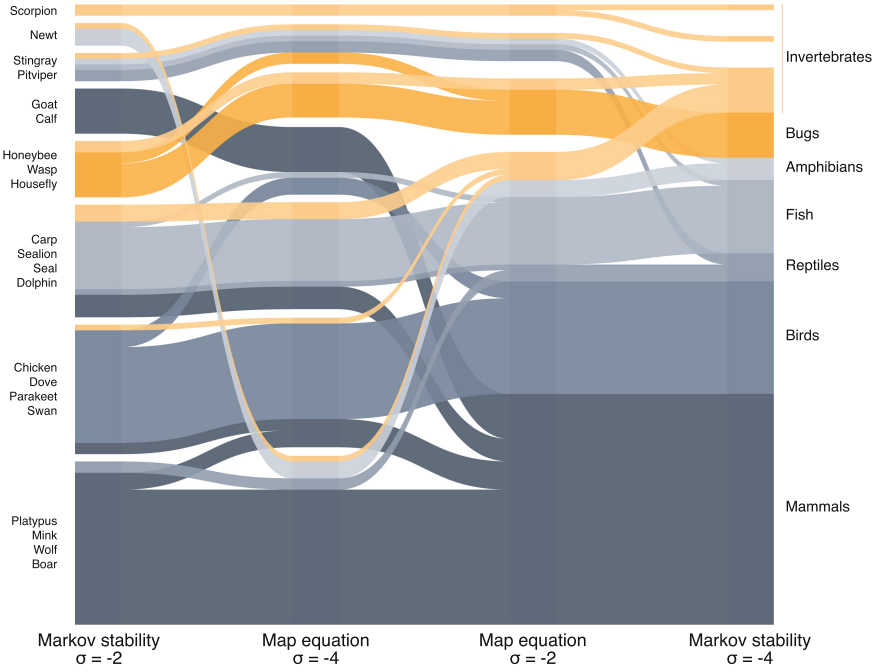


Fig. 4.4 Alluvial diagram of the zoo hypergraph solutions for Markov time $t = 1$ and different hyperedge biases σ. Each block represents a community with block height proportional to the number of nodes in the community. Stream fields connect blocks with shared nodes. Colours indicate animal classes

4.4.2 A Collaboration Hypergraph

We used a collaboration hypergraph compiled from the 734 references in the review paper *Networks beyond pairwise interactions: Structure and dynamics* [9, 35]. Hyperedges represent referenced articles and nodes represent their authors. Authors with multiple articles form connections between the hyperedges. We analysed the largest connected component with 361 author nodes in 220 hyperedges.

Again, changes in Markov time have a more gradual effect on Markov stability than on the map equation. But for the collaboration network, the map equation transitions from the singleton to the all-in-one solution over three orders of magnitude and identifies non-trivial solutions for Markov times over 100 (Fig. 4.5). Different Markov times highlight different hierarchical levels in the hypergraph's hierarchical community structure with smaller communities nested in larger communities. The largest communities remain clear despite long walks that multiply flows tenfold between the communities.

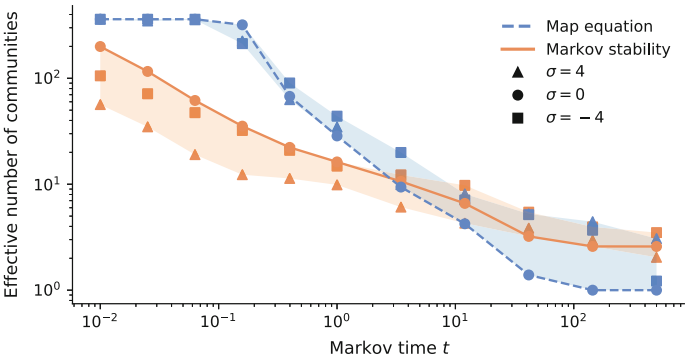


Fig. 4.5 Effective number of communities in the collaboration hypergraph for different hyperedge-size bias σ and Markov time t. Filled areas between the highest and lowest number of effective communities for Markov stability (orange) and the map equation (blue)

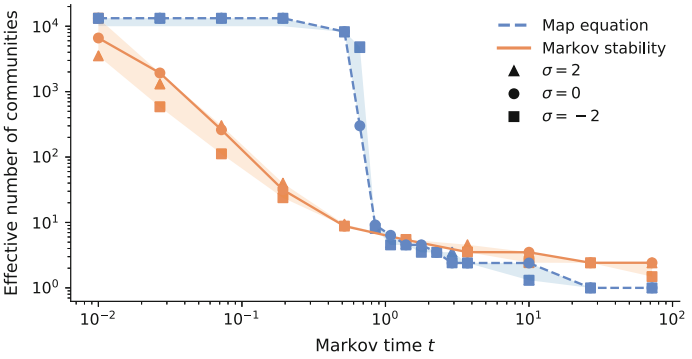


Fig. 4.6 Effective number of communities in the fossil hypergraph for different hyperedge-size bias σ and Markov time t. Filled areas between the highest and lowest number of effective communities for Markov stability (orange) and the map equation (blue)

4.4.3 A Fossil-Record Hypergraph

We analysed a hypergraph representation of marine fossil animals from Cambrian (541 MY) to Cretaceous (66 MY) [50]. Geological stages in the underlying sample-based occurrence data form hyperedges connecting all genera occurring at each stage. Genera occurring in multiple geological stages connect hyperedges. We weighted the hyperedges by dividing the number of samples where a genus occurs in a given geological stage by the total number of samples recorded at the stage. The assembled hypergraph comprises 77 geological stage hyperedges and 13,276 fossil genera nodes [35].

Once more, the effective number of identified communities decreases gradually for the Markov stability without reaching the all-in-one solution even at high Markov times (Fig. 4.6). In contrast, the map equation solutions quickly transition from sin-

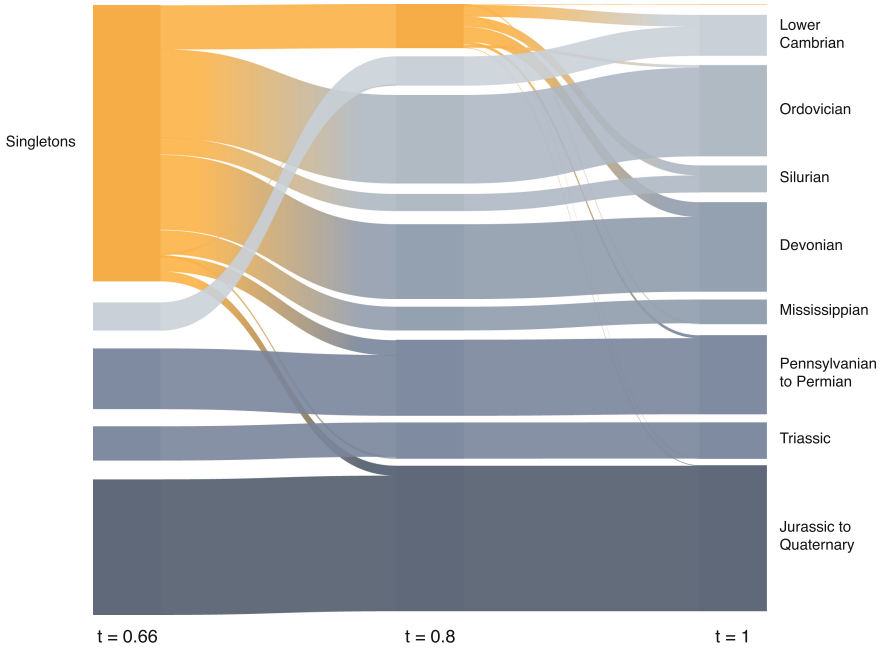


Fig. 4.7 Alluvial diagram of the map equation's fossil hypergraph solutions for hyperedge-size bias $\sigma = 0$. Each block represents a community with block height proportional to the number of nodes in the community. Stream fields connect blocks with shared nodes. Singleton communities aggregated in orange blocks for clarity

gleton to all-in-one communities around Markov time 1. In the narrow range of Markov times with non-trivial solutions, the effective number of communities is two to five. For this substantially larger hypergraph, the extensive hyperedges form a weak community structure that dissolves for long Markov times and prevents modular compression.

Non-trivial Infomap solutions reproduce the underlying temporal structure of the paleontological data with faunas organised into units of geological time (Fig. 4.7). Although fossil genera can occur in more than one of these large-scale temporal units, Infomap identifies successive global faunas that replace each other at their boundaries. Overall, faunas from geological periods are clustered together or combined into coarser temporal units at Markov time 1. These salient structures appear from singletons without intermediate smaller structures for shorter Markov times.

Markov stability also delineates temporal faunas limited by geological units at Markov time around 1. However, the temporal structure of the data cannot explain the numerous communities Markov stability identifies for short Markov times (Fig. 4.8).

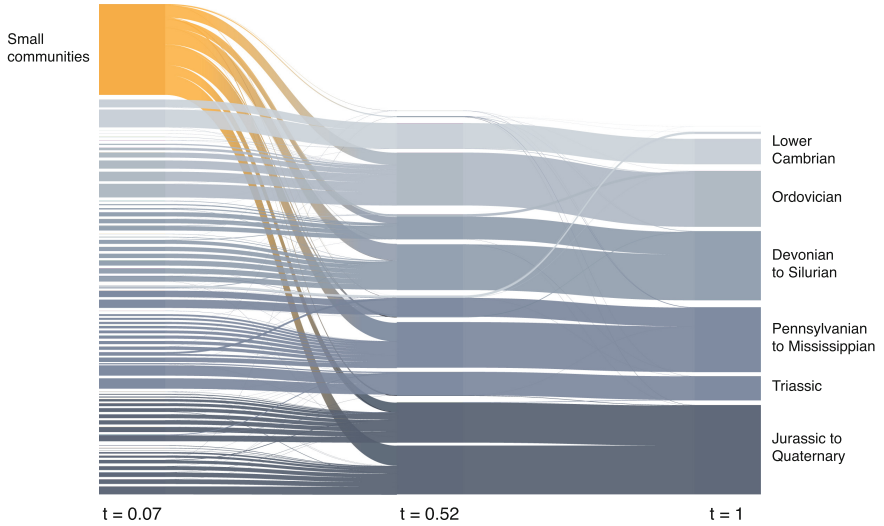


Fig. 4.8 Alluvial diagram of Markov stability's fossil hypergraph solutions for hyperedge-size bias $\sigma = 0$. Each block represents a community with block height proportional to the number of nodes in the community. Stream fields connect blocks with shared nodes. Markov times chosen to match the effective number of communities in Fig. 4.7 and small communities with fewer than five nodes aggregated in an orange block for clarity

4.5 Conclusions

We have derived Markov stability and the map equation for a random-walk process on hypergraphs with hyperedge-size bias. Both methods identify communities where flows stay for long times. Still, with disparate machinery—Markov stability measures overrepresentation of random walkers in communities where they started, whereas the map equation measures the modular description length in bits—they capitalise on distinct flow-based structures. By comparing with the stationary expectation, Markov stability is more sensitive to hyperedge-size biases and gradually finds larger communities for longer Markov times. When increasing the Markov time, the map equation instead transitions more abruptly from identifying many small to few large communities. Compared with the influence from their disparate machineries, the map equation's approximation of a continuous-time process has a negligible effect. Whereas the map equation identifies salient structures, Markov stability can identify communities of any size irrespective of the large-scale hypergraph structure. The question and hypergraph at hand decide which method identifies the flow-based communities that best elucidate the hypergraph's studied function.

Acknowledgements A.E was supported by the Swedish Foundation for Strategic Research, Grant No. SB16-0089. A.R. and M.R. were supported by the Swedish Research Council, Grant No. 2016-00796. R.L. was supported by the EPSRC Grants No. EP/V013068/1 and EP/V03474X/1.

References

1. S. Brin, L. Page, The anatomy of a large-scale hypertextual web search engine. Comput. Netw. **30**(1–7), 107–117 (1998)
2. J. Delvenne, S. Yaliraki, M. Barahona, Stability of graph communities across time scales. Proc. Natl. Acad. Sci. U.S.A. **107**(29), 12755–12760 (2010)
3. R. Guimera, L.A.N. Amaral, Functional cartography of complex metabolic networks. Nature **433**(7028), 895–900 (2005)
4. G. Mangioni, G. Jurman, M. De Domenico, Multilayer flows in molecular networks identify biological modules in the human proteome. IEEE Trans. Netw. Sci. Eng. **7**(1), 411–420 (2018)
5. M. Rosvall, C.T. Bergstrom, Maps of random walks on complex networks reveal community structure. Proc. Natl. Acad. Sci. U.S.A. **105**(4), 1118–1123 (2008)
6. I. Simonsen, K.A. Eriksen, S. Maslov, K. Sneppen, Diffusion on complex networks: a way to probe their large-scale topological structures. Physica A **336**(1–2), 163–173 (2004)
7. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Complex networks: structure and dynamics. Phys. Rep. **424**(4–5), 175–308 (2006)
8. S. Fortunato, Community detection in graphs. Phys. Rep. **486**(3–5), 75–174 (2010)
9. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, *Networks Beyond Pairwise Interactions: Structure and Dynamics* (Phys, Rep, 2020)
10. R. Lambiotte, M. Rosvall, I. Scholtes, From networks to optimal higher-order models of complex systems. Nat. Phys. **15**(4), 313–320 (2019)
11. M. De Domenico, C. Granell, M.A. Porter, A. Arenas, The physics of spreading processes in multilayer networks. Nat. Phys. **12**(10), 901–906 (2016)
12. M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M.A. Porter, S. Gómez, A. Arenas, Mathematical formulation of multilayer networks. Phys. Rev. X **3**(4), 041022 (2013)
13. M. Kivelä, A. Arenas, M. Barthelemy, J.P. Gleeson, Y. Moreno, M.A. Porter, Multilayer networks. J. Complex Netw. **2**(3), 203–271 (2014)
14. P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, J.P. Onnela, Community structure in time-dependent, multiscale, and multiplex networks. Science **328**(5980), 876–878 (2010)
15. M. Rosvall, A.V. Esquivel, A. Lancichinetti, J.D. West, R. Lambiotte, Memory in network flows and its effects on spreading dynamics and community detection. Nat. Commun. **5**(1), 1–13 (2014)
16. I. Scholtes, N. Wider, R. Pfitzner, A. Garas, C.J. Tessone, F. Schweitzer, Causality-driven slow-down and speed-up of diffusion in non-Markovian temporal networks. Nat. Commun. **5**(1), 1–9 (2014)
17. J. Xu, T.L. Wickramarathne, N.V. Chawla, Representing higher-order dependencies in networks. Sci. Adv. **2**(5), e1600028 (2016)
18. I. Iacopini, G. Petri, A. Barrat, V. Latora, Simplicial models of social contagion. Nat. Commun. **10**(1), 1–9 (2019)
19. O. Parzanchevski, R. Rosenthal, Simplicial complexes: spectrum, homology and random walks. Random Struct. Algorithms **50**(2), 225–261 (2017)
20. V. Salnikov, D. Cassese, R. Lambiotte, Simplicial complexes and complex systems. Eur. J. Phys. **40**(1), 014001 (2018)
21. M.T. Schaub, A.R. Benson, P. Horn, G. Lippner, A. Jadbabaie, Random walks on simplicial complexes and the normalized Hodge 1-Laplacian. SIAM Rev. Soc. Ind. Appl. Math. **62**(2), 353–391 (2020)
22. T. Carletti, F. Battiston, G. Cencetti, D. Fanelli, Random walks on hypergraphs. Phys. Rev. E **101**(2), 022308 (2020)
23. T. Carletti, D. Fanelli, R. Lambiotte, Random walks and community detection in hypergraphs. J. Phys. Complex. **2**(1), 015011 (2021)
24. U. Chitra, , B.J. Raphael, Random walks on hypergraphs with edge-dependent vertex weights, in *36th International Conference on Machine Learning, ICML 2019* (International Machine Learning Society (IMLS), 2019), pp. 2002–2011

25. Zhou, D., Huang, J., Schölkopf, B.: Learning with hypergraphs: clustering, classification, and embedding, in *Advances in Neural Information Processing Systems* (2007), pp. 1601–1608
26. M. De Domenico, A. Lancichinetti, A. Arenas, M. Rosvall, Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. Phys. Rev. X **5**(1), 011027 (2015)
27. L.G. Jeub, M.W. Mahoney, P.J. Mucha, M.A. Porter et al., A local perspective on community structure in multilayer networks. Netw. Sci. **5**(2), 144–163 (2017)
28. M.C. Angelini, F. Caltagirone, F. Krzakala, L. Zdeborová, Spectral detection on sparse hypergraphs, in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (IEEE, 2015), pp. 66–73
29. Chien, I., Lin, C.Y., Wang, I.H.: Community detection in hypergraphs: optimal statistical limit and efficient algorithms, in *International Conference on Artificial Intelligence and Statistics* (PMLR, 2018), pp. 871–879
30. P.S. Chodrow, N. Veldt, A.R. Benson, Hypergraph clustering: From blockmodels to modularity. Sci. Adv. **7**(28) abh 1303 (2021)
31. B. Kamiński, V. Poulin, P. Prałat, P. Szufel, F. Théberge, Clustering via hypergraph modularity. PloS one **14**(11), e0224307 (2019)
32. Z.T. Ke, F. Shi, D. Xia, Community detection for hypergraph networks via regularized tensor power iteration. arXiv:1909.06503 (2019)
33. P. Li, O. Milenkovic, Inhomogeneous hypergraph clustering with applications, in *Advances in Neural Information Processing Systems*, vol. 30, ed. by I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Curran Associates, Inc., 2017). https://proceedings.neurips.cc/paper/2017/file/a50abba8132a77191791390c3eb19fe7-Paper.pdf
34. D. Edler, A. Eriksson, M. Rosvall, The Infomap Software Package (2021). https://www.mapequation.org
35. A. Eriksson, D. Edler, A. Rojas, M. de Domenico, M. Rosvall, How choosing random-walk model and network representation matters for community detection in hypergraphs. Commun. Phys. **4**, 133 (2021)
36. J.C. Delvenne, S.N. Yaliraki, M. Barahona, Stability of graph communities across time scales. Proc. Nat. Acad. Sci. **107**, 12755 (2010)
37. R. Lambiotte, J.C. Delvenne, M. Barahona, Random walks, Markov processes and the multiscale modular organization of complex networks. IEEE Trans. Netw. Sci. Eng. **1**, 76 (2014)
38. J.C. Delvenne, M. Schaub, S.N. Yaliraki, M. Barahona, The stability of a graph partition: a dynamics-based framework for community detection, in *Dynamics On and Of Complex Networks (II)* (2013), p. 221
39. J. Reichardt, S. Bornholdt, Statistical mechanics of community detection. Phys. Rev. E **74**(1), 016110 (2006)
40. M. Kheirkhahzadeh, A. Lancichinetti, M. Rosvall, Efficient community detection of network flows for varying Markov times and bipartite networks. Phys. Rev. E **93**(3), 032309 (2016)
41. D. Edler, L. Bohlin et al., Mapping higher-order network flows in memory and multilayer networks with infomap. Algorithms **10**(4), 112 (2017)
42. V.D. Blondel, J.L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. J. Stat. Mech. **2008**(10), P10008 (2008)
43. V.A. Traag, L. Waltman, N.J. Van Eck, From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. **9**(1), 1–12 (2019)
44. M.E. Newman, M. Girvan, Finding and evaluating community structure in networks. Phys. Rev. E **69**(2), 026113 (2004)
45. M. Rosvall, C.T. Bergstrom, Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. PloS one **6**(4), e18209 (2011)
46. C. Blöcker, M. Rosvall, Mapping flows on bipartite networks. Phys. Rev. E **102**(5), 052305 (2020)
47. J. Smiljanić, D. Edler, M. Rosvall, Mapping flows on sparse networks with missing links. Phys. Rev. E **102**(1), 012302 (2020)

48. D. Dua, C. Graff, UCI machine learning repository (2017). https://archive.ics.uci.edu/ml/datasets/Zoo
49. A. Delmotte, M. Schaub, Community detection using the stability of a graph partition (2012). http://wwwf.imperial.ac.uk/~mpbara/Partition_Stability/
50. A. Rojas, J. Calatayud, M. Kowalewski, M. Neuman, M. Rosvall, A multiscale view of the Phanerozoic Fossil record reveals the three major biotic transitions. Commun. Biol. **4**(1), 309 (2021)

Chapter 5
Pattern Formation on Hypergraphs

Timoteo Carletti and Duccio Fanelli

Abstract We present a general framework that enables one to model high-order interactions among entangled dynamical systems, via hypergraphs. Several relevant processes can be ideally traced back to the proposed scheme. We shall here solely elaborate on the conditions that seed the spontaneous emergence of patterns, spatially heterogeneous solutions resulting from the many-body interaction between fundamental units. In particular we will focus, on two relevant settings. First, we will assume long-ranged mean field interactions between populations, and then turn to considering diffusive-like couplings. Two applications are presented, respectively to a generalised Volterra system and the Brusselator model.

5.1 Introduction

The study of many body interactions has a long history in science and technology, and relevant results have been obtained under the assumption of regularity of the underlying substrates, where the dynamics eventually develops. When regularity gets lost, general results are scarce and simplifying assumptions, which implement dedicated approximations, need to be put forward. It is for instance customary to reduce the many body exchanges within a pool of simultaneously interacting entities to a vast collection of pairwise contacts, a working ansatz which drastically reduces the intimate complexity of the scrutinised dynamics. Governing dynamical systems are hence cast on top of networks [1, 2] with diverse and variegated topologies: each node contains a replica of the original system, and the strength of interaction is set by the weight of the associated link.

T. Carletti (✉)
naXys, Namur Institute for Complex Systems, University of Namur, rue de Bruxelles 61, B5000 Namur, Belgium
e-mail: timoteo.carletti@unamur.be

D. Fanelli
Dipartimento di Fisica e Astronomia, CSDC and INFN, Università degli Studi di Firenze, via G. Sansone 1, 50019 Sesto Fiorentino, Italy
e-mail: duccio.fanelli@unifi.it

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_5

Despite this crude approximation, relevant results have been obtained which bear general interest [3–5]. At the same time many examples of systems exist for which the above assumption holds true just as a first order approximation [6, 7]. To overcome this intrinsic limitation, the effect of aggregated structures of nodes, such as cliques, modules or communities [5, 8] has been recently addressed in the literature. This implies analysing the cooperative interference within bunches of tightly connected nodes and assessing their role in shaping the ensuing dynamics, in the framerwok of a generalized picture which accounts for multiple pairwise exchanges.

There are however several examples where the interactions among individuals, being them neurons [9, 10], proteins [11], animals [12, 13] or authors of scientific papers [14, 15], cannot be reduced to binary interactions. The group action is indeed the real driver of the dynamics. Starting from this observation, higher-order models have been developed so as to capture the many body interactions among individual units. We hereby focus on hypergraphs [16–18], versatile tools with a broad potential that is still being fully elucidated. Hypergraphs have been applied to different fields from social contagion model [19, 20], to the modelling of random walks [14], from the study of synchronisation [21–23] and diffusion [20], to non-linear consensus [24], via the emergence of Turing patterns [21]. It is also worth mentioning an alternative approach to high-order interactions which exploits the notion of simplicial complexes [25–27]. Largely used in the past to tackle optimisation or algebraic problems, they have been recently invoked to address problems in epidemic spreading [28, 29] or synchronisation phenomena [30–32]. In this work we will however adopt the viewpoint of hypergraphs, to represent high-order interactions.

Hypergraphs constitute indeed a very flexible paradigm. An arbitrary number of agents are allowed to interact: an hyperedge grouping all the involved agents encodes for the many body interaction, thus extending conventional network models beyond the limit of binary contacts. A hypergraph can reproduce, in a proper limit, a simplicial complex and, in this respect, provides a more general tool for addressing many body simultaneous interactions.

Based on the above, it can be claimed that many body interactions constitute a relevant and transversal research field that is still in its embryonic stage, in particular as concerns studies that relate to hypergraphs. Our contribution is positioned in this context and aims at systematising the study of dynamical systems coupled via a hypergraph. For a sake of definitiveness, we will hereby consider the interactions to be mediated by the hyperedges, that is by the (hyper)adjacency matrix (see Sect. 5.2), or by a diffusive-like process, that is implemented via a properly engineered Laplace matrix (see Sect. 5.3). In both cases, we will be interested in the emergence of spatially heterogeneous solutions, i.e., coherent and extended patterns.

5.2 Hypergraphs and High-Order Interactions

The aim of this section is to introduce the formalism of (hyper) adjacency matrix which enables us to account for the high-order interaction among several identical dynamical systems. We will then present a first study on the emergence of spatial

heterogeneous solutions, i.e., patterns, for systems interacting via a hypergraph, by assuming that uncoupled individual units do converge to a (spatially) homogeneous stable solution.

5.2.1 Hypergraphs

An hypergraph $\mathcal{H}(V, E)$ is defined by a set of nodes, $V = \{v_1, \ldots, v_n\}$, and a set of m hyperedges $E = \{E_1, \ldots, E_m\}$, such that for all $\alpha = 1, \ldots, m : E_\alpha \subset V$. If all hyperedges have size 2 then the hypergraph reduces to a network. A simplicial complex is recovered if each hyperedge contains all its subsets.

One can encode the information on how the nodes are shared among hyperedges, by using the *incidence matrix of the hypergraph*,[1] $e_{i\alpha}$, namely

$$e_{i\alpha} = \begin{cases} 1 & \mathrm{v_i} \in E_\alpha \\ 0 & \text{otherwise}\,. \end{cases} \tag{5.1}$$

Given the latter, one can construct the $n \times n$ hypergraph adjacency matrix,

$$\mathbf{A} = \mathbf{e}\,\mathbf{e}^\top\,, \quad A_{ij} = \sum_\alpha e_{i\alpha} e_{j\alpha}\,, \tag{5.2}$$

thus A_{ij} represents the number of hyperedges containing both nodes i and j. Let us observe that often in the literature the adjacency matrix is defined by imposing a null diagonal. In the following we will adopt a different notation by defining its diagonal to contain all 1's. This in turn amounts to assume the hypergraph to contain all the trivial hyperedges made of just a single node. Finally we define the $m \times m$ hyperedges matrix

$$\mathbf{C} = \mathbf{e}^\top \mathbf{e}\,, \quad C_{\alpha\beta} = \sum_i e_{i\alpha} e_{i\beta}, \tag{5.3}$$

$C_{\alpha\beta}$ counts the number of nodes in $E_\alpha \cap E_\beta$, hence $C_{\alpha\alpha}$ is the size of the hyperedge E_α.

5.2.2 High-Order Coupling

Let us consider a d-dimensional dynamical system described by the ODE:

$$\frac{d\mathbf{x}}{dt}(t) = \mathbf{f}(\mathbf{x}(t)), \tag{5.4}$$

[1] We will adopt the convention of using roman indexes for nodes and greek ones for edges.

where $\mathbf{x}(t) = (x_1(t), \ldots, x_d(t))^\top$ denotes the state of the system at time t and $\mathbf{f}$ is a generic nonlinear function which describes the rate of variation of $\mathbf{x}$. Assume now to replicate system (5.4) into n independent copies, hence yielding a (tensorial) system

$$\frac{d\mathbf{x}^{(i)}}{dt}(t) = \mathbf{f}(\mathbf{x}^{(i)}(t)) \quad \forall i = 1, \ldots, n, \tag{5.5}$$

where $\mathbf{x}^{(i)}(t) = (x_1^{(i)}(t), \ldots, x_d^{(i)}(t))^\top$ denotes the state of the i-th copy of the generalised system. The whole system will thus be described by the $n \times d$ vector $\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)})^\top$. Finally we allow each system (5.5) to simultaneously interact with many others, and specifically belonging to the same hyperedge.

Let thus E_α be an hyperedge containing the i-th system. Then the growth rate associated to this latter will depend on all the systems $j \neq i$, belonging to the same hyperedge; moreover we assume such interaction to depend also on the hyperedge size, $\varphi(C_{\alpha\alpha})$, for a generic function φ. The system i may belong to several hyperedges E_α and thus all these contributions should be taken into account to determine its growth rate. In formula

$$\frac{d\mathbf{x}^{(i)}}{dt}(t) = \frac{\sum_\alpha e_{i\alpha} \sum_j e_{j\alpha} \varphi(C_{\alpha\alpha}) \mathbf{F}(\mathbf{x}^{(i)}(t), \mathbf{x}^{(j)}(t))}{\sum_\alpha e_{i\alpha} \sum_j e_{j\alpha} \varphi(C_{\alpha\alpha})} \quad \forall i = 1, \ldots, n, \tag{5.6}$$

where we introduced the function $\mathbf{F}$ such that $\mathbf{F}(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}) = \mathbf{f}(\mathbf{x}^{(i)})$ and the term at the denominator acts as a normalisation factor. We will show later on, that different functions $\mathbf{F}$ can be used to return the same function $\mathbf{f}$.

Let us define the $m \times m$ diagonal matrix Φ such that $\Phi_{\alpha\alpha} = \varphi(C_{\alpha\alpha})$ and zero otherwise. Then we can rewrite Eq. (5.6) as follows

$$\frac{d\mathbf{x}^{(i)}}{dt}(t) = \frac{1}{d_i} \sum_j D_{ij} \mathbf{F}(\mathbf{x}^{(i)}(t), \mathbf{x}^{(j)}(t)) \quad \forall i = 1, \ldots, n, \tag{5.7}$$

where we introduced the matrix $\mathbf{D} = \mathbf{e}\, \Phi\, \mathbf{e}^\top$ whose elements read

$$D_{ij} = \sum_\alpha e_{i\alpha} \Phi_{\alpha\alpha} e_{j\alpha} \quad \forall i \neq j \text{ and } D_{ii} = \varphi(1). \tag{5.8}$$

Let us observe that the different definition for the diagonal elements is due to the inclusion of the trivial hyperedges containing each single node and thus having size 1. Finally let use define $d_i = \sum_j D_{ij}$.

Remark 5.1 *(Isolated systems)* In the case n systems are isolated, i.e., all the hyperedges have size 1, then $C_{\alpha\alpha} = 1$ for all α. Observing that a single α' (the one associated to the unique hyperedge containing i) does satisfy $e_{i\alpha'} = 1$ (all the other ones being zero, $e_{i\beta} = 0$ for all $\beta = \alpha'$), we can rewrite equation (5.6) by remarking that the sum over j is restricted to $j = i$:

$$\frac{d\mathbf{x}^{(i)}}{dt}(t) = \frac{\varphi(1)\mathbf{F}(\mathbf{x}^{(i)}(t), \mathbf{x}^{(i)}(t))}{\varphi(1)} = \mathbf{f}(\mathbf{x}^{(i)}(t)) \quad \forall i = 1, \dots, n,$$

where use has been made of the relation $\mathbf{F}(\mathbf{x}^{(i)}, \mathbf{x}^{(i)}) = \mathbf{f}(\mathbf{x}^{(i)})$. Because our formalism contains the trivial case of isolated systems (5.5), it results thus a natural extension of the latter.

Remark 5.2 *(Pairwise interacting systems)* In case of systems interacting in pairs, i.e., when all hyperedges have size $C_{\alpha\alpha} = 2$ for all α (but the ones associated to the trivial hyperedges containing each node), we can show that equation (5.6) converges back to the usual setting of a dynamical model anchored on a conventional network [33], once we assume $\varphi \equiv 1$, namely the same unitary weight is associated to each link.

First of all, let us observe that $D_{ii} = (\mathbf{e}\,\Phi\,\mathbf{e}^\top)_{ii} = \varphi(1)A_{ii}$ while for $i \neq j$ we have $D_{ij} = (\mathbf{e}\,\Phi\,\mathbf{e}^\top)_{ij} = \varphi(2)A_{ij}$, where we used the definition of the adjacency matrix that includes self-loops. Then Eq. (5.7) can be rewritten as

$$\frac{d\mathbf{x}^{(i)}}{dt}(t) = \frac{\sum_j A_{ij}\mathbf{F}(\mathbf{x}^{(i)}(t), \mathbf{x}^{(j)}(t))}{k_i} \quad \forall i = 1, \dots, n\,,$$

where use has been made of the definition $k_i = \sum_j A_{ij}$.

5.2.3 Dynamical Behaviour

Assume $\mathbf{s}(t)$ to be a solution of the initial system (5.4), then $\mathbf{x}^{(i)}(t) = \mathbf{s}(t)$, $i = 1, \dots, n$, is trivially also a homogeneous solution of Eq. (5.5) but also of Eq. (5.7). Indeed, for all $i = 1, \dots, n$ one has

$$\begin{aligned}\frac{d\mathbf{x}^{(i)}}{dt}(t) &= \frac{1}{d_i}\sum_j D_{ij}\mathbf{F}(\mathbf{x}^{(i)}(t), \mathbf{x}^{(j)}(t))\Big|_{\mathbf{x}^{(i)}(t)=\mathbf{s}(t)} = \frac{1}{d_i}\sum_j D_{ij}\mathbf{F}(\mathbf{s}(t), \mathbf{s}(t)) \\ &= \frac{1}{d_i}\sum_j D_{ij}\mathbf{f}(\mathbf{s}(t)) = \mathbf{f}(\mathbf{s}(t)),\end{aligned} \tag{5.9}$$

where we used the property $\mathbf{F}(\mathbf{s}, \mathbf{s}) = \mathbf{f}(\mathbf{s})$ and the definition of d_i. By definition of $\mathbf{s}$ the rightmost term equals $\dot{\mathbf{s}}$ which thus coincides also with the leftmost term.

Consider now a spatially dependent perturbation, i.e., a node depending one, about the homogeneous solution, $\mathbf{x}^{(i)}(t) = \mathbf{s}(t) + \mathbf{u}^{(i)}(t)$. Insert this ansatz into Eq. (5.7) and determine the evolution of $\mathbf{u}^{(i)}(t)$ by assuming it to be small (i.e., using a first order expansion), $\forall i = 1, \dots, n$:

$$\frac{d\mathbf{u}^{(i)}}{dt}(t) + \frac{d\mathbf{s}}{dt}(t) = \frac{1}{d_i}\sum_j D_{ij}\mathbf{F}(\mathbf{s} + \mathbf{u}^{(i)}, \mathbf{s} + \mathbf{u}^{(j)})$$

$$
\begin{aligned}
&= \mathbf{f}(\mathbf{s}) + \frac{1}{d_i}\sum_j D_{ij}\left(\sum_\ell \partial_{x_\ell^{(i)}}\mathbf{F}(\mathbf{s},\mathbf{s})u_\ell^{(i)} + \sum_\ell \partial_{x_\ell^{(j)}}\mathbf{F}(\mathbf{s},\mathbf{s})u_\ell^{(j)}\right)\\
&= \mathbf{f}(\mathbf{s}) + \sum_\ell \partial_{x_\ell^{(i)}}\mathbf{F}(\mathbf{s},\mathbf{s})u_\ell^{(i)} + \frac{1}{d_i}\sum_j D_{ij}\sum_\ell \partial_{x_\ell^{(j)}}\mathbf{F}(\mathbf{s},\mathbf{s})u_\ell^{(j)}\\
&= \mathbf{f}(\mathbf{s}) + \mathbf{J}_1\mathbf{u}^{(i)} + \frac{1}{d_i}\sum_j D_{ij}\mathbf{J}_2\mathbf{u}^{(j)}\,,
\end{aligned}
$$

where we defined the Jacobian matrices $\mathbf{J}_1 = \partial_{\mathbf{x}_1}\mathbf{F}(\mathbf{s},\mathbf{s})$, i.e., the derivatives are computed with respect to the first group of variables, and $\mathbf{J}_2 = \partial_{\mathbf{x}_2}\mathbf{F}(\mathbf{s},\mathbf{s})$, i.e., the derivatives are performed with respect to the second group of variables. In both cases the derivatives are evaluated at the reference solution $\mathbf{s}$.

By using the fact that $\dot{\mathbf{s}} = \mathbf{f}(\mathbf{s})$ and by slightly rewriting the previous equation, we obtain

$$
\begin{aligned}
\frac{d\mathbf{u}^{(i)}}{dt}(t) &= \mathbf{J}_1\mathbf{u}^{(i)} + \frac{1}{d_i}\sum_j D_{ij}\mathbf{J}_2\mathbf{u}^{(j)} = \mathbf{J}_1\mathbf{u}^{(i)} + \mathbf{J}_2\mathbf{u}^{(i)} + \sum_j\left(\frac{D_{ij}}{d_i} - \delta_{ij}\right)\mathbf{J}_2\mathbf{u}^{(j)}\\
&= (\mathbf{J}_1 + \mathbf{J}_2)\,\mathbf{u}^{(i)} + \sum_j \mathcal{L}_{ij}\mathbf{J}_2\mathbf{u}^{(j)},
\end{aligned}
$$

where we defined the matrix operator

$$
\mathcal{L}_{ij} = \frac{D_{ij}}{d_i} - \delta_{ij}. \tag{5.10}
$$

By introducing the $n \times d$ vector $\mathbf{u} = (\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(n)})^\top$ we can rewrite the latter equation in a compact form as:

$$
\frac{d\mathbf{u}}{dt}(t) = [(\mathbf{J}_1 + \mathbf{J}_2) \otimes \mathbf{I}_n + \mathbf{J}_2 \otimes \mathcal{L}]\,\mathbf{u}, \tag{5.11}
$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix and $\otimes$ is the Kronecker product of matrices.

One can prove that $\mathcal{L}$ is a novel (consensus) high-order Laplace matrix[2], i.e., it is nonpositive definite, the largest eigenvalue is $\Lambda^{(1)} = 0$ and its is associated to the uniform eigenvector $\phi^{(1)} \sim (1, \ldots, 1)^\top$.

[2] Let us introduce $\mathcal{L}^{sym} = \mathbf{d}^{-1/2}\mathbf{L}^H\mathbf{d}^{-1/2}$, where $\mathbf{d}$ is the diagonal matrix containing the d_i's on the diagonal and $\mathbf{L}^H$ is the high-order (combinatorial) Laplace matrix defined in[21]. Then $\mathcal{L}^{sym} = D_{ij}/\sqrt{d_i d_j} - \delta_{ij}$ from which it immediately follows that $\mathcal{L}^{sym}$ is symmetric and nonpositive definite; indeed take any $\mathbf{x} \in \mathbb{R}^N \setminus \{0\}$, N standing for the dimension of the matrices, then $(\mathbf{x}, \mathcal{L}^{sym}\mathbf{x}) = (\mathbf{d}^{-1/2}\mathbf{x}, \mathbf{L}^H\mathbf{d}^{-1/2}\mathbf{x}) \leq 0$ where the last inequality follows from the fact that $\mathbf{L}^H$ is nonpositive definite. Finally let us observe that $\mathcal{L} = \mathbf{d}^{-1}\mathbf{L}^H = \mathbf{d}^{-1/2}\mathcal{L}^{sym}\mathbf{d}^{1/2}$, hence, $\mathcal{L}$ is similar to $\mathcal{L}^{sym}$ and, thus they display the same non-positive spectrum. Moreover this implies also that $-2 \leq \Lambda^{(\alpha)} \leq 0$.

Recalling the relation $\mathbf{f}(\mathbf{x}) = \mathbf{F}(\mathbf{x}, \mathbf{x})$ one can prove that:

$$\partial_{\mathbf{x}}\mathbf{f} := \mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2,$$

and thus rewrite Eq. (5.11) as

$$\frac{d\mathbf{u}}{dt}(t) = [\mathbf{J} \otimes \mathbf{I}_n + \mathbf{J}_2 \otimes \mathcal{L}]\,\mathbf{u}. \tag{5.12}$$

This is a linear system involving matrices with size $nd \times nd$. To progress with the analytical understanding, we employ the eigenbasis of $\mathcal{L}$, to project the former equation onto each eigendirection

$$\frac{d\mathbf{u}^{(\alpha)}}{dt}(t) = \left[\mathbf{J}(\mathbf{s}(t)) + \mathbf{J}_2(\mathbf{s}(t))\Lambda^{(\alpha)}\right]\mathbf{u}^{(\alpha)}, \tag{5.13}$$

where $\Lambda^{(\alpha)}$ is the eigenvalue relative to the eigenvector $\phi^{(\alpha)}$. The above equation enables us to infer the stability of the homogeneous solution, $\mathbf{s}(t)$, by studying the Master Stability Function, namely the real part of the largest Lyapunov exponent of Eq. (5.13). To illustrate the potentiality of the theory we shall turn to considering a specific application that we will introduce in the following.

5.2.4 Results

In the above analysis we have obtained a one-parameter family (indexed by the eigenvalues $\Lambda^{(\alpha)}$) of linear but (in general) time dependent systems (5.13). For the sake of simplicity we will hypothesise the homogenous solution to be stationary and stable, $\mathbf{s}(t) = \mathbf{s}_0$. In this way we will hence assume each isolated system to converge to the same stationary point. This simplifies the study of Eq. (5.13), by allowing us to deal with a constant linear system. Let us observe that one could in principle study the more general setting of a time dependent solution, by using the Floquet theory in case of a periodic orbit or the full Master Stability Function in the case of irregular oscillators.

As a concrete application we will consider a Volterra model [34] which describes the interaction of preys and predators in an ecological setting:

$$\begin{cases} \dot{x} = -dx + c_1 xy \\ \dot{y} = ry - sy^2 - c_2 xy, \end{cases} \tag{5.14}$$

here x denotes the concentration of predators, while y stands for the preys and $\dot{}$ the time derivative. All the parameters are assumed to be positive; in the following we will make use of the choice $c_1 = 2, c_2 = 13, r = 1, s = 1$ and $d = 1/2$, but of course our results hold true in general. The Volterra model (5.14) admits a nontrivial fixed-

point, $x^* = \frac{c_1 r - sd}{c_1 c_2}$, $y^* = \frac{d}{c_1}$, which is positive and stable, provided $c_1 r - sd > 0$. In the case under scrutiny, we have $x^* = 3/52 \sim 0.0577$ and $y^* = 1/4$.

Following the above presented scheme, let us now considering n replicas of the model (5.14), each associated to a different ecological niche and indexed by the node index i. Assume also that species can sense the remote interaction with other communities populating neighbouring nodes. For instance, the competition of preys for food and resources can be easily extended so as to account for a larger habitat which embraces adjacent patches. At the same time, predators can benefit from a coordinated action to hunt in team. For a sake of definitiveness we will study in the following the high-order coupling (let us stress once again that several "microscopic" high-order models can give rise to the same network-aggregate model) defined by:

$$\begin{cases} \dot{x}_i = -dx_i + ac_1 y_i \frac{1}{d_i} \sum_j D_{ij} x_j + (1-a) c_1 x_i \frac{1}{d_i} \sum_j D_{ij} y_j \\ \dot{y}_i = r y_i - s y_i \frac{1}{d_i} \sum_j D_{ij} y_j - c_2 y_i \frac{1}{d_i} \sum_j D_{ij} x_j, \end{cases} \tag{5.15}$$

where the matrix D_{ij} encodes for the high-order interaction among nodes i and j, taking into account the number and size of the hyperedges containing both nodes (see (5.8)). The parameters $a \in [0, 1]$ describes the relative strength with which the predators in node i increase because of the "in-node" predation or because of the interaction among predators in the hyperedges. The case $a = 1$ corresponds to a purely in-node process while if $a = 0$ a coordinated action to hunt in team is assumed to rule the dynamics. Preys feel the competition for the resources with preys living in nodes belonging to the same hyperedge (second term on the right hand side of the second equation of (5.15)) as well from predators in the same hyperedge (rightmost terms in the same equation). Birth and death of both species are local, i.e., due to resources available in-node.

By using the new Laplace matrix (5.10) we can rewrite the previous model (5.15) as:

$$\begin{cases} \dot{x}_i = -dx_i + c_1 y_i x_i + ac_1 y_i \sum_j \mathcal{L}_{ij} x_j + (1-a) c_1 x_i \sum_j \mathcal{L}_{ij} y_j \\ \dot{y}_i = r y_i - s y_i^2 - c_2 y_i x_i - s y_i \sum_j \mathcal{L}_{ij} y_j - c_2 y_i \sum_j \mathcal{L}_{ij} x_j \,, \end{cases} \tag{5.16}$$

where one can easily recognise the in-node Volterra model (5.14) and the corrections stemming from high-order contributions.

As previously shown, in the general setting (see (5.9)) the homogenous solution (x^*, y^*) is also a solution of the coupled system (5.15), that is $x_i = x^*$ and $y_i = y^*$ solves the latter. In the following we will prove that such solution can be destabilised due to the high-order coupling so driving the system towards a new heterogenous, spatially dependent, solution. To prove this claim, we will linearise system (5.14) about the homogeneous equilibrium by setting $u_i = x_i - x^*$ and $v_i = y_i - y^*$ and then make use of the eigenbasis of the Laplace matrix $\mathcal{L}$, $(\Lambda^{(\alpha)}, \phi^{(\alpha)})$, to project the linear system onto each eigenmode, that is $u_i = \sum_\alpha u^\alpha \phi_i^{(\alpha)}$ and $v_i = \sum_\alpha v^\alpha \phi_i^{(\alpha)}$:

$$\frac{d}{dt}\begin{pmatrix} u^\alpha \\ v^\alpha \end{pmatrix} = \left[\begin{pmatrix} 0 & c_1 x^* \\ -c_2 y^* & -s y^* \end{pmatrix} + \Lambda^{(\alpha)} \begin{pmatrix} a c_1 y^* & (1-a) c_1 x^* \\ -c_2 y^* & -s y^* \end{pmatrix}\right] \begin{pmatrix} u^\alpha \\ v^\alpha \end{pmatrix}$$
$$= \left(\mathbf{J} + \Lambda^{(\alpha)} \mathbf{J}_2\right) \begin{pmatrix} u^\alpha \\ v^\alpha \end{pmatrix} =: \mathbf{J}^{(\alpha)} \begin{pmatrix} u^\alpha \\ v^\alpha \end{pmatrix}. \tag{5.17}$$

The homogenous solution will prove unstable if (at least) one eigenmode $\bar{\alpha}$ exists for which the largest real part of the eigenvalues of $\mathbf{J}^{(\bar{\alpha})}$ is positive. The real part of the largest eigenvalue λ as function of $\Lambda^{(\alpha)}$ is called the *dispersion relation*. One can easily realise that λ is the solution with the largest real part of the second order equation

$$\lambda^2 - \mathrm{tr}\mathbf{J}^{(\alpha)}\lambda + \det \mathbf{J}^{(\alpha)} = 0.$$

Hence the required condition for the instability is

$$\mathrm{tr}\mathbf{J}^{(\alpha)} > 0 \text{ or } \mathrm{tr}\mathbf{J}^{(\alpha)} < 0 \text{ and } \det \mathbf{J}^{(\alpha)} < 0. \tag{5.18}$$

A straightforward computation returns

$$\mathrm{tr}\mathbf{J}^{(\alpha)} = -s y^* + \Lambda^{(\alpha)}(-s + a c_1) \text{ and } \det \mathbf{J}^{(\alpha)}$$
$$= c_1 y^* \left(1 + \Lambda^{(\alpha)}\right)\left[\Lambda^{(\alpha)}\left(c_2 x^*(1-a) - a s y^*\right) + c_2 x^*\right].$$

Let us recall that the homogenous equilibrium is stable for the decoupled system corresponding to setting $\Lambda^{(1)} = 0$. Hence $\mathrm{tr}\mathbf{J}^{(1)} = -s y^* < 0$ and $\det \mathbf{J}^{(1)} = c_1 c_2 x^* y^* > 0$. We have thus to determine the existence of (at least one) $\bar{\alpha} \geq 2$ for which the conditions for instability (5.18) hold true, allowing us to prove the positivity of $\lambda\left(\Lambda^{(\bar{\alpha})}\right)$. In Fig. 5.1 we report a case where the high-order coupling is able to destabilise the homogenous solution (panel b), thus returning a patchy solution (panels c and d) for the involved species. Finally let us observe that interestingly some niches (6 over 20) become empty, that is deprived of any species.

Another even more interesting case is reported in Fig. 5.2. In this case the uncoupled homogeneous equilibrium yields $\tilde{x} = 0$ and $\tilde{y} = r/s$. When extending the study to account for multi body interactions, predators do survive in each niche while the preys go through extinction in a few location (8 nodes over 20). Generally the density of preys is lower than the equilibrium value found in the isolated case.

5.3 Hypergraph and High-Order Diffusive-Like Coupling

In the previous section we have introduced and studied the problem of the emergence of a spatially heterogenous solution in a system of several identical dynamical units coupled together via the (hyper) adjacency matrix of the hypergraph. In particular the microscopic units defining the system are constrained to stay anchored to the node where they interact with those sharing the same location and those belonging to nodes

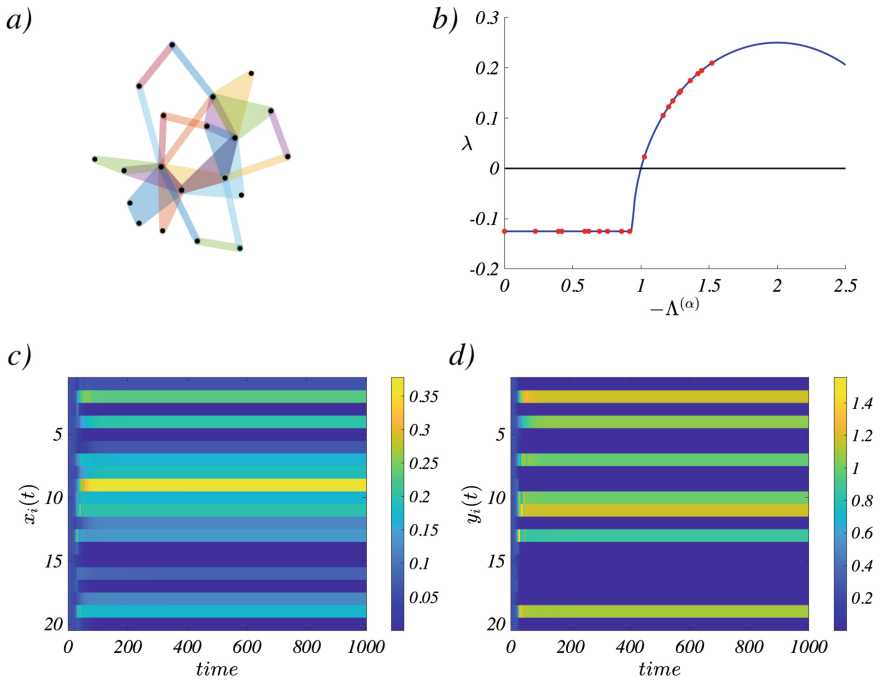


Fig. 5.1 Patterns in the Volterra model with high-order interactions (I). In panel **a** we represent the hypergraph used to model the high-order interactions among species living in different niches. The hypergraph is composed of $n = 20$ nodes, it has been generated using a random attachment process and it is composed by 20 trivial hyperedges of size 1, 11 hyperedges of size 2, 10 hyperedges of size 3 and 1 hyperedge of size 4. In panel **b** we report the dispersion relation for the Volterra model (5.15), the red symbols refer to $\lambda\left(\Lambda^{(\alpha)}\right)$, $\alpha \in \{1, \ldots, n\}$, while the blue line denotes the dispersion relation for the Volterra model reformulated on a continuous support. In panel **c** we show the time evolution of the predator density in each node as a function of time, $x_i(t)$; let us observe that in (almost) each node the density of predators is much larger than the corresponding homogenous equilibrium $x^* \sim 0.0577$ (blue). Panel **d** report the time evolution of the preys density in each node as a function of time, $y_i(t)$; let us observe that in (almost) each node the density of preys is much lower than the corresponding homogenous equilibrium $y^* = 1$ (green). The model parameters have been set to $c_1 = 2$, $c_2 = 13$, $r = 1$, $s = 1$, $d = 1/2$ and $a = 1/2$. We fix $\varphi(c) = c^\sigma$ with $\sigma = 1.5$

of the incident hyperedges. In this section we will present a modified framework based on the assumption that the basic units can travel across the hypergraph jumping from node to node via the available hyperedges.

Starting from the definition of hyper adjacency matrix, Eq. (5.2), the notion of (combinatorial) Laplace matrix for networks can be straightforwardly generalised to the case of hypergraphs [23, 35], by defining $k_i\delta_{ij} - A_{ij}$, where $k_i = \sum_j A_{ij}$. Let us however observe that the latter does not account in full for the higher-order structures encoded in the hypergraph. Notably, the sizes of the incident hyperedges are neglected.

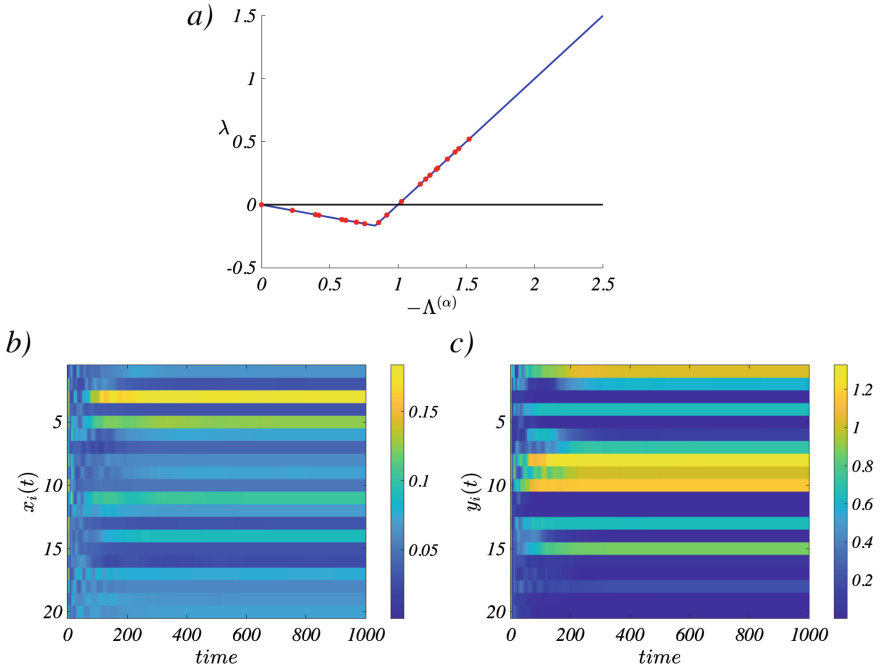


Fig. 5.2 Patterns in the Volterra model with high-order interactions (II). Using the same hypergraph shown in Fig. 5.1 we study the emergence of patterns close to the homogeneous equilibrium $\tilde{x} = 0$ and $\tilde{y} = r/s = 1$. We report in panel **a** the dispersion relation for the Volterra model (5.15), the red symbols refer to $\lambda\left(\Lambda^{(\alpha)}\right), \alpha \in \{1, \ldots, n\}$, while the blue line denotes the dispersion relation for the Volterra model computed on a continuous support. In panel **b** we show the time evolution of the predator density in each node as a function of time, $x_i(t)$; let us observe that in each node the density of predators is positive in striking contrast with it happens for the uncoupled system. Panel **c** reports the time evolution of the preys density in each node as a function of time, $y_i(t)$; let us observe that in each node the density of preys is much lower than the homogenous equilibrium $y^* = 1$ (green) and in 8 niches the preys have gone through extinction. The model parameters have been set to $c_1 = 2$, $c_2 = 13$, $r = 1$, $s = 1$, $d = 1/2$ and $a = 1/2$. We fix $\varphi(c) = c^\sigma$ with $\sigma = 1.5$

To overcome this limitation, authors of [14] studied a random walk process defined on a generic hypergraph using a new (random walk) Laplace matrix. It is worth mentioning that the transition rates of the associated process, linearly correlates with the size of the involved hyperedges. Stated differently, exchanges are favoured among nodes belonging to the same hyperedge (weighted according to its associated size). Note that a similar construction has been proposed in [36] to extract a n-clique graph from a network. The main difference in the present case is that hyperedges can have an heterogeneous size distribution and thus provide a more flexible framework for tackling a wide range of problems.

For the sake of completeness, let us briefly recall the construction of the random walk process on a hypergraph and invite the interested reader to consult [14] for further details. The agents are located on the nodes and hop between them. In a general

setting, the walkers may weight hyperedges depending on their size, introducing a bias in their moves that we shall encode into a function φ of the hyperedge size. This yields the weighted adjacency matrix $\mathbf{D} = \mathbf{e}\,\Phi\,\mathbf{e}^\top$, already defined in Eq. (5.8) and hereby recalled:

$$D_{ij} = \sum_\alpha e_{i\alpha}\Phi_{\alpha\alpha}e_{j\alpha} \quad \forall i \neq j \text{ and } D_{ii} = \varphi(1)\,,$$

where Φ is the diagonal matrix whose elements read $\varphi(C_{\alpha\alpha})$. The transition probabilities of the examined process are then obtained by normalising the columns of the weighted adjacency matrix $T_{ij} = \frac{D_{ij}}{d_i}$ for all i, where again $d_i = \sum_j D_{ij}$.

Let us briefly observe that assuming $\varphi(c) = c^\sigma$ allows to cover several existing models of random walks on hypergraphs. For $\sigma = 1$, we get the random walk defined in [14], while for $\sigma = -1$ we obtain the one introduced by Zhou [37]. Finally, the case $\sigma = 0$ returns a random walk on the so called *clique reduced multigraph*. The latter is a multigraph where each pair of nodes is connected by a number of edges equal to the number of hyperedges containing that pair in the hypergraph.

From the above introduced transition probabilities one can define the *random walk* Laplacian generalising that of standard networks, $L_{ij} = \delta_{ij} - T_{ij}$, and eventually derive the (combinatorial) Laplace matrix,

$$\mathbf{L}^H = \mathbf{D} - \mathbf{d}, \tag{5.19}$$

this latter will be employed in the following to model diffusion on higher-order structures. In the above equation, matrix $\mathbf{d}$ displays, on the diagonal, the values $d_i = \sum_j D_{ij}$ and zeros otherwise. It is clear from its very definition that $\mathbf{D}$ takes into account both the number and the size of the hyperedges incident with the nodes. It can also be noted that $\mathbf{D}$ can be considered as a *weighted adjacency matrix* whose weights have been self-consistently defined so as to account for the higher-order structures encoded in the hypergraph.

Consider again the d-dimensional system Eq. (5.4) described by local, i.e., aspatial, equations:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}) \quad \mathbf{x} \in \mathbb{R}^d, \tag{5.20}$$

and assume further n identical copies of the above system coupled through a hypergraph. In this way each copy of the system attached to a node of a hypergraph belonging to one (or more) hyperedge. Units sharing the same hyperedge are tightly coupled, due to existing many body interactions. In formulas:

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{f}(\mathbf{x}_i) + \varepsilon \sum_{\alpha: i\in E_\alpha} \sum_{j\in E_\alpha} \varphi(C_{\alpha\,\alpha}) \left(\mathbf{G}(\mathbf{x}_j) - \mathbf{G}(\mathbf{x}_i)\right),$$

where $\mathbf{x}_i$ denotes the state of the i-th unit, i.e., anchored to the i-th node, ε the strength of the coupling, φ is the function encoding the bias due to the hyperedge size and $\mathbf{G}$

a generic nonlinear coupling function. From the definition of $e_{i\alpha}$ one can rewrite the previous formula as

$$\begin{aligned}\frac{d\mathbf{x}_i}{dt} &= \mathbf{f}(\mathbf{x}_i) + \varepsilon \sum_{\alpha,j} e_{i\alpha} e_{j\alpha} \varphi(C_{\alpha\alpha}) \left(\mathbf{G}(\mathbf{x}_j) - \mathbf{G}(\mathbf{x}_i)\right) \\ &= \mathbf{f}(\mathbf{x}_i) + \varepsilon \sum_{j} D_{ij} \left(\mathbf{G}(\mathbf{x}_j) - \mathbf{G}(\mathbf{x}_i)\right) = \mathbf{f}(\mathbf{x}_i) + \varepsilon \sum_{j} \left(D_{ij} - d_i \delta_{ij}\right) \mathbf{G}(\mathbf{x}_j) \\ &= \mathbf{f}(\mathbf{x}_i) + \varepsilon \sum_{j} L^H_{ij} \mathbf{G}(\mathbf{x}_j), \end{aligned} \tag{5.21}$$

where we have used the above definitions for d_i and L^H_{ij}. Let us stress once again that the whole high-order structure is encoded in a $n \times n$ matrix. Hence there is no need for tensors and this simplifies the resulting analysis.

By exploiting the fact that $\sum_j L^H_{ij} = 0$ for all $i = 1, \ldots, n$, it is immediate to conclude that the aspatial reference solution $\mathbf{s}(t)$, i.e., the time dependent function solving Eq. (5.20), is also a solution of Eq. (5.21). A natural question hence arises: what can we say of the stability of the homogeneous solution for the system in its diffusive-like coupled variant?

To answer to this question one introduces again the deviations from the reference orbit, i.e., $\mathbf{u}_i = \mathbf{x}_i - \mathbf{s}$. Assuming this latter to be small, one can derive a self-consistent set of linear differential equations for tracking the evolution of the perturbation in time. To this end, we make use of the expression in the above Eq. (5.21) and perform a Taylor expansion to the linear order of approximation, to eventually get:

$$\frac{d\mathbf{u}_i}{dt} = \mathbf{J}(\mathbf{s}(t))\mathbf{u}_i + \varepsilon \sum_j L^H_{ij} \mathbf{J}_{\mathbf{G}}(\mathbf{s}(t))\mathbf{u}_j, \tag{5.22}$$

where $\mathbf{J}(\mathbf{s}(t))$ (resp. $\mathbf{J}_{\mathbf{G}}(\mathbf{s}(t))$) denotes the Jacobian matrix of the function $\mathbf{f}$ (resp. $\mathbf{G}$) evaluated on the trajectory $\mathbf{s}(t)$.

We can improve on our analytical understanding of the problem by employing again the eigenbasis of the Laplace matrix $\mathbf{L}^H$. Being the latter symmetric there exists a basis of orthonormal eigenvectors, $\phi^{(\alpha)}_H$, associated to the eigenvalues $\Lambda^{(\alpha)}_H$. We can then project $\mathbf{u}_i$ on this basis and obtain, for all α:

$$\frac{d\mathbf{y}_\alpha}{dt} = \left[\mathbf{J}(\mathbf{s}(t)) + \varepsilon \Lambda^{(\alpha)}_H \mathbf{J}_{\mathbf{G}}(\mathbf{s}(t))\right] \mathbf{y}_\alpha, \tag{5.23}$$

where $\mathbf{y}_\alpha$ is the projection of $\mathbf{u}_i$ on the α-th eigendirection.

The (in)stability of the homogenous solution $\mathbf{s}(t)$ can be checked by looking at the eigenvalue of the linear system (5.23), and more specifically the eigenvalue with the largest real part. In a general framework, where i.e., $\mathbf{s}(t)$ depends on time, we are dealing with a time dependent eigenvalue problem that can be tackled by using the Master Stability Function [38, 39]. For simplicity we will hereby solely consider

the case of a stationary reference orbit, i.e., $\mathbf{s}(t) = \mathbf{s}_0$. In this way Eq. (5.23) can be directly solved by using spectral methods. We invite the interested reader to refer to [21] where the general case of a periodic or even a chaotic $\mathbf{s}(t)$ has been analysed.

5.3.1 Turing Patterns on Hypergraphs

The problem introduced in the previous section opens up the perspective to address the notion of a Turing instability on hypergraphs. Indeed, according to the Turing instability mechanism, a stable homogeneous equilibrium becomes unstable upon injection of a heterogeneous, i.e., spatially dependent, perturbation once diffusion and reaction terms are simultaneously at play. The Turing phenomenon is exemplified with reference to 2 dimensional systems. In the following we will consequently assume $d = 2$ and rewrite $\mathbf{x}_i = (u_i, v_i)$ as well as $\mathbf{f}(\mathbf{x}_i) = (f(u_i, v_i), g(u_i, v_i))$, where the index $i = 1, \ldots, n$ refers to the specific node to which the dynamical variables are bound. Hence Eq. (5.21) becomes

$$\begin{cases} \dot{u_i} &= f(u_i, v_i) + D_u \sum_j L^H_{ij} u_j \\ \dot{v_i} &= g(u_i, v_i) + D_v \sum_j L^H_{ij} v_j \end{cases}, \tag{5.24}$$

where D_u and D_v replace the diffusion coefficients of species u and v in the case of network and can thus be called generalised diffusion coefficients. At first sight, the above model seems to solely account for binary interactions. However, higher-order interactions are also present, as encoded in the matrix $\mathbf{L}^H$. Finally, let us observe that if the hypergraph is a network, then $\mathbf{L}^H$ reduces to the standard Laplace matrix and thus Eq. (5.24) converges to the usual reaction-diffusion system defined on a network.

The condition for the emergence of a Turing instability can be assessed by performing a linear stability analysis about the homogeneous equilibrium [40–43], as previously shown. Assuming $\mathbf{G}$ to be the identity function and the reference orbit to coincide with a stable stationary equilibrium $\mathbf{s}_0 = (u_0, v_0)$, Eq. (5.22) simplifies into:

$$\begin{cases} \delta\dot{u}_i &= \partial_u f(u_0, v_0)\delta u_i + \partial_v f(u_0, v_0)\delta v_i + D_u \sum_j L^H_{ij} \delta u_j \\ \delta\dot{v}_i &= \partial_u g(u_0, v_0)\delta u_i + \partial_v g(u_0, v_0)\delta v_i + D_v \sum_j L^H_{ij} \delta v_j, \end{cases}$$

where $\delta u_i = u_i - u_0$ and $\delta v_i = v_i - v_0$. By exploiting again the eigenbasis of the Laplace matrix we can write $\delta u_i(t) = \sum_\alpha \hat{u}^\alpha(t)\phi^\alpha_i$ and $\delta v_i(t) = \sum_\alpha \hat{v}^\alpha(t)\phi^\alpha_i$. Finally the ansatz, $\hat{u}^\alpha(t) \sim e^{\lambda_\alpha t}$ and $\hat{v}^\alpha(t) \sim e^{\lambda_\alpha t}$, allows us to compute the dispersion relation, i.e., the linear growth rate $\lambda_\alpha = \lambda(\Lambda^\alpha_H)$ of the eigenmode α, as a function of the Laplacian eigenvalue Λ^α_H.

As it can be straightforwardly proved, the linear growth rate is the largest real part of the roots of the second order equation

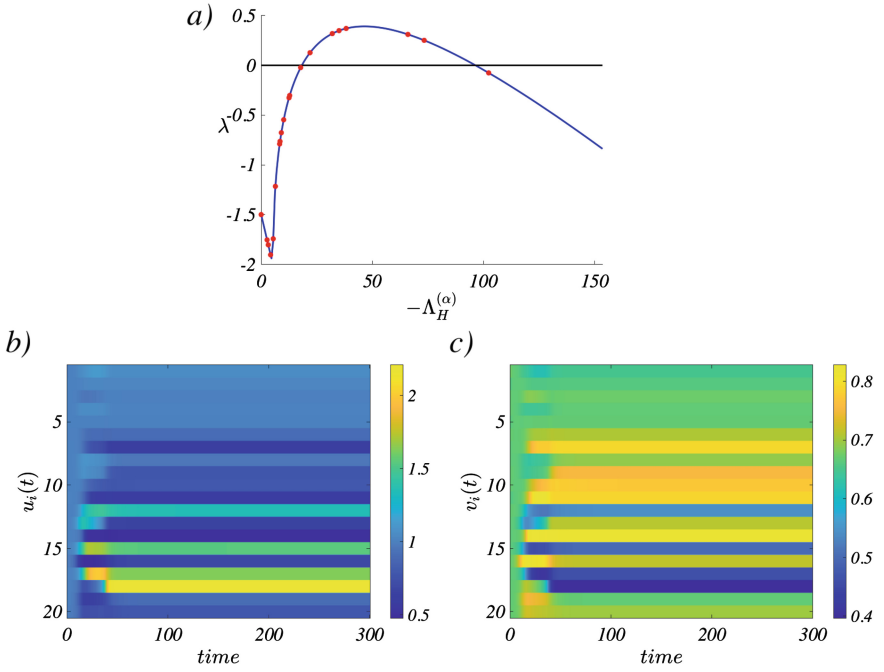


Fig. 5.3 Turing patterns in the Brusselator model with high-order diffusive-like couplling. Using the same hypergraph shown in Fig. 5.1 we study the Turing patterns emerging from the homogeneous equilibrium (u_0, v_0). We report in panel **a** the dispersion relation for the Brusselator model defined by the reaction terms $f(u, v) = 1 - (b + 1)u + cu^2 v$ and $g(u, v) = bu - cu^2 v$; the red symbols refer to $\lambda\left(\Lambda_H^{(\alpha)}\right)$, $\alpha \in \{1, \ldots, n\}$, while the blue line denotes the dispersion relation for the Brusselator model defined on a continuous support. In panel **b** we show the time evolution of the u variable in each node as a function of time, $u_i(t)$. Panel **c** reports the time evolution of the v variable in each node as a function of time, $v_i(t)$. The model parameters have been set to $b = 4$, $c = 6$, $D_u = 0.02$ and $D_v = 0.17$. Hence $u_0 = 1$ and $v_0 = b/c = 2/3$. We fix $\varphi(c) = c^\sigma$ with $\sigma = 1.5$

$$\lambda_\alpha^2 - \lambda_\alpha \left[\mathrm{tr}\mathbf{J}_0 + \Lambda_H^\alpha (D_u + D_v)\right] + \det \mathbf{J}_0 + \Lambda_H^\alpha (D_u \partial_v g + D_v \partial_u f) + D_u D_v (\Lambda_H^\alpha)^2 = 0, \tag{5.25}$$

where $\mathbf{J}_0 = \begin{pmatrix} \partial_u f & \partial_v f \\ \partial_u g & \partial_v g \end{pmatrix}$ is the Jacobian matrix of the reaction part evaluated at the equilibrium $(u_i, v_i) = (u_0, v_0)$. In Eq. (5.25), $\mathrm{tr}(\cdot)$ and $\det(\cdot)$ stand respectively for the trace and the determinant. The existence of at least one eigenvalue $\Lambda_H^{\tilde{\alpha}}$ for which the dispersion relation takes positive values, implies that the system goes unstable via a typical path first identified by Alan Turing in his seminal work. At variance, if the dispersion relation is negative the system cannot undergo a Turing instability: any tiny perturbation fades away and the system settles back to the homogeneous equilibrium.

To proceed further with a concrete example we selected the Brusselator reaction system [44, 45]. This is a nonlinear model defined by $f(u, v) = 1 - (b + 1)u +$

cu^2v and $g(u, v) = bu - cu^2v$, where b and c act as tunable parameters. In Fig. 5.3 we report the results for a choice of the model parameters giving rise to Turing patterns ($b = 4, c = 6, D_u = 0.02$ and $D_v = 0.17$) and the same hypergraph previously used in Figs. 5.1 and 5.2. The dispersion relation (panel a) is clearly positive for a selection of $\Lambda_H^{(\alpha)}$ (red points). The homogeneous solution becomes hence unstable and the ensuing patterns are displayed in panels b) and c).

5.4 Conclusions

Complex systems are composed of a large number of simple units, mutually interacting via nonlinear exchanges. Many-body interactions sit hence at the root of a large plethora of spontaneously emerging phenomena, as exhibited by complex systems. The former are often reduced to a vast collection of pairwise interactions, involving agents interacting in pairs. This enables one to model the inspected problem as a dynamical system flowing on a conventional binary network, a powerful approximation that allows for progresses to be made. In many cases of interest, this reductionist choice constitutes a rough first order approximation to the examined dynamics and more precise models are to be invoked which encompass for the high-order interactions being at play.

In this work, we presented a general framework which allows one to account for multi-body interacting systems coupled via a hypergraph. This materialises in a natural extension of the conventional network paradigm. More specifically, we considered the problem of the emergence of heterogeneous stable solutions in interconnected systems, under the assumption that, once isolated, all units converge to the same, and thus globally homogenous, solution. The high-order interaction is the driver of the resulting patchy states, which emerge as follow a symmetry breaking instability caused by the injection of a tiny non homogeneous perturbation. This can be though as a generalisation of the Turing instability on hypergraphs. In particular, we considered the interaction mediated by the number of interacting neighbouring units, namely the size of the hyperedge, and a diffusive-like process, again biased by the number of neighbours. In both cases we provided sufficient conditions for the emergence of spatial patterns.

Our findings have been corroborated by numerical simulations applied to two reference models. A Volterra model that describes the interaction among predators and preys in ecological niches, and the Brusselator model, a prototype model of nonlinear dynamics, that describes the interaction among reacting and diffusing chemicals.

The proposed framework goes beyond the examples hereby presented and, because of its generality, it could prove useful in tackling those problems were simultaneous many-body interactions within a complex environment are to be properly accounted for.

References

1. R. Albert, A. Barabási, Statistical mechanics of complex networks. Rev Mod. Phys. **74**(1), 47 (2002)
2. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: structure and dynamics. Phys. Rep. **424**(4–5), 175–308 (2006)
3. A.-L. Barabási et al., *Network Science* (Cambridge University Press, 2016)
4. V. Latora, V. Nicosia, G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, 2017)
5. M.E.J. Newman, *Networks: An Introduction* (Oxford University Press, 2010)
6. A.R. Benson, D.F. Gleich, J. Leskovec, Higher-order organization of complex networks. Science **353**(6295), 163–166 (2016)
7. R. Lambiotte, M. Rosvall, I. Scholtes, From networks to optimal higher-order models of complex systems. Nat. Phys. **15**, 313 (2019)
8. S. Fortunato, D. Hric, *Community Detection in Networks: A User Guide* (Phys, Rep, 2016)
9. L.-D. Lord, P. Expert, H.M. Fernandes, G. Petri, T.J. Van Hartevelt, F. Vaccarino, G. Deco, F. Turkheimer, M.L. Kringelbach, Insights into brain architectures from the homological scaffolds of functional connectivity networks. Front. Syst. Neurosci. **10**, 85 (2016)
10. G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer, F. Vaccarino, Homological scaffolds of brain functional networks. J. Royal Soc. Interface **11**(101), 20140873 (2014)
11. E. Estrada, G.J. Ross. Centralities in simplicial complexes. applications to protein interaction networks. J. Their. Biol. **438**, 46 (2018)
12. P.A. Abrams, Arguments in favor of higher order interactions. Am. Nat. **121**, 887 (1983)
13. J. Grilli, G. Barabás, M.J. Michalska-Smith, S. Allesina, Higher-order interactions stabilize dynamics in competitive network models. Nature **548**(7666), 210 (2017)
14. T. Carletti, F. Battiston, G. Cencetti, D. Fanelli, Random walks on hypergraphs. Phys. Rev. E **101**, 022308 (2020)
15. A. Patania, G. Petri, F. Vaccarino, The shape of collaborations. EPJ Data Sci. **6**, 18 (2017)
16. C. Berge, *Graphs and Hypergraphs* (North-Holland Pub. Co. American Elsevier Pub. Co, 1973)
17. E. Estrada, J.A. Rodríguez-Velázquez. Complex networks as hypergraphs. arXiv preprint physics/0505137 (2005)
18. G. Ghoshal, V. Zlatić, G. Caldarelli, M.E.J. Newman, Random hypergraphs and their applications. Phys. Rev. E **79**(6), 066118 (2009)
19. G. Ferraz de Arruda, G. Petri, Y. Moreno, Social contagion models on hypergraphs. Phys. Rev. Res. **2**, 023032 (2020)
20. G. Ferraz de Arruda, M. Tizzani, Y. Moreno, Phase transitions and stability of dynamical processes on hypergraphs. arXiv preprint arXiv:2005.10891 (2020)
21. T. Carletti, D. Fanelli, S. Nicoletti, Dynamical systems on hypergraphs. J. Phys.: Complex. **1**(3), 035006 (2020)
22. A. Krawiecki, Chaotic synchronization on complex hypergraphs. Chaos, Solitons and Fractals **65**, 44 (2014)
23. R. Mulas, C. Kuehn, J. Jost, Coupled dynamics on hypergraphs: Master stability of steady states and synchronization. Phys. Rev. E **101**, 062313 (2020)
24. L. Neuhäuser, A. Mellor, R. Lambiotte, Multibody interactions and nonlinear consensus dynamics on networked systems. Phys. Rev. E **101**(3), 032310 (2020)
25. O.T. Courtney, G. Bianconi, Generalized network structures: the configuration model and the canonical ensemble of simplicial complexes. Phys. Rev. E **93**(6), 062311 (2016)
26. K. Devriendt, P. Van Mieghem, The simplex geometry of graphs. J. Complex Netw. **7**(4), 469–490 (2019)
27. G. Petri, A. Barrat, Simplicial activity driven model. Phys. Rev. Lett. **121**(22), 228301 (2018)
28. Á. Bodó, G.Y. Katona, P.L. Simon, Sis epidemic propagation on hypergraphs. Bull. Math. Biol. **78**(4), 713 (2016)
29. I. Iacopini, G. Petri, A. Barrat, V. Latora, Simplicial models of social contagion. Nat. Commun. **10**(1), 2485 (2019)

30. L.V. Gambuzza, F. Di Patti, L. Gallo , S. Lepri, M. Romance, R. Criado, M. Frasca, V. Latora, S. Boccaletti, The master stability function for synchronization in simplicial complexes. arXiv preprint arXiv:2004.03913v1 (2020)
31. M. Lucas, G. Cencetti, F. Battiston, A multi-order laplacian framework for the stability of higher-order synchronization. arXiv preprint arXiv: 2003.09734v1 (2020)
32. A.P. Millán, J.J. Torres, G. Bianconi, Explosive higher-order kuramoto dynamics on simplicial complexes. Phys. Rev. Lett. **124**, 218301 (2020)
33. G. Cencetti, F. Battiston, T. Carletti, D. Fanelli, Generalized patterns from local and non local reactions. Chaos, Solitons and Fractals **134**, 109707 (2020)
34. A.J. McKane, T.J. Newman, Predator-prey cycles from resonant amplification of demographic stochasticity. Phys. Rev. Lett. **94**(21), 218102 (2005)
35. J. Jost, R. Mulas, Hypergraph laplace operators for chemical reaction networks. Adv. Math. **351**, 870 (2019)
36. T.S. Evans, Clique graphs and overlapping communities. J. Stat. Mech.: Theory Exp. **2010**(12), P12037 (2010)
37. D. Zhou, J. Huang, . Schölkopf, Learning with hypergraphs: clustering, classification, and embedding, in *Advances in Neural Information Processing Systems* (2007), pp. 1601–1608
38. L. Huang, Q. Chen, Y.-C. Lai, L.M. Pecora, Generic behavior of master-stability functions in coupled nonlinear dynamical systems. Phys. Rev. E **80**, 036204 (2009)
39. L.M. Pecora, T.L. Carroll, Master stability functions for synchronized coupled systems. Phys. Rev. Lett. **80**(10), 2109 (1998)
40. M. Asllani, D.M. Busiello, T. Carletti, D. Fanelli, G. Planchon, Turing patterns in multiplex networks. Phys. Rev. E **90**, 042814 (2014)
41. M. Asllani, T. Biancalani, D. Fanelli, A.J. McKane, The linear noise approximation for reaction-diffusion systems on networks. Eur. Phys. J. B **86**(11), 476 (2013)
42. M. Asllani, J.D. Challenger, F.S. Pavone, L. Sacconi, D. Fanelli, The theory of pattern formation on directed networks. Nat. Commun. **5**(1), 4517 (2014)
43. H. Nakao, A.S. Mikhailov, Turing patterns in network-organized activator-inhibitor systems. Nat. Phys. **6**, 544 (2010)
44. I. Prigogine, R. Lefever, Symmetry breaking instabilities in dissipative systems. J. Chem. Phys. **48**, 1695 (1968)
45. I. Prigogine, G. Nicolis, Symmetry breaking instabilities in dissipative systems. J. Chem. Phys. **46**, 3542 (1967)

Chapter 6
Non-pairwise Interaction in Oscillatory Ensembles: from Theory to Data Analysis

Arkady Pikovsky and Michael Rosenblum

Abstract In this chapter, we briefly review several cases when non-pairwise interaction naturally appears in oscillatory networks. First, we analyze globally coupled ensembles of phase oscillators. We demonstrate that nonlinear high-order mean-field coupling is equivalent to a hypernetwork with non-pairwise interactions of the population elements. Next, we consider small networks of limit-cycle oscillators. We show that pairwise interaction in the state variables description results in non-pairwise interaction on the level of phase dynamics, if one goes beyond the first order in the weak-coupling phase reduction. Finally, we discuss the implications for recovery of the network connectivity in terms of the phase dynamics from observations.

6.1 Introduction

Phase reduction is widely and efficiently exploited to investigate dynamics of interacting self-sustained oscillators [1, 2]. The main results of this approach can be summarized as follows. Consider a unit with frequency ω driven by a force with close frequency $\nu \approx \omega$. If the forcing is weak enough so that deviations of the state space trajectory from the limit cycle are small, then in the first approximation in the forcing amplitude, the dynamics of the phase is decoupled of that of the amplitudes, and obeys the equation

$$\dot{\phi} = \omega + Q(\phi, \psi), \tag{6.1}$$

where ϕ, ψ are the phases of the oscillator and the force, respectively, and Q is the coupling function. If the norm of Q is much smaller than ω, another approximation—averaging over the oscillation period—provides a description in terms of phase differences:

A. Pikovsky · M. Rosenblum (✉)
University of Potsdam, Potsdam, Germany
e-mail: mros@uni-potsdam.de

A. Pikovsky
e-mail: pikovsky@uni-potsdam.de

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_6

$$\dot{\phi} = \omega + \bar{Q}(\psi - \phi). \tag{6.2}$$

In a particular case when the forcing $F(t)$ of the oscillator is scalar and independent of the state of the oscillator, one can write the phase equation (6.1) in the Winfree form [3]:

$$\dot{\phi} = \omega + S(\phi)F(t), \tag{6.3}$$

where $S(\phi)$ is the phase sensitivity function, also known as the phase response curve.

A generalization to large oscillatory networks typically implies that interaction is pairwise and additive. Another standard assumption is that the form of the coupling function is the same for all pairs. The simplest case of equally strong sine-like coupling represents the famous Kuramoto-Sakaguchi model [4, 5]:

$$\dot{\phi}_k = \omega_k + \frac{\epsilon}{N}\sum_{j=1}^{N}\sin(\phi_j - \phi_k + \beta), \tag{6.4}$$

where ϵ is the interaction strength, N is the population size, k is the oscillator index, and β is a phase shift. In terms of the complex order parameter $Z = N^{-1}\sum_{k=1}^{N} e^{i\phi_k}$, often called the mean field, the model reads

$$\dot{\phi}_k = \omega_k + \epsilon \mathrm{Im}\left(Ze^{i(-\phi_k+\beta)}\right). \tag{6.5}$$

Numerous generalizations of the Kuramoto model [6–8] also rely on the assumption of pairwise interaction. In this Chapter, we go beyond this assumption and consider general mean-field coupled systems. We demonstrate that higher-order mean-field coupling naturally results in the interaction of oscillatory triplets, quadruplets, etc. In this sense the network of the interactions is hypernetwork. Furthermore, we show that such high-order interaction naturally appears in the phase description of pairwise- but non-weakly coupled oscillators. Finally, we discuss the implication for data analysis, namely for the reconstruction of the network connectivity from observations.

6.2 Theory of Higher-Order Mean-Field Phase Coupling

6.2.1 General Nonlinear Mean-Field Coupling

Here we follow [9, 10] and outline a general approach to nonlinear mean-field coupling in the context of phase dynamics. To simplify the presentation, we assume that there are M populations of oscillators, each population labeled by index $n = 1, \ldots, M$. All elements of a population have the same natural frequencies ω_n; all other properties of the oscillators are assumed to be identical across populations. The dynamics of the phase in the first-order approximation in coupling is, according to (6.3)

$$\dot{\phi}_n[a] = \omega_n + \epsilon F_n S(\phi_n[a]), \tag{6.6}$$

where $\phi_n[a]$ denotes the phase of the ath oscillator within the nth population, $S(\phi) = \sum_p s_p \exp[ip\phi]$ is the phase sensitivity function, and F_n is the force from all other oscillators acting on elements $\phi_n[a]$ [11]. The equations are valid if parameter ϵ is small. We assume a mean-field coupling, what means that the force F is a function of complex mean fields of the populations. These complex mean fields (a.k.a. Kuramoto-Daido order parameters) are defined as

$$Z_k^{(n)} = \langle e^{ik\phi_n[a]} \rangle\big|_a = N_n^{-1} \sum_{a=1}^{N_n} e^{ik\phi_n[a]} \tag{6.7}$$

where averaging is over all N_n units in population n. We assume that the force F_n can be represented as power series in $Z_k^{(n)}$:

$$F_n = \sum_{k,m} h_{k,m}^{(n)} Z_k^{(m)} + \sum_{k,m,l,s} g_{k,m;l,s}^{(n)} Z_k^{(m)} Z_l^{(s)} + \sum_{k,m;l,s;j,r} d_{k,m;l,s;j,r}^{(n)} Z_k^{(m)} Z_l^{(s)} Z_j^{(r)} + \ldots \tag{6.8}$$

where we explicitly write linear, quadratic, and cubic terms. Substitution of (6.8) and (6.7) in (6.6) already yields equations with coupling terms containing complex combinations of phase variables. However, not all of these terms are really relevant and lead to essential effects; to reveal important terms one has to perform an averaging over the basic oscillation period.

To perform averaging, it is convenient to introduce slow phases according to $\varphi_n = \phi_n - \omega_n t$. (For brevity, we omit the index for individual oscillators.) Complex order parameters $z_k^{(n)}$ for these variables are also slow. They are expressed in terms of original order parameters (6.7) as follows

$$z_k^{(n)} = \langle e^{ik\varphi_n} \rangle\big| = Z_k^{(n)} e^{-ik\omega_n t}. \tag{6.9}$$

The equations for the slow phases read

$$\begin{aligned} \dot{\varphi}_n = & \epsilon \sum_p s_p e^{ip\varphi_n} e^{ip\omega_n t} \\ & \Big[\sum_{k,m} h_{k,m}^{(n)} z_k^{(m)} e^{ik\omega_m t} + \sum_{k,m,l,s} g_{k,m;l,s}^{(n)} z_k^{(m)} z_l^{(s)} e^{i(k\omega_m + l\omega_s)t} + \\ & + \sum_{k,m;l,s;j,r} d_{k,m;l,s;j,r}^{(n)} z_k^{(m)} z_l^{(s)} z_j^{(r)} e^{i(k\omega_m + l\omega_s + j\omega_r)t} + \ldots \Big]. \end{aligned} \tag{6.10}$$

One can see that some terms on the r.h.s. contain explicit time dependence - these are fast terms. Other terms either do not depend explicitly on time, or contain a small frequency mismatch in the exponent—these are slow terms. It is convenient to work with exact resonances, therefore one shifts slightly natural frequencies (by terms of order ϵ), so that the combinations of these modified frequencies (like $k\omega_m +$

$p\omega_n$) vanish exactly. Due to these modifications, small terms $\sim \epsilon\delta\omega_n$ appear in the dynamics. Averaging means omitting fast terms, what leads to generic averaged phase coupled equations

$$\begin{aligned}\dot{\varphi}_n &= \epsilon\delta\omega_n + \epsilon \sum_{p;k,m} s_p h^{(n)}_{k,m} z^{(m)}_k \exp[ip\varphi_n]\Delta(k\omega_m + p\omega_n) + \\ &\epsilon \sum_{p;k,m,l,s} s_p g^{(n)}_{k,m;l,s} z^{(m)}_k z^{(s)}_l \exp[ip\varphi_n]\Delta(k\omega_m + l\omega_s + p\omega_n) + \\ &\epsilon \sum_{p;k,m;l,s;j,r} s_p d^{(n)}_{k,m;l,s;j,r} z^{(m)}_k z^{(s)}_l z^{(r)}_j \exp[ip\varphi_n]\Delta(k\omega_m + l\omega_s + j\omega_r + p\omega_n) + \ldots\end{aligned} \tag{6.11}$$

where $\Delta(\omega) = 1$ if $\omega = 0$ and $\Delta(\omega) = 0$ otherwise.

Below we discuss different cases leading to pairwise and multiple couplings of the network elements in the phase approximation.

6.2.2 One Population of Oscillators

The simplest case is when there exist only one population of oscillators. In the context of Eq. (6.11) this means that all frequencies ω_n are equal (or nearly equal, as mentioned above, small frequency differences can be straightforwardly included in the consideration by adding deviations from the central frequency to the r.h.s.). The famous Kuramoto setup [12] belongs to this class. Below, to describe it, we omit index n.

6.2.2.1 Linear Coupling

In the case of linear coupling only terms $\sim hZ$ are present in (6.8). The function $\Delta(k\omega + p\omega)$ picks up from the sum in (6.11) only terms with $p = -k$. In this case the phase dynamics corresponds to a famous Kuramoto-Daido model [13]

$$\dot{\varphi} = \epsilon \sum_k s_{-k} h_k z_k \exp[-ik\varphi]. \tag{6.12}$$

Substitution here of the definition of the order parameters (6.9) reveals terms $\dot{\varphi}[a] \sim \exp[ik(\varphi[b] - \varphi[a])]$, where a and b are indexes of different oscillators in the population. This means that the couplings are pairwise.

6.2.2.2 First Harmonic Phase Sensitivity Function

In some cases, like e.g., the Stuart-Landau oscillators, the phase sensitivity function contains only the first harmonics of the phase, i.e., in (6.11) index p takes values ± 1 only (the term $p = 0$ would lead to a phase-independent frequency shift). Let us consider two cases.

Coupling via the main Kuramoto order parameter

In this case nonlinear terms in (6.11) contain only z_1. It is easy to see that the terms with quadratic nonlinearity vanish, and the only non-trivial, phase-dependent terms with qubic nonlinearity are $\sim z_1^2 z_1^* e^{-i\varphi}$ and its complex conjugate. Together with linear term in (6.11) this yields a Kuramoto model with nonlinearly corrected coupling: acting mean field is modified $z_1 \to z_1 + \alpha|z_1|^2 z_1$, with a complex parameter α. This model has been introduced in [14] and studied in more details in [15], for experimental realization see [16, 17]. The cubic nonlinearity in the order parameter leads to terms in the phase coupling containing four phases (quadruplets): $\dot{\varphi}[a] \sim \exp(i(\varphi[b] + \varphi[c] - \varphi[d] - \varphi[a]))$. If both linear and cubic terms are present, the coupling is a combination of pairs and quadruplets.

Coupling contains many order parameters

In this case already the quadratic terms in (6.11) contribute, provided the relations $k + l = \pm 1$ hold. The simplest case is where $k = 2$, $l = 1$. This corresponds to the coupling term $\sim z_2 z_1^* e^{-i\varphi}$. This coupling is organized in triplets $\dot{\varphi}[a] \sim \exp(i(2\varphi[b] - \varphi[c] - \varphi[a]))$.

6.2.2.3 Second-Harmonic Phase Sensitivity Function

This case is described by the phase sensitivity function possessing the terms $p = \pm 2$ only. The simplest resonance here is provided by the quadratic terms in (6.11) satisfying condition $k + l + p = 0$. One can see that the main complex order parameter ($k = l = 1$) will contribute; the coupling term is $\sim z_1^2 e^{-i\varphi}$. In terms of the phases, the coupling is arranged in triplets $\dot{\varphi}[a] \sim \exp(i(\varphi[b] + \varphi[c] - 2\varphi[a]))$. The dynamics of populations of oscillators with this triplet coupling has been studied in details in [18].

6.2.3 Several Populations of Oscillators

Here we consider a situation where several populations of oscillators with different frequencies ω_n interact. The novel aspect compared to the above-studied case is interaction *across* populations. Because the resonance conditions have to be fulfilled, essential are relations between the basic frequencies.

6.2.3.1 Incommensurate Basic Frequencies

Let us start with the simplest case of two populations with incommensurate frequencies ω_1 and ω_2. This means that nontrivial resonances (i.e., resonances with

nonvanishing p, k, l, s) in Eq. (6.11) are impossible in linear and quadratic terms. The first possible nontrivial term appears in the third order; it corresponds to the condition $k\omega_{1,2} + l\omega_{2,1} - s\omega_{1,2} - p\omega_{2,1} = 0$. The simplest term of this form with $k = l = s = p = 1$ corresponds to the quadruplet coupling $\dot{\varphi}_{1,2}[a] \sim \exp(i(\varphi_{1,2}[b] + \varphi_{2,1}[c] - \varphi_{2,1}[d] - \varphi_{1,2}[a]))$. This is in fact a non-resonant coupling because it does not depend on the values of the frequencies. In the case of three incommensurate frequencies, the lowest-order term involving all three populations would be a six-plet $\dot{\varphi}_{1,2,3}[a] \sim \exp(i(\varphi_{1,2,3}[b] + \varphi_{2,3,1}[c] + \varphi_{3,1,2}[d] - \varphi_{3,1,2}[e] - \varphi_{2,3,1}[f] - \varphi_{1,2,3}[a]))$. Effects of such a nonresonant interaction in several globally coupled populations have been explored in [9].

6.2.3.2 Commensurate Basic Frequencies

The simplest case of resonance between two populations is $\omega_2 = 2\omega_1$. Inspecting Eq. (6.11) one can see that there is already a possibility for nontrivial interaction via linear in order parameters terms. This corresponds to the phase coupling terms $\dot{\varphi}_1[a] \sim \exp[i(\varphi_2[b] - 2\varphi_1[a])]$, $\dot{\varphi}_2[a] \sim \exp[i(2\varphi_1[b] - \varphi_2[a])]$. Such a coupling has been treated in [10, 19]. Additionally, there can exist quadratic in the order parameters resonant terms corresponding to triplet couplings $\dot{\varphi}_1[a] \sim \exp[i(\varphi_2[b] - \varphi_1[c] - \varphi_1[a])]$, $\dot{\varphi}_2[a] \sim \exp[i(\varphi_1[b] + \varphi_1[c] - \varphi_2[a])]$.

The next nontrivial case is of three populations with basic frequencies in resonance $\omega_3 = \omega_1 + \omega_2$. In this case there is no linear (pairwise) coupling between populations, and the first nontrivial terms are triplets $\dot{\varphi}_1[a] \sim \exp[i(\varphi_3[b] - \varphi_2[c] - \varphi_1[a])]$, $\dot{\varphi}_2[a] \sim \exp[i(\varphi_3[b] - \varphi_1[c] - \varphi_2[a])]$, $\dot{\varphi}_3[a] \sim \exp[i(\varphi_1[b] + \varphi_2[c] - \varphi_3[a])]$. The dynamics of three resonant populations has been studied in [20].

6.3 Multiple Effective Phase Coupling Appearing in Higher Orders of Phase Reduction

Section 6.2 discussed how couplings, nonlinear in the mean-field order parameters, result in hypernetworks with triplets, quadruplets, etc., of interacting phases. These nonlinear terms may be intrinsic for the problem, like in physical situations discussed in [14] and experimentally implemented in [16, 17]. Here we demonstrate that such terms also appear in high orders in the phase reduction from the original nonlinear equations having only pairwise linear interactions. Here we only sketch the derivation; for the complete analysis, we refer the reader to Ref. [21].

We consider interacting nonlinear oscillators with variables $\mathbf{y}_k$ possessing stable limit cycles $\mathbf{y}_k^0(t) = \mathbf{y}_k^0(t + T_k)$. For each of these cycles a phase φ_k can be introduced, satisfying $\dot{\varphi}_k = \omega_k = 2\pi/T_k$. The phases are functions of the variables $\varphi_k = \Phi_k(\mathbf{y}_k)$, but only on the limit cycle the variables $\mathbf{y}$ are unique functions of the phases. In the basin of attraction of the limit cycles one has to account for amplitude deviations $\delta\mathbf{y}$. We write the system of coupled oscillators as

$$\dot{\mathbf{y}}_k = \mathbf{F}_k(\mathbf{y}_k) + \epsilon \sum_{m \neq k} \mathbf{G}_{mk}(\mathbf{y}_m, \mathbf{y}_k), \tag{6.13}$$

so that only pairwise couplings are present. For the phases, the equations read

$$\dot{\varphi}_k = \frac{d}{dt}\Phi_k(\mathbf{y}_k) = \omega_k + \epsilon \frac{\partial \Phi_k}{\partial \mathbf{y}_k} \sum_{m \neq k} \mathbf{G}_{mk}(\mathbf{y}_m, \mathbf{y}_k). \tag{6.14}$$

This equation is, of course, not a closed equation for the phases, and one solves it iteratively in powers of ϵ. In parallel, one also represents the deviations from the limit cycle $\delta\mathbf{y}_k$ in powers of ϵ.

In the first order in the small parameter ϵ, one can neglect the deviations $\delta\mathbf{y}_k$, then the coupling terms

$$\left.\frac{\partial \Phi_k}{\partial \mathbf{y}_k}\right|_{\mathbf{y}_k^0} \mathbf{G}_{mk}(\mathbf{y}_m^0, \mathbf{y}_k^0)$$

depend on two phases φ_k, φ_m and one obtains pairwise interactions in the phase dynamics in form of Eqs. (6.1). In this order $\sim \epsilon$ also $\delta\mathbf{y}_k$ can be represented as a sum of terms depending on two phases only.

In the second order in the small parameter ϵ, when one substitutes the expressions of the first approximation $\delta\mathbf{y}_k = \epsilon \sum Q_{km}(\varphi_k, \varphi_m)$ in (6.14), one obtains terms containing three phases $\varphi_k, \varphi_m, \varphi_l$, i.e., an effective triplet interaction. In higher-order approximations in ε also the quadruplet, etc., interactions appear in the phase dynamics equations. One can complete this analysis in an exceptional case of the Stuart-Landau oscillators, where the phases and their derivatives are known explicitly. Reference [21] derives phase equations for three Stuart-Landau units organized in a chain, $1 \leftrightarrow 2 \leftrightarrow 3$. As expected, already the second-order phase approximation provides the terms depending on the phases of all three oscillators. Thus, on the level of the phase dynamics, unit 1 interacts with unit 3, though there is no direct link between them, and the simple motif $1 \leftrightarrow 2 \leftrightarrow 3$ becomes a hypernetwork.

For general oscillators, the high-order phase reduction can be performed only numerically. The interested reader can find the corresponding techniques for computation of phases and instantaneous frequencies in Ref. [21]. For three van der Pol oscillators, also coupled in a $1 \leftrightarrow 2 \leftrightarrow 3$ motif, the analysis yields phase reduction equations in different orders, similar to the theoretical findings. Like in the Stuart-Landau systems, the second-order reduction already represents a hypernetwork, with the coupling terms depending on three phases. Contrary to the Stuart-Landau case, the phase dynamic equations for the van der Pol model also contain the terms with the phases' sums.

6.4 Non-pairwise Interactions in the Network Reconstruction Problem

Reconstruction of the coupled oscillatory models from data is an efficient tool for experimental studies of interacting rhythmical objects. A particular example is the recovery of brain connectivity from multichannel measurements of brain activity [22–24]. Another example is the analysis of mutual influences of the cardiac, respiratory, and brain rhythms [25–27]. In this approach, one assumes that the registered time series represent outputs of interacting self-sustained oscillating units. These series allow for estimating of phases and instantaneous frequencies of all oscillators. Typically, one computes these quantities exploiting the Hilbert transform. Finally, one uses these estimates to construct the observed network's phase dynamics model and exploits this model to quantify the strength and directionality of all connections. For technical details of phase estimation and equation reconstruction, we refer to Refs. [27–29]. An essential issue is that the described approach yields the effective phase connectivity that generally differs from the structural connectivity. The latter is determined by physical connections between the oscillators, while the former represents the approximately equivalent phase model's connections. Below we illustrate that the difference is precisely due to the appearance of the non-pairwise interaction on the level of phase reduction.

This approach's main idea is that the dynamics of N interacting oscillators are represented by a torus in the N-dimensional space if the coupling is not too strong. Since the coupling function Q (cf. Eq. (6.1)) is 2π-periodic with respect to its arguments, it can be written as an N-dimensional Fourier series, and the coefficients of this series can be determined by fit.

Consider a simple motif of three pairwise coupled oscillators, described by Eqs. (6.13). Our goal is to determine the network structure, i.e., to quantify all connecting links' strength. While doing this, we shall distinguish between direct or indirect links. It is convenient to quantify first all incoming connections to one of the units and then repeat it for other network elements. Without loss of generality, we can consider the first oscillator. Its phase equation reads:

$$\dot{\phi}_1 = \omega_1 + Q(\phi_1, \phi_2, \phi_3), \tag{6.15}$$

where Q does not contain the constant term. The simplest and straightforward approach, used in many studies, is to perform a pairwise analysis of the network. It means that to quantify the link $1 \leftarrow 2$, we neglect the third unit entirely and reconstruct the equation in the form $\dot{\phi}_1 = \omega_1 + Q(\phi_1, \phi_2)$ with a two-dimensional coupling function $Q(\phi_1, \phi_2)$. Then we compute the norm of the coupling function $\|Q\|$ and use it as a measure of the action exerted by the second oscillator on the first one. To emphasize, that this quantity comes from a pairwise analysis, we denote it as $\mathcal{P}_{1\leftarrow 2}$.

However, this estimation may yield spurious effective phase connections. Indeed, suppose three oscillators are organized in a chain, $1 \leftrightarrow 2 \rightarrow 3$. Because φ_1 is corre-

lated with φ_2 and φ_2 acts on φ_3, pairwise analysis for φ_1, φ_3 will yield spurious non-zero coupling for the $1 \to 3$ link. Only the full phase dynamics given by Eq. (6.15) would reveal the absence of a direct connection between the nodes 1 and 3 and the presence of indirect coupling $1 \to 2 \to 3$. Indeed, as argued in the previous section, the second-order phase approximation yields the terms depending on all three phases. It means that the reconstructed from data coupling function Q in Eq. (6.15) generally contains Fourier components depending on all three phases, and these components describe the indirect connection $1 \to 2 \to 3$. The direct (pairwise) interaction can be quantified by its total strength

$$\mathcal{T}_{1\leftarrow 2} = \left[\sum_{k,l\neq 0} \left| F_{k,l,0} \right|^2 \right]^{1/2}, \tag{6.16}$$

where $F_{k,l,m}$ are Fourier coefficients of $Q(\phi_1, \phi_2, \phi_3)$ and the summation is performed over the terms which do not depend on the third phase φ_3. Correspondingly, the joint action of the second and third oscillators on the first one, i.e., the triplet interaction, that appears in the higher-order approximation can be quantified by the triplet norm

$$\mathcal{T}_{j\leftarrow 2,3} = \left[\sum_{k,l,m\neq 0} \left| F_{k,l,m} \right|^2 \right]^{1/2}, \tag{6.17}$$

where summation is performed over terms depending on three phases. Numerical experiments in Refs. [30] demonstrate that coefficient $\mathcal{P}_{1\leftarrow 2}$ that describes direct, structurally existing, connections scales linearly with coupling strength ϵ, see Eq. (6.15). On the contrary, the scaling of $\mathcal{T}_{j\leftarrow 2,3}$ reveals high-order dependence on ϵ, in full agreement with the theory outlined in the previous section.

Extension of the connectivity analysis through partial norms to the case of $N > 3$ oscillators seems to be straightforward; in Sect. 6.5 we provide such an example. However, reconstruction of the coupling function for more than three variables requires very long data sets. As shown in [29], the triplet analysis performed for moderate lengths of time series, can eliminate this difficulty. Suppose the goal is to quantify the link $j \leftarrow k$. The solution is to consider all possible $N-2$ triplets of oscillators j, k, m, where $m = 1, 2, \ldots, N$, $m \neq j, k$. For each triplet one reconstructs the coupling function $Q_j(\varphi_j, \varphi_k, \varphi_m)$, ignoring all other phases, and computes the partial norm $\mathcal{T}^{(m)}_{j\leftarrow k}$ like in Eq. (6.17). The minimal value of these estimates yields the final triplet-based measure of the binary (pairwise) effective phase connectivity $\mathcal{T}_{j\leftarrow k} = \min_m \mathcal{T}^{(m)}_{j\leftarrow k}$. We illustrate this approach to reconstruction of phase dynamics hypernetworks by an example of $N = 5$ and $N = 9$ randomly coupled van der Pol oscillators [29]:

$$\ddot{x}_k - \mu(1 - x_k^2)\dot{x}_k + \omega_k^2 x_k = \varepsilon \sum_l \sigma_{kl}(x_l \cos\Theta_{kl} + \dot{x}_l \sin\Theta_{kl}). \tag{6.18}$$

The data was generated in many runs, and then the coupling structure was reconstructed by estimating the strength of all links as already discussed. For each run, the random frequencies ω_k were taken from the uniform distribution $0.5 < \omega < 1.5$. The asymmetric connection matrix σ_{kl} composed of zeros and ones was also randomly generated, with four incoming connections. Another coupling parameter Θ was taken from a uniform distribution $0 \le \Theta < 2\pi$. The results, presented in Figs. 4–7 of Ref. [29] demonstrate that the phase dynamics reconstruction using hypernetworks provides enhanced separation between truly existing and absent structural connections. For application of this approach to a hypernetwork with triplet interactions of 12 phase oscillators see Ref. [31].

6.5 Example of Phase Dynamics Reconstruction in a Network with Triplet Couplings

Here we consider a simple example of a hypernetwork of oscillators with triplet coupling. It consists of four FitzHugh-Nagumo units [32], and the force acting on a unit is a product of observables of two other units:

$$\begin{aligned}
\dot{u}_1 &= u_1 - \frac{u_1^3}{3} - v_1 + 0.9 + 0.4\epsilon u_2 u_3,\\
\dot{v}_1 &= 0.35(u_1 - 0.8v_1 + 0.7),\\
\dot{u}_2 &= u_2 - \frac{u_2^3}{3} - v_2 + 0.8 - 0.6\epsilon u_1 u_3 + 0.7\epsilon u_3 u_4,\\
\dot{v}_2 &= 0.5(u_1 - 0.8v_1 + 0.7),\\
\dot{u}_3 &= u_3 - \frac{u_3^3}{3} - v_3 + 1.1 + 0.8\epsilon u_1 u_2 - 0.3\epsilon u_2 u_4,\\
\dot{v}_3 &= 0.42(u_1 - 0.8v_1 + 0.7),\\
\dot{u}_4 &= u_1 - \frac{u_1^3}{3} - v_1 + 1 + 0.5\epsilon u_2 u_3,\\
\dot{v}_4 &= 0.45(u_1 - 0.8v_1 + 0.7).
\end{aligned} \tag{6.19}$$

This configuration is schematically presented in Fig. 6.1. It should be noted, that although the coupling looks like a pure triplet one, because the interaction terms are products of the corresponding variables (cf. [33]), in fact the pairwise coupling is also present, because average values of observables u_i are generally nonzero. Thus, if one separates these average values by writing $u_i = \overline{u}_i + \tilde{u}_i$, then, e.g., the forcing for the first unit will be written as $\overline{u}_2\overline{u}_3 + \overline{u}_2\tilde{u}_3 + \overline{u}_3\tilde{u}_2 + \tilde{u}_2\tilde{u}_3$, i.e. with terms which can be effectively considered as a pairwise coupling.

For the model (6.19) we performed numerically the phase reduction analysis, as described above in Sect. 6.3 and in Ref. [21], for a range of values of the coupling parameter $0.001 \le \epsilon \le 0.05$. We have looked for the phase dynamics equations in the form

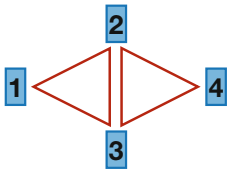


Fig. 6.1 A schematic representation of network (6.19), with two triplet couplings

$$\dot{\varphi}_s = \omega_s + \sum_{jklm} \left[A^{(s)}_{jklm} \cos(j\varphi_1 + k\varphi_2 + l\varphi_3 + m\varphi_4) + B^{(s)}_{jklm} \sin(j\varphi_1 + k\varphi_2 + l\varphi_3 + m\varphi_4) \right] \tag{6.20}$$

In the reconstruction we took into account all terms with $|j, k, l, m| \leq 2$. To represent the strength of each coupling mode, we calculated $C^{(s)}_{jklm} = \left[(A^{(s)}_{jklm})^2 + (B^{(s)}_{jklm})^2 \right]^{1/2}$. Altogether, for each oscillator this gives 310 coupling terms.

After finding the coupling modes, we sorted them according to dependence on ϵ. We performed a fit $C \sim \epsilon^p$, and interpreted terms with $|p - 1| < 0.1$ as linear in ϵ, and those with $|p - 2| < 0.1$ as quadratic in ϵ. We present the dependencies on ϵ of all these terms for oscillators 1 and 2 in Fig. 6.2.

It is instructive to see, which effective phase coupling terms appear in the first and the second orders in the coupling strength ϵ. Here are all 34 first-order terms for the oscillator 1, sorted in descending order in their strength:

$$\begin{gathered}(1,1,1,0),\ (1,-1,1,0),\ (1,1,-1,0),\ (1,-1,-1,0),\ (1,1,0,0),\ (1,-1,0,0),\\ (1,-1,-2,0),\ (1,1,2,0),\ (1,1,-2,0),\ (1,-1,2,0),\ (1,0,-1,0),\ (1,0,1,0),\\ (1,-2,1,0),\ (1,2,-1,0),\ (1,2,1,0),\ (1,-2,-1,0),\ (2,1,1,0),\ (2,-1,-1,0),\\ (2,1,-1,0),\ (0,1,1,0),\ (0,1,-1,0),\ (2,-1,1,0),\ (1,0,0,0),\ (1,2,0,0),\\ (1,0,-2,0),\ (1,-2,0,0),\ (1,0,2,0),\ (2,1,0,0),\ (1,2,-2,0),\ (1,-2,-2,0),\\ (1,2,2,0),\ (2,1,-2,0),\ (0,1,0,0),\ (2,-1,-2,0).\end{gathered} \tag{6.21}$$

One can see that the largest terms describe triplet coupling $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 1$ and the pairwise couplings. There is no term that includes the phase of oscillator 4. Such terms appear in the second order in ϵ. Altogether, there are 63 terms:

$$\begin{gathered}(1,2,0,2),\ (1,2,-2,2),\ (1,-2,0,-2),\ (1,-2,2,-2),\ (1,0,-2,-2),\ (1,2,-2,-2),\\ (1,0,2,2),\ (1,-2,2,2),\ (1,2,-1,2),\ (1,-2,1,-2),\ (1,2,1,2),\ (1,-2,-1,-2),\\ (1,-2,1,-1),\ (1,2,-1,1),\ (1,-2,-1,-1),\ (2,2,2,0),\ (1,2,1,1),\ (1,2,2,1),\\ (1,-2,-2,-1),\ (1,0,0,1),\ (1,0,0,-1),\ (1,0,1,1),\ (1,-2,-2,-2),\ (1,2,1,-1),\\ (1,2,2,2),\ (1,0,-1,-1),\ (1,-2,-1,1),\ (1,-1,2,2),\ (1,0,-1,1),\ (1,2,0,1),\\ (1,-2,0,-1),\ (1,2,-2,-1),\ (1,0,1,-1),\ (1,-2,2,1),\ (1,2,2,-1),\ (1,0,-2,1),\\ (1,-1,-2,-2),\ (1,0,2,-1),\ (1,0,-2,-1),\ (1,1,2,2),\ (1,0,2,1),\ (2,-2,0,-2),\\ (1,-2,-2,1),\ (1,0,0,-2),\ (1,0,0,2),\ (2,2,-2,-2),\ (0,2,0,2),\ (1,-2,0,1),\\ (0,0,2,2),\ (0,2,-2,-2),\ (1,-2,-2,2),\ (1,-1,-2,1),\ (2,2,1,1),\ (1,0,-2,2),\\ (1,1,-2,1),\ (1,2,-1,-1),\ (2,-2,1,-2),\ (1,-2,0,2),\ (1,0,1,2),\ (1,0,1,-2),\\ (1,0,-1,-2),\ (1,-2,-1,2),\ (0,0,1,-1).\end{gathered} \tag{6.22}$$

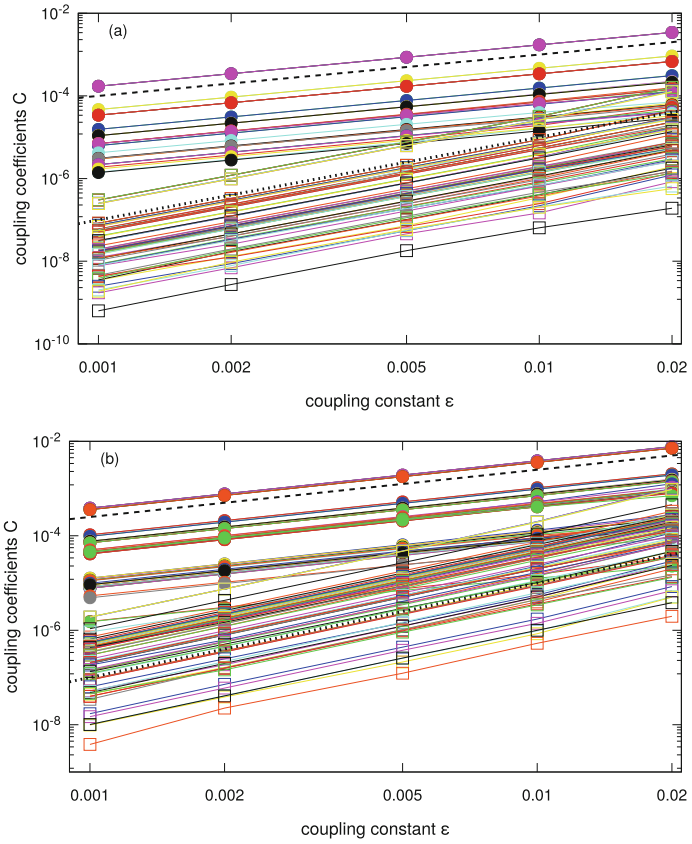


Fig. 6.2 Scaling of coupling terms with ϵ for oscillators 1 (panel (a)) and 2 (panel (b)). The dashed and the dotted lines show scalings $\sim \epsilon$ and $\sim \epsilon^2$, respectively. The corresponding coupling coefficients are depicted with filled circles and open squares

Thirty-three of them include all four phases and therefore describe effective quadruplet coupling.

The second oscillator participates in two triplet couplings, and therefore the number of phase interaction terms in the first and in the second order is larger, 65 and 69, respectively. The first-order terms are

$$(0,1,1,1),\ (0,1,-1,1),\ (0,1,-1,-1),\ (0,1,1,-1),\ (1,-1,-1,0),\ (1,1,-1,0),$$
$$(1,-1,1,0),\ (1,1,1,0),\ (0,1,0,-1),\ (0,1,0,1),\ (1,-1,0,0),\ (1,1,0,0),$$
$$(0,1,2,1),\ (0,1,-2,-1),\ (0,1,2,-1),\ (0,1,-2,1),\ (1,-1,-2,0),\ (1,1,2,0),$$
$$(1,-1,2,0),\ (1,1,-2,0),\ (0,2,-1,1),\ (0,0,1,-1),\ (0,1,1,0),\ (1,0,-1,0),$$
$$(0,2,1,-1),\ (0,0,1,1),\ (1,0,1,0),\ (0,1,-1,0),\ (0,2,-1,-1),\ (1,-2,-1,0),$$
$$(1,-2,1,0),\ (0,1,1,2),\ (1,2,-1,0),\ (1,2,1,0),\ (0,1,-1,2),\ (0,1,1,-2),$$
$$(0,1,-1,-2),\ (0,2,1,1),\ (1,0,0,0),\ (0,2,0,-1),\ (0,1,0,2),\ (0,1,0,0),$$
$$(1,2,0,0),\ (0,1,2,2),\ (0,1,-2,2),\ (1,0,2,0),\ (1,-2,0,0),\ (0,1,2,-2),$$
$$(1,-2,2,0),\ (2,-1,-1,0),\ (0,2,2,-1),\ (2,1,-1,0),\ (0,1,0,-2),\ (1,0,-2,0),$$
$$(1,2,2,0),\ (2,-1,1,0),\ (0,0,2,-1),\ (2,1,1,0),\ (1,2,-2,0),\ (0,1,2,0),$$
$$(0,1,-2,0),\ (0,2,-1,0),\ (0,2,-1,-2),\ (2,-1,-2,0),\ (2,1,-2,0) \tag{6.23}$$

and the second-order terms are

$$(1,-2,2,-2),\ (1,2,0,2),\ (1,2,-2,2),\ (1,-2,0,-2),\ (1,-2,-1,1),\ (1,-2,-1,-1),$$
$$(1,-2,1,-2),\ (1,0,2,2),\ (1,2,-2,-2),\ (1,-2,2,2),\ (1,2,-1,2),\ (0,2,-2,0),$$
$$(1,0,-2,-2),\ (1,-2,0,1),\ (1,-2,-1,-2),\ (1,-2,2,1),\ (1,2,1,1),\ (1,-2,-2,-1),$$
$$(1,2,0,1),\ (1,-2,-2,1),\ (1,0,0,1),\ (1,-2,2,-1),\ (1,0,2,1),\ (1,-2,0,-1),$$
$$(1,2,1,2),\ (1,2,0,-1),\ (1,2,-2,1),\ (1,0,-2,1),\ (1,2,-2,-1),\ (1,0,1,1),$$
$$(1,2,2,1),\ (2,2,0,0),\ (2,-2,-2,0),\ (1,0,0,-1),\ (1,-2,1,-1),\ (1,0,2,-1),$$
$$(1,2,2,2),\ (1,0,-2,-1),\ (1,0,1,-1),\ (1,0,-1,1),\ (1,2,2,-1),\ (1,1,0,1),$$
$$(1,2,1,-1),\ (1,1,2,1),\ (1,2,-1,-1),\ (2,2,-2,0),\ (1,-1,2,1),\ (1,-2,1,1),$$
$$(1,0,-1,-1),\ (1,1,-2,-1),\ (1,-1,2,-1),\ (1,-2,0,2),\ (1,-2,-1,2),\ (1,1,2,-1),$$
$$(1,-1,1,1),\ (1,2,1,-2),\ (1,0,0,2),\ (1,0,1,2),\ (1,-1,-2,1),\ (2,2,-2,2),$$
$$(1,1,0,-1),\ (2,-2,2,-2),\ (1,1,1,-1),\ (2,2,1,-1),\ (2,0,1,-1),\ (2,0,-2,-2),$$
$$(2,2,1,2),\ (2,2,-2,-2),\ (2,2,2,2) \tag{6.24}$$

6.6 Conclusions

This mini-review demonstrates that hypernetworks naturally appear in the phase dynamics description of ensembles of coupled oscillators. There are two main scenarios. First, the hypernetworks arise due to nonlinear mean-field coupling. Second, simple pairwise connections on the level of state variables result in hypernetworks of phase oscillators in the process of high-order phase reduction. This fact is significant for a practical problem, namely, to determine the network connectivity from measurements. Fitting a hypernetwork of phase oscillators to experimental data essentially improves the recovery of the structural connectivity.

Acknowledgements The work was supported by the Russian Science Foundation (grant Nr. 17-12-01534).

References

1. H. Nakao, Phase reduction approach to synchronisation of nonlinear oscillators. Contemp. Phys. **57**, 188–214 (2016)
2. B. Pietras, A. Daffertshofer, Network dynamics of coupled oscillators and phase reduction techniques. Phys. Rep. **819**, 1–105 (2019)
3. A.T. Winfree, Biological rhythms and the behavior of populations of coupled oscillators. J. Theor. Biol. **16**, 15 (1967)
4. Y. Kuramoto, *Chemical Oscillations*. Waves and Turbulence (Springer, Berlin, 1984)
5. H. Sakaguchi, Y. Kuramoto, A soluble active rotator model showing phase transition via mutual entrainment. Prog. Theor. Phys. **76**(3), 576–581 (1986)
6. J. A. Acebrón, L. L. Bonilla, C. J. Pérez Vicente, F. Ritort, R. Spigler, The Kuramoto model: a simple paradigm for synchronization phenomena. Rev. Mod. Phys. **77**(1), 137–175 (2005)
7. A. Pikovsky, M. Rosenblum, Dynamics of globally coupled oscillators: Progress and perspectives. Chaos **25**, 097616 (2015)
8. S. Chandra, M. Girvan, E. Ott, Continuous versus discontinuous transitions in the d-dimensional generalized kuramoto model: odd d is different. Phys. Rev. X **9**, 011002 (2019)
9. M. Komarov, A. Pikovsky, Effects of nonresonant interaction in ensembles of phase oscillators. Phys. Rev. E **84**(1), 016210 (2011)
10. M. Komarov, A. Pikovsky, Intercommunity resonances in multifrequency ensembles of coupled oscillators. Phys. Rev. E **92**, 012906 (2015)
11. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization. A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, 2001)
12. Y. Kuramoto, Self-entrainment of a population of coupled nonlinear oscillators, in *International Symposium on Mathematical Problems in Theoretical Physics*, ed. by H. Araki, p. 420, New York, 1975. Springer Lecture Notes Phys, v. 39
13. H. Daido, Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function. Physica D **91**, 24–66 (1996)
14. M. Rosenblum, A. Pikovsky, Self-organized quasiperiodicity in oscillator ensembles with global nonlinear coupling. Phys. Rev. Lett. **98**, 064101 (2007)
15. A. Pikovsky, M. Rosenblum, Self-organized partially synchronous dynamics in populations of nonlinearly coupled oscillators. Physica D **238**(1), 27–37 (2009)
16. A.A. Temirbayev, ZZh. Zhanabaev, S.B. Tarasov, V.I. Ponomarenko, M. Rosenblum, Experiments on oscillator ensembles with global nonlinear coupling. Phys. Rev. E **85**, 015204 (2012)
17. A.A. Temirbayev, Y.D. Nalibayev, ZZh. Zhanabaev, V.I. Ponomarenko, M. Rosenblum, Autonomous and forced dynamics of oscillator ensembles with global nonlinear coupling: an experimental study. Phys. Rev. E **87**, 062917 (2013)
18. M. Komarov, A. Pikovsky, Finite-size-induced transitions to synchrony in oscillator ensembles with nonlinear global coupling. Phys. Rev. E **92**, 020901 (2015)
19. S. Lück, A. Pikovsky, Dynamics of multi-frequency oscillator ensembles with resonant coupling. Phys. Lett. A **375**(28–29), 2714–2719 (2011)
20. M. Komarov, A. Pikovsky, Dynamics of multifrequency oscillator communities. Phys. Rev. Lett. **110**, 134101 (2013)
21. E. Gengel, E. Teichmann, M. Rosenblum, A. Pikovsky, High-order phase reduction for coupled oscillators. J. Phys.: Complexity **2**, 015005 (2021)
22. K. Lehnertz, Assessing directed interactions from neurophysiological signals—an overview. Physiol. Measurement **32**, 1715–1724 (2011)

23. T.B. Leergaard, C.C. Hilgetag, O. Sporns, Mapping the connectome: multi-level analysis of brain connectivity. Front. Neuroinformatics, **6**(14) (2012)
24. O. Sporns, Making sense of brain network data. Nature Methods **10**(6), 491–493 (2013)
25. R. Mrowka, L. Cimponeriu, A. Patzak, M.G. Rosenblum, Directionality of coupling of physiological subsystems—age related changes of cardiorespiratory interaction during different sleep stages in babies. Ame. J. Physiol. Regul. Comp. Integr. Physiol. **145**, R1395–R1401 (2003)
26. B. Musizza, A. Stefanovska, P.V.E. McClintock, M. Paluš, J. Petrovčič, S. Ribarič, F.F. Bajrović, Interactions between cardiac, respiratory, and EEG-δ oscillations in rats during anæsthesia. J. Physiol. (Londn.) **580**(1), 315–326 (2007)
27. B. Kralemann, M. Frühwirth, A. Pikovsky, M. Rosenblum, T. Kenner, J. Schaefer, M. Moser, In vivo cardiac phase response curve elucidates human respiratory heart rate variability. Nature Commun. **4**, 2418 (2013)
28. B. Kralemann, L. Cimponeriu, M. Rosenblum, A. Pikovsky, R. Mrowka, Phase dynamics of coupled oscillators reconstructed from data. Phys. Rev. E **77**, 066205 (2008)
29. B. Kralemann, A. Pikovsky, M. Rosenblum, Reconstructing effective phase connectivity of oscillator networks from observations. New J. Phys.**16**, 085013 (2014)
30. B. Kralemann, A. Pikovsky, M. Rosenblum, Reconstructing phase dynamics of oscillator networks. Chaos **21**, 025104 (2011)
31. A. Pikovsky, Reconstruction of a random phase dynamics network from observations. Phys. Lett. A **382**(4), 147–152 (2018)
32. E.M. Izhikevich, R. FitzHugh, FitzHugh-Nagumo model. Scholarpedia, **1**(9), 1349 (2006). revision #123664
33. J. Petereit, A. Pikovsky, Chaos synchronization by nonlinear coupling. Comm. Nonl. Sci. Num. Simul. **44**, 344–351 (2017)

Chapter 7
From Symmetric Networks to Heteroclinic Dynamics and Chaos in Coupled Phase Oscillators with Higher-Order Interactions

Peter Ashwin, Christian Bick, and Ana Rodrigues

Abstract We highlight some results from normal form theory for symmetric bifurcations that give a rational way to organize higher-order interactions between phase oscillators in networks with fully symmetric coupling. For systems near Hopf bifurcation the lowest order (pairwise) interactions correspond to the system of Kuramoto and Sakaguchi. At next asymptotic order one must generically include higher-order interactions of up to four oscillators. We discuss some dynamical consequences of these interactions in terms of heteroclinic attractors, chaos, and chimeras for related systems.

7.1 Introduction

Network dynamical systems consists of individual dynamical units (nodes) that evolve under mutual interaction. Examples include coupled neural oscillators, flashing fireflies, and power grid networks. Such dynamical systems often give rise to intriguing collective behavior, such as synchronization where nodes eventually behave in unison [1, 2]. Mathematical descriptions of such network dynamical systems often make the assumption that nodes interact in a pairwise fashion: The network interactions are determined by the joint state of pairs of nodes, that is, there is an underlying (directed) graph and such that if (j, k) is an edge from node j to node k

P. Ashwin · C. Bick (✉) · A. Rodrigues
Centre for Systems, Dynamics and Control and Department of Mathematics, University of Exeter, Exeter EX4 4QF, UK
e-mail: c.bick@vu.nl

P. Ashwin
e-mail: P.Ashwin@exeter.ac.uk

A. Rodrigues
e-mail: A.Rodrigues@exeter.ac.uk

C. Bick
Department of Mathematics, Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV, Amsterdam, The Netherlands

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_7

then the influence of j onto k does not depend on any other nodes. As an example, the interactions in the classical Kuramoto model [3, 4] where the phase $\theta_k \in \mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$ of oscillator $k \in \{1, \ldots, N\}$ evolves according to

$$\dot{\theta}_k := \frac{\mathrm{d}}{\mathrm{d}t}\theta_k = \omega_k + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_k), \tag{7.1}$$

with intrinsic frequency $\omega_k \in \mathbb{R}$ and subject to coupling strength K. In the Kuramoto model, the interactions are all-to-all (i.e., the underlying graph is the complete graph) but pairwise, that is, the influence of node j onto node k is determined by $\sin(\theta_j - \theta_k)$ which does not depend on the state of other nodes. This property allows to generalize the Kuramoto model to arbitrary graphs [5]. Sakaguchi generalized the Kuramoto model by incorporating a phase-shift parameter $\alpha \in \mathbb{T}$ in the interaction function [6].

Recently, the dynamics of networks with nonpairwise interactions—interactions containing nonlinear terms of more than two nodes—have attracted significant attention; cf. [7, 8] for recent reviews as well the other chapters in this book. Such network dynamical systems have been studied in their own right as generalizations of dynamics on graphs to "higher-order" combinatorial objects such as simplicial complexes or hypergraphs. Intuitively speaking, a simplicial complex or hypergraph is an object on a number of nodes that may not only contain edges between pairs of nodes but also simplices that are spanned by three or more nodes. For a network dynamical system on a simplex or hypergraph, the interactions along such a simplex corresponds to a nonlinear term in the state variables of the nodes that span it. For example, Skardal and Arenas [9, 10] considered a generalization of the Kuramoto model

$$\begin{aligned}\dot{\theta}_k = \omega_k + \frac{K_2}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_k) + \frac{K_3}{N^2}\sum_{j,l=1}^{N}\sin(2\theta_l - \theta_j - \theta_k)\\ + \frac{K_4}{N^3}\sum_{j,l,m=1}^{N}\sin(\theta_j + \theta_l - \theta_m - \theta_k),\end{aligned} \tag{7.2}$$

where K_2 and K_3, K_4 are the coupling strength of pairwise and nonpairwise interactions, respectively. Here terms such as $\sin(2\theta_l - \theta_j - \theta_k)$ describe the nonadditive joint influence of nodes l, j onto node k. These nonadditive terms can change the properties of the collective dynamics as one may expect [11]: For (7.2) they lead to a change in the criticality of the synchronization transition [10].

Nonadditively coupled phase oscillator networks—such as (7.2)—also arise as phase approximations of weakly coupled nonlinear oscillator networks. In other words, they can be *derived* from more general oscillator networks through phase reduction [12, 13]. In this case, the phase dynamics (7.2) reflect the effective dynamics of the corresponding nonlinear oscillator network [14] and nonadditive terms can reflect the effect of the nonlinearities as the dynamics deviate from the original limit cycle. For example, a globally coupled network of oscillatory nodes close to a Hopf bifurcation has the effective phase dynamics

$$\begin{aligned}\dot{\theta}_k = \omega &+ \sum_{j=1}^{N} g_2(\theta_j - \theta_k) + \sum_{j,l=1}^{N} g_3(\theta_j + \theta_l - 2\theta_k) \\ &+ \sum_{j,l=1}^{N} g_4(2\theta_j - \theta_l - \theta_k) + \sum_{j,l,m=1}^{N} g_5(\theta_j + \theta_l - \theta_m - \theta_k)\end{aligned} \tag{7.3}$$

up to some order of approximation, as shown in [15], where g_2, g_3, g_4, g_5 are 2π-periodic coupling functions. Thus, the dynamics of the phase reduction (7.3) reflect the effective dynamics of the underlying nonlinear oscillator networks and can reveal the possibility for chaotic phase dynamics [16]. Note that phase dynamics with nonpairwise interaction terms can arise independent of whether the nonlinear oscillator network has pairwise or nonpairwise coupling [14, 17].

In this chapter, we review recent progress on phase reductions in symmetric systems and their effective phase dynamics. We will also explicitly discuss these systems from the perspective of symmetry. First, we will outline the phase reduction of generically coupled symmetric systems close to a Hopf bifurcation [15]; equation (7.3) yields the resulting phase dynamics to higher order. The phase reduction is based on the calculation of the equivariants of the system. Second, we analyze the phase dynamics (7.3) and show that due to the inclusion of higher-order terms, chaotic dynamics can arise; see [16]. These dynamics arise in globally coupled networks. Third, we will analyze a variation of (7.3) that allows to introduce a nontrivial network structure. The resulting equations determine the dynamics of coupled populations of phase oscillator networks, where the coupling within populations and between populations is distinct. We summarize results from a series of papers [18–20] showing that the network dynamics can not only show localized frequency synchrony (i.e., frequencies are synchronized for some populations but not for others) akin to chimeras [21, 22] but the location of synchrony can also wander around the network through heteroclinic connections. We conclude with some remarks in the final section.

7.2 Symmetric Normal Forms and Higher-Order Interactions

An important tool to understand and classify bifurcations of dynamical systems is transformation to a normal form: This is the simplest form of nonlinear equation that locally explains the dynamics for all generic cases. In the next subsection we briefly recall relevant ideas from symmetric Hopf bifurcation before applying it to the problem of phase reduction near such a Hopf bifurcation; more details are in [15]. The main result of this section is to show that phase equations (7.3) with nonpairwise interactions arise as higher-order approximations of the dynamics for symmetric coupled oscillator networks with generic interaction close to a Hopf bifurcation.

7.2.1 Hopf Bifurcation With $\mathbf{S}_N$ Symmetry

In the general theory of symmetric (equivariant) dynamical systems [23] we study a system of ordinary differential equations (ODEs)

$$\dot{x} = f(x, \lambda) \tag{7.4}$$

with $x \in V, \lambda \in \mathbb{R}$, where V is a finite-dimensional space, λ is the bifurcation parameter, and f is a symmetric function.

We say that an invertible $n \times n$ matrix γ is a *symmetry* of (7.4) if $f(\gamma x, \lambda) = \gamma f(x, \lambda)$ for all $x \in V, \lambda \in \mathbb{R}$. A consequence of this is that if $x(t)$ is a solution to (7.4), then so is $\gamma x(t)$. For periodic solutions, if $x(t)$ is a T-periodic solution of (7.4) then so is $\gamma x(t)$. Uniqueness of solutions to the initial problem for (7.4) implies that the trajectory of $x(t)$ and $\gamma x(t)$ are either disjoint, in which case we have a new periodic solution, or identical, in which case they differ only by a phase shift, that is, $x(t) = \gamma x(t - t_0)$ for some t_0. In this case we say that the pair (γ, t_0) is a symmetry of the periodic solution $x(t)$. Symmetries of periodic solutions have both a spatial component γ and a temporal (phase shift) component t_0.

Bifurcation Theory investigates how solutions to differential equations can branch as a parameter is varied. Assume that $x = 0$ is an equilibrium of (7.4) for any λ. When convenient and there is confusion with subindices, we also write $f_\lambda(x) = f(x, \lambda)$. The symmetry of f imposes restrictions on the bifurcations that can occur as λ is varied. These can be a steady-state bifurcation, when an eigenvalue of the Jacobian $\mathrm{d}f_\lambda(0)$ of f at $x = 0$ passes through 0 (without loss of generality at $\lambda = 0$) or a Hopf bifurcation, when a pair of complex conjugate eigenvalues of $\mathrm{d}f_\lambda(0)$ crosses the imaginary axis with nonzero speed at $\pm\omega i$, $\omega \neq 0$ where $i = \sqrt{-1}$.

The problem of N identical and identically interacting smooth (C^∞) dynamical systems on $x_k \in \mathbb{R}^d$ ($d \geq 2$) that simultaneously undergo a Hopf bifurcation is considered in [15]. In such a case the dynamics close to the Hopf bifurcation can be approximated (beyond first order) by a phase oscillator system of the form (7.3). Specifically, consider the coupled ordinary differential equations

$$\begin{aligned} \dot{x}_1 &= H_\lambda(x_1) + \epsilon h_{\lambda,\epsilon}(x_1; x_2, \ldots, x_N) \\ &\vdots \quad \vdots \quad \vdots \\ \dot{x}_N &= H_\lambda(x_N) + \epsilon h_{\lambda,\epsilon}(x_N; x_1, \ldots, x_{N-1}). \end{aligned} \tag{7.5}$$

Note the parameter $\epsilon \in \mathbb{R}$ is such that the system completely decouples for $\epsilon = 0$. We now assume that each system undergoes a Hopf bifurcation of $x = 0$ when $\lambda \in \mathbb{R}$ passes through zero for $\epsilon = 0$. We assume that the uncoupled system for $x \in \mathbb{R}^d$ given by $\dot{x} = H_\lambda(x)$ has a linearly stable fixed point at $x = 0$ for $\lambda < 0$ that undergoes supercritical Hopf bifurcation at $\lambda = 0$, in particular $\mathrm{d}H_\lambda(0)$ has a complex pair of eigenvalues $\lambda \pm i\omega$, where $\omega \neq 0$ and all other eigenvalues μ of

$\mathrm{d}H_\lambda(0)$ satisfy $Re(\mu) < -r < 0$. Without loss of generality we can assume 0 is an equilibrium in some neighborhood of $(\lambda, \epsilon) = (0, 0)$.

7.2.2 Normal Forms for Symmetric Hopf Bifurcations With $\mathbf{S}_N$ Symmetry

System (7.5) describes a population of N identical, symmetrically coupled dynamical systems with state $x_k \in \mathbb{R}^d$ ($d \geq 2$) close to a Hopf bifurcation. We assume that the coupling respects the fact that the uncoupled systems can be permuted arbitrarily, i.e., that the system is equivariant under the action of $\mathbf{S}_N$ on $\mathbb{R}^{dN}$ by permutation of coordinates. Since the system is close to a bifurcation, the dynamics can now be reduced to a center manifold using equivariant bifurcation theory [24]: We explain how this can be used as a basis for a phase oscillator description as in [15].

Note that the action of the symmetry group $\mathbf{S}_N$ means that for $\epsilon > 0$ a generic Hopf bifurcation will have center manifold of dimension either 2 or $2N - 2$. For the uncoupled case $\lambda = \epsilon = 0$ the center manifold will be $2N$ dimensional with each coordinate x_k parametrized by $z_k \in \mathbb{C}$. That is, for $\lambda = \epsilon = 0$ points on the center manifold are parametrized by $(z_1, \ldots, z_N) \in \mathbb{C}^N$. The system on the center manifold is

$$\dot{z}_1 = f_\lambda(z_1) + \epsilon g_\lambda(z_1; z_2, \ldots, z_N) + O(\epsilon^2) \tag{7.6}$$

etc., where $z \in \mathbb{C}^N$ and we have changed coordinates so that for $z_k = 0$ is an equilibrium that undergoes generic supercritical Hopf bifurcation at $\lambda = 0$. Note that for $N > 1$ this will not be a generic Hopf bifurcation, but still we can assume $f_0(0) = 0$ and $\mathrm{d}f_0(0)$ has a pair of purely imaginary eigenvalues $\pm i\omega$ that pass transversely through the imaginary axis with non-zero speed on changing λ.

The reduced system (7.6) has symmetries. First, the action of $\gamma \in \mathbf{S}_N$ on $\mathbb{C}^N$ is by permutation of coordinates

$$\gamma(z_1, \ldots, z_N) = (z_{\gamma^{-1}(1)}, \ldots, z_{\gamma^{-1}(N)}), \tag{7.7}$$

where $(z_1, \ldots, z_N) \in \mathbb{C}^N$ meaning that $g_\lambda(z_1; z_2, \ldots, z_N)$ is symmetric under all permutations of the last $N - 1$ arguments. Second, Poincaré–Birkhoff normal form theory [23] means that to all polynomial orders we can assume there is a normal form symmetry given by the action of $\mathbb{T}$ on $\mathbb{C}^N$, where $\theta \in \mathbb{T}$ acts by

$$\theta(z_1, \ldots, z_N) = e^{i\theta}(z_1, \ldots, z_N). \tag{7.8}$$

The symmetries (7.7) and (7.8) restrict the possible terms that can appear in the normal form; we can characterize these by finding the possible equivariants, one order at a time. Suppose $N \geq 4$. Let $f : \mathbb{C}^N \to \mathbb{C}^N$ be $\mathbf{S}_N \times \mathbb{T}$-equivariant with respect to

the action (7.7), (7.8) with polynomial components of degree lower or equal than 3. From results in [24, Section 2.1.2] we can write $f = (f_1, f_2, \ldots, f_N)$ where

$$f_1(z_1, z_2, \ldots, z_N) = \sum_{i=-1}^{11} a_i h_i(z_1, z_2, \ldots, z_N) \tag{7.9}$$

with the other equations obtained by permutation, where the h_i are equivariants listed in [15] and $a_j \in \mathbb{C}$ are constants.

Following [15], this means we can write the equation for $\dot{z}_1$ from (7.6) in Poincaré-Birkhoff normal form [23] as the $\mathbf{S}_N \times \mathbb{T}$-equivariant system

$$\dot{z}_1 = U(z_1) + \epsilon F_1(z_1, \ldots, z_N, \epsilon), \tag{7.10}$$

where the third order truncated expression for F_1 is given in (7.35) and the other derivatives $\dot{z}_j$ are obtained by permutation of the indices.

7.2.3 *Perturbations from the Uncoupled Limit*

Note the Hopf bifurcation of (7.6) at $\lambda = 0$ has special structure: Following [15] we assume there is an "uncoupled limit" corresponding to $\epsilon = 0$. This extra structure means that

$$\dot{z_1} = U(z_1) := V(z_1)z_1 := \left[\lambda + i\omega + a_1|z_1|^2 + \tau(z_1)\right] z_1, \tag{7.11}$$

and we write $V(z_1) = V_R(z_1) + iV_I(z_1)$. We assume the uncoupled Hopf is supercritical, meaning $a_{1R} < 0$. We seek solutions of (7.11) of the form

$$z_1(t) = R_1(t)e^{i\phi_1(t)} = R_1(t)e^{i[\Omega t + \psi_1(t)]} \tag{7.12}$$

for some $R_1(t)$, $\psi_1(t)$ and constant Ω. Substituting this into (7.11), we require

$$\dot{R}_1 + iR_1\left[\Omega + \dot{\psi}_1\right] = R_1 V_R(R_1) + iR_1 V_I(R_1)$$

where

$$V_R(R_1) = \lambda + a_{1R}R_1^2 + \tau_R(R_1^2), \qquad V_I(R_1) = \omega + a_{1I}R_1^2 + \tau_I(R_1).$$

From this, it is clear that for small enough $\lambda > 0$ and $\epsilon = 0$ there is a stable periodic orbit at fixed $R_1 = R_* > 0$ such that $V_R(R_*) = 0$, with angular frequency $\Omega = V_I(R_*)$ and arbitrary but fixed phase ψ_1. More precisely, [15] shows that on solving $V_R(R_*) = 0$, we obtain

$$\begin{aligned} R_*^2 &= \frac{\lambda}{-a_{1R}} + O(\lambda^2), \\ \Omega = V_I(R_*^2) &= \omega + a_{1I}R_*^2 + \tau(R_*) = \omega + \frac{a_{1I}}{-a_{1R}}\lambda + O(\lambda^2). \end{aligned} \tag{7.13}$$

This implies there is a $\lambda_0 > 0$ such that for any $0 < \lambda < \lambda_0$ there is a stable periodic orbit (7.12) satisfying (7.13).

7.2.4 Reduction to Phase Oscillators

The final stage of the reduction undertaken by [15] is to show that, even though the uncoupled limit cycles for $\lambda > 0$ are weakly stable, the normal form can gives an explicit reduction to coupled phase oscillators as long as $\epsilon = o(\lambda)$. This involves some coordinate changes to ensure that standard results from normally hyperbolic invariant manifolds can be applied, followed by an averaging approximation. Since we will be dealing with multiple timescales here, we will write out the temporal derivatives $\frac{\mathrm{d}}{\mathrm{d}t}$ explicitly in this section.

For $\epsilon = 0$ and any $0 < \lambda < \lambda_0$ there is a stable invariant torus given by

$$(z_1, \ldots, z_N) = (R_* e^{i(\Omega t + \psi_1)}, \ldots, R_* e^{i(\Omega t + \psi_N)}), \tag{7.14}$$

parametrized by the phases $(\psi_1, \ldots \psi_N) \in \mathbb{T}^N$. This invariant torus is foliated by neutrally stable periodic orbits with period $2\pi/\Omega$ and so for each $0 < \lambda < \lambda_0$, the torus is normally hyperbolic. The theory of normal hyperbolicity [25] implies there is an ϵ_0 such that for $0 < \epsilon < \epsilon_0$ the invariant torus persists and is C^r-smooth for arbitrarily large r. Note that reducing r will restrict the ϵ_0: We will need $r \geq 5$ for the approximation to be valid.

We write $a_k = \alpha_k e^{i\theta_k} = a_{kR} + i a_{kI}$ and $z_k(t) = R_k(t) e^{i(\Omega t + \psi_k(t))} = [R_* + \rho_k(t)]$ $e^{i(\Omega t + \psi_k(t))}$ In particular, we seek solutions such that ρ_k is small and ψ_k varies slowly with t. Re-writing (7.10), note that

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_1 + i R_1 \left[\Omega + \frac{\mathrm{d}}{\mathrm{d}t}\psi_1\right] = U(R_1) + \epsilon F_1(z_1, \ldots, z_N, 0) e^{-i(\Omega t + \psi_1)} + O(\epsilon^2). \tag{7.15}$$

Writing U in real and imaginary parts and expanding for small ρ_1, [15] show that $A(\lambda) := U_R'(R_*)/\lambda$, $\quad B(\lambda) := V_I'(R_*)/(\lambda^{1/2})$, so that $U(R_1) = \lambda A(\lambda)\rho_1 + i R_1[\Omega + \lambda^{1/2} B(\lambda)\rho_1] + O(\rho_1^2)$. This implies that (7.15) can be expressed as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\rho_1 + i R_1 \left[\Omega + \frac{\mathrm{d}}{\mathrm{d}t}\psi_1\right] &= \lambda A(\lambda)\rho_1 + i R_1[\Omega + \lambda^{1/2} B(\lambda)\rho_1] \\ &\quad + \epsilon F_1(z_1, \ldots, z_N) e^{-i(\Omega t + \psi_1)} + O(\epsilon^2) \end{aligned} \tag{7.16}$$

Recalling from (7.13) that $R_*^2 = \lambda/(-a_{1R}) + O(\lambda^2)$, $U(R_*) = U_R(R_*) + iV_I(R_*)$ $R_* = (\lambda + a_{1R}R_*^2 + \tau(R_*))R_*$, $\tau(z) = O(z^4)$, and $\tau'(z) = O(z^3)$ so one can show $A(\lambda) = 1 + 3a_{1R}/(-a_{1R}) + O(\lambda) = -2 + O(\lambda)$. Similarly, one can show $B(\lambda) = 2a_{1I}/\sqrt{-a_{1R}} + O(\lambda)$. In particular, for $\lambda \to 0$ there are finite limits $A(0) = -2$, $B(0) = 2a_{1I}/\sqrt{-a_{1R}}$. Careful expansion of the terms in F_1 and taking real parts of (7.16) gives the expression (7.36). The equivalent equation for ψ_1 is obtained by taking imaginary parts of (7.15) and after cancellation and dividing by R_1, gives (7.37).

In terms of slow time $T = \lambda t$, calculations in [15] show that (7.15, 7.16) can be written as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}T} r_j &= A(\lambda) r_j + f_j + O(\epsilon) \\ \frac{\mathrm{d}}{\mathrm{d}T} \psi_j &= \epsilon \lambda^{-1} \left[C(\lambda) r_j + h_j \right] + O(\epsilon^2) \end{aligned} \tag{7.17}$$

for $j = 1, \ldots, N$. Note that f_j and h_j are trigonometric polynomials and A, C, f_j and h_j have finite limits as $\lambda \to 0$. Hence (7.17) gives a slow timescale for evolution of ψ_j as long as $\epsilon = o(\lambda)$. Defining scaled amplitude variables $\sigma_j := r_j + \frac{f_j(\psi_1, \ldots, \psi_{N-1})}{A(\lambda)}$, system (7.17) can be expressed as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}T} \sigma_j &= A(\lambda) \sigma_j + O(\epsilon) \\ \frac{\mathrm{d}}{\mathrm{d}T} \psi_j &= \epsilon \lambda^{-1} \left[C(\lambda) \sigma_j + H_j \right] + O(\epsilon^2), \end{aligned} \tag{7.18}$$

where $H_j = h_j - f_j C(\lambda)/A(\lambda)$. We write $H_j = H_j^0 + \lambda H_j^1 + O(\lambda^2)$, where $H_j^0 = h_j^0 - C(0)/A(0) f_j^0$, $H_j^1 = R_*^2(\lambda)\left[h_j^1 - f_j^1 C(0)/A(0)\right]/\lambda - f_j^0[C'(0)A(0) - A'(0)$ $C(0)]/A(0)^2$, which is a trigonometric polynomial in $\psi_k - \phi_j$. It can be shown that H_j^0 only involves pairwise coupling while H_j^1 includes coupling of up to four phases (and on $\alpha_2, \ldots, \alpha_{11}$).

After further manipulations [15], the reduced equations for ϕ_j can be written in the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_j = \Omega + \epsilon \left[H_j^0 + \lambda H_j^1 \right] \tag{7.19}$$

where the phase differences $\psi_j - \psi_k = \phi_j - \phi_k$ for all j and k, and the approximation will be close for times $0 < t < \tilde{t}$ with $\tilde{t} = O(\epsilon^{-1}\lambda^{-2})$. For $k = -1, 1, \ldots, 11$ we define β_k and γ_k such that for all θ we have $\beta_k \cos(\gamma_j + \theta) := \alpha_k \sin(\theta_k + \theta) - \frac{C(0)}{A(0)} \alpha_k \cos(\theta_k + \theta)$. Then we can write (7.19) in the form

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi_j &= \Omega + \epsilon H_1 \\ &= \tilde{\Omega}(\phi,\epsilon) + \frac{\epsilon}{N}\sum_{k=1}^{N} g_2(\phi_k - \phi_j) + \frac{\epsilon}{N^2}\sum_{k,\ell=1}^{N} g_3(\phi_k + \phi_\ell - 2\phi_j) \\ &\quad + \frac{\epsilon}{N^2}\sum_{k,\ell=1}^{N} g_4(2\phi_k - \phi_\ell - \phi_j) + \frac{\epsilon}{N^3}\sum_{k,\ell,m=1}^{N} g_5(\phi_k + \phi_\ell - \phi_m - \phi_j)\end{aligned} \tag{7.20}$$

where the various coupling functions have the form

$$\begin{aligned}\tilde{\Omega}(\phi,\epsilon) &= \Omega + R_*^2\varepsilon\left[\beta_4\cos\gamma_4 + \frac{\beta_5}{N^2}\sum_{j,k}\cos(\gamma_5 + \phi_j - \phi_k)\right] \\ g_2(\varphi) &= \beta_{-1}\cos(\gamma_{-1}+\varphi) + R_*^2\left[\beta_2\cos(\gamma_2-\varphi) + \beta_3\cos(\gamma_3+\varphi)\right. \\ &\quad \left.+\beta_6\cos(\gamma_6+2\varphi) + \beta_8\cos(\gamma_8+\varphi) + \beta_{10}\cos(\gamma_{10}+\varphi)\right] \\ &\quad - \lambda\frac{C'(0)A(0) - A'(0)C(0)}{A(0)^2}\alpha_{-1}\cos(\theta_{-1}+\varphi) \\ g_3(\varphi) &= R_*^2\left[\beta_7\cos(\gamma_7+\varphi)\right] \\ g_4(\varphi) &= R_*^2\left[\beta_9\cos(\gamma_9+\varphi)\right] \\ g_5(\varphi) &= R_*^2\left[\beta_{11}\cos(\gamma_{11}+\varphi)\right].\end{aligned} \tag{7.21}$$

To summarize, we have illustrated how the reduction of [15] demonstrates that, to first order, the generic dynamics of N weakly coupled coupled identical oscillators close to a Hopf bifurcation are approximated by the Kuramoto equations (7.1) with an additional phase-shift parameter α, i.e., the Kuramoto–Sakaguchi equations [6]. Moreover, at second order in the bifurcation parameter λ we have phase dynamics given by (7.20), a system very similar to (7.3): The phase dynamics are determined by

$$\dot{\theta}_k = \tilde{\Omega}(\theta,\epsilon) + \varepsilon\left(F_k^{(2)}(\theta) + F_k^{(3)}(\theta) + F_k^{(4)}(\theta)\right) \tag{7.22}$$

for $k \in \{1,\ldots,N\}$ with

$$F_k^{(2)}(\theta) = \frac{1}{N}\sum_{j=1}^{N} g_2(\theta_j - \theta_k) \tag{7.23a}$$

$$F_k^{(3)}(\theta) = \frac{1}{N^2}\sum_{j,\ell=1}^{N} g_3(\theta_j + \theta_\ell - 2\theta_k) + \frac{1}{N^2}\sum_{j,\ell=1}^{N} g_4(2\theta_j - \theta_\ell - \theta_k) \tag{7.23b}$$

$$F_k^{(4)}(\theta) = \frac{1}{N^3}\sum_{j,\ell,m=1}^{N} g_5(\theta_j + \theta_\ell - \theta_m - \theta_k) \tag{7.23c}$$

and coupling functions

$$\begin{aligned} g_2(\phi) &= \xi_1^0 \cos(\phi + \chi_1^0) + \lambda \xi_1^1 \cos(\phi + \chi_1^1) + \lambda \xi_2^1 \cos(2\phi + \chi_2^1) \\ g_3(\phi) &= \lambda \xi_3^1 \cos(\phi + \chi_3^1) \\ g_4(\phi) &= \lambda \xi_4^1 \cos(\phi + \chi_4^1) \\ g_5(\phi) &= \lambda \xi_5^1 \cos(\phi + \chi_5^1) \end{aligned} \tag{7.24}$$

for coefficients ξ_i^j and χ_i^j determined from (7.21). In particular, this next order includes pairwise, triplet and quadruplet interactions of phases.

7.3 Coupled Phase Oscillators Networks with Nonpairwise Interactions

In this section, we recall some results from [16] and related literature that explores the phase equations (7.22) with higher-order interactions. For concreteness, we set $\tilde{\Omega}(\theta, \epsilon) = \omega$ and fix $\lambda = \varepsilon = 1$. That is, we consider (7.3) with the coupling functions

$$\begin{aligned} g_2(\phi) &= \xi_1 \cos(\phi + \chi_1) + \xi_2 \cos(2\phi + \chi_2) \\ g_3(\phi) &= \xi_3 \cos(\phi + \chi_3) \\ g_4(\phi) &= \xi_4 \cos(\phi + \chi_4) \\ g_5(\phi) &= \xi_5 \cos(\phi + \chi_5) \end{aligned} \tag{7.25}$$

such that for general N the function g_2 determines pairwise, g_3, g_4 triplet and g_5 quadruplet interaction.

7.3.1 Symmetric Phase Oscillator Networks

The symmetries of the phase equations (7.3) have consequences for the dynamics. Here the phase equations "inherit" symmetries from the generically coupled system (7.5): First, the phase equations are symmetric with respect to the rotation by a common angle. As a consequence, we may assume—without loss of generality—that the phase of the first oscillator θ_1 is always equal to zero by going into a co-rotating reference frame that moves with oscillator $k = 1$. Second, the $\mathbf{S}_N$-symmetry acts by permuting oscillators. By using the permutational symmetry, we may assume that the phases are in ascending order. Note that these properties are due to the symmetry alone, independent of whether the phase oscillators are subject to pairwise or nonpairwise interactions; cf. [26].

Because of the symmetries, we do not need to consider the dynamics of (7.3) on the entire phase space $\mathbb{T}^N$ but can restrict the analysis to a smaller but still representative subset. Specifically, define the *canonical invariant region* (CIR) [26] as the set of phases

$$\mathcal{C} = \left\{ \theta \in \mathbb{T}^N \mid 0 = \theta_1 < \theta_2 < \cdots < \theta_N < 2\pi \right\}. \tag{7.26}$$

The CIR is a $(N-1)$-simplex whose boundary consists of *cluster configurations* where the phases of two or more oscillators are equal. The intersection of all cluster configurations is the fully synchronized phase configuration

$$\mathrm{S} := \left\{ (\theta_1, \ldots, \theta_N) \in \mathbb{T}^N \mid \theta_k = \theta_{k+1} \right\} \tag{7.27a}$$

where the phases of all oscillators are equal. At the centroid of the CIR is the *splay phase configuration*

$$\mathrm{D} := \left\{ (\theta_1, \ldots, \theta_N) \in \mathbb{T}^N \;\middle|\; \theta_{k+1} = \theta_k + \frac{2\pi}{N} \right\}, \tag{7.27b}$$

where the oscillator phases are uniformly distributed on the circle. As fixed point subsets of symmetries—e.g., S is invariant under any permutation of the oscillator indices—the cluster configurations are also dynamically invariant.

The CIR for $N = 3$ and $N = 4$ is illustrated in Fig. 7.1. For $N = 3$ the CIR is a two-dimensional simplex and we cannot expect any chaotic dynamics [27]. For

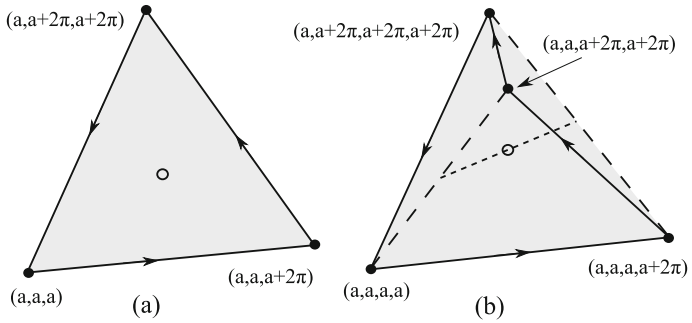


Fig. 7.1 Structure of the canonical invariant region $\mathcal{C}$ for $N = 3$ and $N = 4$ (see [26]). Panels **a**, **b** show $\mathcal{C}$ as an orthogonal projection of into $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. The edges of $\mathcal{C}$ for, **a** and the faces of $\mathcal{C}$ for, **b** are points where two oscillators have the same phase. The filled circles represent the fully synchronous phase configuration S; the open circle represents the splay phase configuration D in $\mathcal{C}$. In **b** the solid lines correspond to 3:1 cluster configurations where three oscillators have the same phase and one is distinct while the long-dashed lines correspond to 2:2 cluster configurations of two clusters of two oscillators. The short-dashed lines are points $(a, b, a+\pi, b+\pi)$. For any N there is a residual $\mathbb{Z}/N\mathbb{Z}$ symmetry that "rotates" the canonical invariant region (the direction of rotation is indicated by the arrows in **b**). Overall $(N-1)!$ symmetric copies of $\mathcal{C}$ pack a generating region for the torus. Reprinted from [16]

$N = 4$ the CIR is three-dimensional which does not preclude chaotic dynamics. If the coupling is pairwise with a single harmonic as in the Kuramoto–Sakaguchi model, there is additional degeneracy that prevent chaotic attractors to emerge [28]. If the coupling is pairwise but one allows for higher harmonics in the coupling function (cf. [29]) one may observe chaotic dynamics for a pairwise coupling function with four harmonics [30]. But no further examples of coupling functions with fewer harmonics are known for fully symmetric systems with pairwise interactions.

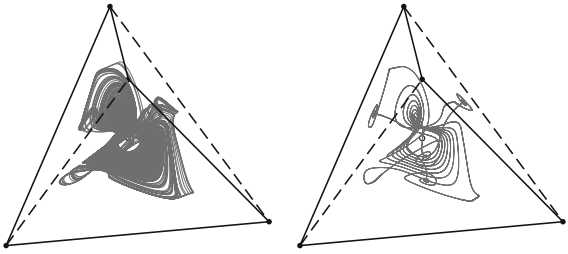
7.3.2 *Chaos in Globally Coupled Phase Oscillator Networks with Higher-Order Interactions*

The phase dynamics of (7.3) with nonpairwise coupling mediated by the functions (7.25) can give rise to chaotic dynamics. Following [16] we fix Fourier coefficients

$$\xi = (-0.3, 0.3, 0.02, 0.8, 0.02) \tag{7.28}$$

while varying the phase shifts χ. Calculating the maximal Lyapunov exponent $\lambda_{\max}$ reveals a region in parameter space where $\lambda_{\max} > 0$ and chaotic attractors appear in the canonical invariant region. Figure 7.2a shows a solution $\theta(t)$ in $\mathcal{C}$ for $\chi = (0.154, 0.318, 0, 1.74, 0)$. While the attracting set lies in the interior of $\mathcal{C}$, the trajectories on the chaotic attractor come close to its boundary that consist where oscillators are clustered. Indeed, a small variation shows periodic dynamics that appear to be close to a heteroclinic network: Fig. 7.2b shows a stable periodic orbit close to such a heteroclinic network for parameters $\chi = (0.2, 0.316, 0, 1.73, 0)$.

Since the equilibria on the boundary include a saddle-focus, the chaotic dynamics appear to arise through a nonstandard Shilnikov saddle-focus scenario [31]. Indeed,



(a) Chaotic attractor (b) Periodic orbit near heteroclinic network

Fig. 7.2 Heteroclinic networks organize chaotic behavior in $\mathcal{C}$ for networks of $N = 4$ oscillators; line styles on the boundary of $\mathcal{C}$ are as in Fig. 7.1. The right panel shows a trajectory with positive maximal Lyapunov exponents for phase shift parameters $\chi = (0.154, 0.318, 0, 1.74, 0)$ that comes close to the boundary of $\mathcal{C}$. For nearby parameter values $\chi = (0.2, 0.316, 0, 1.73, 0)$ there is an attracting periodic orbit close to a heteroclinic network involving two saddle equilibria, one a saddle-focus, on the boundary of $\mathcal{C}$. Reprinted from [16]

modulo the residual $\mathbb{Z}_N$ symmetry on $\mathcal{C}$ the heteroclinic network on the boundary of $\mathcal{C}$ involves two equilibria. Grines and Osipov [32] took this observation as a starting point to determine what homoclinic and heteroclinic trajectories are possible in (7.3) for $N = 4$ oscillators. More specifically, the symmetries of the system restrict the saddle connections that are possible between equilibria that lie on the boundary of $\mathcal{C}$ such as those in Fig. 7.2b.

While $N = 4$ is the smallest number of oscillators for which chaos can arise in the phase equations, chaotic dynamics also arise in networks with $N > 4$ phase oscillators. In [16] we gave explicit parameter values for which $\lambda_{\max} > 0$ but a detailed analysis of these larger phase oscillator networks is still an outstanding problem.

7.4 Chimeras and Other Creatures for Multiple Populations

The dynamics of a globally coupled network (7.3) of N identical oscillators with nonpairwise interactions is constrained by the symmetries of the system. Since the system is $\mathbf{S}_N$-equivariant, the asymptotic average frequencies $\Omega_k(\theta(0)) := \lim_{T\to\infty} \theta(T)/T$ for any initial condition $\theta(0) \in \mathbb{T}^N$ are identical[1]: We have $\Omega_k = \Omega_j$ for all $k, j \in \{1, \ldots, N\}$ independent of the initial condition and the oscillators are *frequency synchronized* [21]. This restriction breaks down if the $\mathbf{S}_N$ symmetry is broken. In this section we discuss the dynamics of a generalization of (7.29) where the phase θ_k evolves according to

$$\begin{aligned}\dot{\theta}_k = \omega &+ \sum_{j=1}^{N} a_2^{(jk)} g_2(\theta_j - \theta_k) + \sum_{j,l=1}^{N} a_3^{(jlk)} g_3(\theta_j + \theta_l - 2\theta_k) \\ &+ \sum_{j,l=1}^{N} a_4^{(jlk)} g_4(2\theta_j - \theta_l - \theta_k) + \sum_{j,l,m=1}^{N} a_5^{(jlmk)} g_5(\theta_j + \theta_l - \theta_m - \theta_k)\end{aligned} \tag{7.29}$$

where $a_2^{(jk)} \in \mathbb{R}$ and $a_3^{(jlk)}, a_4^{(jlk)}, a_5^{(jlmk)} \in \mathbb{R}$ are the coupling strength of pairwise and nonpairwise interactions. For nonhomogeneous choice of these coupling coefficients, the system (7.29) can describe coupled populations of phase oscillators that allow for frequency synchrony to be localized in one or more populations.

[1] Here we assume that the limit exists; for a generalization to frequency intervals see [22].

7.4.1 Frequency Synchrony in Coupled Oscillator Populations

Suppose that the N oscillators are grouped into M populations (assuming $N = MQ$) indexed by $\sigma \in \{1, \ldots, M\}$. The first Q oscillators belong to population $\sigma = 1$, oscillators $k \in \{Q+1, \ldots, 2Q\}$ to population $\sigma = 2$ etc., and we write $k = (\sigma, q)$ if oscillator k (in the linear ordering as above) corresponds to oscillator q in population σ and $\theta_{\sigma,q}$ denote its phase and $\Omega_{\sigma,q}$ the asymptotic average frequency. We now consider networks with coupling coefficients $a_2^{(jk)} = K_{\mathrm{p}}^{(\sigma)}/Q$ if k, j belong to population σ and $a_2^{(jk)} = 0$ otherwise so that interactions within populations are pairwise and $a_3^{(jlk)} = a_4^{(jlk)} = 0$, and $a_5^{(jlmk)} = K_{\mathrm{np}}^{(\sigma\tau)}/Q^3$ if and only if oscillators m, k belong to population σ and oscillators j, l to population τ and $a_5^{(jlmk)} = 0$ otherwise determine the nonpairwise interactions. With this choice of coefficients the M populations are globally and identically coupled through pairwise interactions while the nonpairwise interactions mediate the coupling between distinct populations.

The specific form of network coupling induces symmetries: The dynamical system is $(\mathbf{S}_Q \times \mathbb{T})^M$-equivariant where, for each population, $\mathbf{S}_Q$ acts by permuting the oscillators and $\mathbb{T}$ acts by shifting all oscillators of the given population by a constant. Note that there is one phase-shift symmetry for each population. For population σ, write $\theta_\sigma = (\theta_{\sigma,1}, \ldots, \theta_{\sigma,Q})$ to denote the state of the population. Recall that S and D, as defined in (7.27), denote the synchronized and splay configurations in a network consisting of a single population. For the network of interacting populations, write

$$\theta_1 \cdots \theta_{\sigma-1} \mathrm{S} \theta_{\sigma+1} \cdots \theta_M = \left\{ \theta \in \mathbb{T}^N \;\middle|\; \theta_\sigma \in \mathrm{S} \right\} \tag{7.30a}$$

$$\theta_1 \cdots \theta_{\sigma-1} \mathrm{D} \theta_{\sigma+1} \cdots \theta_M = \left\{ \theta \in \mathbb{T}^N \;\middle|\; \theta_\sigma \in \mathrm{D} \right\} \tag{7.30b}$$

to indicate that population σ is fully phase synchronized or in splay phase. Because of the symmetry these sets are dynamically invariant. We extend this notation to intersections of the sets (7.30), so that $\mathrm{S} \cdots \mathrm{S}$ (M times) denotes cluster states where all populations are fully phase synchronized and $\mathrm{D} \cdots \mathrm{D}$ (M times) the set where all populations are in splay phase.

These invariant sets can display frequency synchrony that is *localized* in a specific part of the network: The oscillators within one populations are frequency synchronized while oscillators in different populations are not. This is a characterizing feature of a *weak chimera* [21, 22]. To see this take $K_{\mathrm{np}}^{(\sigma\tau)} = 0$, that is, there is no coupling between different populations. If population σ is phase synchronized, that is, $\theta_\sigma(0) = (\theta_{\sigma,1}(0), \ldots, \theta_{\sigma,Q}(0)) \in \mathrm{S}$ we have

$$\Omega_{\sigma,k}(\theta(0)) = \omega + K_{\mathrm{p}}^{(\sigma)} g_2(0). \tag{7.31}$$

Similarly, if population σ in splay phase, that is, $\theta_\sigma(0) \in \mathrm{D}$, we have

$$\Omega_{\sigma,k}(\theta(0)) = \omega + \sum_{j=1}^{Q} \frac{K_{\mathrm{p}}^{(\sigma)}}{Q} g_2\left(\frac{2\pi j}{Q}\right). \tag{7.32}$$

Since these two values are distinct for a generic pairwise coupling function g_2, we have that any set of the form DS$\cdots$S has populations with distinct frequency. Moreover, this property is preserved for sufficiently small $\left|K_{\mathrm{np}}^{(\sigma,\tau)}\right| > 0$.

7.4.2 Heteroclinic Cycles and Networks

While much attention has focused on localized frequency and chimeras to be attractors in network dynamical systems [33], the nonpairwise interactions also allow for heteroclinic dynamics that connect different localized frequency synchrony patterns. For us, a heteroclinic cycle consists of a finite number of normally hyperbolic invariant sets ξ_s, $s \in \{1, \ldots, S\}$, together with trajectories $[\xi_s \to \xi_{s+1}]$ (indices are taken modulo S) that lie in the intersection of the unstable manifold of ξ_s and the stable manifold of ξ_{s+1}; cf. [34, 35]. Trajectories close to a heteroclinic cycle show "switching dynamics": The trajectory will spend time close to one of the invariant sets ξ_s before a rapid transition to the next set.

For small networks that consist of $M = 3$ populations of $Q \in \{2, 3\}$ we can explicitly give conditions for the existence of robust heteroclinic cycles that are asymptotically stable. Here we outline the results for $Q = 2$ oscillators and refer to [18–20] for more detailed results.

Theorem *Consider $M = 3$ populations of $Q = 2$ oscillators with coupling functions $g_2(\vartheta) = \sin(\vartheta + \alpha_2) + r\sin(2(\vartheta + \alpha_2))$ and $g_4(\vartheta) = \sin(\vartheta + \alpha_4)$ and nonpairwise coupling parameters $K_{\mathrm{np}}^{(\sigma\tau)} = -K$ if $\tau = \sigma - 1$, $K_{\mathrm{np}}^{(\sigma\tau)} = K$ if $\tau = \sigma + 1$, and $K_{\mathrm{np}}^{(\sigma\tau)} = 0$ if $\tau = \sigma$. Then there exists an open set of parameter values K, r, α_2, α_4 such that the coupled phase oscillator network (7.29) with higher-order interactions has an asymptotically stable robust heteroclinic cycle.*

The main ideas of the proof is as follows. First, note that because of the $\mathbf{S}_Q^M$ symmetry we can reduce the 6-dimensional dynamics to a system of 3 phase difference variables $\psi_\sigma = \theta_{\sigma,2} - \theta_{\sigma,1}$ for each population $\sigma \in \{1, 2, 3\}$. In the reduced coordinates invariant sets of the form SSS, DSS, … are equilibrium points. Second, we can linearize the equations close to these equilibria. This allows to write down conditions that ensure that the equilibria have the right (local) stability properties. For example, we can impose that DSS is stable in the invariant subspaces DSθ_3 and θ_1SS but unstable in the invariant subspace Dθ_2S. Moreover, we want that DDS is stable in Dθ_2S and DDθ_3 but unstable in θ_1DS. The stability conditions for the other equilibria are similar. Third, we have to ensure that there are heteroclinic connections between the equilibria: There is a connection [DSS $\to$ DDS] if there are no other equilibria in the one-dimensional invariant set Dθ_2S. This condition—as well as conditions for the other heteroclinic connections—can be explicitly expressed in terms

of the coupling parameters. Fourth, we have that the resulting heteroclinic cycle is in the class of quasi-simple heteroclinic cycles; see [36]. This allows to write down explicit conditions for the stability of the resulting cycle [20]. Heteroclinic structures organize the dynamics even if these structures are broken by perturbations: Typically, periodic or chaotic dynamics appear that closely mimic the switching dynamics of the cycle.

For a larger number of populations, such heteroclinic cycles can be part of larger networks of heteroclinic connections. Existence of a heteroclinic network in $M = 4$ coupled populations of $Q = 2$ oscillators each is proved in [20]. This network consists of two cycles of the form discussed above with the difference that from the equilibrium SDSS there are two distinct heteroclinic connections [SDSS $\to$ SDDS] and [SDSS $\to$ SDSD] resulting in a network that contains two distinct heteroclinic cycles. In other words, the second population can desynchronize either the third or the fourth population. If weak noise is added to the system nearby trajectories exhibit dynamics that can follow either of the two cycles in the network. As quasi-simple heteroclinic cycles—one can calculate their stability properties explicitly.

7.5 Outlook

In this chapter, we reviewed results from [15, 16] and related literature [18–20] that discuss how nonpairwise interactions in phase oscillator networks arise naturally in phase reductions and their consequences for the phase dynamics. The framework of symmetric Hopf bifurcation theory helps organize and understand the importance these nonpairwise interactions of the phase dynamics in a rigorous manner. We discussed the dynamics of the resulting phase oscillator networks and a generalization thereof that allows for a more general network structure other than global and identical coupling.

One of the more puzzling aspects of higher order interactions in phase oscillator networks is that it seems to be hard to characterize the dynamical restrictions imposed by having only pairwise interactions. With a few exceptions (e.g., the scenarios for cluster state stabilities considered in [15]), pairwise coupled systems are remarkably rich in their dynamics. This may be the reason why higher order interactions have only recently become of interest. In another approach, Komarov and Pikovsky [37] consider a phase oscillator system of the form

$$\dot{\phi}_k = \Omega + \omega + S(\phi_k)F \tag{7.33}$$

where F depends on the mean fields. They show that the second order phase dynamics are given by

$$\dot{\theta}_k = \omega + \varepsilon \left(F_k^{(2)}(\theta) + F_k^{(3)}(\theta) \right) \tag{7.34}$$

with $F_k^{(2)}(\theta) = \frac{1}{N}\sum_{j=1}^{N} g_2(\theta_j - \theta_k)$, $F_k^{(3)}(\theta) = \frac{1}{N^2}\sum_{j,\ell=1}^{N} g_3(\theta_j + \theta_\ell - 2\theta_k)$ and the interactions between the phases are given by $g_2(\phi) = \xi_1 \cos(\phi + \chi_1) + \xi_2 \cos(2\phi + \chi_2)$, $g_3(\phi) = \xi_3 \cos(\phi + \chi_3)$. This is a special case of (7.29) where the coupling functions g_4, g_5 are zero. Similarly, the phase oscillator network (7.2) considered by Skardal and Arenas [9, 10] is a special case of (7.29) as mentioned above.

While phase oscillators with nonpairwise interactions can be analyzed in their own right, it is instructive to remember that such interaction terms arise in phase reductions as discussed here. The nonpairwise interactions capture the nonlinearities of the (unreduced) nonlinear oscillator system and their interactions. Thus, phase oscillator networks with nonpairwise interactions can capture some properties of their dynamics. It seems natural to assume that it is especially when one moves away from the weakly coupled limit that higher-order interactions will become decisive: For example, the discontinuous synchronization transitions in [38] appear in a strongly-coupled oscillator network, while [39] also consider effects that can be viewed as associated with higher-order interactions.

Truncated expressions for phase and amplitude dynamics

For completeness, the expression for the cubic truncated Hopf normal form from [15] is

$$\begin{aligned} F_1 = \Bigg[& a_{-1}\frac{1}{N}\sum_j z_j + a_2\frac{z_1^2}{N}\sum_j \overline{z}_j + a_3\frac{|z_1|^2}{N}\sum_j z_j \\ & + a_4\frac{z_1}{N}\sum_j |z_j|^2 + a_5\frac{z_1}{N^2}\sum_{j,k} z_j\overline{z}_k + a_6\frac{\overline{z}_1}{N}\sum_j z_j^2 \\ & + a_7\frac{\overline{z}_1}{N^2}\sum_{j,k} z_j z_k + a_8\frac{1}{N}\sum_j |z_j|^2 z_j + a_9\frac{1}{N^2}\sum_{j,k} z_j^2\overline{z}_k \\ & + a_{10}\frac{1}{N^2}\sum_{j,k} z_j|z_k|^2 + a_{11}\frac{1}{N^3}\sum_{j,k,\ell} z_j z_k \overline{z}_\ell \Bigg] + \tilde{F}_1 + O(\epsilon). \end{aligned} \tag{7.35}$$

where the $\epsilon = 0$ error term is $\tilde{F}_1 = O(|z|^5)$, $\sum_i$ denotes $\sum_{i=1}^{N}$, $\sum_{i,j}$ denotes $\sum_{i=1}^{N}\sum_{j=1}^{N}$ and $\sum_{i,j,k}$ denotes $\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}$. This can be recovered from [24].

The radial dynamics for phase reduction is

$$\begin{aligned}\dot{\rho}_1(t) = \lambda A(\lambda)\rho_1 + \epsilon\Big[&\alpha_{-1}\textstyle\sum_j' R_j\cos(\theta_{-1}+\psi_j-\psi_1)\\ &+\alpha_2\textstyle\sum_j' R_1^2R_j\cos(\theta_2+\psi_1-\psi_j)\\ &+\alpha_3\textstyle\sum_j' R_1^2R_j\cos(\theta_3+\psi_j-\psi_1)\\ &+\alpha_4\textstyle\sum_j' R_1R_j^2\cos\theta_4\\ &+\alpha_5\textstyle\sum_{j,k}' R_1R_jR_k\cos(\theta_5+\psi_j-\psi_k)\\ &+\alpha_6\textstyle\sum_j' R_1R_j^2\cos(\theta_6+2\psi_j-2\psi_1)\\ &+\alpha_7\textstyle\sum_{i,j}' R_1R_iR_j\cos[\theta_7+(\psi_i-\psi_1)+(\psi_j-\psi_1)]\\ &+\alpha_8\textstyle\sum_j' R_j^3\cos(\theta_8+\psi_j-\psi_1)+\\ &+\alpha_9\textstyle\sum_{j,k}' R_j^2R_k\cos(\theta_9+2\psi_j-\psi_k-\psi_1)\\ &+\alpha_{10}\textstyle\sum_{j,k}' R_jR_k^2\cos(\theta_{10}+\psi_j-\psi_1)\\ &+\alpha_{11}\textstyle\sum_{i,j,k}' R_iR_jR_k\cos(\theta_{11}+\psi_i+\psi_j-\psi_k-\psi_1)\Big]\\ &+O(\rho^2,\epsilon^2)\end{aligned} \tag{7.36}$$

where $\rho^2 = \max_j(\rho_j^2)$ and $\sum_j' a_j := \frac{1}{N}\sum_{j=1}^N a_j$, $\sum_{j,k}' a_{j,k} := \frac{1}{N^2}\sum_{j,k=1}^N a_{j,k}$, etc. are the normalized sums. Similarly the phase dynamics are given by

$$\begin{aligned}\dot{\psi}_1(t) = \lambda^{1/2}B(\lambda)\rho_1 + \epsilon\Big[&\alpha_{-1}\textstyle\sum_j'(R_j/R_1)\sin(\theta_{-1}+\psi_1-\psi_j)\\ &+\alpha_2\textstyle\sum_j' R_1R_j\sin(\theta_2+\psi_1-\psi_j)\\ &+\alpha_3\textstyle\sum_j' R_1R_j\sin(\theta_3+\psi_j-\psi_1)\\ &+\alpha_4\textstyle\sum_j' R_j^2\sin\theta_4\\ &+\alpha_5\textstyle\sum_{j,k}' R_jR_k\sin(\theta_5+\psi_j-\psi_k)\\ &+\alpha_6\textstyle\sum_j' R_j^2\sin(\theta_6+2(\psi_j-\psi_1))\\ &+\alpha_7\textstyle\sum_{i,j}' R_iR_j\sin[\theta_7+(\psi_i-\psi_1)+(\psi_j-\psi_1)]\\ &+\alpha_8\textstyle\sum_j'(R_j^3/R_1)\sin(\theta_8+\psi_j-\psi_1)\\ &+\alpha_9\textstyle\sum_{j,k}'(R_j^2R_k/R_1)\sin(\theta_9+2\psi_j-\psi_k-\psi_1)\\ &+\alpha_{10}\textstyle\sum_{j,k}'(R_jR_k^2/R_1)\sin(\theta_{10}+\psi_j-\psi_1)\\ &+\alpha_{11}\textstyle\sum_{i,j,k}'(R_iR_jR_k/R_1)\sin(\theta_{11}+\psi_i+\psi_j-\psi_k-\psi_1)\Big]\\ &+\tfrac{1}{R_1}O(\rho^2,\epsilon^2)\end{aligned} \tag{7.37}$$

References

1. A. Pikovsky, M. Rosenblum, J. Kurths, *A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Synchronization, 2003)
2. S.H. Strogatz. *Sync: The Emerging Science of Spontaneous Order* (Penguin, 2004)
3. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, vol. 19. (Springer, Berlin, 1984)
4. J. Acebrón, L. Bonilla, Conrad P. Vicente, F. Ritort, R. Spigler. The Kuramoto model: A simple paradigm for synchronization phenomena. Rev. Mod. Phys. **77**(1), 137–185 (2005)
5. F.A. Rodrigues, T.K.D.M. Peron, P. Ji, J. Kurths, The Kuramoto model in complex networks. Phys. Rep. **610**, 1–98 (2016)
6. H. Sakaguchi, Y. Kuramoto, A soluble active rotator model showing phase transitions via mutual entrainment. Prog. Theor. Phys. **76**, 576–581 (1986)

7. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J. Young, G. Petri, Networks beyond pairwise interactions: structure and dynamics. Phys. Rep. **874**, 1–92 (2020)
8. C. Bick, E. Gross, H.A. Harrington, M.T. Schaub, What are higher-order networks? (2021) arXiv:2104.11329
9. P.S. Skardal, A. Arenas, Abrupt desynchronization and extensive multistability in globally coupled oscillator simplexes. Phys. Rev. Lett. **122**(24), 248301 (2019)
10. P.S. Skardal, A. Arenas, Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching. Commun. Phys. **3**(1), 218 (2020)
11. C. Kuehn, C. Bick, A universal route to explosive phenomena. Sci. Adv. **7**(16), eabe3824 (2021)
12. H. Nakao, Phase reduction approach to synchronisation of nonlinear oscillators. Contem. Phys. **57**(2), 188–214 (2016)
13. B. Pietras, A. Daffertshofer, Network dynamics of coupled oscillators and phase reduction techniques. Phys. Rep. **819**, 1–105 (2019)
14. B. Kralemann, A. Pikovsky, M. Rosenblum, Reconstructing effective phase connectivity of oscillator networks from observations. New J. Phys. **16**(8), 085013 (2014)
15. P. Ashwin, A. Rodrigues, Hopf normal form with $\mathbf{S}_N$ symmetry and reduction to systems of nonlinearly coupled phase oscillators. Physica D **325**, 14–24 (2016)
16. C. Bick, P. Ashwin, A. Rodrigues, Chaos in generically coupled phase oscillator networks with nonpairwise interactions. Chaos **26**(9), 094814 (2016)
17. I. León, D. Pazó, Phase reduction beyond the first order: the case of the mean-field complex Ginzburg-Landau equation. Phys. Rev. E **100**(1), 012211 (2019)
18. C. Bick, Heteroclinic switching between chimeras. Phys. Rev. E **97**(5), 050201(R) (2018)
19. C. Bick, Heteroclinic dynamics of localized frequency synchrony: heteroclinic cycles for small populations. J. Nonlinear Sci. **29**(6), 2547–2570 (2019)
20. C. Bick, A. Lohse, Heteroclinic dynamics of localized frequency synchrony: stability of heteroclinic cycles and networks. J. Nonlinear Sci. **29**(6), 2571–2600 (2019)
21. P. Ashwin, O. Burylko, Weak chimeras in minimal networks of coupled phase oscillators. Chaos **25**, 013106 (2015)
22. C. Bick, P. Ashwin, Chaotic weak chimeras and their persistence in coupled populations of phase oscillators. Nonlinearity **29**(5), 1468–1486 (2016)
23. M. Golubitsky, D.G. Schaeffer, I.N. Stewart. *Singularities and Groups in Bifurcation Theory, Vol. II*, vol. 69 of Appl. Math. Sci. (Springer, New York, 1988)
24. A.P. Dias, A. Rodrigues, Secondary bifurcations in systems with all-to-all coupling. ii. Dyn. Syst. **21**, 439 – 463 (2006)
25. N. Fenichel, Geometric singular perturbation theory for ordinary differential equations. J. Differ. Equ. **31**, 53–98 (1979)
26. P. Ashwin, J.W. Swift, The dynamics of n weakly coupled identical oscillators. J. Nonlinear Sci. **2**(1), 69–108 (1992)
27. A.J. Schwartz, A generalization of a Poincaré-Bendixson theorem to closed two-dimensional manifolds. Am. J. Math. **85**(3), 453 (1963)
28. S. Watanabe, S.H. Strogatz, Constants of motion for superconducting Josephson arrays. Physica D **74**(3–4), 197–253 (1994)
29. H. Daido, Onset of cooperative entrainment in limit-cycle oscillators with uniform all-to-all interactions: bifurcation of the order function. Physica D **91**(1–2), 24–66 (1996)
30. C. Bick, M. Timme, D. Paulikat, D. Rathlev, P. Ashwin, Chaos in symmetric phase oscillator networks. Phys. Rev. Lett. **107**(24), 244101 (2011)
31. L.P. Shilnikov, A case of the existence of a denumerable set of periodic motions. Soviet Math.-Doklady **6**, 163–166 (1965)
32. E.A. Grines, G.V. Osipov, Heteroclinic and Homoclinic Structures in the System of Four Identical Globally Coupled Phase Oscillators with Nonpairwise Interactions. Regul. Chaotic Dyn. **23**(7–8), 974–982 (2018)
33. O.E. Omel'chenko, The mathematics behind chimera states. Nonlinearity **31**(5), R121–R164 (2018)

34. J.M. Guckenheimer, P. Holmes, Structurally stable heteroclinic cycles. Math. Proc. Cambridge Philos. Soc. **103**(01), 189–192 (1988)
35. O. Weinberger, P. Ashwin, From coupled networks of systems to networks of states in phase space. Discrete Continuous Dyn. Syst. B **23**(5), 2043–2063 (2018)
36. L. Garrido-da Silva, S.B.S.D. Castro, Stability of quasi-simple heteroclinic cycles. Dyn. Syst. 1–26 (2018)
37. M. Komarov, A. Pikovsky, Finite-size-induced transitions to synchrony in oscillator ensembles with nonlinear global coupling. Phys. Rev. E **92**(2), 020901 (2015)
38. D. Călugăru, J.F. Totz, E.A. Martens, H. Engel, First-order synchronization transition in a large population of strongly coupled relaxation oscillators. Sci. Adv. **6**(39), eabb2637 (2020)
39. M. Rosenblum, A. Pikovsky, Nonlinear phase coupling functions: a numerical study. Philos. Trans. Roy. Soc. A **377**(2160), 20190093 (2019)

Chapter 8
Explosive Synchronization and Multistability in Large Systems of Kuramoto Oscillators with Higher-Order Interactions

Per Sebastian Skardal and Alex Arenas

Abstract Since its original formulation, the Kuramoto model and its many variants have served as critical tools for uncovering and understanding the emergence of nonlinear collective behavior. However, recent evidence suggests that in such phase-reduced systems, interactions beyond the typical pair-wise angle differences need to be considered to develop a full picture of the dynamics. In particular, *higher-order interactions*, namely non-additive, nonlinear interactions that take place between three or more oscillators are required. Here we explore these interactions and their effect on the macroscopic dynamics of coupled phase oscillator systems. The analysis for these systems begins with all-to-all coupled systems where a range of techniques including dimensionality reductions and self-consistency analyses may be employed. The effects of the various higher-order coupling terms on the macroscopic dynamics may then be explored, revealing a natural mechanism for nonlinear phenomena that includes abrupt (i.e., explosive) synchronization transitions and extensive multistability. These dynamics are qualitatively preserved under more heterogeneous network topologies. Moreover, the high degree of multistability in such networks allows for the system to store information and memory.

8.1 Introduction

The Kuramoto model of coupled phase oscillators has guided researchers interested in studying collective behavior in nonlinear dynamical systems for decades [1]. The classical formulation involves N all-to-all coupled phase oscillators with states θ_n, $n = 1 \ldots, N$, whose dynamics are governed by

P. S. Skardal (✉)
Department of Mathematics, Trinity College, 300 Summit St., Hartford, CT 06106, USA
e-mail: persebastian.skardal@trincoll.edu

A. Arenas
Department d'Enginyeria Informática i Matemátiques, Universitat Rovira i Virgili, Tarragona 43007, Spain
e-mail: alexandre.arenas@urv.cat

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_8

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \tag{8.1}$$

where ω_n is the *natural frequency* of oscillator n, K is the global coupling strength, and the sine of the respective phase angles represents the effect of oscillator m on oscillator n. The degree of synchronization of the system is typically measured using Kuramoto's order parameter,

$$z = re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_m}, \tag{8.2}$$

which gives a complex number that represents the centroid of all phases when placed appropriately on the complex unit circle. In particular, the magnitude $r = |z|$ takes values in between zero and one with $r \approx 0$ and $r \approx 1$ representing incoherent and synchronized states, respectively. In cases where multiple clusters emerge a collection of higher-order variants of the order parameters, also known as the Daido order parameters, are used to capture the macroscopic dynamics and are given by

$$z_q = r_q e^{i\psi_q} = \frac{1}{N}\sum_{j=1}^{N} e^{qi\theta_m}, \tag{8.3}$$

where q is a natural number and $q = 1$ recovers the classical Kuramoto order parameter given by Eq. (8.2). As we will see in Sect. 8.3, the collection of order parameters $z_1, \ldots, z_m$ are useful for characterizing the synchronization in a system with m clusters. In addition to the all-to-all coupled system described given in Eq. (8.1) the dynamics of the network-coupled analogue, given by

$$\dot{\theta}_i = \omega_i + K\sum_{j=1}^{N} A_{ij}\sin(\theta_j - \theta_i), \tag{8.4}$$

has been used to explore the many effects of network structure on collective behavior.

It is important to note that the original system given by Eq. (8.1) is the result of a phase-reduction from a system of globally-coupled limit cycle oscillators, where the state of a limit-cycle oscillator is assumed to be well-approximated by a single phase angle. Interestingly, recent work that aims to study these phase reductions beyond first-order contributions have identified higher-order interactions, i.e., non-additive, nonlinear interactions between three or more oscillators, as important components for describing the full system dynamics [2, 3]. These theoretical studies are complemented by empirical evidence in neuroscience that suggest higher-order interactions may play an important role in brain function [4–6]. In particular, three such higher-order interactions emerge and as an analogue to Kuramoto sine-coupling in Eq. (8.1) take the following forms:

$$\text{Type I}: \sum_{j=1}^{N}\sum_{l=1}^{N}\sin(2\theta_j-\theta_l-\theta_i), \tag{8.5}$$

$$\text{Type II}: \sum_{j=1}^{N}\sum_{l=1}^{N}\sin(\theta_j+\theta_l-2\theta_i), \tag{8.6}$$

$$\text{Type III}: \sum_{j=1}^{N}\sum_{l=1}^{N}\sum_{m=1}^{N}\sin(\theta_j+\theta_l-\theta_m-\theta_i), \tag{8.7}$$

where type I and II coupling naturally take place between sets of three oscillators (i.e., triangles or 2-simplexes) and type III coupling takes place between sets of four oscillators (i.e., tetrahedra or 3-simplexes). In the remainder of this chapter we explore the dynamics that emerge in phase oscillator systems as a result of incorporating these higher-order interaction terms, both together with the classical pairwise coupling term and on their own. In Sect. 8.2 we study a system that includes pairwise coupling along with type I and III coupling and show that the higher-order coupling terms give rise to nonlinear behavior that allows for abrupt synchronization transitions, hysteresis, and stable synchronized states even for negative pairwise-coupling. In Sect. 8.3 we study type II coupling in isolation and show that it leads to abrupt desynchronization transitions (without a complementary mechanism for synchronization) and an extensive multistability that allows the system to store memory. In Sect. 8.4 we provide some initial investigation into the effects of nontrivial network structures. Lastly, we briefly provide an outlook for future work in Sect. 8.5.

8.2 Coupling Types I and III: Abrupt Synchronization

When incorporating type I and III interactions into the classical Kuramoto model we obtain the following higher-order Kuramoto model originally studied in Ref. [7]:

$$\begin{aligned}\dot{\theta}_i = \omega_i &+ \frac{K_1}{N}\sum_{j=1}^{N}\sin(\theta_j-\theta_i) + \frac{K_2}{N^2}\sum_{j=1}^{N}\sum_{l=1}^{N}\sin(2\theta_j-\theta_l-\theta_i)\\ &+ \frac{K_3}{N^3}\sum_{j=1}^{N}\sum_{l=1}^{N}\sum_{m=1}^{N}\sin(\theta_j+\theta_l-\theta_m-\theta_i),\end{aligned} \tag{8.8}$$

where coupling strengths K_1, K_2, and K_3 measure the strength of interactions between pairs, triangles, and tetrahedra of oscillators (i.e., 1-simplexes, 2-simplexes, and 3-simplexes, respectively). The choice of coupling terms used in Eq. (8.8) conveniently allows us to solve exactly for the macroscopic system dynamics in the continuum limit $N \to \infty$ using the Ott-Antonsen ansatz [8–10]. In particular, we note that by defining

$$H = K_1 z + K_2 z_2 z^* + K_3 z^2 z^* \tag{8.9}$$

where $*$ denotes the complex conjugate (and recall that $z = z_1$), we may rewrite Eq. (8.8) as

$$\dot{\theta}_i = \omega_i + \frac{1}{2i}\left(He^{-i\theta_i} - H^* e^{i\theta_i}\right). \tag{8.10}$$

Next, in the continuum limit we may describe the state of the system using the density function $f(\theta, \omega, t)$, where $f(\theta, \omega, t)d\theta d\omega$ gives the fraction of oscillators with phase in $[\theta, \theta + d\theta)$ and natural frequency in $[\omega, \omega + d\omega)$ at time t. The conservation of oscillators and the static natural frequencies implies that f satisfies the continuity equation $0 = \frac{\partial f}{\partial t} + \frac{\partial}{\partial \theta}\left(f\dot{\theta}\right)$, i.e.,

$$0 = \frac{\partial f}{\partial t} + \frac{\partial}{\partial \theta}\left\{f\left[\omega_i + \frac{1}{2i}\left(He^{-i\theta_i} - H^* e^{i\theta_i}\right)\right]\right\}. \tag{8.11}$$

Since θ-domain of f is the circle it is natural to consider the Fourier expansion of f as

$$f(\theta, \omega, t) = \frac{g(\omega)}{2\pi}\left[1 + \sum_{n=1}^{\infty} \hat{f}_n(\omega, t)e^{in\theta} + \sum_{n=1}^{\infty} \hat{f}_n^*(\omega, t)e^{-in\theta}\right]. \tag{8.12}$$

Remarkably, Ott and Antonsen discovered that the choice of geometrically-decaying Fourier coefficients, i.e., $\hat{f}_n(\omega, t) = \alpha^n(\omega, t)$ where α is assumed to be an analytic function in the ω-complex plane, drastically reduces the system dynamics. In particular, inserting this choice into Eq. (8.12) and then into Eq. (8.11) yields a single ordinary differential equation for α, in this case given by

$$\dot{\alpha} = -i\omega\alpha + \frac{1}{2}\left(H^* - H\alpha^2\right). \tag{8.13}$$

It is worth noting that this family of solutions with geometrically-decaying Fourier coefficients yields density functions $f(\theta, \omega, t)$ that are Poison kernels and describes a stable manifold to which all solutions tend [9, 10]. To close the system dynamics we note that the order parameter satisfies

$$z^* = \iint f(\theta, \omega, t)e^{i\theta} d\theta d\omega = \int \alpha(\omega, t)g(\omega)d\omega. \tag{8.14}$$

A convenient choice for the frequency distribution g is the Lorentzian with mean ω_0 and width Δ, i.e., $g(\omega) = \Delta/\pi[\Delta^2 + (\omega - \omega_0)^2]$. The integral in Eq. (8.14) can then be evaluated by closing the contour with the infinite-radius semi-circle in the negative-half complex plane and using Cauchy's integral theorem, yielding $z^* = \alpha(\omega_0 - i\Delta, t)$. Similarly, for the order parameter z_2 we have that $z_2^* =$

$\alpha^2(\omega_0 - i\Delta) = z^{*2}$, and by evaluating Eq. (8.13) at $\omega = \omega_0 - i\Delta$ and taking a complex conjugate we find

$$\dot{z} = -\Delta z + i\omega_0 z + \frac{1}{2}\left[\left(K_1 z + K_{2+3} z^2 z^*\right) - \left(K_1 z^* + K_{2+3} z^{*2} z\right) z^2\right], \tag{8.15}$$

where we have defined the combined higher-order coupling strength $K_{2+3} = K_2 + K_3$. Finally, it is convenient to decompose the complex dynamics given in Eq. (8.15) into amplitude and angle components, yielding

$$\dot{r} = -\Delta r + \frac{K_1}{2} r(1 - r^2) + \frac{K_{2+3}}{2} r^3(1 - r^2), \tag{8.16}$$

$$\dot{\psi} = \omega_0. \tag{8.17}$$

Equations (8.16) and (8.17) close the dynamics for the macroscopic system dynamics and allow us to now perform a bifurcation and stability analysis of the synchronization dynamics of the higher-order Kuramoto model in the continuum limit. Unsurprisingly, the amplitude and angle dynamics decoupling, leaving states to rotate with a constant frequency equal to the mean natural frequency of the system. Inspecting Eq. (8.16) more closely, we also observe that the contribution of the higher-order interaction terms only contribute higher-order nonlinear terms implying that they have no effect on the stability of the incoherent state $r = 0$. In fact, $r = 0$ is always a steady-state solution to Eq. (8.16) and is stable for $K_1 < 2\Delta$, and loses stability at the critical value $K_1 = 2\Delta$ in a pitchfork bifurcation that may be supercritical or subcritical, depending on the amplification of the nonlinear terms by K_{2+3}. The nature of this bifurcation is most easily seen by solving for synchronized steady-states, namely, $r > 0$ and $\dot{r} = 0$, yielding

$$r = \sqrt{\frac{K_{2+3} - K_1 \pm \sqrt{(K_1 + K_{2+3})^2 - 8\Delta K_{2+3}}}{2K_{2+3}}}. \tag{8.18}$$

The transition between incoherence and synchronization can be identified by investigating how this synchronized branch meets the incoherent branch $r = 0$ in particular, as r tends towards zero K_1 tends towards 2Δ regardless of K_{2+3}, but for $K_{2+3} \leq 2\Delta$ and $K_{2+3} > 2\Delta$ it tends towards 2Δ from above and below, respectively, indicating a supercritical to subcritical shift. This can be seen most easily in Fig. 8.1a where we plot the synchronized branch as a function of K_1 for K_{2+3} values 0, 2, 5, 8, and 10 (blue ranging to red) and $\Delta = 1$. The curves represent the exact prediction given in Eq. (8.18) while circles correspond to simulations of $N = 10^4$ oscillators. In particular, the synchronized branch begins to fold over itself at $K_{2+3} = 2\Delta$, indicating supercritical and subcritical pitchfork bifurcations for $K_{2+3} \leq 2\Delta$ and $K_{2+3} > 2\Delta$, respectively. A straight forward analysis reveals that, when the synchronized branch folds over itself the plus and minus signs in Eq. (8.18) give stable and unstable solutions, respectively, which are depicted in solid and dashed curves, respectively.

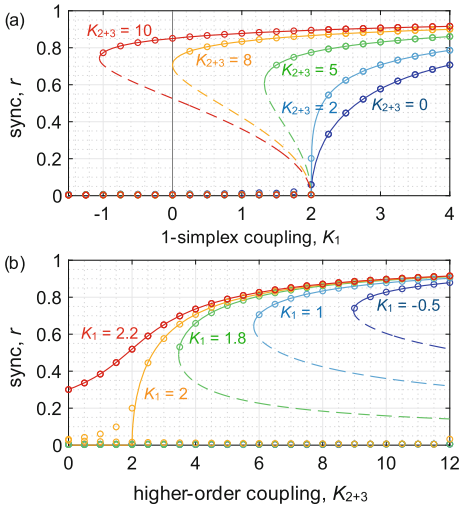


Fig. 8.1 Abrupt synchronization in the higher-order Kuramoto model. Synchronization profiles describing the macroscopic system state: the order parameter r **a** as a function of 1-simplex coupling K_1 for higher-order coupling $K_{2+3} = 0$, 2, 5, 8, and 10 (blue to red) and **b** as a function of higher-order coupling K_{2+3} for 1-simplex coupling $K_1 = -0.5$, 1, 1.8, 2, and 2.2. Solid and dashed curves represent stable and unstable solutions given by Eq. (8.18), respectively, and circles denote results taken from direct simulations of equation (8.8) with $N = 10^4$ oscillators with $\Delta = 1$ and $\omega_0 = 0$. (Modified and adapted from Ref. [7])

To complement Fig. 8.1a we plot in Fig. 8.1b the synchronized branch as a function of K_{2+3} for K_1 values -0.5, 1, 1.8, 2, and 2.2 (blue ranging to red) and $\Delta = 1$.

This reveals, for $K_{2+3} > 2\Delta$, a region of bistability between the incoherent and synchronized states in the form of a hysteresis loop bounded by two critical values. The first of these critical values is the *synchronization* value $K_1^{\text{sync}} = 2\Delta$ where the incoherent state loses stability and the system undergoes an abrupt transition to a synchronized state. The second critical value occurs at a smaller value of K_1 located at the left-most point of the synchronized branch, at which point, as K_1 is decreased, the synchronized branch annihilates in a saddle-node bifurcation and the system undergoes an abrupt transition to the incoherent state. Solving for this second *desynchronization* value, we obtain $K_1^{\text{desync}} = 2\sqrt{2\Delta K_{2+3}} - K_{2+3}$, which collides with $K_1^{\text{sync}} = 2\Delta$ at $K_{2+3} = 2\Delta$, corresponding to the codimension-two point where the pitchfork bifurcation transitions between subcritical and supercritical. Setting $\Delta = 1$ (which can be done without loss of generality by rescaling time and the parameters K_1, K_{2+3} and ω_0) we plot in Fig. 8.2 the stability diagram for the higher-order Kuramoto model indicating incoherent, synchronized, and bistable regions between pitchfork and saddle-node curves (blue and red, respectively) and the codiension-two point at $(K_1, K_{2+3}) = (2, 2)$. Interestingly, as the saddle-node curve crosses into the negative-half plane $K_1 < 0$ the system is able to support synchronized states even for negative pairwise coupling, as was observed in Fig. 8.1a.

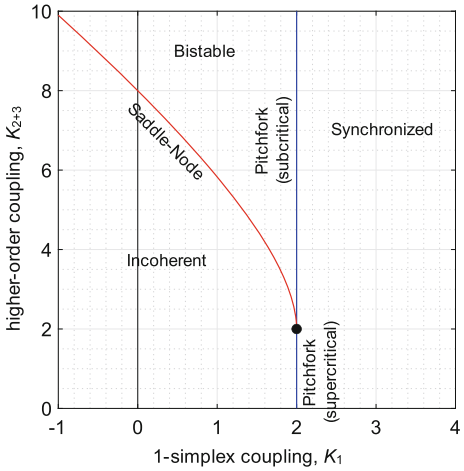


Fig. 8.2 Stability diagram for the higher-order Kuramoto model. The full stability diagram describing incoherent, synchronized, and bistable states as a function of 1-simplex coupling K_1 and higher-order coupling K_{2+3} for $\Delta = 1$. Blue and red curves correspond to pitchfork and saddle-node bifurcations, which collide at a codimension-two point (black circle) at $(K_1, K_{2+3}) = (2, 2)$. For $K_{2+3} < 2$ and $K_{2+3} > 2$ the pitchfork bifurcation is supercritical and subcritical, respectively. (Modified and adapted from Ref. [7]

8.3 Coupling Type II: Abrupt Desynchronization and Multistability

Next we investigate the dynamics in the presence of only type II interactions, yielding the following model originally studied in Ref. [11]:

$$\dot{\theta}_i = \omega_i + \frac{K}{N^2} \sum_{j=1}^{N} \sum_{l=1}^{N} \sin\left(\theta_j + \theta_l - 2\theta_i\right). \tag{8.19}$$

As we will see in what follows, the collective dynamics under type II interactions deserve special attention for two reasons. First, the arrangement of terms in the coupling function (i.e., compared to that in type I coupling) makes complicates the analysis in that the Ott-Antonsen dimensionality reduction does not fully solve the system–instead they provide a partial dimensionality reduction, beyond which a self-consistency approach is required. Second, the effects of type II coupling on the collective dynamics differ significantly from those explored above under type I and III interactions. One such effect is the emergence of cluster synchronization, where oscillators become entrained in two different clusters on the torus centered at opposite angles. We note here that by entering a suitable rotating frame we may set the mean natural frequency to zero, ensuring non-rotating entrained solutions, and

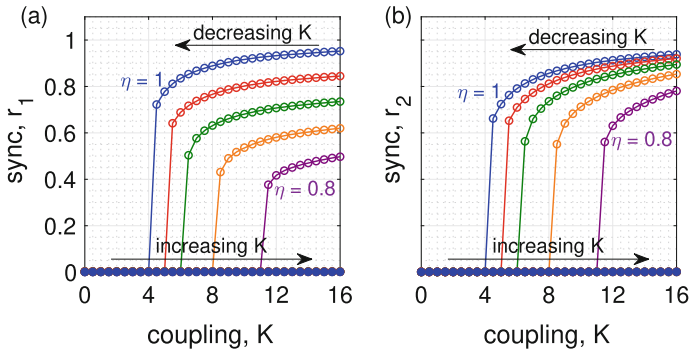


Fig. 8.3 Collective dynamics under type II coupling. The order parameters **a** r_1 and **b** r_2 as a function of the coupling strength K for various asymmetry values η. Blue, red, green, orange, and purple circles represent results obtained from direct simulations of Eq. (8.19) with $N = 10^5$ oscillators with natural frequencies drawn from a Lorentzian with $\omega_0 = 0$ and $\Delta = 1$ for $\eta = 1$, 0.95, 0.9, 0.85, and 0.8, respectively (Modified and adapted from Ref. [11])

moreover a shift in initial conditions ensures that the largest cluster (i.e., that with the largest fraction of entrained oscillators) is centered at $\theta = 0$ with the smaller cluster centered opposite at $\theta = \pi$.

The formation of these two clusters and the particular form of coupling then lead to unexpected nonlinear behaviors which we illustrate now with numerical simulations of $N = 10^5$ oscillators with natural frequencies drawn from a Lorentzian distribution with mean $\omega_0 = 0$ and width $\Delta = 1$. We then introduce an asymmetry parameter η that describes the initial conditions: at time $t = 0$ we start with a randomly chosen fraction of η of all oscillators at phase $\theta = 0$, while a fraction $(1 - \eta)$ of the oscillators start at $\theta = \pi$. We then simulate the system dynamics by adiabatically decreasing the coupling strength from $K = 16$ to 0, then adiabatically increase it again until we reach $K = 16$ again. The results are shown in Fig. 8.3, where the values of r_1 and r_2 are plotted in panels (a) and (b), respectively, as a function of the coupling strength K, for asymmetry parameters $\eta = 1$ (blue), 0.95 (red), 0.9 (green), 0.85 (orange), and 0.8 (purple). As seen in both plots the system begins in a synchronized state and, as K is decreased, the degree of synchronization decreases continuously until, at a critical value of K that depends on the asymmetry η, the system undergoes an abrupt desynchronization transition. The state then remains incoherent until K reaches zero. Then, as K is increased and restored to 16 the system surprisingly remains incoherent, i.e., there is no synchronization transition (abrupt or continuous) that complements the abrupt desynchronization transition. Moreover, further numerical simulations (not shown) and the analysis presented below confirm that no such synchronization transitions exist above $K = 16$. In addition to these abrupt desynchronization transitions, the simulations above show that, for a given coupling strength, many stable entrained states are possible, depending on the asymmetry η. In fact, in the thermodynamic limit of $N \to \infty$ a continuum of entrained states may be stable, indicating an extensive multistability.

We now present an analysis to explain these dynamics. First, we use the definition of the order parameter to rewrite Eq. (8.19) as

$$\dot{\theta}_i = \omega_i + K r_1^2 \sin[2(\psi_1 - \theta_i)]. \tag{8.20}$$

As is the previous section, we consider the continuum limit $N \to \infty$ where the state of the system can be described by a density function $f(\theta, \omega, t)$ that satisfies the same continuity equation $0 = \partial_t f + \partial_\theta (f\dot{\theta})$. In contrast to the case with type I and III interactions, here we must consider the symmetric and asymmetric parts $f_s(\theta, \omega, t)$ and $f_a(\theta, \omega, t)$, respectively, of $f(\theta, \omega, t)$ which satisfy

$$f(\theta, \omega, t) = f_s(\theta, \omega, t) + f_a(\theta, \omega, t), \tag{8.21}$$

with symmetries

$$f_s(\theta + \pi, \omega, t) = f_s(\theta, \omega, t), \quad \text{and} \quad f_a(\theta + \pi, \omega, t) = -f_a(\theta, \omega, t). \tag{8.22}$$

Note that the linearity of the continuity equation implies that if both f_s and f_a are solutions, then so is f. While the asymmetric part f_a may not be simplified by dimensionality reduction, we may simplify the symmetric part f_s. Noting that the Fourier series of f_s is given by the even terms of the Fourier series of f, i.e.,

$$f_s(\theta, \omega, t) = \frac{g(\omega)}{2\pi}\left[1 + \sum_{m=1}^{\infty} \hat{f}_{2m}(\omega, t) e^{2im\theta} + c.c.\right], \tag{8.23}$$

we propose an ansatz similar to that above where these even Fourier coefficients decay geometrically, i.e., $\hat{f}_{2m}(\omega, t) = a^m(\omega, t)$. Inserting this and Eq. (8.23) and then into the continuity equation, we find that the symmetric dynamics collapse onto the same low-dimensional manifold characterized by the condition

$$\partial_t a = -2i\omega a + K\left(z_1^{*2} - z_1^2 a^2\right). \tag{8.24}$$

As was done above, the symmetric dynamics may be closed by expressing $a(\omega, t)$ as an order parameter as follows. In particular, assuming again that that the frequency distribution $g(\omega)$ is Lorentzian, $g(\omega) = \Delta/\{\pi\left[(\omega - \omega_0)^2 + \Delta^2\right]\}$ a similar technique as that used in the previous section below Eq. (8.15) results in $z_2 = a^*(\omega_0 - i\Delta, t)$. We then evaluate Eq. (8.24) at $\omega = \omega_0 - i\Delta$ to obtain

$$\dot{z}_2 = 2i\omega_0 z_2 - 2\Delta z_2 + K\left(z_1^2 - z_1^{*2} z_2^2\right), \tag{8.25}$$

or, in polar form,

$$\dot{r}_2 = -2\Delta r_2 + K r_1^2 (1 - r_2^2)\cos(2\psi_1 - \psi_2), \tag{8.26}$$

$$\dot{\psi}_2 = 2\omega_0 + Kr_1^2 \frac{1+r_2^2}{r_2} \sin(2\psi_1 - \psi_2). \tag{8.27}$$

Equations (8.26) and (8.27) describe the symmetric dynamics captured by the order parameter z_2. However, these equations do not capture the asymmetric part of the dynamics, and moreover they depends on the asymmetric dynamics via z_2's dependence on z_1. To complete the analysis we next need a self consistency analysis. Recalling our assumptions where $\omega_0 = 0$ and clusters set at $\theta = 0$ and π, effectively setting the order parameter phases at $\psi_1, \psi_2 = 0$, we revisit Equation (8.20) which implies that oscillators that become phase-locked satisfy $|\omega_i| \leq Kr_1^2$, in which case they relax to one of the two stable fixed points $\theta_i = \theta^*(\omega_i)$ or $\theta^*(\omega_i) + \pi$, where

$$\theta^*(\omega) = \arcsin(\omega/Kr_1^2)/2. \tag{8.28}$$

These two fixed points correspond to the two clusters to which that the phase-locked oscillators become entrained – specifically, phase-locked oscillators starting near $\theta = 0$ or π will end up at the fixed points $\theta^*(\omega)$ or $\theta^*(\omega) + \pi$, respectively. The phase-locked population is described by the density function

$$f_{\text{locked}}(\theta, \omega) = \eta\delta(\theta - \theta^*(\omega)) + (1-\eta)\delta(\theta - \theta^*(\omega) - \pi), \tag{8.29}$$

where the asymmetry parameter η appears and describes the fraction of phase-locked oscillators in the $\theta = 0$ cluster. On the other hand, oscillators satisfying $|\omega_i| > Kr_1^2$ drift for all time and relax to the following stationary distribution

$$f_{\text{drift}}(\theta, \omega) = \frac{\sqrt{\omega^2 - K^2 r_1^4}}{2\pi\left[\omega + Kr_1^2 \sin(2\psi_1 - 2\theta)\right]}. \tag{8.30}$$

Next, the order parameter z_1 is given by the integral $z_1 = \iint f(\theta, \omega, t) e^{i\theta} d\theta d\omega$, which after inserting the density f as defined by Eqs. (8.29) and (8.30) reduces to

$$r_1 = (2\eta - 1) \int_{-Kr_1^2}^{Kr_1^2} \sqrt{\frac{1 + \sqrt{1 - (\omega/Kr_1^2)^2}}{2}} g(\omega) d\omega, \tag{8.31}$$

where the contribution from the drifting oscillators vanishes due to the symmetry of f_{drift}. Returning to r_2, Eq. (8.26) implies that at steady state we have

$$r_2 = \frac{-\Delta + \sqrt{\Delta + K^2 r_1^4}}{Kr_1^2}. \tag{8.32}$$

Thus, the macroscopic steady-state is described by Eqs. (8.31) and (8.32).

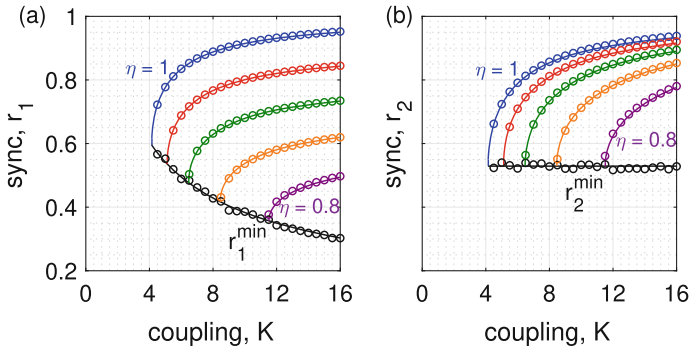


Fig. 8.4 Synchronized states with type II coupling. For order parameters **a** r_1 and **b** r_2, the synchronization branches predicted by Eqs. (8.31) and (8.32) plotted as solid curves and results from simulations plotted in circles. Results correspond to asymmetries $\eta = 1$, 0.95, 0.9, 0.85, and 0.8, as well as the minimum branches [predictions for which are given in Eq. (8.35)]. Simulated results use $N = 10^5$ oscillators (Modified and adapted from Ref. [11])

In Fig. 8.4 we overlay the analytical predictions of Eqs. (8.31) and (8.32) (solid curves) with the synchronized states observed in simulations (circles) for the same asymmetry parameter values η used in Fig. 8.3, noting excellent agreement. [Equation (8.31) was solved numerically to find r_1.] We also investigate the smallest r_1 and r_2 supported by the system by, for each K, decreasing η until the synchronized state is lost. These simulation results are plotted in black circles. Interestingly, while these $r_1^{\min}$ values decrease as K increases, $r_2^{\min}$ appears to remain roughly constant. In fact, we now use this surprising property to uncover the critical desynchronization coupling strength as a function of the asymmetry parameter, i.e., $K_c(\eta)$, or equivalently, the minimum asymmetry parameter that supports a synchronized state as a function of coupling, i.e., $\eta_{\min}(K)$. We begin by first inverting Eq. (8.32), obtaining $Kr_1^2 = 2\Delta r_2/(1 - r_2^2)$, which can be inserted into Eq. (8.31), yielding

$$\sqrt{\frac{2\Delta r_2}{1-r_2^2}} = \sqrt{K}(2\eta - 1) \int_{-2\Delta r_2/(1-r_2^2)}^{2\Delta r_2/(1-r_2^2)} \sqrt{\frac{1 + \sqrt{1 - \left[\omega(1-r_2^2)/2\Delta r_2\right]^2}}{2}}\, g(\omega) d\omega. \tag{8.33}$$

While Eq. (8.11) appears more complicated than Eq. (8.9), we note that the coupling strength K has been entirely scaled out of the integral, appearing outside with $(2\eta - 1)$. We therefore conclude that if the quantities $\sqrt{K}$ and $2\eta - 1$ cancel one another out, i.e., $\sqrt{K}(2\eta - 1)$ is constant, it follows that the solution r_2 in Eq. (8.11) is independent of K. We therefore ansatz that $\sqrt{K}(2\eta - 1) = \text{const.}$ and use the initial condition $\eta_{\min}(K_c(1)) = 1$, where $K_c(1)$ denotes the very first coupling strength where a synchronized state is possible with $\eta = 1$, yielding

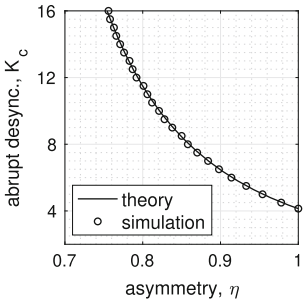


Fig. 8.5 Abrupt desynchronization transition. The critical coupling strength K_c at which the abrupt desynchronization transition occurs as a function of the asymmetry η. The solid curve represents the theoretical prediction given by Eq. (8.36) and black circles represent observations from direct simulation with $N = 10^5$ oscillators (Modified and adapted from Ref. [11])

$$\eta_{\min}(K) = \frac{\sqrt{K_c(1)}}{2\sqrt{K}} + \frac{1}{2}. \tag{8.34}$$

Equation (8.34) implies that along the minimum branch we have that $\sqrt{K}(2\eta - 1) = \sqrt{K_c(1)} \approx 2.034$, which can be used in Eq. (8.32) to compute the minimum branch of r_2, and in turn r_1 via Eq. (8.31), yielding

$$r_1^{\min}(K) \approx \frac{1.2120}{\sqrt{K}} \quad \text{and} \quad r_2^{\min}(K) \approx 0.5290. \tag{8.35}$$

These predictions are plotted as black curves in Fig. 8.4. Lastly, by inverting Eq. (8.34) we find the critical coupling strength K_c as a function of η where the abrupt desynchronization transition occurs, namely

$$K_c(\eta) \approx \frac{4.137}{(2\eta - 1)^2}. \tag{8.36}$$

In Fig. 8.5 we plot the theoretical prediction of the abrupt desynchronization point $K_c(\eta)$ as a solid curve vs observations from direct simulations as black circles, noting excellent agreement. Above this curve we observe extensive multistability and below the curve only the incoherent state is stable. Some additional rigorous results that agree and complement these may be found in Refs. [12, 13].

8.4 Effects of Network Structure

Having presented results on synchronization in the presence of higher-order interactions in globally-coupled systems, we now turn our attention to the effects on network structure. Here we focus on the network analogue of the higher-order Kuramoto model presented in Sect. 8.2. In the presence of a non-trivial network topology the system dynamics are described by

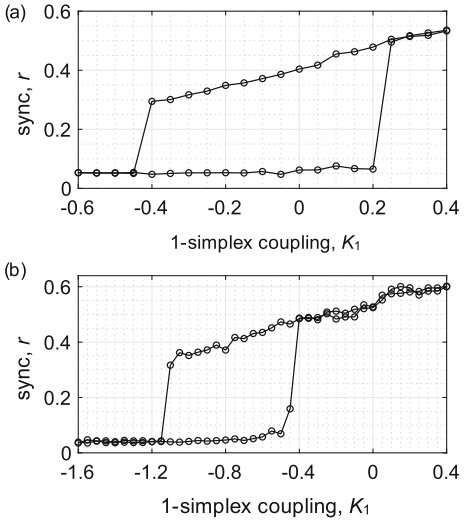


Fig. 8.6 Higher-order Kuramoto model: Real networks. The order parameter r as a function of 1-simplex coupling K_1 for the network analogue of the higher-order Kuramoto model for **a** the Macaque brain network and **b** the UK power grid. In each case K_1 is adiabatically increased then decreased to highlight the bistable region. Higher-order coupling is given by **a** $K_2 = 1.6$ and $K_3 = 1.1$ and **b** $K_2 = 2.2$ and $K_3 = 3.3$ (Modified and adapted from Ref. [7])

$$\dot{\theta}_i = \omega_i + \frac{K_1}{\langle k^1 \rangle} \sum_{j=1}^{N} A_{ij} \sin(\theta_j - \theta_i) + \frac{K_2}{2\langle k^2 \rangle} \sum_{j=1}^{N} \sum_{l=1}^{N} B_{ijl} \sin(2\theta_j - \theta_l - \theta_i)$$
$$+ \frac{K_3}{6\langle k^3 \rangle} \sum_{j=1}^{N} \sum_{l=1}^{N} \sum_{m=1}^{N} C_{ijlm} \sin(\theta_j + \theta_l - \theta_m - \theta_i), \tag{8.37}$$

where all dynamical variables and parameters are the same as in Sect. 8.2 except for those that are network-dependent. In particular, he network structure (assumed to be undirected and unweighted) is encoded in the 1-simplex adjacency matrix A, 2-simplex adjacency tensor B, and 3-simplex adjacency tensor C, where $A_{ij} = 1$ if nodes i and j are connected by a link (and otherwise $A_{ij} = 0$), $B_{ijl} = 1$ if nodes i, j, and l belong to a common 2-simplex (and otherwise $B_{ijl} = 0$), and $C_{ijlm} = 1$ if nodes i, j, l, and m belong to a common 3-simplex (and otherwise $C_{ijlm} = 0$). For each node i we denote the q-simplex degree k_i^q as the number of distinct q-simplexes node i is a part of, and $\langle k^q \rangle$ is the mean q-simplex degree across the network. Importantly, the division of each coupling strength by the respective mean degree $\langle k^q \rangle$ in Eq. (8.37) simply amounts to a rescaling of the respective coupling strength, and moreover ensures that the mean-field approximation of the dynamics of Eq. (8.37) are given precisely by Eq. (8.8).

We begin with direct simulations of Eq. (8.37) on two real-world networks: the Macaque brain network [14] and the UK power grid [15]. In the Macaque brain

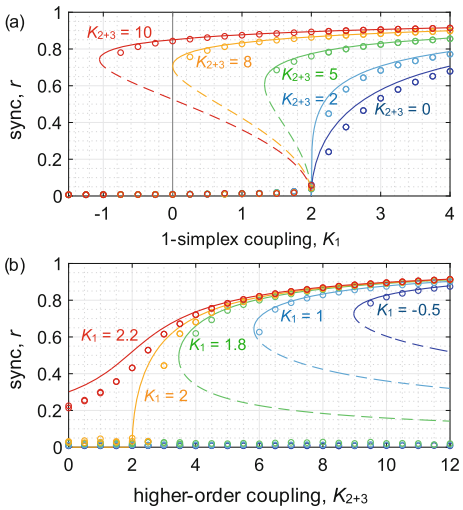


Fig. 8.7 Higher-order Kuramoto model: Multiplex model. The order parameter r **a** as a function of 1-simplex coupling K_1 for higher-order coupling $K_{2+3} = 0$, 2, 5, 8, and 10 (blue to red) and **b** as a function of higher-order coupling K_{2+3} for 1-simplex coupling $K_1 = -0.5$, 1, 1.8, 2, and 2.2. Circles represent direct simulations on a network of $N = 10^4$ nodes with mean degrees $\langle k^1 \rangle = \langle k^2 \rangle = \langle k^3 \rangle = 30$ and solid and dashed curves represent stable and unstable solutions of the mean field approximation. (Modified and adapted from Ref. [7])

network we infer higher-order connection directly as triplets (i, j, l) and quadruplets (i, j, l, m) that are fully-connected cliques. Using natural frequencies drawn from the standard normal distribution, we set higher-order coupling strengths $K_2 = 1.6$ and $K_3 = 1.1$ and adiabatically increase and decrease K_1 and plot the order parameter r as a function of K_1 in Fig. 8.6a. For the UK power grid we follow a similar protocol with $K_2 = 2.2$ and $K_3 = 3.3$, plotting the results in Fig. 8.6b, but given the geographical nature of the network, we incorporate a triplet or quadruplet interaction for any connected three- or four-path in the network. Importantly, in these real-world topologies we see qualitatively similar dynamics as in the globally coupled system presented in Sect. 8.2. We note, however, that while the results are qualitatively similar, more quantitative predictions of the dynamics remain inaccurate.

Next we turn our attention to a generative network model, in particular, a multiplex model where interactions of different order are uncorrelated. Specifically, setting target mean degrees $\langle k^1 \rangle$, $\langle k^2 \rangle$, and $\langle k^3 \rangle$, we place at random $M_1 = N\langle k^1 \rangle/2$ 1-simplexes (i.e., links), $M_2 = N\langle k^1 \rangle/3$ 2-simplexes (i.e., filled triangles), and $M_3 = N\langle k^3 \rangle/4$ 3-simplexes (i.e., filled tetrahedra). Note that in this model the interactions of different orders are uncorrelated, so that, for example, a pair of nodes may be part of a triangle even if there exists no link between them and vice versa. Thus, this model may be thought of as a three-layer multiplex with 1-, 2-, and 3-simplexes belonging to respective layers. In Fig. 8.7 we plot the results of the dynamics of such a multiplex model of $N = 10^4$ oscillators with mean degrees $\langle k^1 \rangle = \langle k^2 \rangle = \langle k^3 \rangle = 30$,

plotting r as a function of K_1 for $K_{2+3} = 0$, 2, 5, 8, and 10 in panel (a) and as a function of K_{2+3} for $K_1 = -0.5$, 1, 1.8, 2, and 2.2 in panel (b). simulation results are plotted in circles with solid curves representing the prediction from the mean-field approximation, i.e., Eq. (8.18), showing strong agreement.

8.5 Outlook and Future Work

We now close with a brief outlook to future work. As we have seen in this chapter, the presence of higher-order interactions in systems of coupled oscillators introduces many new dynamics effects in synchronization dynamics. Even in the globally-coupled case there are many avenues to explore that promise rich dynamics. Examples include (but are certainly not limited to) the incorporation of time delays, phase lags, positive/negative coupling, external forcing—all of which have profound effects on the collective dynamics in the absence of higher-order interactions. For the higher-order Kuramoto model studied in Sect. 8.2 the viability of the Ott-Antonsen ansatz promises to be very useful, but the technical peculiarities of type II coupling complicate the analysis of this future work.

Further progress in the case of nontrivial network topologies is also needed. One such case would be the presence of modular structure in networks and how it affects the dynamics explored above. More generally, however, the generic effects of nontrivial network structures are poorly understood. In Sect. 8.4 we saw that while the real networks displayed qualitatively similar dynamics as the globally-coupled system, quantitative predictions were poor. On the other hand, the dynamics of the multiplex model were well-captured by the mean field analysis, suggesting that correlations between interactions of different orders plays in important role. Another broad question lies in the effect of network (i.e., degree) heterogeneity and its affect on collective dynamics, most importantly the abrupt transitions.

Lastly, the multistability induced by type II coupling studied in Sect. 8.3 deserves special mention. This phenomenon allows an oscillator system to store information and memory by treating each oscillator as a bit: 0 or 1 if the oscillator is in the $\theta = 0$ or π cluster, respectively. Relatively small perturbations and parameter modifications then allow the system to represent a different string, i.e., piece of information. In the presence of a network topology these effects remain, although the supported (i.e., stable) states changes. Preliminary work [16] shows that sparser networks support fewer states, but the overall effect of network structure remains poorly understood.

References

1. Y. Kuramoto, Self-entrainment of a population of coupled nonlinear oscillators, in *International Symposium on Mathematical Problems in Theoretical Physics* (Springer, Heidelberg, 1975), pp. 420–422

2. P. Ashwin, A. Rodrigues, Hopf normal form with S_N symmetry and reduction to systems of nonlinearly coupled phase oscillators. Phys. D **325**, 14 (2016)
3. I. Léon, D. Pazó, Phase reduction beyond the first order: the case of the mean-field complex Ginzburg-Landau equation. Phys. Rev. E **100**, 012211 (2019)
4. G. Petri et al., Homological scaffolds of brain functional networks. J. R. Soc. Interface **11**, 20140873 (2014)
5. C. Giusti, R. Ghrist, D.S. Bassett, Two's company, three (or more) is a simplex. J. Comput. Neurosci. **41**, 1 (2016)
6. M.W. Reimann et al., Cliques of neurons bound into cavities provide a missing link between structure and function. Frontiers Comp. Neuro. **11**, 48 (2017)
7. P.S. Skardal, A. Arenas, Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching. Comm. Phys. **3**, 218 (2020)
8. E. Ott, T.M. Antonsen, Low dimensional behavior of large systems of globally coupled oscillators. Chaos **18**, 037113 (2008)
9. E. Ott, T.M. Antonsen, Long time evolution of phase oscillator systems. Chaos **19**, 023117 (2009)
10. J. Engelbrecht, R. Mirollo, Is the Ott-Antonsen manifold attracting? Phys. Rev. Res. **2**, 023057 (2020)
11. P.S. Skardal, A. Arenas, Abrupt desynchronization and extensive metastability in globally coupled oscillator simplexes. Phys. Rev. Lett. **122**, 248301 (2019)
12. C. Xu, X. Wang, P.S. Skardal, Bifurcation analysis and structural stability of simplicial oscillator populations. Phys. Rev. Res. **2**, 023281 (2020)
13. C. Xu, P.S. Skardal, Spectrum of extensive multiclusters in the Kuramoto model with higher-order interactions. Phys. Rev. Res. **3**, 013013 (2021)
14. K. Amunts et al., BigBrain: an ultrahigh-resolution 3D human brain model. Science **340**, 1472–1475 (2013)
15. M. Rohden, A. Sorge, M. Timme, D. Witthaut, Self-organized synchronization in decentralized power grids. Phys. Rev. Lett. **109**, 064101 (2012)
16. P.S. Skardal, A. Arenas, Memory selection and information switching in oscillator networks with higher-order interactions. J. Phys. Complex. **2**, 015003 (2020)

Chapter 9
Multiorder Laplacian for Kuramoto Dynamics with Higher-Order Interactions

Maxime Lucas, Giulia Cencetti, and Federico Battiston

Abstract Many real-world systems are characterised by higher-order interactions, where influences among units involve more than two nodes at a time, and which can significantly affect the emergence of collective behaviors. A paradigmatic case is that of synchronization, occuring when oscillators reach coherent dynamics through their mutual couplings, and which is known to display richer collective phenomena when connections are not limited to simple dyads. Here, we consider an extension of the Kuramoto model with higher-order interactions, where oscillators can interact in groups of any size, arranged in any arbitrary complex topology. We present a new operator, the multiorder Laplacian, which allows us to treat the system analytically and that can be used to assess the stability of synchronization in general higher-order networks. Our spectral approach, originally devised for Kuramoto dynamics, can be extended to a wider class of dynamical processes beyond pairwise interactions, advancing our quantitative understanding of how higher-order interactions impact network dynamics.

9.1 Introduction

Synchronization, i.e. the emergence of order in populations of oscillators, is ubiquitous in nature [1, 2], from the brain [3] to circadian rhythms [4] and the cardio-

This chapter is built on the exposition of [27], in particular with respect to the theoretical formalism. Our presentation here follows a different order and has a more didactic style. We also added new examples such as the discussion of synchronization in random simplicial complexes.

M. Lucas (✉)
Aix-Marseille University, Marseille, France
e-mail: maxime.lucas@isi.it

G. Cencetti
Bruno Kessler Foundation, Trento, Italy
e-mail: gcencetti@fbk.eu

F. Battiston
Department of Network and Data Science, Central European University, Vienna, Austria
e-mail: battistonf@ceu.edu

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_9

vascular system [5, 6]. Over the years network science has proven crucial to understand how the emergence of collective behavior is governed by the underlying network of interactions [7, 8], from full synchrony to more complex phenomena such as the appearance of explosive [9, 10] and cluster synchronization [11]. In networks the systems units are represented by nodes and their interactions encoded in links. For this reason, couplings and feedbacks among the different agents are intrinsically limited to dyadic relationship. Yet, in many real-world social and biological systems interactions take place among three or more units at a time, highlighting a need for different mathematical frameworks.

Networks beyond pairwise interactions have recently been the focus on great attention in the literature [12]. Higher-order interactions can significantly impact the emergent behaviors of diverse dynamical processes [13], including diffusion [14], social contagion [15], ecological [16] and cooperative [17] dynamics. For systems of oscillators, non-dyadic couplings were shown to affect synchronization, for instance promoting explosive transitions [18–20] and cluster synchronization [21, 22]. Higher-order interactions terms were also shown to arise naturally from higher-order phase reduction [23, 24], and were considered in the inverse approach [25, 26]. Yet, a general analytical framework for synchronization in higher-order networks has long been missed.

In this chapter, following the ideas originally proposed in [27], we introduce a novel operator which allows us to characterize the stability of synchronization in presence of higher-order interactions of any order and arbitrary complex topology. To this scope, we present a generalization of the Kuramoto model of identical phase oscillators on hypergraphs. Traditionally, the stability of the fully synchronized state is assessed by spectral analysis of the Laplacian matrix. Here, we extend such procedure in a natural way by defining a multiorder Laplacian which takes into account all orders of interactions and whose eigenvalues unequivocally determine the stability of the synchronized state. We apply our framework to two case studies: (i) an hypergraph with all-to-all interactions at all orders, for which the eigenvalues can be derived fully analytically, and (ii) a generalization of Erdős-Rényi graphs where each node is randomly connected to the other nodes via both first and second order hyperedges. For each case, we show how the stability of synchronisation is altered by the presence of higher-order interactions. Given the wealth of applications of Laplacians operators, our multiorder Laplacian opens up new possibilities for formal analyses of the structure and dynamics of higher-order systems.

9.2 Kuramoto with Higher-Order Interactions

We study the effect of generalized higher-order interactions in a population of N identical phase oscillators. Each node i is associated to a phase θ_i, whose dynamical evolution reads:

$$\begin{aligned}\dot{\theta}_i = \omega &+ \frac{\gamma_1}{\langle K^{(1)}\rangle}\sum_{j=1}^{N} A_{ij}\sin(\theta_j-\theta_i)\\ &+ \frac{\gamma_2}{2!\langle K^{(2)}\rangle}\sum_{j,k=1}^{N} B_{ijk}\sin(\theta_j+\theta_k-2\theta_i)\\ &+ \frac{\gamma_3}{3!\langle K^{(3)}\rangle}\sum_{j,k,l=1}^{N} C_{ijkl}\sin(\theta_j+\theta_k+\theta_l-3\theta_i)\\ &+ \ldots\\ &+ \frac{\gamma_D}{D!\langle K^{(D)}\rangle}\sum_{j_1,\ldots,j_D=1}^{N} M_{ij_1\ldots j_D}\sin\left(\sum_{m=1}^{D}\theta_{j_m}-D\,\theta_i\right). \end{aligned} \tag{9.1}$$

System (9.1) is a generalization of the canonical Kuramoto model [28] to higher-order interactions of any order $d = 1, \ldots, D$, Indeed, the first two terms on the right-hand side correspond to the traditional Kuramoto, where pairs of oscillators are coupled, with strength γ_1, via a sinusoidal and diffusive function. The (undirected) adjacency matrix A determines the structure of 2-oscillator interactions: $A_{ij} = 1$ if there i and j are coupled, i.e. there is a link, between them, and $A_{ij} = 1$ otherwise, ω is the natural frequency of the oscillators, and $\langle K^{(1)}\rangle$ is the pairwise average degree. The following terms refer instead to the higher-order interactions taking place in the system, with $\gamma_2, \ldots, \gamma_D$ the coupling strengths at each higher order, while $\langle K^{(2)}\rangle, \ldots, \langle K^{(D)}\rangle$ represent the average degrees at order $2, \ldots, D$ (explicitly defined in the next section as natural extensions of the pairwise degree). The topology is encoded in the adjacency tensors $\mathbf{M} = \{M_{ij_1\ldots j_D}\}$ just like $A_{ij} = 1$ if there is a pairwise interaction (i, j) but 0 otherwise, $B_{ijk} = 1$ if there is a triplet interaction (i, j, k) but 0 otherwise, and similarly for all orders. Note that the interactions are assumed undirected, such that the adjacency tensors are invariant under any permutation of their indices. These adjacency tensors encode the most general topology of higher-order interactions which can be formalized as simplicial complexes, or more general hypergraphs.

Let us emphasize two features of model (9.1). First, the coupling function at each order d is symmetric with respect to i, i.e. it is invariant for any permutation of the remaining indices. This is the most natural generalization of the pairwise Kuramoto coupling function. Other possible choices, e.g. $\sin(2\theta_j - \theta_k - \theta_i)$ at order 2, can also be treated with the multiorder Laplacian framework we will discuss in the following (see [27] for details). Second, the normalization factors $\langle K^{(d)}\rangle$ at each order d ensure that the total contributions of each order d can be compared across orders. In other words, even if a given hypergraph contains many more 2-hyperedges than 1-hyperedges, the contribution of orders 1 and 2 can be compared, and will be of equal relative strength if $\gamma_1 = \gamma_2$. Finally, when we deal with a hypergraph it is sometimes convenient to look at the interactions at each order d (d-hyperedge) separately, rather than at the full system (9.1). This can be achieved by turning on one order of interaction at a time, while all others are turned off. For example, a classical network with pairwise interaction corresponds to the *pure 1-hyperedge* version of the

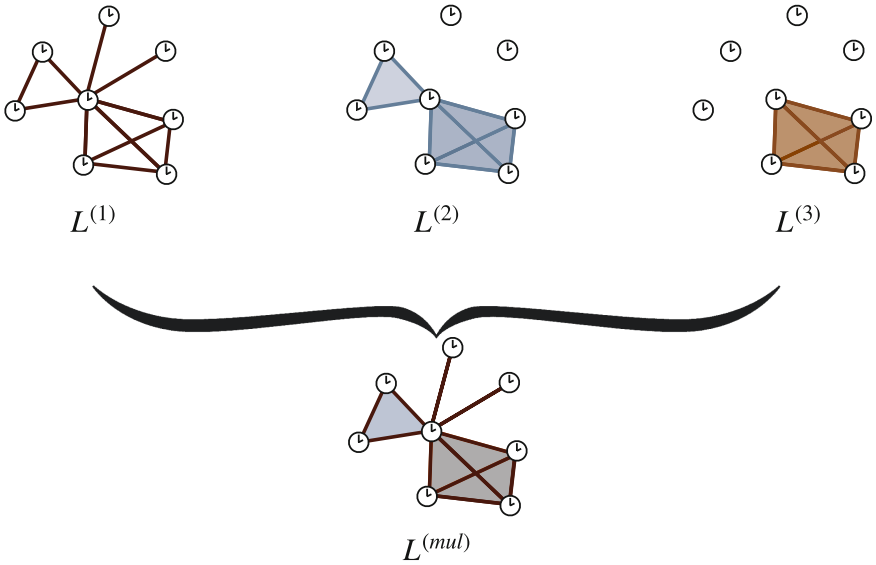


Fig. 9.1 Oscillators with higher-order interactions on simplicial complexes. Example of a simplicial complex and its building blocks: in the top left sketch only edges (1-simplices) are depicted, which can be encoded in Laplacian $L^{(1)}$. The top middle and top right sketches respectively represent the triangles (2-simplices) and the tetrahedra (3-simplices) of the same simplicial complex, described by $L^{(2)}$ and $L^{(3)}$. We recall that a d-simplex represents a $(d+1)$-oscillator interaction. The bottom sketch shows all orders combined, as encoded in the multiorder Laplacian $L^{(mul)}$

full hypergraph, i.e. $\gamma_1 \neq 0$ but $\gamma_d = 0$ for $d \neq 1$. Similarly, the pure 2-hyperedge case is that with only 3-oscillator interactions, i.e. $\gamma_2 \neq 0$ but $\gamma_d = 0$ for $d \neq 2$. In general, we have a pure d^*-hyperedge when $\gamma_{d^*} \neq 0$ but $\gamma_d = 0$ for $d \neq d^*$. All these cases are graphically represented in Fig. 9.1. We will use such a decomposition in the rest of the paper.

Model (9.1) admits a fully synchronized solution where each oscillator i evolves according to $\theta_i(t) = \omega t$. We are interested in analyzing the stability of this fully synchronized state. It is convenient to move to the rotating reference frame $\psi_i = \theta_i - \omega t$, which is equivalent to applying the transformation $\theta_i \mapsto \psi_i$ and $\omega \mapsto 0$ to the original system (9.1). In this new reference frame, the synchronized solution is given by $\psi_i(t) = 0$ for all $i = 1, \ldots, N$. By considering a heterogeneous and time-dependent perturbation $\delta\psi_i(t)$ to this state and inserting it in system (9.1) with the new reference frame, we obtain the linearized dynamics:

$$\begin{aligned}\delta\dot{\psi}_i = &+ \frac{\gamma_1}{\langle K^{(1)}\rangle}\sum_{j=1}^{N} A_{ij}(\delta\psi_j - \delta\psi_i)\\ &+ \frac{\gamma_2}{2!\langle K^{(2)}\rangle}\sum_{j,k=1}^{N} B_{ijk}(\delta\psi_j + \delta\psi_k - 2\delta\psi_i)\\ &+ \frac{\gamma_3}{3!\langle K^{(3)}\rangle}\sum_{j,k,l=1}^{N} C_{ijkl}\,(\delta\psi_j + \delta\psi_k + \delta\psi_l - 3\delta\psi_i)\\ &+ \ldots\\ &+ \frac{\gamma_D}{D!\langle K^{(D)}\rangle}\sum_{j_1,\ldots,j_D=1}^{N} M_{ij_i\ldots j_D}\left(\sum_{m=1}^{D}\delta\psi_{j_m} - D\,\delta\psi_i\right). \end{aligned} \tag{9.2}$$

Let us observe that the same equations could also be derived by assuming a different (not necessarily sinusoidal) shape of the coupling functions in the original system, via the linearization mechanism. The linear stability of the synchronized state can be determined by studying the solution of this system of equations. In a traditionally pairwise framework, the conditions under which the perturbations tend to zero (stable synchronized state) are typically assessed by using the so-called Laplacian formalism. We therefore need to generalize this approach by introducing a Laplacian operator which takes into account all orders of interactions, which will allow us to fully characterize the stability of system (9.1).

9.3 Multiorder Laplacian Framework

In systems of oscillators coupled in a dyadic way only, the linearized dynamics such as Eq. (9.2) can be written using the pairwise Laplacian. The solutions to this linear system can then be obtained in terms of the eigenvalues of that Laplacian. Various generalizations of the pairwise Laplacian to higher-order interactions have been proposed in the literature, from the simplest versions for uniform hypergraphs [29, 30], to more complicated ones associated to simplicial complexes [31–33] and Hodge Laplacians [34, 35], to mention a few. These models include higher-order interactions by transferring the role that nodes have in the traditional setting to higher-order structures (edges, triangles, tetrahedra, etc.). Hence the higher-order Laplacians can describe not only connections among nodes but also among edges, among triangles, and so on.

By contrast, here we study the synchronization of oscillators present on nodes only, which requires a different solution. We hence present an alternative generalization of the Laplacian which allows to mimic as closely as possible the standard procedure to solve the linearized system (9.2). This is a multiorder Laplacian, since it is able to simultaneously take into account all different orders of interactions. As described

in the next section, it is defined as a weighted sum of the Laplacians associated to the single orders of interactions. For this reason, before introducing the multiorder Laplacian, we first need to define a Laplacian operator for each order d.

9.3.1 Laplacians of Order d and Multiorder Laplacian

We start by recalling the classical Laplacian formalism for pairwise interactions, i.e. 1-simplex interactions, defined as $L_{ij}^{(1)} = K_i^{(1)}\delta_{ij} - A_{ij}^{(1)}$. Here, we write $A_{ij}^{(1)}$ and A_{ij} interchangeably to refer to the first-order adjacency matrix and $K_i^{(1)} = \sum_{j=1} A_{ij}$ is the degree of order 1 of oscillator i, i.e. the number of pairwise (1-simplex) interactions it has with other oscillators, and δ_{ij} denotes the Kronecker delta. For example, in Fig. 9.1, pairwise interactions are shown in the top left sketch, represented as dark brown links. We focus on pure first-order interactions in system (9.1), which is equivalent to imposing $\gamma_d = 0$ for $d \geq 2$. We can use the definitions above to rewrite system (9.2):

$$\delta\dot{\psi}_i = -\frac{\gamma_1}{\langle K^{(1)}\rangle}\sum_{j=1}^{N} L_{ij}^{(1)}\delta\psi_j. \tag{9.3}$$

Let us now generalize this formalism to the case of 3-oscillator interactions, i.e. those interactions represented by filled blue triangles in Fig. 9.1. By switching off all interactions of other orders, $\gamma_d = 0$ for all $d \neq 2$ in Eq. (9.2), we can write

$$\begin{aligned}\delta\dot{\psi}_i &= \frac{\gamma_2}{2!\langle K^{(2)}\rangle}\sum_{j,k=1}^{N} B_{ijk}(\delta\psi_j + \delta\psi_k - 2\delta\psi_i)\\ &= \frac{\gamma_2}{\langle K^{(2)}\rangle}\sum_{j,k=1}^{N} B_{ijk}(\delta\psi_j - \delta\psi_i),\end{aligned} \tag{9.4}$$

by using the symmetry of the adjacency tensor $\mathbf{B}$ and of the coupling function with respect to indices j and k. In fact, this symmetry allows us to write the 2-simplex interaction of 3 phases as 2 identical terms of the difference of only 2 phases. The phase difference in expression (9.4) is similar to that of the pairwise case, and this allows us to naturally generalize the Laplacian formalism to 2-simplex interactions.

In parallel, we need to generalize two important 1-simplex quantities to 2-simplex interactions: first, the 2-degree $K_i^{(2)}$ of node i, i.e. the number of distinct 2-simplices node i is part of, which can be written as $K_i^{(2)} = \frac{1}{2!}\sum_{j,k=1}^{N} B_{ijk}$, where the factor 2! ensures that each 2-simplex is counted only once. For example, in the middle left sketch of Fig. 9.1, the central node has a degree of order 2 of $K_i^{(2)} = 4$. Then, we define the adjacency matrix of order 2, whose entries $A_{ij}^{(2)}$ represent the number of 2-simplices shared by the pair (i, j), as $A_{ij}^{(2)} = \sum_{k=1}^{N} B_{ijk}$. This is a natural gener-

alization of the usual pairwise adjacency matrix A. Indeed, $A^{(2)}_{ij} = n$ if i and j are part of n common (but distinct) 2-simplex interactions, and 0 otherwise. With these definitions in hand, we can re-write the 2-simplex interaction term (9.4) as follows

$$\delta\dot{\psi}_i = \frac{\gamma_2}{\langle K^{(2)}\rangle}\left[\sum_{j=1}^{N} A^{(2)}_{ij}\delta\psi_j - \delta\psi_i 2! K^{(2)}_i\right] = -\frac{\gamma_2}{\langle K^{(2)}\rangle}\sum_{j=1}^{N} L^{(2)}_{ij}\delta\psi_j, \tag{9.5}$$

where we obtained the last member by defining the Laplacian of order 2: $L^{(2)}_{ij} = 2K^{(2)}_i\delta_{ij} - A^{(2)}_{ij}$, as a natural generalization of the usual Laplacian of order 1.

Now, we briefly show how to rewrite the 3-simplex interactions in a similar way, i.e. those interactions represented by filled light brown tetrahedra in Fig. 9.1 (bottom left sketch). Although it is very similar to the 2-simplex case, the case of 3-simplices better prepares us for the next step: the general d-simplex case. We switch off all interactions of other orders, i.e. $\gamma_d = 0$ for all $d \neq 3$ in Eq. (9.2), and it is simple to rewrite the 3-simplex interactions in terms of differences of two phases: $\delta\dot{\psi}_i = \frac{\gamma_3}{2!\langle K^{(3)}\rangle}\sum_{j,k,l=1}^{N} C_{ijkl}(\delta\psi_j - \delta\psi_i)$, where the factor $2! = 3!/3$ comes from the intermediate step where the expression is written as the sum of 3 identical terms. Similarly to the 2-simplex case, we can define the degree of order 3 for node i: $K^{(3)}_i = \frac{1}{3!}\sum_{j,k,l=1}^{N} C_{ijkl}$ as the number of distinct 3-simplex interactions it is part of; the adjacency matrix, $A^{(3)}_{ij} = \frac{1}{2!}\sum_{k,l=1}^{N} C_{ijkl}$ as the number of shared 3-simplices including nodes i and j, and finally the Laplacian $L^{(3)}_{ij} = 3K^{(3)}_i\delta_{ij} - A^{(3)}_{ij}$. It is therefore possible to use the 3-Laplacian to rewrite the 3-simplex term of the linearized system: $\delta\dot{\psi}_i = -\frac{\gamma_3}{\langle K^{(3)}\rangle}\sum_{j=1}^{N} L^{(3)}_{ij}$.

We are now ready to show that our generalized Laplacian formalism can be used to describe higher-order interactions of any order d. We switch off all interactions of other orders, $\gamma_{d'} = 0$ for all $d' \neq d$ in Eq. (9.2), and rewrite the d-simplex interactions as the differences of two phases

$$\begin{aligned}\delta\dot{\psi}_i &= +\frac{\gamma_d}{d!\langle K^{(d)}\rangle}\sum_{j_1,\ldots,j_d=1}^{N} M_{ij_1\ldots j_d}\left(\sum_{m=1}^{d}\delta\psi_{j_m} - d\,\delta\psi_i\right)\\ &= \frac{\gamma_d\, d}{d!\langle K^{(d)}\rangle}\sum_{j_1,\ldots,j_d=1}^{N} M_{ij_1\ldots j_d}(\delta\psi_j - \delta\psi_i).\end{aligned} \tag{9.6}$$

Similarly to the previous lower orders, we now define at order d the degree, $K^{(d)}_i = \frac{1}{d!}\sum_{j_1,\ldots,j_d=1}^{N} M_{ij_1\ldots j_d}$; the adjacency matrix $A^{(d)}_{ij} = \frac{1}{(d-1)!}\sum_{j_2,\ldots,j_d=1}^{N} M_{ijj_2\ldots j_d}$; and the Laplacian $L^{(d)}_{ij} = d\,K^{(d)}_i\delta_{ij} - A^{(d)}_{ij}$. Hence, similarly as before, we can simply rewrite the d-simplex interaction term

$$\delta\dot{\psi}_i = \frac{\gamma_d\, d}{d!\langle K^{(d)}\rangle}\sum_{j_1,\ldots,j_d=1}^{N} M_{ij_1\ldots j_d}(\delta\psi_j - \delta\psi_i) = -\frac{\gamma_d}{\langle K^{(d)}\rangle}\sum_{j=1}^{N} L^{(d)}_{ij}\delta\psi_j. \tag{9.7}$$

This shows that, at any order d, the dynamics caused by interactions between $d+1$ oscillators is determined by the matrix $\mathbf{L}^{(d)}$. Such a matrix fulfills the requirements necessary to be a Laplacian: it is positive semi-definite, its rows sum to zero and it is symmetric for undirected networks. Its eigenvalues are hence non-negative and include at least one zero, as we shall see in next section, and they are real for undirected networks. The definitions of the degree, adjacency matrix, and Laplacian matrix of order d are natural generalizations of their usual pairwise counterparts. Indeed, $K_i^{(d)}$ is the number of d-simplex interactions node i is part of, the adjacency matrix $A_{ij}^{(d)} = n$ where n is the number of distinct d-simplex interactions shared by nodes i and j. Finally, $L_{ij}^{(d)}$ is defined in terms of $A_{ij}^{(d)}$ and $K_i^{(d)}$ mimicking how $L_{ij}^{(1)}$ is defined in terms of $A_{ij}^{(1)}$ and $K_i^{(1)}$, up to factor d. Note that the usual pairwise definitions are recovered for $d = 1$.

We note that all single order Laplacians are defined as matrices of the same size $N \times N$, with N the number of nodes, as they all focus on node dynamics, separating the contributions to the couplings according to the order of interactions. Now that we can cast each order d into this Laplacian framework, we show that the full system (9.1) can also be rewritten with a single multiorder Laplacian operator.

Let us consider now the original system (9.1) with all interactions at all orders switched on, $\gamma_d \neq 0$ for all d. We are now able to write the linearized system as

$$\delta\dot{\psi}_i = -\sum_{j=1}^{N} L_{ij}^{(\mathrm{mul})} \delta\psi_j, \tag{9.8}$$

with

$$L_{ij}^{(\mathrm{mul})} = \sum_{d=1}^{D} \frac{\gamma_d}{\langle K^{(d)} \rangle} L_{ij}^{(d)} \tag{9.9}$$

where D is the maximum order of interactions in the hypergraph. We have defined the multiorder Laplacian $L_{ij}^{(\mathrm{mul})}$ as a weighted sum of the Laplacian matrices of order d. The weight given to each order is proportional to γ_d, and normalized by the average degree of order d. As we discuss in the following, the spectrum of the multiorder Laplacian controls the stability of a dynamical systems of coupled oscillators with higher-order interactions.

9.3.2 Stability and Lyapunov Exponents

The synchronized state of our system of oscillators is stable if the perturbation $\delta\psi_i$ on each node i converges to zero. To assess the overall stability, let us first consider interactions at the pure order d. The temporal evolution of the perturbation is given by solving Eq. (9.7) and this can be done by introducing Laplacian eigenvalues $\Lambda_\alpha^{(d)}$ and eigenvectors $\phi_\alpha^{(d)}$ defined by $\sum_{j=1}^{N} L_{ij}^{(d)} (\phi_\alpha^{(d)})_j = \Lambda_\alpha^{(d)} (\phi_\alpha^{(d)})_i$, with

$\alpha = 1, \ldots, N$. Indeed, this eigenbasis can be used to project the perturbation vector $\delta\psi_i(t) = \sum_{\alpha=1}^{N} c_\alpha \exp(\lambda_\alpha^{(d)} t)\ \phi_\alpha^{(d)}$, where the c_α are real constants and $\lambda_\alpha^{(d)}$ the Lyapunov exponents. Those define the condition for stability since the perturbations tend to zero over time if all values $\lambda_\alpha^{(d)}$ are negative. The Lyapunov exponents of the synchronized state can be found by plugging this solution into system (9.7), which allows to decouple our system of N equations, and can be hence written as:

$$\lambda_\alpha^{(d)} = -\frac{\gamma_d}{\langle K^{(d)} \rangle} \Lambda_\alpha^{(d)}. \tag{9.10}$$

The Laplacian eigenvalues are non-negative by definition so that if all $\gamma_d > 0$ (attractive coupling), the synchronized state is always stable, with $0 = \lambda_1^{(d)} > \lambda_2^{(d)} \geq \cdots \geq = \lambda_N^{(d)}$.

In the multiorder system (9.8) the stability of the synchronized equilibrium depends on the interplay of all different orders, as encoded in the multiorder Laplacian. Notice that even though the multiorder Laplacian is the weighted sum of each d-Laplacian, its eigenvalues cannot be obtained analytically from the singular spectra and we will need to numerically compute them. An exception to this is represented by the case with all-to-all interactions, where the d-Laplacian eigenvalues can instead be summed to obtain the multiorder spectrum and the system can be characterised in a fully analytical manner. We present such case in Sect. 9.4.1. However in general we need to compute the spectrum of $\mathbf{L}^{(\text{mul})}$ directly from the matrix to obtain the N Lyapunov exponents:

$$\lambda_\alpha^{(\text{mul})} = -\Lambda_\alpha^{(\text{mul})}. \tag{9.11}$$

Since the multiorder Laplacian has all the characteristics of a traditional Laplacian, also being semi-positive definite if the coupling is attractive, again all the eigenvalues are non-negative. In the multiorder framework too the synchronized state is hence stable. The Lyapunov exponent that determines the long-term behavior is the second Lyapunov exponent, $\lambda_2^{(\text{mul})}$, i.e. the smallest non-zero one. Its value determines the resilience of the system to perturbations, i.e. how fast the system comes back to the stable state after a perturbation. In particular, the more negative the $\lambda_2^{(\text{mul})}$, the more stable the synchronized state.

9.4 Full Synchrony of Kuramoto in Simplicial Complexes

In this section we apply the general framework exposed above to two specific simplicial complex topologies. The first case is a simplicial complex where all the nodes are connected to each other by means of all the possible simplices. This special case, that we call all-to-all higher-order network, is analytically solvable and we will provide a formula for the Lyapunov exponents. The second case is the simplicial generalization of an Erdős-Rényi network where each node is randomly connected to the other nodes via 1-simplexes and 2-simplexes.

9.4.1 All-to-all Higher-Order Networks

All-to-all coupling, also called *global* coupling, is a traditional network topology where all possible pairwise interactions take place. Here, we consider a generalization of this coupling scheme to higher-order interactions; we refer to it as *higher-order all-to-all*. In this scheme, all possible higher-order interactions occur among the nodes. The homogeneity of the higher-order network with that coupling scheme makes it mathematically tractable. In fact, it has been investigated analytically in a few studies, often focusing on cluster states [18, 19, 22, 36]. Here, instead, we consider the case of full synchrony and present a fully analytical description of its stability.

In the higher-order all-to-all setting, the linearized version of system (9.1) reads

$$\begin{aligned}\delta\dot{\psi}_i = &+ \frac{\gamma_1}{\langle K^{(1)}\rangle}\sum_{j=1}^{N}(\delta\psi_j - \delta\psi_i)\\ &+ \frac{\gamma_2}{2!\langle K^{(2)}\rangle}\sum_{j,k=1}^{N}(\delta\psi_j + \delta\psi_k - 2\delta\psi_i)\\ &+ \dots\\ &+ \frac{\gamma_d}{d!\langle K^{(d)}\rangle}\sum_{\{j_2,\dots,j_{d+1}\}=1}^{N}\left(\sum_{j_m=2}^{d+1}\delta\psi_{j_m} - d\,\delta\psi_i\right).\end{aligned} \tag{9.12}$$

Now that we have defined the coupling scheme, we can explicitly write the multiorder Laplacian.

9.4.1.1 Higher-Order Degree and Adjacency Matrices

Before explicitly writing down the multiorder Laplacian, we need to derive the higher-order degrees and adjacency matrices at order d. In a traditional all-to-all setting, the pairwise adjacency matrix has all entries equal to one, but 0 on the diagonal, which can be written $A_{ij}^{(1)} = 1 - \delta_{ij}$. At order d, the generic entry (i, j) of the adjacency matrix $A_{ij}^{(d)}$ is equal to the number of distinct $d+1$-oscillator interactions including both oscillators i and j, as described above. Hence, the matrix is simply given by the number of ways to pick $d-1$ oscillators among the $N-1$ oscillators left. This number is given by the combinatorics formula $A_{ij}^{(d)} = \binom{N-2}{d-1}(1-\delta_{ij})$, where $(1-\delta_{ij})$ ensures that entries with $i = j$ are equal to zero. Similarly, the degree of order d is equal to the number of $d+1$-oscillator interactions that oscillator i is part of. Hence, it is given by the number of ways to pick d oscillators out of the $N-1$ left, which can be written $K_i^{(d)} = \binom{N-1}{d}$ in combinatorics notation.

Additionally, the following identities will prove useful in the next section, to relate the quantities at order d to their traditional counterparts (at order 1)

$$A_{ij}^{(d)} = [(N-2)\cdots(N-d)/(d-1)!]\, A_{ij}^{(1)}, \tag{9.13}$$

$$K_i^{(d)} = [(N-2)\cdots(N-d)/d!]\, K_i^{(1)}. \tag{9.14}$$

In the higher-order all-to-all setting, all nodes have the same degree of order d. Hence, in this section, we write $K_i^{(d)}$ and $K^{(d)}$ interchangeably.

9.4.1.2 Higher-Order Laplacians Are Proportional to the Traditional Pairwise Laplacian

The pairwise Laplacian ($d = 1$) of the higher-order all-to-all case can be obtained from its definition: $L_{ij}^{(1)} = K_{ij}^{(1)}\delta_{ij} - A_{ij}^{(1)} = N\delta_{ij} - 1$. We can now write the higher-order Laplacian at order d in terms of this pairwise Laplacian, as follows:

$$\begin{aligned} L_{ij}^{(d)} &= d\,K_i^{(d)}\delta_{ij} - A_{ij}^{(d)} \\ &= [(N-2)\cdots(N-d)/(d-1)!]\; L_{ij}^{(1)} \\ &= \frac{d\,K^{(d)}}{N-1}\; L_{ij}^{(1)} \end{aligned} \tag{9.15}$$

With this equation, we have shown that the Laplacian at order d is proportional to the pairwise all-to-all Laplacian. More specifically, it grows with the order d but decreases with the number of nodes N. This also holds true for the multiorder Laplacian, which now reduces to

$$L_{ij}^{(\mathrm{mul})} = \left(\sum_{d=1}^{N-1} \frac{\gamma_d\, d}{N-1}\right)\; L_{ij}^{(1)}. \tag{9.16}$$

We stress that this linear relation between the higher-order Laplacians and that at order 1 is due to the homogeneity of the coupling scheme. Even though complex coupling scheme yield more complex forms of higher-order Laplacian, this analytical formula can serve as a limit case to help us understand more complex settings for which only numerical insights can be obtained.

9.4.1.3 Spectrum and Stability

The simplicity of the mathematical formulation of the higher-order all-to-all Laplacian at each order allows to find an analytical expression for the eigenvalues, and hence the Lyapunov exponents which determine the stability of the synchronized solution. In this particular case the traditional pairwise Laplacian shapes the Laplacian spectrum at each order d, and for higher-order all-to-all networks the spectrum of $\mathbf{L}^{(1)}$ is simply given by

$$\Lambda_1^{(1)} = 0 \qquad \Lambda_{2,\dots,N}^{(1)} = N. \tag{9.17}$$

It is easy to derive from this the Lyapunov exponents of each order d by simply using Eqs. (9.10) and (9.15), and the fact that the eigenvalues linearly scale with the Laplacian. We obtain:

$$\lambda_1^{(d)} = 0 \qquad \lambda_{2,\dots,N}^{(d)} = -\gamma_d \, d \, \frac{N}{N-1}. \tag{9.18}$$

This expression allows to describe the stability around synchronization, represented by the Lyapunov exponents and in particular by $\lambda_2^{(d)}$. We first of all observe that, as expected, the Lyapunov exponents are more negative for larger values of γ_d, enhancing stability. Then, this expression also says that the stability does not depend on the size N of the system if it is large enough. Lastly, it appears clear that interactions of higher orders d stabilize the synchronized state more (more negative Lyapunov exponents as d increases). This means that a system with pure 2-simplex interactions converges faster to synchronization than one with pure pairwise ones. This is true despite each order being given an equal weight through the normalization in system (9.12).

Then, combining all orders, we can obtain the multiorder Lyapunov exponents:

$$\lambda_1^{(\mathrm{mul})} = 0 \qquad \lambda_{2,\dots,N}^{(\mathrm{mul})} = -\frac{N}{N-1} \sum_{d=1}^{N-1} \gamma_d \, d. \tag{9.19}$$

Let us notice that the Lyapunov spectrum of the full system reduces to a linear combination of those at each pure order. We remark that this is possible because of the absence of a complex topology, for which instead the aggregated spectrum would be determined by a nonlinear combination of each order.

9.4.2 Higher-Order Interactions in Erdős-Rényi networks

As a second example, we consider system (9.1) on randomly generated simplicial complexes, with interactions up to order 2. The 2-parameter mechanism to generate them, introduced in [15], is as follows. First, generate an Erdős-Rényi pairwise network of N nodes, where edges are created between any two nodes with probability p_1. Second, for any three nodes, generate a 2-simplex (filled triangle) with probability p_2. Then, if some edges of these 2-simplices were not already created in step 1, add them to make the structure a simplicial complex. Conveniently, Iacopini et al. [15] derived a formula that links the average degrees of order 1, $\langle K^{(1)} \rangle$, and 2, $\langle K^{(2)} \rangle$, to the two parameters p_1 and p_2. For this reason, it is possible to directly choose the these two average degrees to generate such simplicial complexes.

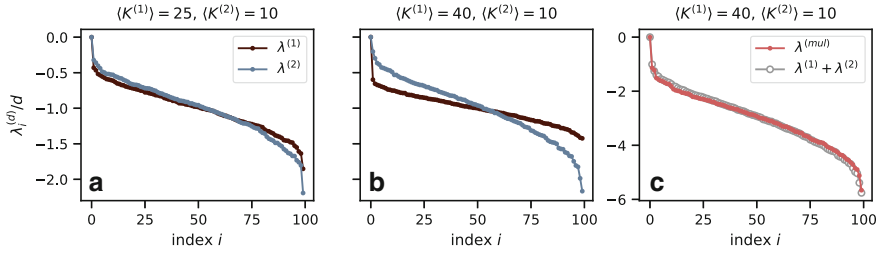


Fig. 9.2 Lyapunov spectra in random simplicial complexes with $(\langle K^{(1)}\rangle, \langle K^{(2)}\rangle)$ values of **(a)** (25, 10) and (**b**, **c**) (40, 10). (**a**, **b**) Respective Lyapunov spectra of order 1 (brown) and 2 (blue), divided by their respective order. **(c)** Lyapunov spectrum of the multiorder Laplacian (red) compared to the sum (grey) of the Lyapunov exponents of orders 1 and 2 from panel **(b)**

Figure 9.2 shows the various Lyapunov spectra associated to two instances of random simplicial complexes. The first has been assigned a first order degree $\langle K^{(1)}\rangle = 25$, in Fig. 9.2a, and the second $\langle K^{(1)}\rangle = 40$, in Fig. 9.2b, while both have second order degree $\langle K^{(2)}\rangle = 10$. In Fig. 9.2a, b, we show the Lyapunov spectra of order $d = 1$ (brown) and $d = 2$ (blue), normalized by their respective order d. These spectra are computed semi-analytically: they are defined by the generic analytical formulae (9.10) and (9.11) in which we numerically compute the eigenvalues of the Laplacian matrices. They appear to scale with their order d just as in the all-to-all case of last section. We also see that changing the number of pairwise interactions between Figs. 9.2a and 9.2b changes the Laplacian at order 1 but not at order 2, as expected. The tiny differences between the two spectra at order 2 (blue) exist just because they are two independent random simplicial complexes. Finally, we complete Fig. 9.2b by showing in Fig. 9.2c the multiorder Lyapunov spectrum (red) for the same simplicial complex, i.e. combining all orders. First, note that the exponents of the multiorder Laplacian are larger (in absolute value) than those obtained when only taking pairwise interactions into account (brown in Fig. 9.2b). This is because the (attractive) interactions at each order contribute to stabilizing the synchronized state. Besides, their respective contributions do not in general add linearly, as shown by the difference with the sum of the exponents at order 1 and 2 (red) and the multiorder exponents (red), except from special cases like the one previously described.

9.5 Summary and Conclusions

In this chapter, we presented the multiorder Laplacian: a framework to assess the stability of full synchrony in systems of coupled oscillators with higher-order interactions at any order and with arbitrary topology. We showed how it can be applied to real cases, considering two particular synthetic higher-order topologies. First, we discussed an all-to-all scheme, where all computations can be derived analytically, which gives us insight into the mechanisms at play. Second, we investigated ran-

dom simplicial complexes, where the spectra of the multiorder Laplacian receives non-trivial contributions from the single orders. Our code to compute the multiorder Laplacian and Laplacians at each order in a hypergraph is available at https://github.com/maximelucas/multiorder_Laplacian. In the next chapter, the multiorder Laplacian is applied to different dynamical systems, showing how it can be used to extend the ideas of the Master Stability Function approach to higher-order systems [37].

Finally, an adjusted version of the multiorder Laplacian with a fixed coupling budget distributed between pairwise and higher-order interactions has been introduced in Ref. [38], showing that in such a case hyperedges tend to enhance synchronization in random hypergraphs, but have the opposite effect in simplicial complexes.

References

1. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, Cambridge, UK)
2. S.H. Strogatz, *Sync: The Emerging Science of Spontaneous Order* (Penguin UK, 2004)
3. D. Cumin, C. Unsworth, Phys. D **226**(2), 181 (2007). https://doi.org/10.1016/j.physd.2006.12.004
4. J.C. Leloup, D. Gonze, A. Goldbeter, J. Biol. Rhythms **14**(6), 433 (1999). https://doi.org/10.1177/074873099129000948
5. M.B. Lotrič, A. Stefanovska, Phys. A **283**(3–4), 451 (2000). https://doi.org/10.1016/S0378-4371(00)00204-1
6. Y.F. Suprunenko, P.T. Clemson, A. Stefanovska, Phys. Rev. Lett. **111**(2), 024101 (2013). https://doi.org/10.1103/PhysRevLett.111.024101
7. A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Phys. Rep. **469**(3), 93 (2008). https://doi.org/10.1016/j.physrep.2008.09.002
8. F.A. Rodrigues, T.K.D.M. Peron, P. Ji, J. Kurths, Phys. Rep. **610**, 1 (2016). https://doi.org/10.1016/j.physrep.2015.10.008
9. J. Gómez-Gardenes, S. Gómez, A. Arenas, Y. Moreno, Phys. Rev. Lett. **106**(12) (2011). https://doi.org/10.1103/PhysRevLett.106.128701
10. D. Pazó, Phys. Rev. E **72**(4), 046211 (2005). https://doi.org/10.1103/PhysRevE.72.046211
11. L.M. Pecora, F. Sorrentino, A.M. Hagerstrom, T.E. Murphy, R. Roy, Nat. Comm. **5**, 4079 (2014). https://doi.org/10.1038/ncomms5079
12. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, Phys. Rep. **874**, 1 (2020). https://doi.org/10.1016/j.physrep.2020.05.004
13. F. Battiston, E. Amico, A. Barrat, G. Bianconi, G.F. de Arruda, B. Franceschiello, I. Iacopini, S. Kefi, V. Latora, Y. Moreno, M. Murray, T. Peixoto, F. Vaccarino, G. Petri, Nat. Phys. **17**(10), 1093–1098 (2022). https://doi.org/10.1038/s41567-021-01371-4
14. T. Carletti, F. Battiston, G. Cencetti, D. Fanelli, Phys. Rev. E **101**(2), 022308 (2020). https://doi.org/10.1103/PhysRevE.101.022308
15. I. Iacopini, G. Petri, A. Barrat, V. Latora, Nat. Commun. **10**(1), 2485 (2019). https://doi.org/10.1038/s41467-019-10431-6
16. J. Grilli, G. Barabás, M.J. Michalska-Smith, S. Allesina, Nature **548**(7666), 210 (2017)
17. U. Alvarez-Rodriguez, F. Battiston, G.F. de Arruda, Y. Moreno, M. Perc, V. Latora, Nat. Human Behav. **5**(5), 586 (2021)
18. P.S. Skardal, A. Arenas, Phys. Rev. Lett. **122**(24), 248301 (2019). https://doi.org/10.1103/PhysRevLett.122.248301
19. P.S. Skardal, A. Arenas, Commun. Phys. **3**(1), 1 (2020)
20. A.P. Millán, J.J. Torres, G. Bianconi, Phys. Rev. Lett. **124**(21), 218301 (2020)

21. T. Tanaka, T. Aoyagi, Phys. Rev. Lett. **106**(22), 224101 (2011). https://doi.org/10.1103/PhysRevLett.106.224101
22. M. Komarov, A. Pikovsky, Phys. Rev. E **92**(2), 020901 (2015). https://doi.org/10.1103/PhysRevE.92.020901
23. C. Bick, P. Ashwin, A. Rodrigues, Chaos **26**(9), 094814 (2016). https://doi.org/10.1063/1.4958928
24. P. Ashwin, C. Bick, O. Burylko, Front. Appl. Math. Stat. **2**, 7 (2016). https://doi.org/10.3389/fams.2016.00007
25. B. Kralemann, A. Pikovsky, M. Rosenblum, Phys. Rev. E **87**(5), 052904 (2013). https://doi.org/10.1103/PhysRevE.87.052904
26. B. Kralemann, A. Pikovsky, M. Rosenblum, New J. Phys. **16**(8), 085013 (2014). https://doi.org/10.1088/1367-2630/16/8/085013
27. M. Lucas, G. Cencetti, F. Battiston, Phys. Rev. Res. **2**(3), 033410 (2020). https://doi.org/10.1103/PhysRevResearch.2.033410
28. Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer, Tokyo, 1984)
29. F. Chung, Expanding Graphs (DIMACS series) (1993), pp. 21–36
30. L. Lu, X. Peng, in *International Workshop on Algorithms and Models for the Web-Graph* (Springer, 2011), pp. 14–25. https://doi.org/10.1007/978-3-642-21286-4_2
31. A. Muhammad, M. Egerstedt, in *Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems* (Citeseer, 2006), pp. 1024–1038
32. R. Rosenthal, arXiv:1412.5406 (2014)
33. O. Parzanchevski, R. Rosenthal, Random Struct. Algorithms **50**(2), 225 (2017). https://doi.org/10.1002/rsa.20657
34. L.H. Lim, *Proc Sympos. Appl, Math* (2015)
35. M.T. Schaub, A.R. Benson, P. Horn, G. Lippner, A. Jadbabaie, arXiv:1807.05044 (2018)
36. C.C. Gong, A. Pikovsky, Phys. Rev. E **100**(6), 062210 (2019). https://doi.org/10.1103/PhysRevE.100.062210
37. L. Gambuzza, F. Di Patti, L. Gallo, S. Lepri, M. Romance, R. Criado, M. Frasca, V. Latora, S. Boccaletti, Nat. Commun. **12**(1), 1 (2021)
38. Y. Zhang, M. Lucas, F. Battiston, Do higher-order interactions promote synchronization? arXiv:2203.03060 (2022)

Chapter 10
The Master Stability Function for Synchronization in Simplicial Complexes

Lucia Valentina Gambuzza, Francesca Di Patti, Luca Gallo, Stefano Lepri, Miguel Romance, Regino Criado, Mattia Frasca, Vito Latora, and Stefano Boccaletti

Abstract The collective behaviors of a complex system are determined by the intricate way in which its components interact. In this chapter we discuss a novel and general analytical framework to study synchronized states in systems of many dynamical units with many-body interactions, which allows to account for the microscopic structure of the interactions at any possible order. In such a framework, the N dynamical units of a system are associated to the N nodes of a D dimensional ($D \geq 1$) simplicial complex, whose simplices represent the structure of the different types of coupling. Namely, 1-simplices (links) describe pairwise interactions, 2-simplices (triangles) describe three-body interactions, 3-simpliced (tetrahedra) four-body interactions, and so on. Such a description generalizes that of a complex network of dynamical units, and reduces to it in the particular case of $D = 1$ simplicial complexes. Within this framework, we study the onset of full synchronization and the conditions for

L. V. Gambuzza · L. Gallo · M. Frasca (✉) · V. Latora
University of Catania, Catania, Italy
e-mail: mattia.frasca@dieei.unict.it

L. V. Gambuzza
e-mail: lucia.gambuzza@dieei.unict.it

L. Gallo
e-mail: luca.gallo@phd.unict.it

V. Latora
e-mail: v.latora@qmul.ac.uk

F. Di Patti · S. Lepri · S. Boccaletti
CNR-Institute of Complex Systems, Florence, Italy
e-mail: stefano.lepri@isc.cnr.it

M. Romance · R. Criado
University Rey Juan Carlos, Madrid, Spain
e-mail: miguel.romance@urjc.es

R. Criado
e-mail: regino.criado@urjc.es

V. Latora
Queen Mary University of London, London, UK

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_10

the stability of a synchronized state in systems of identical dynamical units. We show that, under certain assumptions on the network topology or on the form of the coupling, these conditions can be written in terms of a Master Stability Function that generalizes the existing results valid for pairwise interactions (i.e. networks) to the case of complex systems with the most general possible architecture. As an example of the potential utility of the proposed method we study the dynamics of $D = 3$ simplicial complexes of chaotic systems (Rössler oscillators) and we investigate how the stability of synchronized states depends on the interplay between the control parameters of the chaotic units and the structural properties of the simplicial complex.

10.1 Introduction

Synchronization is an ubiquitous phenomenon in natural and engineered systems. It corresponds to the emergence of a collective behavior wherein the system components eventually adjust themselves into a common evolution in time. [1, 2] Various studies have shed light on the intimate relations between the topology of a networked system, its synchronizability, and the properties of the synchronized states. In particular, synchronous behaviors have been observed and characterized in small-world [3], weighted [4], multilayer [5], and adaptive networks [6, 7]. Outside complete synchronization, moreover, other types of synchronization have been revealed to emerge in networked systems, including remote synchronization [8, 9], cluster states [10] and synchronization of group of nodes [11], chimera [12, 13], Bellerophon states [14, 15], and Benjamin-Feir instabilities [16–18]. Finally, the transition to synchronization has been shown to be either smooth and reversible, or abrupt and irreversible (as in the case of explosive synchronization, resembling a first-order like phase transition [19]).

While attempts of extending to $p-$uniform hypergraphs the analysis of complete synchronization of dynamical systems have been recently made [20], most studies of systems interplaying through higher order interactions in simplicial complexes have focused on the case of the Kuramoto model [21, 22]. This is, in fact, a specific model, wherein each unit of the ensemble $i = 1, \ldots, N$ is a phase oscillator and is characterized by the evolution of its real valued phase $\theta_i(t) \in [0, 2\pi]$. The model has been studied in all different sorts of network topologies with possible applications to biological and social systems [2, 21], and recently extensions of it have been proposed that include higher-order interactions. Namely, it has been shown that the Kuramoto model may exhibit abrupt desynchronization when three-body interactions among all the oscillators are added to [23], or completely replace [24], the all-to-all pairwise interactions of the original model. Similar results have been obtained with a non-symmetric variation of the Kuramoto model in which the microscopic details of the interactions among the phase oscillators are described in the form of a simplicial complex [25]. A different approach has been proposed by Millán et al., who have formulated a higher-order Kuramoto model in which the oscillators are placed not

on the nodes but on higher-order simplices, such as links, triangles, and so on, of a simplicial complex [26].

In Chap. 9, an extension of the Kuramoto model to interactions of any order, which is still analytically tractable because all the oscillators have identical frequencies, has been discussed [27]. In this chapter, we move from the analysis of a specific model to the study of the most general ensemble of, yet identical, dynamical systems, organized on the nodes of a simplicial complex of any order, and interacting via any coupling functions. In such a general context, we show that complete synchronization in systems of identical units exists as an invariant solution, as far as the coupling functions cancel out. Furthermore, we give the necessary condition for it to be observed as a stable state and then we show that such condition can be written in terms of a Master Stability Function, a method initially developed in Ref. [28] for pairwise coupled systems, and later extended in many ways to complex networks [29] and to time-varying interactions [30–33]. Therefore, the framework discussed here is valid for a large number of situations, and, as so, it is applicable to a very wide range of experimental and/or practical circumstances. We will show, indeed, that all the theoretical predictions that our method entitles us to make are fully verified in simulations of synthetic and real-words networked systems.

10.2 Networks and Higher-Order Structures

A network is a collection of nodes and of edges connecting pairs of nodes. Mathematically, it is represented by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, which consists of a set $\mathcal{V}$ with $N = |\mathcal{V}|$ elements called vertices (or nodes), and a set $\mathcal{E}$ whose K elements, called edges or links, are pairs of nodes (i, j) ($i, j = 1, 2, \ldots, N$ and $i \neq j$). As graphs explicitly refer to pairwise interactions, networks have been very successful in capturing the properties of coupled dynamical systems in all such cases in which the interactions can be expressed (or approximated) as a sum of two-body terms [34]. Conversely, their limits emerge when it comes to model higher-order interactions. In fact, the presence of a triangle of three nodes i, j, k in a network, e.g. the presence of the three links (i, j), (i, k), (j, k) in the corresponding graph, is not able to capture the difference between a three-body interaction of the three individuals, from the sum of three pairwise interactions. Notice that these are two completely different situations, with completely different social mechanisms and dynamics at work [35].

Simplicial complexes are instead the proper mathematical structures for describing high order interactions. A simplicial complex is an aggregate of simplices, objects that generalize links and can in general be of different dimension. A *d-simplex*, or simplex of dimension d, σ is, in its simplest definition, a collection of $d + 1$ nodes. In this way, a 0-simplex is a node, a 1-simplex is a link, a 2-simplex (i, j, k) is a two-dimensional object made by three nodes, usually called a (full) triangle, a 3-simplex is a tetrahedron, i.e. a three-dimensional object and so on (Fig. 10.1a). It is now possible to differentiate between a three-body interaction, and three bodies in pairwise interactions: the first case will be represented by a complete triangle, a two-dimensional

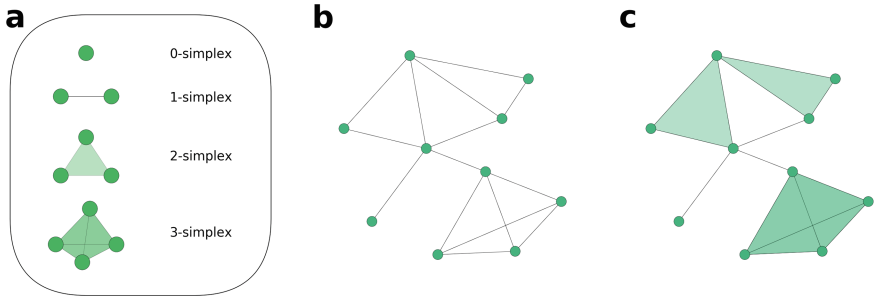


Fig. 10.1 From networks to simplicial complexes. **a** Simplices of different orders d: the node $d = 0$, the edge $d = 1$, the triangle $d = 2$, the tetrahedron $d = 3$. **b** A network only consists of 0− and 1−simplices. **c** A simplicial complex consists of simplices of any order $d = 0, 1, \ldots, D$ (in this case $D = 3$)

simplex, while the second case will consist of three one-dimensional objects. Hence, in the following of this chapter, simplices of dimension d will be used to describe the structure of $(d + 1)$-body interactions. Finally, a *simplicial complex* $\mathcal{S}$ on a given set of nodes $\mathcal{V}$, with $|\mathcal{V}| = N$, is a collection of M simplices, $\mathcal{S} = \{\sigma_1, \sigma_2, \ldots, \sigma_M\}$, with the extra requirement that, for any simplex $\sigma \in \mathcal{S}$, all the simplices σ' with $\sigma' \subset \sigma$, i.e. all the simplices built from subsets of σ, are also contained in $\mathcal{S}$. Due to this requirement, simplicial complexes are a very particular type of hypergraphs [36]. Simplicial complexes have shown to be appropriate in the context of social systems [35, 37, 38] and, as we will see in the next Section, they will turn very useful to study coupled dynamical systems. In the following, we will indicate as M_d, $d = 1, 2, \ldots D$ the number of d-simplices present in $\mathcal{S}$ (where D, the order of the simplicial complex, is the dimension of the largest simplex in $\mathcal{S}$), and we have the constraint $\sum_{d=1}^{D} M_d = M$. Note that a network is a particular case of a simplicial complex with $D = 1$ (Fig. 10.1b), whereas for $D > 1$ a truly higher-order structure is obtained (Fig. 10.1c).

As a mathematical representation of simplicial complexes, we will use here a formalism which generalises directly the concept of adjacency matrix for a network. For each dimension d, we can define the $\underbrace{N \times N \times \cdots \times N}_{d+1}$ *adjacency tensor* $\mathrm{A}^{(d)}$, whose entry $a^{(d)}_{i_1,\ldots,i_{d+1}}$ is equal to 1 if the d-simplex $(i_1, \ldots, i_{d+1})$ belongs to the simplicial complex $\mathcal{S}$, and is 0 otherwise [39]. Notice that each tensor is symmetric with respect to its $d + 1$ indices, which means that the value of a given entry $a^{(d)}_{i_1,\ldots,i_{d+1}}$ is equal to the value of the entries corresponding to any permutation of the indices.

With the definition above, $\mathrm{A}^{(1)}$ coincides with the standard adjacency matrix A, while the $N \times N \times N$ adjacency tensor $\mathrm{A}^{(2)}$ characterizes two-dimensional objects: one has $a^{(2)}_{ijk} = 1$ if the three nodes i, j, k form a full triangle, and otherwise $a^{(2)}_{ijk} = 0$. As a conclusion, it is possible to map completely the connectivity structure of a simplicial complex $\mathcal{S}$ into the entire set of D adjacency tensors $\mathrm{A}^{(d)}, d = 1, 2, \ldots D$.

A node i of a simplicial complex $\mathcal{S}$ cannot be, therefore, characterized only by giving its degree $k_i = \sum_{j=1}^{N} a_{ij}^{(1)}$, but one needs instead to account for the number of simplices of any dimension, incident in i. It is therefore extremely useful to define the generalized *d-degree*, $k_i^{(d)}$, of a node i as:

$$k_i^{(d)} = \frac{1}{d!} \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \cdots \sum_{i_d=1}^{N} a_{i,i_1,i_2,\ldots,i_d}^{(d)}, \tag{10.1}$$

with $d = 1, 2, \ldots, D$ so that $k_i^{(1)}$ coincides with the standard degree of node i, $k_i^{(2)}$ counts the number of triangles (2-simplices) to which i participates:

$$k_i^{(2)} = \frac{1}{2} \sum_{j=1}^{N} \sum_{k=1}^{N} a_{ijk}^{(2)} \tag{10.2}$$

$k_i^{(3)}$ the number of tetrahedrons, and so on.

Analogously, we can also define the generalized *d-degree* $k_{ij}^{(d)}$ of a link (i, j) as the number of d-simplices to which link (i, j) is part of. We can write its expression in terms of the adjacency tensor $\mathrm{A}^{(d)}$ of dimension d, with $d = 1, 2, \ldots, D$, as[39]:

$$k_{ij}^{(d)} = \frac{1}{(d-1)!} \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \cdots \sum_{i_{d-1}=1}^{N} a_{i,j,i_1,i_2,\ldots,i_{d-1}}^{(d)}, \tag{10.3}$$

so that $k_{ij}^{(1)} = a_{ij}^{(1)}$, while $k_{ij}^{(2)}$ counts the number of triangles (2-simplices) to which (i, j) participates:

$$k_{ij}^{(2)} = \sum_{k=1}^{N} a_{ijk}^{(2)}, \tag{10.4}$$

and so on.

Finally, we introduce here a generalized Laplacian describing the case of systems with high-order interactions. The *generalized Laplacian* of order d, with $d = 1, 2, \ldots, D$, is a matrix $\mathcal{L}^{(d)}$ whose elements are defined as:

$$\mathcal{L}_{ij}^{(d)} = \begin{cases} 0 & \text{for } i \neq j \quad \text{and} \quad a_{ij}^{(1)} = 0 \\ -(d-1)! k_{ij}^{(d)} & \text{for } i \neq j \quad \text{and} \quad a_{ij}^{(1)} = 1 \\ d!\, k_i^{(d)} & \text{for } i = j, \end{cases} \tag{10.5}$$

where $k_{ij}^{(d)}$ is the generalized d-degree of the link (i, j), and $k_i^{(d)}$ is the generalized d-degree of node i. Replacing (10.1) and (10.3) in (10.5), in the case $D = 2$, we get an equivalent expression for the generalized Laplacian:

$$\mathcal{L}^{(2)}_{ij} = \begin{cases} -\sum_k a^{(2)}_{ijk} & \text{for } i \neq j \\ -\sum_{\ell \neq i} \mathcal{L}^{(2)}_{i\ell} & \text{for } i = j, \end{cases} \tag{10.6}$$

Notice that $\mathcal{L}^{(1)}$ recovers exactly the classical Laplacian matrix. This definition of generalized Laplacian will turn useful in the following sections.

10.3 Dynamical Systems with Higher-Order Interactions

The object of our study is the most general simplicial complex of N coupled dynamical systems, such that the nodes are subject not only to pairwise interactions, but also to three-body interactions, four-body interactions and so on. We write the equations of motion governing the dynamics of our D-dimensional simplicial complex as follows

$$\begin{aligned} \dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i) &+ \sigma_1 \sum_{j_1=1}^{N} a^{(1)}_{ij_1}\, \mathbf{g}^{(1)}(\mathbf{x}_i, \mathbf{x}_{j_1}) + \sigma_2 \sum_{j_1=1}^{N} \sum_{j_2=1}^{N} a^{(2)}_{ij_1j_2}\, \mathbf{g}^{(2)}(\mathbf{x}_i, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}) + \dots \\ &+ \sigma_D \sum_{j_1=1}^{N} \dots \sum_{j_D=1}^{N} a^{(D)}_{ij_1\dots j_D}\, \mathbf{g}^{(D)}(\mathbf{x}_i, \mathbf{x}_{j_1}, \dots, \mathbf{x}_{j_D}), \end{aligned} \tag{10.7}$$

where $\mathbf{x}_i(t)$ is the m-dimensional vector state describing the dynamics of unit i, $\sigma_1, \dots, \sigma_D$ are real valued parameters describing coupling strengths, $\mathbf{f} : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ describes the local dynamics (which is assumed identical for all units), while $\mathbf{g}^{(d)} : \mathbb{R}^{(d+1)\times m} \longrightarrow \mathbb{R}^m$ $(d = 1, \dots, D)$ are synchronization non-invasive functions (i.e. $\mathbf{g}^{(d)}(\mathbf{x}, \mathbf{x}, \dots, \mathbf{x}) \equiv 0\ \forall d$) ruling the interaction forms at different orders. Furthermore, for $d = 1, \dots, D$, $a^{(d)}_{ij_1\dots j_d}$ are the entries of the adjacency tensor $\mathrm{A}^{(d)}$. This is the most general type of system we can consider, as there are no further specific restrictions on both the adjacency tensors of the simplicial complex and the functions $\mathbf{f}$ and $\mathbf{g}^{(d)}$.

For the sake of clarity in what follows we illustrate our study for the case of $D = 2$ and then summarize the steps needed to generalize the results to any order D. Let us then consider the following set of coupled differential equations

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i) + \sigma_1 \sum_{j=1}^{N} a^{(1)}_{ij}\, \mathbf{g}^{(1)}(\mathbf{x}_i, \mathbf{x}_j) + \sigma_2 \sum_{j=1}^{N} \sum_{k=1}^{N} a^{(2)}_{ijk}\, \mathbf{g}^{(2)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k), \tag{10.8}$$

where σ_1 and σ_2 are the coupling strengths associated to two- and three-body interactions.

Notice that existence and invariance of the synchronized solution $\mathbf{x}^s(t) = \mathbf{x}_1(t) = \dots = \mathbf{x}_N(t)$ are warranted by the non-invasiveness of the coupling functions.

10.4 Linear Stability Analysis

Here we study the stability of the synchronous solution via linearization around the synchronous state $\mathbf{x}^s$. Let us then consider considers small perturbations around the synchronous state $\mathbf{x}^s$, i.e., $\delta\mathbf{x}_i = \mathbf{x}_i - \mathbf{x}^s$, and write the dynamics of these variables as follows

$$\begin{aligned}\dot{\delta\mathbf{x}_i} = & J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i + \sigma_1 \sum_{j=1}^{N} a_{ij}^{(1)} \left[\left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_i + \left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_j \right] \\ & + \sigma_2 \sum_{j=1}^{N}\sum_{k=1}^{N} a_{ijk}^{(2)} \left[\left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_i + \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_j \right. \\ & \left. + \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_k \right],\end{aligned} \tag{10.9}$$

where $J\mathbf{f}(\mathbf{x}^s)$ denotes the $m \times m$ Jacobian matrix of the function $\mathbf{f}$, evaluated at the synchronous state $\mathbf{x}^s$.

Now, let us make our first, very important, conceptual step, noticing that all coupling functions are synchronization non-invasive, i.e. $\mathbf{g}^{(1)}(\mathbf{x}, \mathbf{x}) \equiv 0$ and $\mathbf{g}^{(2)}(\mathbf{x}, \mathbf{x}, \mathbf{x}) \equiv 0$. As their value is then constant (equal to zero) at the synchronization manifold, it immediately follows that their total derivative vanishes as well, which implies on its turn that

$$\begin{aligned} & \left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s)} + \left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s)} = 0, \\ & \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} + \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} + \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} = 0. \end{aligned} \tag{10.10}$$

Then, one can factor out the terms $\left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_i$ and $\left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_i}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)} \delta\mathbf{x}_i$ in the summations (both of them, indeed, do not depend on the indices of the summations). Furthermore, one has that $\sum_{j=1}^{N} a_{ij}^{(1)} = k_i^{(1)}$ and $\sum_{j=1}^{N}\sum_{k=1}^{N} a_{ijk}^{(2)} = 2k_i^{(2)}$. Plugging back the resulting terms inside the summations, and using Eq. (10.10), one eventually obtains

$$\begin{aligned}\dot{\delta\mathbf{x}_i} = & J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i - \sigma_1 \sum_{j=1}^{N} \mathcal{L}_{ij}^{(1)}\, J\mathbf{g}^{(1)}(\mathbf{x}^s, \mathbf{x}^s)\delta\mathbf{x}_j \\ & -\sigma_2 \sum_{j=1}^{N}\sum_{k=1}^{N} \tau_{ijk} \left[J_1\mathbf{g}^{(2)}(\mathbf{x}^s, \mathbf{x}^s, \mathbf{x}^s)\delta\mathbf{x}_j + J_2\mathbf{g}^{(2)}(\mathbf{x}^s, \mathbf{x}^s, \mathbf{x}^s)\delta\mathbf{x}_k \right],\end{aligned} \tag{10.11}$$

where we introduced a tensor T whose elements are $\tau_{ijk} = 2k_i^{(2)}\delta_{ijk} - a_{ijk}^{(2)}$ for $i, j, k = 1, \ldots, N$, and simplified the notation as

$$
\begin{aligned}
J\mathbf{g}^{(1)}(\mathbf{x}^s,\mathbf{x}^s) &= \left.\frac{\partial \mathbf{g}^{(1)}(\mathbf{x}_i,\mathbf{x}_j)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s)},\\
J_1\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s) &= \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_j}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)},\\
J_2\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s) &= \left.\frac{\partial \mathbf{g}^{(2)}(\mathbf{x}_i,\mathbf{x}_j,\mathbf{x}_k)}{\partial \mathbf{x}_k}\right|_{(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)}.
\end{aligned}
\tag{10.12}
$$

Already at this stage, it is fundamental to remark that our approach even extends the validity of the classical Master Stability Function theory [the case $\sigma_2 = 0$ in Eq. (10.11)], in that *we do not require* a diffusive functional form for the interplay among the network nodes, and therefore we are actually encompassing a much broader class of coupling functions. For instance, our approach allows the formal treatment of the Kuramoto model [21], where $m = 1$, each network unit i is identified by the instantaneous phase θ_i of an oscillator, and the coupling between nodes i and j is given by the function $\sin(\theta_j - \theta_i)$, which is not diffusive.

Let us now make our second, conceptual, step, which will allow us to greatly simplify the last term on the right hand side of Eq.(10.11). Such a term refers to three-body interactions, and we now show how to map it into a single summation involving the generalized Laplacian matrix. This is done by remarking that the two Jacobian matrices $J_1\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)$ and $J_2\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)$ are both independent on k and j. Accordingly, Eq.(10.11) becomes

$$
\begin{aligned}
\dot{\delta\mathbf{x}}_i = {} & J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i - \sigma_1 \sum_{j=1}^{N} \mathcal{L}_{ij}^{(1)}\, J\mathbf{g}^{(1)}(\mathbf{x}^s,\mathbf{x}^s)\delta\mathbf{x}_j \\
& -\sigma_2\Bigg[\sum_{j=1}^{N} J_1\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)\delta\mathbf{x}_j \sum_{k=1}^{N}\tau_{ijk} + \sum_{k=1}^{N} J_2\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)\delta\mathbf{x}_k \sum_{j=1}^{N}\tau_{ijk}\Bigg].
\end{aligned}
\tag{10.13}
$$

Then, using the symmetric property of T, namely $\sum_k \tau_{ijk} = \sum_k \tau_{ikj}$, we obtain

$$
\begin{aligned}
\dot{\delta\mathbf{x}}_i = {} & J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i - \sigma_1 \sum_{j=1}^{N} \mathcal{L}_{ij}^{(1)}\, J\mathbf{g}^{(1)}(\mathbf{x}^s,\mathbf{x}^s)\delta\mathbf{x}_j \\
& -\sigma_2 \sum_{j=1}^{N} \mathcal{L}_{ij}^{(2)}\Big[J_1\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s) + J_2\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)\Big]\delta\mathbf{x}_j.
\end{aligned}
\tag{10.14}
$$

Equations (10.14) can be rewritten in block form by introducing the stack vector $\delta\mathbf{x} = [\delta\mathbf{x}_1^T, \delta\mathbf{x}_2^T, \ldots, \delta\mathbf{x}_N^T]^T$ and denoting by $\mathrm{JF} = J\mathbf{f}(\mathbf{x}^s)$, $\mathrm{JG}^{(1)} = J\mathbf{g}^{(1)}(\mathbf{x}^s,\mathbf{x}^s)$ and $\mathrm{JG}^{(2)} = J_1\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s) + J_2\mathbf{g}^{(2)}(\mathbf{x}^s,\mathbf{x}^s,\mathbf{x}^s)$. One obtains

$$
\dot{\delta\mathbf{x}} = \left[\mathrm{I}_N \otimes \mathrm{JF} - \sigma_1\mathcal{L}^{(1)} \otimes \mathrm{JG}^{(1)} - \sigma_2\mathcal{L}^{(2)} \otimes \mathrm{JG}^{(2)}\right]\delta\mathbf{x}. \tag{10.15}
$$

The third, and final, conceptual step is to remark that all generalized Laplacians $\mathcal{L}^{(d)}$ are symmetric real-valued zero-row-sum matrices. Therefore: (i) they are all diagonalizable; (ii) for each one of them the set of eigenvalues is made of real non-negative numbers, and the corresponding set of eigenvectors constitutes a orthonor-

mal basis of $\mathbb{R}^N$; (iii) they all share, as the smallest of their eigenvalues, $\lambda_1 \equiv 0$, whose associated eigenvector $\frac{1}{\sqrt{N}}\,(1, 1, 1, \ldots, 1)^T$ is aligned along the synchronization manifold; (iv) as in general they do not commute, the sets of eigenvectors corresponding to all others of their eigenvalues are different from one another, and yet any perturbation to the synchronization manifold (which, by definition, lies in the tangent space) can be expanded as linear combination of one whatever of such eigenvector sets (the relevant consequence is that one can arbitrarily select any of the generalized Laplacians as the reference for the choice of the basis of the transverse space, and all other eigenvector sets will map to such a basis by means of unitary matrix transformations).

We are then fully entitled to take, as reference basis, the one constituted by the eigenvectors of the classic Laplacian $\mathcal{L}^{(1)}$ ($\mathrm{V} = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_N]$), and consider new variables $\delta\eta = (\mathrm{V}^{-1} \otimes \mathrm{I}_m)\delta\mathbf{x}$. We get

$$\dot{\delta\eta} = \left[\mathrm{I}_N \otimes \mathrm{JF} - \sigma_1 \Lambda^{(1)} \otimes \mathrm{JG}^{(1)} - \sigma_2 \tilde{\mathcal{L}}^{(2)} \otimes \mathrm{JG}^{(2)}\right] \delta\eta. \tag{10.16}$$

where we have used the fact that $\mathrm{V}^{-1}\mathcal{L}^{(1)}\mathrm{V} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N) = \Lambda^{(1)}$, where $0 = \lambda_1 < \lambda_2 \leq \ldots \lambda_N$ are the eigenvalues of $\mathcal{L}^{(1)}$, and we have indicated with $\tilde{\mathcal{L}}^{(2)} = \mathrm{V}^{-1}\mathcal{L}^{(2)}\mathrm{V}$ the transformed generalized Laplacian of order 2.

As $\mathcal{L}^{(2)}$ is zero-row sum (i.e. $\mathcal{L}^{(2)}\mathbf{v}_1 = 0$), Eqs. (10.16) may be rewritten as

$$\begin{cases} \dot{\eta}_1 &= \mathrm{JF}\eta_1 \\ \dot{\eta}_i &= (\mathrm{JF} - \sigma_1\lambda_i \mathrm{JG}^{(1)})\eta_i - \sigma_2 \sum_{j=2}^{N} \tilde{\mathcal{L}}_{ij}^{(2)} \mathrm{JG}^{(2)} \eta_j, \end{cases} \tag{10.17}$$

that is, the dynamics of the linearized system is decoupled into two parts: the dynamics of η_1 accounting for the motion along the synchronous manifold, and that of all other variables η_i (with $i = 2, \ldots, N$, representing the different modes transverse to the synchronization manifold) which are coupled each other by means of the coefficients $\tilde{\mathcal{L}}_{ij}^{(2)}$ (all of them being known quantities) given by transforming $\mathcal{L}^{(2)}$ with the matrix that diagonalizes $\mathcal{L}^{(1)}$. The problem of stability is then reduced to: (i) simulating *a single, uncoupled,* nonlinear system; (ii) using the obtained trajectory to feed up the elements of the Jacobians $\mathrm{JG}^{(1)}$ and $\mathrm{JG}^{(2)}$; (iii) simulating the dynamics of a system of $N - 1$ coupled linear equations, and tracking the behavior of the norm $\sqrt{\sum_{i=2}^{N}\sum_{j=1}^{m}(\eta_i^{(j)})^2}$ for the calculation of the maximum Lyapunov exponent (being $\eta_i \equiv (\eta_i^{(1)}, \eta_i^{(2)}, \ldots, \eta_i^{(m)})$).

Stability of the synchronous solution requires, as a necessary condition, that the maximum among the Lyapunov exponents associated to all transverse modes is negative. Therefore, this quantity provides a generalized Master Stability Function, $\Lambda_{\max}$, which, given the node dynamics and the coupling functions, is in general function of the topology of the two body interactions, the topology of the three body interactions, and the two coupling strengths σ_1 and σ_2, i.e., $\Lambda_{\max} = \Lambda_{\max}(\sigma_1, \sigma_2, \mathcal{L}^{(1)}, \mathcal{L}^{(2)})$.

Notice that, in full analogy with the classical MSF approach, also in the case of simplicial complexes one is, therefore, able to separate the motion along the synchronization manifold and that transverse to it. And it is such a crucial separation that ultimately enables the study of stability of the synchronous manifold, retaining the general applicability of the original approach. In the case of simplical complexes, the higher complexity in the structure of the interactions yields a formalism requiring the analysis of a set of coupled differential equations, rather than of a single parametric variational equation (as in the case of network synchronization). In other words, in the fully general case the set of equations describing the motion transverse to the synchronous manifold cannot be further decomposed into independent, decoupled modes, as it happens in the network case; however, the analysis of stability still requires the computation of a single quantity, i.e., the maximum Lyapunov exponent, which has to be performed on such a set of coupled, linear equations. Hence, while in the classical MSF on networks, once fixed the node dynamics and coupling function, one obtains $\Lambda_{\max} = \Lambda_{\max}(\sigma, \mathrm{L})$, which can be further simplified introducing α parametrizing the product of the coupling coefficient and the nonzero eigenvalues of L, i.e., $\Lambda_{\max} = \Lambda_{\max}(\alpha)$, for simplicial complexes, once fixed the node dynamics and the coupling functions, one obtains $\Lambda_{\max} = \Lambda_{\max}(\sigma_1, \sigma_2, \mathcal{L}^{(1)}, \mathcal{L}^{(2)})$. Note, however, that there are still special cases where such an expression can be simplified up to even recover cases where it is formally identical to that of the classical MSF, as we will show explicitly later on in Sec. 10.5.

In analogy with the classification scheme adopted for synchronization in complex networks (Chap. 5 in Ref. [40]), one immediately realizes that, once the local dynamics of each node (i.e. the function $\mathbf{f}$), the various coupling functions $\mathbf{g}^{(1,2)}$, and the structure of the simplicial complex (i.e. $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$) are specified, all possible cases can be divided in three different classes:

1. class I problems, where $\Lambda_{\max}$ is positive in all the half plane ($\sigma_1 \geq 0, \sigma_2 \geq 0$), and therefore synchronization is never stable;
2. class II problems, for which $\Lambda_{\max}$ is negative within a unbounded area of the half plane and
3. class III problems, for which the area of the half plane in which $\Lambda_{\max}$ is negative is instead bounded, and therefore additional instabilities of the synchronous motion may occur at larger values of the coupling strengths.

We conclude this section by showing how the approach described above can be extended to simplicial complexes of any order D. Indeed, each term on the right hand side of Eq. (10.7) can be manipulated following exactly the same steps described above. Defining $\mathrm{JG}^{(d)} = J_1\mathbf{g}^{(d)}(\mathbf{x}^s, \ldots, \mathbf{x}^s) + J_2\mathbf{g}^{(d)}(\mathbf{x}^s, \ldots, \mathbf{x}^s) + \cdots + J_d\mathbf{g}^{(d)}(\mathbf{x}^s, \ldots, \mathbf{x}^s)$, Eq. (10.15) becomes

$$\dot{\delta\mathbf{x}} = \left[\mathrm{I}_N \otimes \mathrm{JF} - \sigma_1\mathcal{L}^{(1)} \otimes \mathrm{JG}^{(1)} - \sigma_2\mathcal{L}^{(2)} \otimes \mathrm{JG}^{(2)} - \ldots - \sigma_D\mathcal{L}^{(D)} \otimes \mathrm{JG}^{(D)}\right]\delta\mathbf{x}. \quad (10.18)$$

Once again, one can select the eigenvector set which diagonalizes $\mathcal{L}^{(1)}$, and to introduce the new variables $\delta\eta = (\mathrm{V}^{-1} \otimes \mathrm{I}_m)\delta\mathbf{x}$. Following the same steps which led us to write Eqs. (10.17), one then obtains

$$\begin{cases} \dot{\eta}_1 = \mathrm{JF}\eta_1, \\ \dot{\eta}_i = (\mathrm{JF} - \sigma_1\lambda_i\mathrm{JG}^{(1)})\eta_i - \sigma_2 \sum_{j=2}^{N} \tilde{\mathcal{L}}_{ij}^{(2)}\mathrm{JG}^{(2)}\eta_j - \ldots - \sigma_D \sum_{j=2}^{N} \tilde{\mathcal{L}}_{ij}^{(D)}\mathrm{JG}^{(D)}\eta_j, \end{cases} \tag{10.19}$$

where the coefficients $\tilde{\mathcal{L}}_{ij}^{(d)}$ result from transforming $\mathcal{L}^{(d)}$ with the matrix that diagonalizes $\mathcal{L}^{(1)}$. As a result, one has the same reduction of the problem to a single, uncoupled, nonlinear system, plus a system of $N-1$ coupled linear equations, from which the maximum Lyapunov exponent $\Lambda_{\max} = \Lambda_{\max}(\sigma_1, \sigma_2, \ldots, \sigma_D, \mathcal{L}^{(1)}, \mathcal{L}^{(2)}, \ldots, \mathcal{L}^{(D)})$ can be extracted and monitored (for each simplicial complex) in the D-dimensional hyper-space of the coupling strength parameters.

10.5 Master Stability Functions

The problem can be greatly simplified in a few special cases in which either the connectivity of the system (i.e. the structure of the simplicial complex), or the coupling functions, allow for a further reduction of complexity. For the sake of illustration, we start considering first the case of $D = 2$, and, then, the extension to any possible order D.

The first special case is the all-to-all coupling, for which every pair of two nodes is connected by a link and every triple of nodes is connected by a 2-simplex, namely all the possible two and three-body interactions are active. In this case, the classical Laplacian matrix is

$$\mathcal{L}_{ij}^{(1)} = \begin{cases} -1 & \text{for } i \neq j \\ N-1 & \text{for } i = j. \end{cases} \tag{10.20}$$

Then, it is easy to rewrite $\mathcal{L}^{(2)}$. First, the off diagonal terms $-\mathcal{L}_{ij}^{(2)}$ $(i \neq j)$ represent the number of triangles formed by the link (i, j) which, in the present case, is simply equal to $N-2$. Second, the terms of the main diagonal $\mathcal{L}_{ii}^{(2)}$ indicates the number of triangles having the node i as a vertex, which is

$$k_i^{(2)} = \binom{N-1}{2} = \frac{(N-1)(N-2)}{2}. \tag{10.21}$$

Consequently, one has that

$$\mathcal{L}^{(2)} = (N-2)\ \mathcal{L}^{(1)}. \tag{10.22}$$

For the linearized dynamics, one gets

$$\dot{\delta \mathbf{x}}_i = \mathrm{JF}\delta \mathbf{x}_i - \sum_{j=1}^{N} \mathcal{L}_{ij}^{(1)} \left[\sigma_1 \, \mathrm{JG}^{(1)} + \sigma_2 \, (N-2) \, \mathrm{JG}^{(2)}\right] \delta \mathbf{x}_j. \tag{10.23}$$

By expanding the perturbation vector $\delta \mathbf{x}$ on the othornormal basis formed by the eigenvectors of the classical Laplacian matrix $\mathcal{L}^{(1)}$, and after noticing that in the all-to-all configuration $\lambda_2 = \ldots \lambda_N = N$, for each η_i (with $i \in \{2, \ldots, N\}$) one has

$$\dot{\eta}_i = [\mathrm{JF} - \sigma_1 N \, \mathrm{JG}^{(1)} - \sigma_2 N(N-2) \, \mathrm{JG}^{(2)}]\eta_i. \tag{10.24}$$

In other words, in the all-to-all case, the variables η_i come out to be all uncoupled to each other, so that the MSF uniquely depends on σ_1, σ_2 and N, i.e., $\Lambda_{\max} = \Lambda_{\max}(\sigma_1, \sigma_2, N)$.

In the more general case of a D-dimensional simplicial complex, it is easy to write the generalized Laplacian of order d as a function of the classical Laplacian matrix. In fact, the number of d-simplices having node i as a vertex and the number of d-simplices formed by the link (i, j) are respectively

$$k_i^{(d)} = \binom{N-1}{d} = \frac{(N-1)(N-2)\ldots(N-d)}{d!} \tag{10.25}$$

and

$$k_{ij}^{(d)} = \binom{N-2}{d-1} = \frac{(N-2)\ldots(N-d)}{(d-1)!}. \tag{10.26}$$

Given the definition of the generalized Laplacian in Eq. (10.5), we find that

$$\mathcal{L}^{(d)} = (N-d) \, \mathcal{L}^{(d-1)} = (N-2)(N-3)\ldots(N-d)\mathcal{L}^{(1)}. \tag{10.27}$$

Once again, one can derive a parametric equation analogous to Eq. (10.24), with the MSF (once fixed both the node dynamics and the coupling functions) which solely depends on the coupling coefficients and the number of nodes, i.e. $\Lambda_{\max} = \Lambda_{\max}(\sigma_1, \sigma_2, \ldots, \sigma_D, N)$

$$\dot{\eta}_i = [\mathrm{JF} - \sigma_1 N \, \mathrm{JG}^{(1)} - \cdots - \sigma_D N(N-2)\ldots(N-D) \, \mathrm{JG}^{(D)}]\eta_i. \tag{10.28}$$

Another interesting case is that of generalized diffusion interactions with *natural* coupling functions. This amounts to consider diffusive coupling functions, given by

$$\begin{aligned} \mathbf{g}^{(1)}(\mathbf{x}_i, \mathbf{x}_j) &= \mathbf{h}^{(1)}(\mathbf{x}_j) - \mathbf{h}^{(1)}(\mathbf{x}_i), \\ \mathbf{g}^{(2)}(\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k) &= \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) - \mathbf{h}^{(2)}(\mathbf{x}_i, \mathbf{x}_i), \end{aligned} \tag{10.29}$$

where $\mathbf{h}^{(1)} : \mathbb{R}^m \longrightarrow \mathbb{R}^m$ and $\mathbf{h}^{(2)} : \mathbb{R}^{2m} \longrightarrow \mathbb{R}^m$. In addition, a condition of natural coupling is considered

$$\mathbf{h}^{(2)}(\mathbf{x}, \mathbf{x}) = \mathbf{h}^{(1)}(\mathbf{x}). \tag{10.30}$$

Equation (10.30) expresses, indeed, the fact that the coupling to node i from two-body and three-body interactions is essentially similar, in that a three-body interaction where two nodes are on the same state is equivalent to a two-body interaction. Here, the MSF assumes a particularly convenient form, as it can be written as a function of a single parameter.

The consequence of Eq. (10.30) is that $J_1\mathbf{h}^{(2)}(\mathbf{x}^s, \mathbf{x}^s) + J_2\mathbf{h}^{(2)}(\mathbf{x}^s, \mathbf{x}^s) = J\mathbf{h}^{(1)}(\mathbf{x}^s)$. Therefore, one has

$$\dot{\delta\mathbf{x}_i} = J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i - \sum_{j=1}^{N}\left[\sigma_1\mathcal{L}_{ij}^{(1)} + \sigma_2\mathcal{L}_{ij}^{(2)}\right] J\mathbf{h}^{(1)}(\mathbf{x}^s)\delta\mathbf{x}_j. \tag{10.31}$$

Alternatively, one can consider the zero-row-sum, symmetric, *effective matrix* $\mathcal{M}$, given by

$$\mathcal{M} = \mathcal{L}^{(1)} + r\mathcal{L}^{(2)}, \qquad r = \frac{\sigma_2}{\sigma_1}. \tag{10.32}$$

The eigenvalues of $\mathcal{M}$ depend on the ratio r of the coupling coefficients, and one has that

$$\dot{\delta\mathbf{x}_i} = J\mathbf{f}(\mathbf{x}^s)\delta\mathbf{x}_i - \sigma_1\sum_{j=1}^{N}\mathcal{M}_{ij}\, J\mathbf{h}^{(1)}(\mathbf{x}^s)\delta\mathbf{x}_j. \tag{10.33}$$

Equation (10.33) allows to establish a formal full analogy between the case of a simplicial complex and that of a network with weights given by the coefficients of the effective matrix $\mathcal{M}$. In this case, by diagonalizing the effective matrix $\mathcal{M}$, the transverse modes can be fully decoupled and a single-parameter MSF can be defined, starting from the following m-dimensional linear parametric variational equation

$$\dot{\eta} = \left[J\mathbf{f}(\mathbf{x}^s) - \alpha J\mathbf{h}^{(1)}(\mathbf{x}^s)\right]\eta \tag{10.34}$$

from which the maximum Lyapunov exponent is calculated: $\Lambda_{\max} = \Lambda_{\max}(\alpha)$ with $\alpha = \lambda(\sigma_1\mathcal{L}^{(1)} + \sigma_2\mathcal{L}^{(2)})$ or $\alpha = \sigma_1\lambda(\mathcal{L}^{(1)} + r\mathcal{L}^{(2)}) = \sigma_1\lambda(\mathcal{M})$. The situation is therefore conceptually equivalent to that of synchronization in complex networks, with the effective matrix $\mathcal{M}$ playing the same role of the classical Laplacian: given the dynamical system $\mathbf{f}$, the coupling functions $\mathbf{h}^{(1)}$ and $\mathbf{h}^{(2)}$, and the structure of connection of the simplicial complex (i.e. $\mathcal{L}^{(1)}$ and $\mathcal{L}^{(2)}$) one can define three possible classes of problems: 1) class I problems, for which the curve $\Lambda_{\max} = \Lambda_{\max}(\alpha)$ does not intercept the abscissa and it is always positive. In this case synchronization is always forbidden, no matter which simplicial complex is used for connecting the dynamical systems; 2) class II problems, for which the curve $\Lambda_{\max} = \Lambda_{\max}(\alpha)$ intercepts the abscissa only once at α_c, and for which, therefore, the synchroniza-

tion threshold is given by the self consistent equation $\sigma_1^c = \alpha_c/\lambda_2[\mathcal{M}(\sigma_1^c, \sigma_2^c)]$, i.e. it scales with the inverse of the second smallest eigenvalue of the effective matrix; 3) class III problems, for which the curve $\Lambda_{\max} = \Lambda_{\max}(\alpha)$ intercepts the abscissa twice at α_1 and $\alpha_2 > \alpha_1$. In this case, synchronization can be observed only if the entire eigenvalue spectrum of the effective matrix is such that $\sigma_1\lambda_2(\mathcal{M}) > \alpha_1$ and, at the same time, $\sigma_1\lambda_N(\mathcal{M}) < \alpha_2$. In this case, the ratio $\lambda_2(\mathcal{M})/\lambda_N(\mathcal{M})$ can be considered as a proxy measure of *synchronizability* of the simplicial complex, in that the closer is such a parameter to unity (the more compact is the spectrum of eigenvalues of $\mathcal{M}$) the larger can be the range of coupling strengths for which the two above synchronization conditions can be satisfied.

We have so far considered the case of $D = 2$. In the fully general scenario, the condition for natural coupling is given by

$$\mathbf{h}^{(D)}(\mathbf{x}, \ldots, \mathbf{x}) = \ldots = \mathbf{h}^{(2)}(\mathbf{x}, \mathbf{x}) = \mathbf{h}^{(1)}(\mathbf{x}). \tag{10.35}$$

The equation for the MSF is formally analogous to Eq. (10.34), where now $\alpha = \sigma_1\lambda_2(\mathcal{M}^{(D)})$ parametrizes the eigenvalues of the effective matrix of order D

$$\mathcal{M}^{(D)} = \mathcal{L}^{(1)} + \frac{\sigma_2}{\sigma_1}\mathcal{L}^{(2)} + \ldots + \frac{\sigma_D}{\sigma_1}\mathcal{L}^{(D)}. \tag{10.36}$$

10.6 Numerical Results

In this section we show some numerical results confirming the validity of the proposed approach. In particular, we focus on a paradigmatic example of three-dimensional ($\mathbf{x} = (x, y, z)^T \in \mathbb{R}^3$) chaotic systems, i.e., the Rössler oscillator [41], and consider the case of natural coupling with $D = 3$.

In this case, as discussed in Sec. 10.5, in full analogy with what occurs for networks, the MSF is a function of a single parameter, i.e., $\Lambda_{\max} = \Lambda_{\max}(\alpha)$ with $\alpha = \lambda(\sigma_1\mathcal{L}^{(1)} + \sigma_2\mathcal{L}^{(2)} + \sigma_3\mathcal{L}^{(3)})$. This enables the study of synchronization stability into two steps, one pertaining only to the node dynamics and coupling functions, providing $\Lambda_{\max} = \Lambda_{\max}(\alpha)$, and a second step, where the condition $\Lambda_{\max}(\alpha) < 0$ is checked at the points $\alpha = \{\sigma_1\lambda_2(\mathcal{M}), \ldots, \sigma_1\lambda_N(\mathcal{M})\}$.

Here, we have calculated the MSF for the Rössler oscillator with several choices of the coupling functions[1]:

[1] For the calculation of the MSFs we have used the algorithm for the computation of the entire spectrum of Lyapunov exponents in Ref. [42] (with parameters: integration step size of the Euler algorithm $\delta t = 10^{-5}$, length of the simulation $L = 2500$, windows of averaging $T = 0.9L$).

$$\begin{aligned}
\mathbf{h}^{(1)}(\mathbf{x}_j) &= [x_j^3, 0, 0]^T, \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) = [x_j^2 x_k, 0, 0]^T, \mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) \\
&= [x_j x_k x_h, 0, 0]^T; \mathbf{h}^{(1)}(\mathbf{x}_j) = [0, x_j^3, 0]^T, \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) \\
&= [0, x_j^2 x_k, 0]^T, \mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) = [0, x_j x_k x_h, 0]^T; \mathbf{h}^{(1)}(\mathbf{x}_j) \\
&= [0, 0, x_j^3]^T, \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) = [0, 0, x_j^2 x_k]^T, \mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) \\
&= [0, 0, x_j x_k x_h]^T; \mathbf{h}^{(1)}(\mathbf{x}_j) = [y_j^3, 0, 0]^T, \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) \\
&= [y_j^2 y_k, 0, 0]^T, \mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) = [y_j y_k y_h, 0, 0]^T \ldots \text{ and } \mathbf{h}^{(1)}(\mathbf{x}_j) \\
&= [0, 0, z_j^3]^T, \mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) = [0, 0, z_j^2 z_k]^T, \mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) \\
&= [0, 0, z_j z_k z_h]^T.
\end{aligned}$$

As an example, when the coupling functions are fixed as $\mathbf{h}^{(1)}(\mathbf{x}_j) = [x_j^3, 0, 0]^T$, $\mathbf{h}^{(2)}(\mathbf{x}_j, \mathbf{x}_k) = [x_j^2 x_k, 0, 0]^T$, and $\mathbf{h}^{(3)}(\mathbf{x}_j, \mathbf{x}_k, \mathbf{x}_h) = [x_j x_k x_h, 0, 0]^T$, the equations governing the simplicial complex of Rössler systems read

$$\begin{aligned}
\dot{x}_i &= -y_i - z_i + \sigma_1 \sum_{j=1}^{N} a_{ij}^{(1)} (x_j^3 - x_i^3) \\
&\quad + \sigma_2 \sum_{j=1}^{N} \sum_{k=1}^{N} a_{ijk}^{(2)} (x_j^2 x_k - x_i^3) + \sigma_3 \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{h=1}^{N} a_{ijkh}^{(3)} (x_j x_k x_h - x_i^3), \qquad (10.37) \\
\dot{y}_i &= x_i + a y_i, \\
\dot{z}_i &= b + z_i (x_i - c),
\end{aligned}$$

where $i = 1, \ldots, N$, and the parameters have been fixed so that the resulting dynamics is chaotic, i.e., $a = b = 0.2$, $c = 9$.

It is interesting to note that the MSFs characterizing simplicial complexes of Rössler oscillators (Fig. 10.2) exhibit a variety of behaviors that actually encompass all possible classes of MSF. In particular, we find one class III example (Fig. 10.2a), one class II example (Fig. 10.2e), while all remaining cases do correspond to class I.

Let us now consider the simplicial complex with $N = 10$ nodes represented in Fig. 10.1c and simulate Eqs. (10.37) on top of this structure. Let us first fix $\sigma_3 = 0.1$ and integrate[2] Eqs. (10.37) for different values of σ_1 and σ_2. For each configurations of the coupling parameters, the state of the system is monitored by the average synchronization error defined as $E = \langle \left(\frac{1}{N(N-1)} \sum_{i,j=1}^{N} \|\mathbf{x_j} - \mathbf{x_i}\|^2 \right)^{\frac{1}{2}} \rangle_T$, where T is a sufficiently large window of time where the synchronization error is averaged, after discarding the transient. As it is shown in Fig. 10.3a, which illustrates $E(\sigma_1, \sigma_2)$ along with the theoretical predictions provided by the MSF, the numerical simulations are in very good agreement with the theoretical predictions for the synchronization thresholds. This is an example of a class III system, with a synchronization region that is bounded. In particular, it is interesting to note that synchronization

[2] Numerical integrations have been performed by means of an Euler algorithm, with integration step $\delta t = 10^{-4}$, in a windows of time equal to $2T$ with $T = 500$.

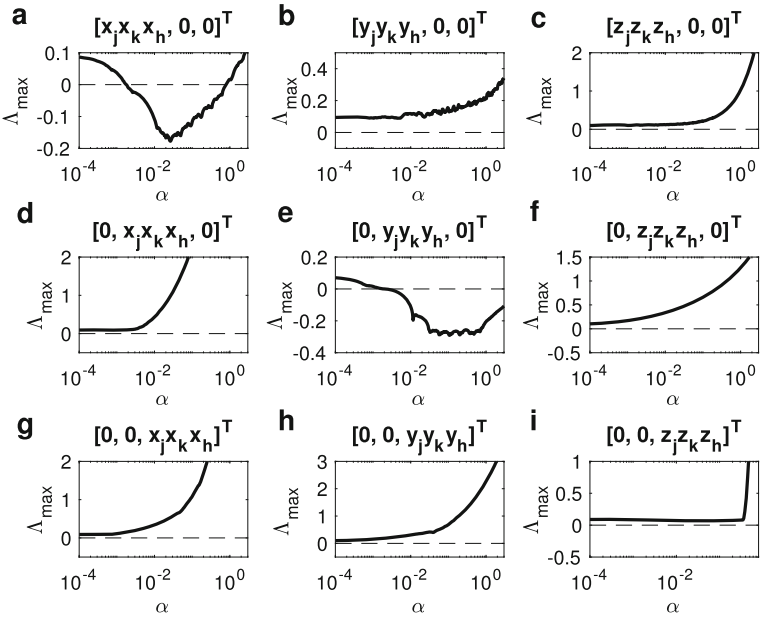


Fig. 10.2 Synchronization in simplicial complexes of Rössler oscillators, in the case of natural coupling, is characterized by a master stability function $\Lambda_{\max}(\alpha)$, here obtained for several coupling functions. On the top of each panel, the expression used for $\mathbf{h}^{(3)}$ is reported

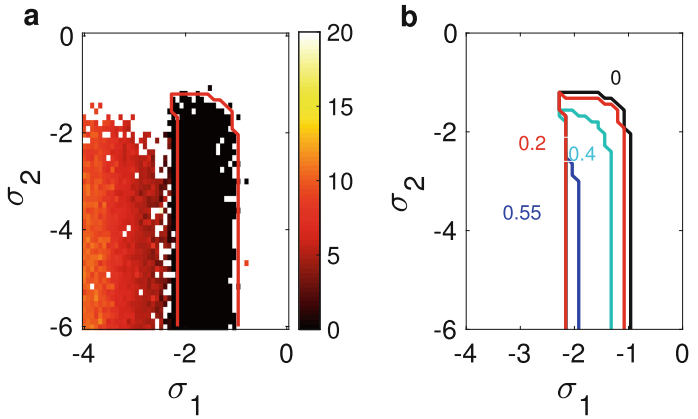


Fig. 10.3 Synchronization in the simplicial complex of Rössler oscillators of Fig. 10.1c. **a** Contour plot of the time averaged (over an observation time $T = 500$) synchronization error E in the plane (σ_1, σ_2) for $\sigma_3 = 0.1$. The red continuous line is the theoretical prediction of the synchronization thresholds obtained from Eq. (10.34). **b** Synchronization region as predicted from Eq. (10.34) at different values of the coupling strength σ_3

is impossible to achieve for very small values of σ_1, suggesting that in this structure interactions through links are essential for synchronization. To study the role of four-bodies interaction we have then varied σ_3 and illustrated the predictions of the MSF obtained from Eq. (10.34) in Fig. 10.3b. We observe that, due to the fact that configuration under analysis is class III, increasing σ_3 actually decreases the area of the synchronization region, in particular, lowering the threshold (for both σ_1 and σ_2) from synchronization to desynchronization.

10.7 Conclusions

In a complex system consisting of many coupled dynamical units, collective behaviors are shaped by the functional form and by the architecture of the interactions. To account for the most general type of interactions, we have here leveraged the mathematical formalism of simplicial complexes to formulate a general model that include many-body interactions of arbitrary order among dynamical systems of arbitrary nature. Assuming the non-invasiveness of the coupling functions, we have shown that is possible to derive necessary conditions for stable synchronous regime in a simplicial complex. Remarkably, these conditions depend on generalized Laplacian matrices that map the effects of high-order interactions. For specific types of structures (e.g., all-to-all interactions) and couplings (that we named natural couplings), this approach ultimately provides a Master Stability Function, which formalizes the interplay between the dynamics of the single units and the topology of the simplicial complex.

The generality of the introduced framework and of the assumptions considered make it applicable in a wide range of scenarios, so that we expect that our method could be used to produce a-priori predictions on the emergence of synchronization in many diverse theoretical and practical cases. In particular, our study only focused on the regime of synchronization where all the units follow the same trajectory. However, many other different forms of synchronization exist, including cluster synchronization, chimera and Bellerophon states, remote synchronization, and so on. It would be particularly intriguing to investigate the emergence of such states, or even of novel ones, in structures that do not include exclusively pairwise couplings, but also incorporate other, types of higher-order interaction mechanisms, such as the simplicial complexes considered here.

Acknowledgements F.D.P., S.L. and S.B. acknowledge funding from the project EXPLICS granted by the Italian Ministry of Foreign Affairs and International Cooperation.
R.C. and M.R. acknowledge funding from the project PGC2018-101625-B-I00, granted by the Spanish Ministry of Science and Innovation.
L.V.G. and M.F. acknowledge the support of the Italian Ministry for Research and Education through the Research Program PRIN 2017 (Grant 2017CWMF93, project VECTORS).
V.L. acknowledges support from the Leverhulme Trust Research Fellowship "CREATE: The Network Components of Creativity and Success".

References

1. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: a universal concept in nonlinear sciences*, vol. 12 (Cambridge University Press, 2003)
2. S. Boccaletti, A.N. Pisarchik, C.I. Del Genio, A. Amann, *Synchronization: From Coupled Systems to Complex Networks* (Cambridge University Press, 2018)
3. M. Barahona, L.M. Pecora, Phys. Rev. Lett. **89**(5), 054101 (2002)
4. M. Chavez, D.U. Hwang, A. Amann, H.G.E. Hentschel, S. Boccaletti, Phys. Rev. Lett. **94** (2005). https://doi.org/10.1103/PhysRevLett.94.218701
5. C.I. del Genio, J. Gómez-Gardeñes, I. Bonamassa, S. Boccaletti, Sci. Adv. **2**(11) (2016). https://doi.org/10.1126/sciadv.1601679
6. R. Gutiérrez, A. Amann, S. Assenza, J. Gómez-Gardenes, V. Latora, S. Boccaletti, Phys. Rev. Lett. **107**(23), 234103 (2011)
7. V. Avalos-Gaytán, J.A. Almendral, I. Leyva, F. Battiston, V. Nicosia, V. Latora, S. Boccaletti, Phys. Rev. E **97**, 042301 (2018). https://doi.org/10.1103/PhysRevE.97.042301
8. L.V. Gambuzza, A. Cardillo, A. Fiasconaro, L. Fortuna, J. Gómez-Gardenes, M. Frasca, Chaos Interdisc. J. Nonlinear Sci. **23**(4), 043103 (2013)
9. V. Nicosia, M. Valencia, M. Chavez, A. Díaz-Guilera, V. Latora, Phys. Rev. Lett. **110**(2013). https://doi.org/10.1103/PhysRevLett.110.174102
10. L.M. Pecora, F. Sorrentino, A.M. Hagerstrom, T.E. Murphy, R. Roy, Nat. Commun. **5**(1), 1 (2014)
11. L.V. Gambuzza, M. Frasca, V. Latora, IEEE Trans. Autom. Control **64**(1), 365 (2018)
12. D.M. Abrams, S.H. Strogatz, Phys. Rev. Lett. **93**(17), 174102 (2004)
13. M.J. Panaggio, D.M. Abrams, Nonlinearity **28**(3), R67 (2015)
14. H. Bi, X. Hu, S. Boccaletti, X. Wang, Y. Zou, Z. Liu, S. Guan, Phys. Rev. Lett. **117**(20), 204101 (2016)
15. C. Xu, S. Boccaletti, S. Guan, Z. Zheng, Phys. Rev. E **98**(5), 050202 (2018)
16. F. Di Patti, D. Fanelli, F. Miele, T. Carletti, Chaos. Solitons & Fractals **96**, 8 (2017)
17. F. Di Patti, D. Fanelli, F. Miele, T. Carletti, Commun. Nonlinear Sci. Numer. Simul. **56**, 447 (2018)
18. G. Cencetti, F. Bagnoli, G. Battistelli, L. Chisci, F. Di Patti, D. Fanelli, Euro. Phys. J. B **90**, 9 (2017)
19. S. Boccaletti, J. Almendral, S. Guan, I. Leyva, Z. Liu, I. Sendiña-Nadal, Z. Wang, Y. Zou, Phys. Rep. **660**, 1 (2016)
20. A. Krawiecki, Chaos. Solitons & Fractals **65**, 44 (2014)
21. J.A. Acebrón, L.L. Bonilla, C.J.P. Vicente, F. Ritort, R. Spigler, Rev. Mod. Phys. **77**(1), 137 (2005). https://doi.org/10.1103/RevModPhys.77.137
22. F.A. Rodrigues, T.K.D. Peron, P. Ji, J. Kurths, Phys. Rep. **610**, 1 (2016)
23. T. Tanaka, T. Aoyagi, Phys. Rev. Lett. **106**(22) (2011). https://doi.org/10.1103/PhysRevLett.106.224101
24. P.S. Skardal, A. Arenas, Phys. Rev. Lett. **122**(24), 248301 (2019)
25. P.S. Skardal, A. Arenas, Commun. Phys. **3**, 218 (2020). https://doi.org/10.1038/s42005-020-00485-0
26. A.P. Millán, J.J. Torres, G. Bianconi, Phys. Rev. Lett. **124**, 218301 (2020). https://doi.org/10.1103/PhysRevLett.124.218301
27. M. Lucas, G. Cencetti, F. Battiston, Phys. Rev. Res. **2**, 033410 (2020). https://doi.org/10.1103/PhysRevResearch.2.033410
28. L.M. Pecora, T.L. Carroll, Phys. Rev. Lett. **80**(10), 2109 (1998)
29. J. Sun, E.M. Bollt, T. Nishikawa, EPL (Europhysics Letters) **85**(6), 60011 (2009)
30. D.J. Stilwell, E.M. Bollt, D.G. Roberson, SIAM J. Appl. Dyn. Syst. **5**(1), 140 (2006)
31. M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, S. Boccaletti, Phys. Rev. Lett. **100**(4), 044102 (2008)
32. J. Zhou, Y. Zou, S. Guan, Z. Liu, S. Boccaletti, Sci. Rep. **6**(1), 1 (2016)

33. Y. Zhang, V. Latora, A.E. Motter, Commun. Phys. **4**, 195 (2021). https://doi.org/10.1038/s42005-021-00695-0
34. V. Latora, V. Nicosia, G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, 2017)
35. I. Iacopini, G. Petri, A. Barrat, V. Latora, Nat. Commun. **10**(1), 2485 (2019)
36. C. Berge, *Graphs and hypergraphs*. North-Holl Math. Libr. (North-Holland, Amsterdam, 1973)
37. K.F. Kee, L. Sparks, D.C. Struppa, M. Mannucci, Commun. Q. **61**(1), 35 (2013)
38. U. Alvarez-Rodriguez, F. Battiston, G.F. de Arruda, Y. Moreno, M. Perc, V. Latora, Nat. Hum. Behav. **5**, 586–595 (2021). https://doi.org/10.1038/s41562-020-01024-1
39. O.T. Courtney, G. Bianconi, Phys. Rev. E **93**(2016). https://doi.org/10.1103/PhysRevE.93.062311
40. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Phys. Rep. **424**(4–5), 175 (2006)
41. O.E. Rössler, Phys. Lett. A **57**(5), 397 (1976)
42. A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Physica D Nonlinear Phenomena **16**(3), 285 (1985)

Chapter 11
Geometry, Topology and Simplicial Synchronization

Ana Paula Millán, Juan G. Restrepo, Joaquín J. Torres, and Ginestra Bianconi

Abstract Simplicial synchronization reveals the role that topology and geometry have in determining the dynamical properties of simplicial complexes. Simplicial network geometry and topology are naturally encoded in the spectral properties of the graph Laplacian and of the higher-order Laplacians of simplicial complexes. Here we show how the geometry of simplicial complexes induces spectral dimensions of the simplicial complex Laplacians that are responsible for changing the phase diagram of the Kuramoto model. In particular, simplicial complexes displaying a non-trivial simplicial network geometry cannot sustain a synchronized state in the infinite network limit if their spectral dimension is smaller or equal to four. This theoretical result is here verified on the Network Geometry with Flavor simplicial complex generative model displaying emergent hyperbolic geometry. On its turn simplicial topology is shown to determine the dynamical properties of the higher-order Kuramoto model. The higher-order Kuramoto model describes synchronization of topological signals, i.e., phases not only associated to the nodes of a simplicial complexes but associated also to higher-order simplices, including links, triangles and so

A. P. Millán
Department of Clinical Neurophysiology and MEG Center, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, Amsterdam, The Netherlands
e-mail: a.p.millanvidal@amsterdamumc.nl

J. G. Restrepo
Department of Applied Mathematics, University of Colorado at Boulder, Boulder, CO 80309, USA
e-mail: juanga@colorado.edu

J. J. Torres
Departamento de Electromagnetismo y Física de la Materia and Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, 18071 Granada, Spain
e-mail: jtorres@onsager.ugr.es

G. Bianconi (✉)
School of Mathematical Sciences, Queen Mary University of London, The Alan Turing Institute, The British Library, London, UK
e-mail: ginestra.bianconi@gmail.com

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_11

on. This model displays discontinuous synchronization transitions when topological signals of different dimension and/or their solenoidal and irrotational projections are coupled in an adaptive way.

11.1 Introduction

The interplay between structure and dynamics of complex networks [1–3] has been at the forefront of network theory since the beginning of the field. In this context it has been found that the combinatorial and statistical properties of complex networks have unexpected effects on dynamics. For instance, a scale-free degree distribution changes the phase diagram of a wide range of dynamical processes including percolation, epidemic spreading, and the Ising model. The recent interest on higher-order networks [4–7] provides an opportunity to bring a fresh perspective to this subject. Indeed, higher-order networks, and in particular simplicial complexes, constitute the ideal mathematical framework to capture the simplicial network topology and geometry of data. Here we reveal that the network topology and geometry of simplicial complexes can be crucial to define higher-order dynamics. The dynamical process considered in this chapter is synchronization [8–10], captured by the Kuramoto model [11] and the recently introduced higher-order Kuramoto model [12]. The dynamical properties of these dynamical processes will be shown to be highly dependent on the spectral properties [13, 14] of the simplicial complexes [15–19] on which they are defined. The main message of this chapter is summarized in Fig. 11.1, which highlights the role of network topology and network geometry in shaping higher-order network dynamics. In particular, in this chapter we will disclose how the spectral properties of simplicial complexes are foundational to reveal the relation higher-order network geometry, topology and dynamics. Note that while our approach to simplicial synchronization is based on simplicial network geometry and topology, other approaches based on a combinatorial definition of higher-order interactions have been pursued in the literature [20–22], as covered by the Skardal and Arenas chapter of this book.

11.2 Simplicial Complex Models

Simplicial network models are ideal to test, in a well-controlled setting, the interplay between simplicial network geometry, topology and dynamics. Here we focus in particular on two large classes of simplicial complex models with very distinct structural properties: the Network Geometry with Flavor (NGF) [16–19] and the configuration model of simplical complexes [15] (see schematic illustrations of the two models in Fig. 11.2). These models are implemented in codes available at the repository [23].

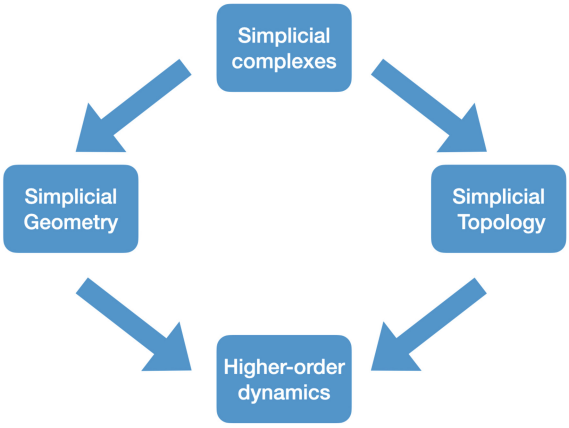


Fig. 11.1 Simplicial complexes encode the rich simplicial network topology and geometry of data and models, which strongly affects the higher-order dynamics. In this chapter we will see how this interplay between structure and dynamics can enrich our understanding of synchronization dynamics defined on simplicial complexes

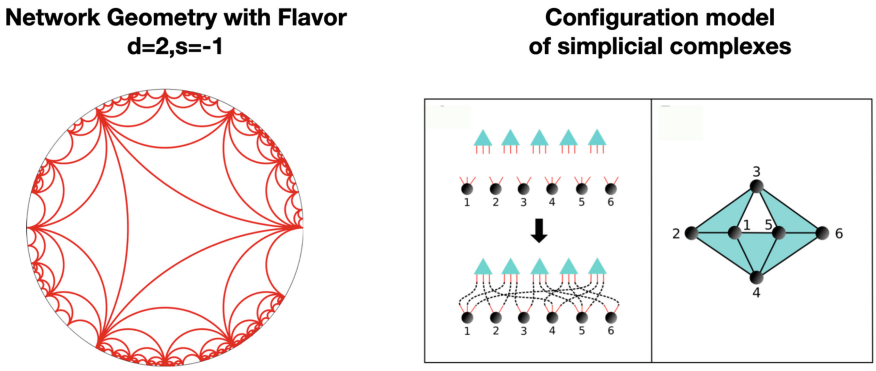


Fig. 11.2 Schematic representation of the two classes of simplicial complex models considered in this work. The Network Geometry with Flavor is a model of growing simplicial complexes describing emergent hyperbolic network geometries. The left panel shows a realization of the NGF simplicial complex of dimension $d = 2$ and flavor $s = -1$. The configuration model of simplicial complexes (right panel) is a maximum entropy model enforcing a given sequence of generalized degrees of the nodes. Right panel reprinted with permission from Ref. [15] ©Copyright (2016) by the American Physical Society

11.2.1 The Network Geometry with Flavor (NGF)

The Network Geometry with Flavor (NGF) [16–19] is a general mathematical framework for growing simplicial complexes that displays emergent hyperbolic network geometry. In other words, the NGF model generates simplicial complexes with hyperbolic geometry that evolve following purely combinatorial and stochastic rules that do not make any use of the natural hyperbolic embedding of the simplicial complexes.

The NGFs are simplicial complexes characterized by two main parameters: the dimension of the simplicial complex d and the *flavor* s which is a parameter that takes values $s \in \{-1, 0, 1\}$. The NGFs are generated by a dynamical process which starting at time $t = 1$ from a single d-dimensional simplex proceeds at each time $t > 1$ by adding a new d-dimensional simplex to the simplicial complex. The new d-dimensional simplex includes one new node and is glued to a $(d-1)$-dimensional face α of the existing simplicial complex chosen according to the probability

$$\Pi_\alpha = \frac{1 - s + s k_{d,d-1}(\alpha)}{\sum_{\alpha'} 1 - s + s k_{d,d-1}(\alpha')}, \tag{11.1}$$

where $k_{d,d-1}(\alpha)$ indicates the generalized degree [15] of a $(d-1)$-dimensional face α, i.e., the number of d-dimensional simplices incident to the $(d-1)$-dimensional face α. This model generates emergent hyperbolic simplicial complexes which satisfy Gromov criteria [24] of hyperbolicity and are δ-hyperbolic with $\delta = 1$ for every value of the flavor s [25]. Moreover, for flavor $s = -1$ the generated simplicial complexes form d-dimensional hyperbolic manifolds [17]. The network skeleton of the NGFs are small world, display hierarchical community structure and are scale-free for $d \geq 2 - s$ [16–18]. Interestingly, in the case $d = 1$ and $s = 1$ the NGF reduces to the Barabási-Albert model, and for $d = 1$ and $s = -1$ the NGF reduces to random Apollonian networks.

This model can be generalized in different directions. Instead of considering simplicial complexes, one can use a similar model to generate cell complexes by gluing together convex regular polytopes [18]. Another possibility is to consider weighted simplicial complexes or to allow any new node to be incident to more than one d-dimensional simplex [19]. Finally, the faces can be assigned a fitness that can be used to modulate the attachment probability Π_α causing topological phase transitions for certain fitness distributions [16, 17].

11.2.2 Configuration Model of Simplicial Complexes

The configuration model of simplicial complexes [15] is a maximum entropy model of pure d-dimensional simplicial complexes. A pure d-dimensional simplicial complex $\mathcal{K}$ can be fully encoded in a $(d+1)$-dimensional adjacency tensor indicating the presence of each d-dimensional facet of the simplicial complex. In particular adjacency tensor $\mathbf{a}$ has elements $a_\alpha = 1$ if the d-dimensional simplex α is present in the simplicial complex, otherwise $a_\alpha = 0$.

The configuration model of simplicial complexes [15] is the least biased ensemble of simplicial complexes with a given generalized degree sequence of the nodes

$$\mathbf{k}_{d,0} = \{k_{d,0}(1), k_{d,0}(2), \dots k_{d,0}(N)\} \tag{11.2}$$

where $k_{d,0}(i)$ indicates the generalized degree of the node i, i.e., the number of d-dimensional complexes of the node $i \in \{1, 2, \dots, N\}$.

The configuration model of simplicial complexes is fully characterized by the probability $P(\mathcal{K})$ assigned to each pure d-dimensional simplicial complex $\mathcal{K}$ of N nodes. The probability $P(\mathcal{K})$ maximizes the entropy S of the simplicial complex ensemble

$$S = -\sum_{\mathcal{K}} P(\mathcal{K}) \ln P(\mathcal{K}), \tag{11.3}$$

given the constrain that each node i has generalized degree $k_{d,0}(i)$, i.e.,

$$\sum_{\alpha \in Q_d(N) | i \subset \alpha} a_\alpha = k_{d,0}(i), \tag{11.4}$$

where $Q_d(N)$ indicates the set of all possible d-dimensional simplices of a simplicial complex formed by N nodes. Therefore the configuration model of simplicial complexes is characterized by the uniform distribution

$$P(\mathcal{K}) = \frac{1}{\mathcal{N}} \prod_{i=1}^{N} \delta \left(\sum_{\alpha \in Q_d(N) | i \subset \alpha} a_\alpha, k_{d,0}(i) \right), \tag{11.5}$$

where here and in the following $\delta(a, b)$ indicates the Kronecker delta. In Ref. [15] the Authors proposed an algorithm for sampling simplicial complexes from this ensemble. This algorithm [23] uses the mapping of simplicial complexes to factor graphs. The configuration model of simplicial complexes is a very valuable null model of simplicial complexes, and provides an ideal benchmark to study dynamical processes on higher-order networks.

11.3 Laplacians

11.3.1 Graph Laplacian

The graph Laplacian describes linear diffusion on a network and it is an important operator that encodes the interplay between network structure and dynamics [3, 26]. The graph Laplacian can also capture the underlying network geometry of the skeleton of a simplicial complex, i.e., of the network obtained from the simplicial complex by retaining only its nodes and links.

The graph Laplacian $\boldsymbol{L}_{[0]}$ is the discrete version of the Laplacian operator defined on continuous functions. The graph Laplacian of a network with $N_{[0]}$ nodes is a $N_{[0]} \times N_{[0]}$ matrix of elements

$$\left[L_{[0]}\right]_{ij} = k_i \delta(i, j) - a_{ij}, \tag{11.6}$$

where k_i is the degree of node i and a_{ij} are the elements of the adjacency matrix of the network. In some cases it is useful to consider a generalization of this operator called the normalized Laplacian $\hat{\boldsymbol{L}}_{[0]}$ that has elements

$$\left[\hat{L}_{[0]}\right]_{ij} = \delta(i, j) - \frac{a_{ij}}{k_i}. \tag{11.7}$$

For instance, the normalized Laplacian is commonly employed to describe random walk dynamics on a network.

Both the standard and the normalized Laplacians have real eigenvalues $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N_{[0]}}$. The density of eigenvalues is described by the spectral density,

$$\rho(\lambda) = \frac{1}{N_{[0]}} \sum_{i=1}^{N_{[0]}} \tilde{\delta}(\lambda - \lambda_i), \tag{11.8}$$

where $\tilde{\delta}\,(x)$ indicates the delta function. The eigenvectors of the normalized Laplacian also encode relevant properties of the underlying network.

11.3.2 Spectral Dimension

For many complex networks, the smallest non-zero eigenvalue (also called Fiedler eigenvalue) λ_2 remains finite as the network size increases. In this case, the network is said to have a *spectral gap*. On the contrary, if $\lambda_2 \to 0$ as $N \to \infty$, and the density of eigenvalues $\rho(\lambda)$ scales as

$$\rho(\lambda) \propto \lambda^{d_S^{[0]}/2-1}, \tag{11.9}$$

for $\lambda \ll 1$, the network is said to present a *spectral dimension* $d_S^{[0]}$ [26–28]. The spectral dimension can be interpreted as the perceived dimension of the network by diffusion processes, and it is a notable feature of networks with an underlying geometrical nature. This is a definition of dimension that is alternative to the Hausdorff dimension d_H characterizing the scaling of the diameter D of a network with the network size N, i.e., $D \propto N^{1/d_H}$.

For Euclidean lattices of dimension d, the spectral dimension coincides with the Hausdorff dimension and we have $d_S^{[0]} = d_H = d$. However, in general the spectral

dimension of a network skeleton of a simplicial complex does not have to coincide with the topological dimension of the simplicial complex, nor with its Hausdorff dimension [25, 27, 29, 30].

While Euclidean lattices display a finite spectral dimension, random graphs and the configuration model of networks are instead characterized by a finite spectral gap. The presence of a finite spectral gap indicates the mean-field nature of the network interactions, and the absence of a clear notion of locality for these networks. Interestingly, also the configuration model of simplicial complexes displays a finite spectral gap.

Remarkably, the emergent network geometry of NGF reveals itself on their significant spectral properties. Indeed the NGF network skeletons, together with other small-world models of simplicial complexes [31], have a finite spectral dimension whose value can be tuned according to the different control parameters [18] although the NGFs are small world, i.e., they have an infinite Hausdorff dimension $d_H = \infty$.

For $s = -1$, NGF networks formed purely by d-dimensional simplices have $d_S^{[0]} \sim d$ for $d \in \{2, 3, 4\}$ as shown in Fig. 11.3 [33]. More generally, for NGFs formed by regular polytopes, $d_S^{[0]}$ increases with the dimension d of the polytopes, and it saturates for hypercubes and orthoplexes ($d_S^{[0]} \leq 3$) [32]. It was shown in Ref. [18] that a similar trend is observed for different flavours of the NGF networks: $d_S^{[0]}$

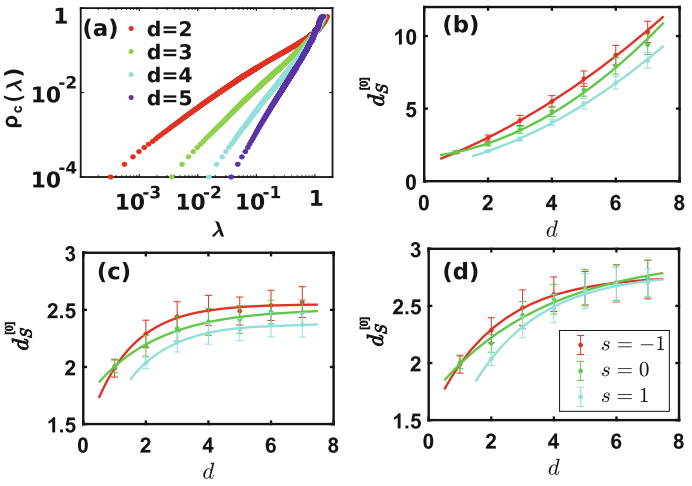


Fig. 11.3 Spectral dimension of NGF networks. Panel (**a**) The cumulative distribution $\rho_c(\lambda)$ of eigenvalues of the NGF with flavor $s = -1$ is shown for dimension $d = 2, 3, 4, 5$. The power-law scaling of $\rho_c(\lambda)$ observed for small values of λ indicates that the skeleton of NGF has a finite spectral dimension. Panels (**b–d**) The fitted spectral dimension $d_S^{[0]}$ of the skeleton of NGF simplicial and cell complexes being formed by simplices (panel **b**), hypercubes (panel **c**) and orthoplexes (panel **d**), is shown for values of the flavor $s \in \{-1, 0, 1\}$ as indicated in the legend of panel (**d**). Lines indicate best fit of the $d_S^{[0]}$ versus d dependence using parabolic (panel **b**) and exponential (panels **c, d**) functional forms. Data from Ref. [32] and Ref.[18]. Details of the fits are described in Ref. [18]

grows quadratically with d for simplicial NGF networks, whereas it saturates at a value $d_S^{[0]} = \bar{d}_S$ with $2 \leq \bar{d}_S \leq 3$ for hypercube and orthoplex NGF networks. Moreover, it was shown recently in Ref. [25] that, in a generalization of the NGF model in which different polytopes are glued together in the same higher-order network, the spectral dimension of the network skeleton can be continuously tuned as a function of the fraction of simplexes in the cell complex.

Thus, not only the dimension of the building blocks shapes the spectral dimension of the networks, but the specific nature and symmetry of these building blocks also play a role in the emerging spectral dimension of the network skeleton (see Fig. 11.3).

11.3.3 Higher-Order Laplacians

Important topological aspects of simplicial complexes are reflected in the spectral properties of the higher-order Laplacians that generalize the graph Laplacian to describe diffusion that occurs among higher-order simplices. The graph Laplacian describes diffusion occurring among nodes connected by links. Similarly the n-th order up-Laplacian describes the diffusion occurring among n-dimensional simplices connected by $(n+1)$-dimensional simplices and the n-th order down-Laplacian describes the diffusion occurring among n-dimensional simplices connected by $(n-1)$-dimensional simplices. The higher-order Laplacians capture the topology of simplicial complexes. For instance their spectrum encodes the Betti numbers, i.e., the number of n-dimensional cavities of the simplicial complex.

The higher-order Laplacians are defined in terms of the incidence matrices of the simplicial complex which represent the boundary operators playing a fundamental role in algebraic topology.

Here we provide a brief introduction to the necessary elements of algebraic topology needed to define higher-order Laplacians.

We consider a d-dimensional simplicial complex formed by $N_{[n]}$ simplices of dimension n. The simplices of the simplicial complexes are associated with an orientation induced by the labelling of the nodes so that the link $[i, j]$ has a positive orientation if $j > i$ and so on (see Fig. 11.4).

We consider algebraic entities called n-chains that are linear combinations of n-dimensional simplices with coefficients in $\mathbb{Z}$. In a less informal definition n-chains are the elements of a free abelian group $\mathcal{C}_n$ with basis on the n-simplices of the simplicial complex. The boundary map is a linear map $\partial_n : \mathcal{C}_n \to \mathcal{C}_{n-1}$ defined by its action on each simplex $\alpha = [i_0, i_1, i_2 \ldots, i_n]$. In particular the boundary map associates to every n-dimensional simplex $\alpha = [i_0, i_1, i_2 \ldots, i_n]$ a linear combination of the $(n-1)$-dimensional oriented faces at its boundary, given by

$$\partial_n [i_0, i_1 \ldots, i_n] = \sum_{p=0}^{n} (-1)^p [i_0, i_1, \ldots, i_{p-1}, i_{p+1}, \ldots, i_n]. \tag{11.10}$$

It follows that the image of the boundary operator ∂_n are the $(n-1)$-chains that are at the boundary of n-chains. For instance we have

$$\partial_2([1,2,3]) = [1,2] + [2,3] - [1,3], \tag{11.11}$$

i.e., the image of a triangle is the linear combination of the links at its boundary with the correct orientation. Additionally, from this definition it is also easy to see that a cyclic n-chain is in the kernel of the boundary map ∂_n independently of whether the cyclic n-chain is the boundary of a $(n+1)$-chain. For instance we have

$$\partial_1([1,2] + [2,3] - [1,3]) = [2] - [1] + [3] - [2] - [3] + [1] = 0, \tag{11.12}$$

whether the simplex [123] belongs to the simplicial complex or not. One important topological property of the boundary operator is that the "boundary of a boundary is null" which implies

$$\text{im } \partial_{n+1} \subseteq \ker \partial_n \tag{11.13}$$

or equivalently

$$\partial_n \partial_{n+1} = 0. \tag{11.14}$$

For instance we have

$$\partial_1 \partial_2 [1,2,3] = \partial_1([1,2] + [2,3] - [1,3]) = 0. \tag{11.15}$$

The boundary map ∂_n can be represented by a $N_{[n-1]} \times N_{[n]}$ incidence matrix $\mathbf{B}_{[n]}$ if we adopt as a basis of the space $\mathcal{C}_n$ an ordered set of the n-dimensional simplices α, and as a basis of the space $\mathcal{C}_{n-1}$ an ordered set of the $(n-1)$-dimensional simplices $\hat{\alpha}$.

If the basis of n-chains $\mathcal{C}_n$ is given by the n-simplices $\{\alpha_1, \alpha_2, \ldots \alpha_s \ldots\}$ and the basis of $(n-1)$-chains $\mathcal{C}_{n-1}$ is given by the $(n-1)$-simplices $\{\hat{\alpha}_1, \hat{\alpha}_2, \ldots \hat{\alpha}_r \ldots\}$ the action of the boundary map over any arbitrary n-dimensional simplex $\alpha_s = [i_0, i_1 \ldots, i_n]$ given by Eq. (11.10) can be expressed as

$$\partial_n \alpha_s = \sum_{r=1}^{N_{n-1}} [B_{[n]}]_{rs} \hat{\alpha}_r. \tag{11.16}$$

This equation fully determines the incidence matrices $\mathbf{B}_{[n]}$. Since we have seen that the "boundary of a boundary is null" then the incidence matrices follow $\mathbf{B}_{[n]}\mathbf{B}_{[n+1]} = \mathbf{0}$ and also $\mathbf{B}^\top_{[n+1]}\mathbf{B}^\top_{[n]} = \mathbf{0}$.

As an example we can consider the simplicial complex shown in Fig. 11.4 whose incidence matrices are given by

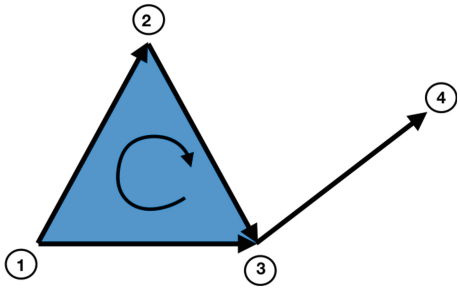


Fig. 11.4 Example of oriented simplicial complex. This illustrates an example of a 2-dimensional oriented simplicial complex having associated incidence matrices given by Eqs. (11.17). Reprinted figure with permission from Ref. [12]. ©Copyright (2020) by the American Physical Society.

$$\mathbf{B}_{[1]} = \begin{array}{c|cccc} & [1,2] & [1,3] & [2,3] & [3,4] \\ \hline [1] & -1 & -1 & 0 & 0 \\ [2] & 1 & 0 & -1 & 0 \\ [3] & 0 & 1 & 1 & -1 \\ [4] & 0 & 0 & 0 & 1 \end{array}, \quad \mathbf{B}_{[2]} = \begin{array}{c|c} & [1,2,3] \\ \hline [1,2] & 1 \\ [1,3] & -1 \\ [2,3] & 1 \\ [3,4] & 0 \end{array}. \tag{11.17}$$

The graph Laplacian can be expressed in terms of the incidence matrix $\mathbf{B}_{[1]}$ of the graph as

$$\mathbf{L}_{[0]} = \mathbf{B}_{[1]}\mathbf{B}_{[1]}^{\top}. \tag{11.18}$$

Similarly, in a simplicial complex the higher-order Laplacian $\mathbf{L}_{[n]}$ (with $n > 0$) [14, 34, 35] is the $N_{[n]} \times N_{[n]}$ matrix defined as

$$\mathbf{L}_{[n]} = \mathbf{L}_{[n]}^{[\text{down}]} + \mathbf{L}_{[n]}^{[\text{up}]}, \tag{11.19}$$

where

$$\mathbf{L}_{[n]}^{[\text{down}]} = \mathbf{B}_{[n]}^{\top}\mathbf{B}_{[n]}, \quad \mathbf{L}_{[n]}^{[\text{up}]} = \mathbf{B}_{[n+1]}\mathbf{B}_{[n+1]}^{\top}. \tag{11.20}$$

The n-Laplacian is positive semi-definite and, therefore, it has $N_{[n]}$ non negative eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_r \leq \ldots \leq \lambda_{N_{[n]}}$. A notable property of the spectrum of the n-th order Laplacian is that the degeneracy of zero eigenvalues is given by the n-th Betti number. Despite the fact that the construction of the higher-order Laplacians described above seems to rely on the choice of the orientations adopted for the simplices of the simplicial complex, it is possible to show that the higher-order Laplacians are independent on the orientation of the simplices as long as such orientation is induced by the labeling of the nodes.

We note that also the up and the down Laplacians are positive semi-definite and that from the definition of these matrices it follows immediately than the non-zero eigenvalues in the spectrum of $\mathbf{L}_{[n]}^{[\text{up}]}$ are the same as the non-zero eigenvalues in the spectrum of $\mathbf{L}_{[n]}^{[\text{down}]}$.

By considering the property of the incidence matrices such that $\mathbf{B}_{[n]}\mathbf{B}_{[n+1]} = 0$, it is possible to derive the Hodge decomposition of the space of n-chains, which reads

$$\mathbb{R}^{D_n} = \text{img}(\mathbf{B}_{[n]}^{\top}) \oplus \ker(\mathbf{L}_{[n]}) \oplus \text{img}(\mathbf{B}_{[n+1]}). \tag{11.21}$$

This implies that the higher-order Laplacian $\mathbf{L}_{[n]}$ can be simultaneously diagonalized with the n-th order up and n-th order down Laplacian and that the non-zero eigenvectors of $\mathbf{L}_{[n]}$ are either non-zero eigenvector of $\mathbf{L}_{[n]}^{[\text{down}]}$ or non-zero eigenvectors of $\mathbf{L}_{[n]}^{[\text{up}]}$. Therefore there is a basis in which $\mathbf{L}_{[n]}$, $\mathbf{L}_{[n]}^{[\text{down}]}$ and $\mathbf{L}_{[n]}^{[\text{up}]}$ have diagonal form given by

$$\mathbf{U}^{-1}\mathbf{L}_{[n]}\mathbf{U} = \begin{pmatrix} \mathbf{D}_{[n]}^{[\text{down}]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{[n]}^{[\text{up}]} \end{pmatrix}, \quad \mathbf{U}^{-1}\mathbf{L}_{[n]}^{[\text{down}]}\mathbf{U} = \begin{pmatrix} \mathbf{D}_{[n]}^{[\text{down}]} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{U}^{-1}\mathbf{L}_{[n]}^{[\text{up}]}\mathbf{U} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{[n]}^{[\text{up}]} \end{pmatrix},$$

where $\mathbf{D}_{[n]}^{[\text{up}]}$ and $\mathbf{D}_{[n]}^{[\text{down}]}$ are diagonal matrices having positive diagonal elements.

11.3.4 Higher Order Spectral Dimension

The notion of spectral dimension (see Sect. 11.3.2) can be generalized to n-order up-Laplacians with important consequences for higher-order simplicial complex dynamics. Here, we will focus on the model of NGF simplicial complexes introduced in Sect. 11.2.1. In Sect. 11.3.2, we have shown that the graph Laplacian of NGFs displays a finite spectral dimension $d_S^{[0]}$ [18, 32, 33]. Interestingly, higher-order up-Laplacians $\mathbf{L}_{[n]}^{[\text{up}]}$ and the higher-order down-Laplacians $\mathbf{L}_{[n]}^{[\text{down}]}$ of NGFs also display a finite spectral dimension.

In particular, the higher-order up-Laplacians of NGFs display a finite spectral dimension $d_S^{[n]}$ depending on the order n, the dimension of the simplicial complex d and the flavor parameter s [13]. Therefore, we can define different spectral dimensions for $0 < n < d - 1$. In order to show this remarkable geometrical property of NGFs in Fig. 11.5 we provide numerical evidence of the scaling of the cumulative density of non-zero eigenvalues $\rho_c^{\text{up}}(\lambda)$ of the $\mathbf{L}_{[n]}^{[\text{up}]}$ with λ for $\lambda \ll 1$, given by

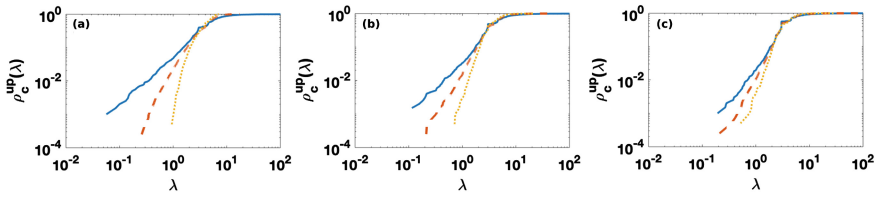


Fig. 11.5 The cumulative density of non-zero eigenvalues $\rho_c^{up}(\lambda)$ of the $\mathbf{L}_{[n]}^{[\text{up}]}$ for NGF of dimension $d = 3$ and flavor $s = -1$ (panel **a**), $s = 0$ (panel **b**), and $s = 1$ (panel **c**) for $n = 0$ (blue solid lines), $n = 1$ (red dashed lines), $n = 2$ (yellow dotted lines) and $n = 3$ (purple dot-dashed lines). Considered NGFs sizes are $N_{[0]} = 2000$ nodes, $N_{[1]} = 5994$ links, $N_{[2]} = 5992$ triangles, and $N_{[3]} = 1997$ tetrahedra. Adapted figure from [13]

$$\rho_c^{[\mathrm{up}]}(\lambda) \propto \lambda^{d_S^{[n]}/2}, \tag{11.22}$$

for different value of the order n of the up-Laplacian, and the flavor s of the NGF. It follows that a NGF model is not characterized by a single spectral dimension, i.e., the spectral dimension $d_S^{[0]}$ of the graph Laplacian, rather the NGF simplicial complexes have a higher-order network geometry encoded in a vector of spectral dimensions

$$\mathbf{d}_S = \left(d_S^{[0]}, d_S^{[1]}, d_S^{[2]} \ldots, d_S^{[d-1]}\right). \tag{11.23}$$

Consequently, the diffusion dynamics defined on simplices of different order n of the same NGF simplicial complex can be significantly different [36]. We finally note that for deterministic Apollonian and pseudo-fractal simplicial complexes that constitute the deterministic counterpart of NGF simplicial complexes, the higher-order spectral dimension can be predicted analytically by the real-space renormalization group [37] showing that the higher-order spectral dimension of these structures depends on the order n and remains finite as long as $0 \leq n < d - 1$.

11.4 Simplicial Synchronization

Synchronization is a fundamental dynamical state observed in a wide variety of complex systems and capturing among other phenomena important aspects of brain dynamics and circadian rhythms. The Kuramoto model [8–11, 38] is a stylized model that explains how coupled oscillators, that in absence of interactions would have different intrinsic frequencies, can start to follow a collective coherent motion when their coupling constant σ, measuring the strength of their interaction, is larger than a critical value σ_c also called *synchronization threshold*.

In order to model the coupling between the oscillators the Kuramoto model considers a network of $N_{[0]}$ nodes and associates a phase θ_i to each node $i \in \{1, 2, \ldots, N_{[0]}\}$ of the network. Therefore in the Kuramoto model, the dynamical state of the network is determined by the vector $\boldsymbol{\theta}$ of phases associated to its nodes given by

$$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{N_{[0]}})^\top. \tag{11.24}$$

Each phase θ_i describes the dynamical state of an oscillator that in absence of interactions oscillates at an intrinsic frequency ω_i drawn independently from a distribution $g(\omega)$. Common choices for $g(\omega)$ are the unimodal Gaussian or Lorentzian distributions.

The equations determining the dynamics of the phases associated to the nodes include an important contribution indicating the coupling among the phases of neighbour nodes. This coupling term has a strength modulated by the coupling constant σ. In particular, it is assumed that this contribution expresses the tendency of the phase of any given node to oscillate together with the phases of its neighbour nodes. The

resulting standard Kuramoto dynamics is captured by the differential equations

$$\dot{\theta}_i = \omega_i + \sigma \sum_{j=1}^{N_{[0]}} a_{ij} \sin\left(\theta_j - \theta_i\right), \tag{11.25}$$

valid for every node i of the network, where a_{ij} is the generic element of the adjacency matrix of the network. The level of synchronization in the system is measured by the Kuramoto order parameter,

$$Z_0 = R_0 e^{i\Theta} = \frac{1}{N} \sum_{j=1}^{N_{[0]}} e^{i\theta_j}, \tag{11.26}$$

where R_0 and Θ are both real and where $0 \leq R_0 \leq 1$ measures the overall coherence and $\Theta = \Theta(t)$ is the phase of global oscillations.

The relation between the Kuramoto model and the graph Laplacian of the underlying network is revealed when the Kuramoto model is linearized for $|\theta_i - \theta_j| \ll 1$ for every pair of neighbour nodes (i, j). In this limit the Kuramoto model can be shown to be described by the system of equations

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega} - \sigma \mathbf{L}_{[0]} \boldsymbol{\theta}, \tag{11.27}$$

where $\boldsymbol{\omega}$ indicates the vector of elements ω_i with $i \in \{1, 2, \ldots, N_{[0]}\}$.

The Kuramoto model has been analytically solved only on a fully connected network, although important progress has been made in understanding the Kuramoto model in random complex networks [9, 39, 40],

In the fully connected network and in random networks the Kuramoto model displays a second order phase transition at the synchronization threshold $\sigma = \sigma_c$ when the number of nodes goes to infinity, i.e., $N_{[0]} \to \infty$. For $\sigma < \sigma_c$ the Kuramoto model is in an incoherent state characterized by having a zero order parameter $R_0 = 0$. For $\sigma > \sigma_c$ the Kuramoto model is in a coherent state characterized by a non-zero order parameter $R_0 > 0$ [8–10, 38].

In this chapter we show how the network geometry and topology of simplicial complexes, directly acting on the spectral properties of the graph Laplacian and the higher-order Laplacians, can dramatically change the dynamical properties of the synchronization process on higher-order networks.

– **Simplicial network geometry and the Kuramoto model.** First we will illustrate how the phase diagram of the Kuramoto model changes when the model is defined on the skeleton of simplicial complexes with distinct higher-order network geometry. We have previously introduced the finite spectral dimension $d_S^{[0]}$ of the graph Laplacian as a fundamental observable of the higher-order network geometry of networks and simplicial complexes. In the following we will discuss how the dynamics of the Kuramoto model depends on $d_S^{[0]}$ revealing that network geometries have a more rich dynamics than random networks.

– **Simplicial network topology and the higher-order Kuramoto model.** Secondly we will reveal how the Kuramoto model can be considered as a limiting case of a much wider class of higher-order Kuramoto models describing coupling of topological signals. Topological signals are phases of oscillators associated not only to the nodes but also to higher-order simplices of a simplicial complex. For instance topological signals can be associated with both nodes and links of a simplicial complex as schematically described by Fig. 11.6. In this case the dynamical state of a simplicial complex is captured by the vector $\boldsymbol{\theta}$ of the phases associated to the nodes (defined in Eq. (11.24)) and by the vector $\boldsymbol{\phi}$ of the phases associated to the links of the simplicial complex given by

$$\boldsymbol{\phi} = \left(\phi_{\ell_1}, \phi_{\ell_2}, \ldots, \phi_{\ell_{N_{[1]}}}\right)^{\top}. \tag{11.28}$$

The phases associated to the links of the simplicial complexes are topological signals that have the potential to capture the dynamics of fluxes in brain networks [41] and biological transportation networks [42]. The newly formulated higher-order Kuramoto model opens new scenarios for characterizing how topology affects dynamics on higher-order networks and simplicial complexes. As we will discuss in Sect. 11.5 the higher-order Kuramoto model has a linearized dynamics described by the higher-order Laplacians of the simplicial complex, and can dis-

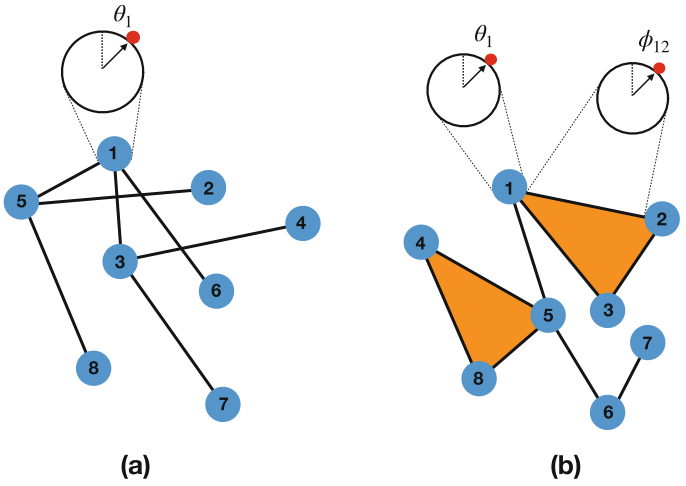


Fig. 11.6 Schematic representation of the Kuramoto model and the higher-order Kuramoto model capturing dynamics of topological signals. The Kuramoto model (panel **a**) captures the emergence of a synchronized state among coupled oscillators described by phases associated to the nodes of a network. The higher-order Kuramoto model (panel **b**) reveals the synchronization of topological signals on simplicial complexes, i.e., oscillators associated not only to the nodes of a simplicial complex, but also to higher-dimensional simplices such as links or triangles. Interestingly, topological signals of different dimension can co-exist and co-evolve and can be non-trivially coupled leading to simultaneous explosive transitions

play the simultaneous explosive synchronization transition of the soleinodal and the irrotational component of topological signals and even the simultaneous explosive synchronization transition of topological signals of different dimensions.

11.5 Kuramoto Model on Simplicial Network Geometry

11.5.1 Synchronization on Simplicial Network Skeletons with Finite Spectral Dimension

In order to investigate the role that simplicial network geometry has on the Kuramoto model we explore the phase diagram of the normalized Kuramoto model on networks with finite spectral dimension [32]. Our theoretical results are then validated by simulations performed over the skeleton of the NGF with flavor $s = -1$. Indeed NGF with flavor $s = -1$ provide a very suitable benchmark to test our theoretical results as they display a spectral dimension $d_S^{[0]}$ that can be changed by tuning the dimension d of the simplicial complex [33].

The normalized Kuramoto model determines the dynamics of the phases $\boldsymbol{\theta}$ associated to the nodes of a network. The only difference with the standard Kuramoto model is that the coupling between the phase of a given node i and the phases of its neighbour nodes is normalized with the node degree k_i. Therefore, the normalized Kuramoto model is dictated by the differential equations

$$\dot{\theta}_i = \omega_i + \sigma \sum_{j=1}^{N_{[0]}} \frac{a_{ij}}{k_i} \sin\left(\theta_j - \theta_i\right), \tag{11.29}$$

where here and in the following we consider internal frequencies ω_i drawn independently from a normal distribution, i.e., $\omega_i \sim \mathcal{N}(0, 1)$. The normalization of the coupling term by the degree of the node i is a very efficient way to screen out the effects of the heterogeneity of the degrees of the nodes and single out only the effects due to the geometrical nature of the network of their interactions.

The linearized equation of the normalized Kuramoto model is therefore determined by the normalized Laplacian $\hat{\mathbf{L}}_{[0]}$ instead of the graph Laplacian $\mathbf{L}_{[0]}$, i.e.,

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega} - \sigma \hat{\mathbf{L}}_{[0]} \boldsymbol{\theta}. \tag{11.30}$$

The analytical investigation of the stability of the synchronized phase indicates that the spectral dimension $d_S^{[0]}$ of the (normalized) graph Laplacian of the network plays a fundamental role for determining the phase diagram of the normalized Kuramoto model in the limit of infinite network size $N_{[0]} \to \infty$ [32]. Interestingly, we note that the spectral dimension of the graph Laplacian and the normalized graph Laplacian

take the same value under very general regularity conditions of the network [28]. For the very heterogeneous NGF, numerical results show that the two spectral dimensions differ by a small amount as long as the topological dimension d is small. Depending on the value of the spectral dimension the phase diagram of the normalized Kuramoto model defined in the infinite network limit changes drastically [32]:

(1) For networks with finite spectral dimension $d_S^{[0]} \leq 2$, the Kuramoto model cannot synchronize and is found in the incoherent state for every value of the coupling constant σ.
(2) For networks with spectral dimension $2 < d_S^{[0]} \leq 4$, global synchronization is not achievable in the infinite network limit but an entrained state can be observed. Therefore it is possible the Kuramoto model can have a transition between an incoherent state and an entrained phase.
(3) Only for networks with spectral dimension $d_S^{[0]} > 4$ it is possible to see a synchronized phase.

These results reveal how the dynamics of the Kuramoto model depends on the simplicial network geometry on which is defined, and extend previous results valid on regular lattices of dimension d [43, 44].

11.5.2 Frustrated Synchronization on Network Geometry with Flavor

The NGF constitutes a perfect model to investigate numerically the role that simplicial network geometry has on the dynamics of the Kuramoto model [32, 33]. Indeed, as we discussed previously, the NGF displays an emergent hyperbolic network geometry and for flavor $s = -1$ generates random hyperbolic manifolds. The simplicial network geometry of NGF is also reflected on their spectral properties. Specifically, NGFs have a finite spectral dimension $d_S^{[0]}$ which for flavor $s = -1$ and $d \in \{2, 3, 4\}$ can be approximated by $d_S^{[0]} \simeq d$. It follows that by changing the dimension d of the NGF with flavor $s = -1$ we can explore the dynamics of the Kuramoto model when the global synchronization state is not stable in the infinite network limit.

A computational finite size analysis of the Kuramoto model [32, 33] reveals, in agreement with the theoretical expectations, that for $d = 2$ and $d_S^{[0]} \simeq 2$, global synchronization is never achieved for large network sizes. On the other hand, synchronization in NGFs with $d = 3$ and $d = 4$ and $d_S^{[0]} \simeq d$ is only possible for finite networks. In fact its onset occurs for higher couplings when the system size increases, revealing, in agreement with the theoretical expectation, that in the limit $N_{[0]} \to \infty$ this state is never achieved.

Interestingly, we observe that for NGF with $d = 3$ and $d = 4$ and $d_S^{[0]} \simeq d$ the Kuramoto model exhibits a phase with entrained synchronization that we call a phase of *frustrated synchronization* for a wide range of coupling values. In the *frustrated synchronization* regime the order parameter R_0 displays strong temporal fluctuations.

This phase is observed on finite NGF between the incoherent state and the globally synchronized state, and for $d = 3$ a much broader regime of large fluctuations is observed than for the $d = 4$ case.

Interestingly, the frustrated synchronization phase of NGFs is not only characterized by strong temporal fluctuations of the global order parameter but displays also strong spatial fluctuations induced by the non-trivial hyperbolic network geometry of the NGF. As such, the frustrated synchronization phase can capture an important mechanism for inducing spatio-temporal fluctuations in brain dynamics.

The hyperbolic network geometry of NGF has a strong hierarchical nature that is responsible for the emergence of a relevant community structure. In order to study how the dynamics of the Kuramoto model is affected by the community structure of the NGFs, we define mesoscopic synchronization order parameters Z_{mod} that

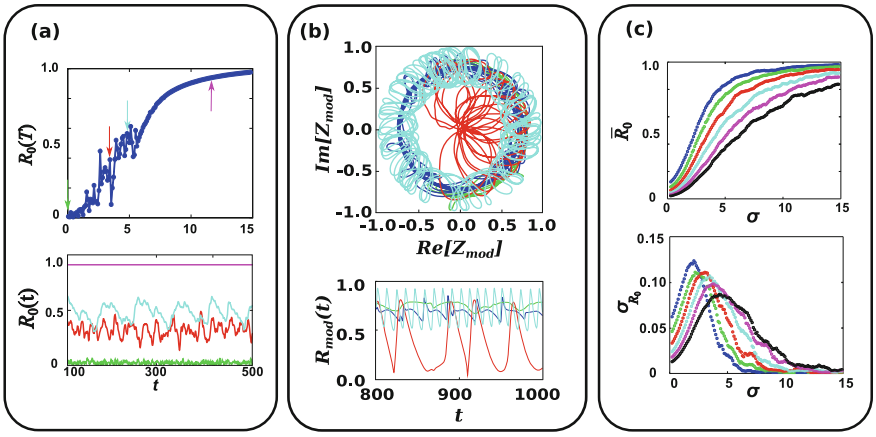


Fig. 11.7 Frustrated synchronization on NGF characterized by spatio-temporal fluctuations of the order parameter. Panel **(a)** The synchronization order parameter $R_0(T)$ calculated at time T is plotted versus the coupling constant σ revealing the regime of frustrated synchronization (top panel). The time series $R_0(t)$ are shown (bottom panel) for the values of the coupling constant σ indicated by arrows in the top panel. These time series reveal the temporal fluctuations of the order parameter in the frustrated synchronization regime. Panel **(b)** The spatial fluctuations of the order parameter in the frustrated synchronization regime are revealed by the local order parameter Z_{mod} of four different communities of the NGF calculated for $\sigma = 5$ (top panel). In this representation, a circular trajectory describes a situation of global oscillations of the nodes in the community, with constant $R_{\text{mod}}(t)$. Random trajectories around the origin describe unsynchronized communities. Partially synchronized communities, on the other hand, may describe more complex trajectories. The bottom panel shows the corresponding time series of $R_{\text{mod}}(t)$. Panel **(c)** Synchronization transition as given by the mean order parameter $\bar{R}_0$ averaged over time and its variance σ_{R_0}, as functions of σ, for different network sizes $N = 100$ (blue), 200, 400, …, 3200 (black). The apparent onset of the synchronized regime is retarded to large values of the coupling constant σ for large network sizes, revealing that the NGF of dimension $d = 3$ cannot sustain a synchronized phase in the limit of infinite network sizes. All simulations reported in the figures are obtained for NGF of flavor $s = -1$, dimension $d = 3$ and number of nodes $N_{[0]} = 1600$. In panel (**a**) $T = 500$. Figure adapted from [33]

characterize the dynamical state of each community:

$$Z_{\text{mod}} = R_{\text{mod}} e^{i\Theta_{\text{mod}}} = \frac{1}{|\mathcal{C}|} \sum_{j \in \mathcal{C}} e^{i\theta_j}, \tag{11.31}$$

where $\mathcal{C}$ is the set of nodes in the community and $|\mathcal{C}|$ the total number of nodes in said community. Figure 11.7b displays the trajectory of $Z_{\text{mod}} = Z_{\text{mod}}(t)$ in the complex plane for some exemplary modules of an NGF with flavour $s = -1$ in $d = 3$, for the coupling that leads to the largest fluctuations of R_0 as a function of time. As shown in the figure, different modules display different synchronization regimes and may oscillate at different frequencies. Due to the underlying geometrical structure of NGFs, these modules correspond to spatially localized regions.

We note here that the frustrated synchronization observed in NGF can be related with analogous phases observed in other hierarchical models [45, 46] where temporal fluctuations of the synchronization order parameter are observed. However, the combination of both temporal and spatial fluctuations is a specific property of the frustrated synchronization in NGF due to their rich simplicial network geometry.

11.6 Higher-Order Kuramoto Model: a Topological Approach to Synchronization

11.6.1 Synchronization of Topological Signals

Simplicial complexes are formed by nodes and higher-order simplices including links, triangles, tetrahedra, and so on. As such, simplicial complexes have the ability to sustain topological signals, i.e., dynamic variables not only associated with the nodes of their network skeleton but also associated to links, triangles, and so on [12, 14, 47]. Topological signals have the ability to capture dynamics associated to links, such as fluxes in brain networks [41] and biological transportation networks [42, 48].

Here we will present the higher-order Kuramoto model [12] that reveals how topological signals can undergo continuous and discontinuous synchronization transitions. Interestingly, we will observe that this synchronization transition can be detected only if the signals are filtered with the appropriate topological operators. Therefore, the model not only captures a new topological critical phenomenon but also prescribes a way to process real data in order to investigate whether this topological synchronization phenomenon can be observed in real systems such as the brain or biological transportation networks.

11.6.2 Higher-Order Kuramoto Dynamics

The higher-order Kuramoto model [30] describes the synchronization of topological signals defined on simplices of dimension n. For instance, one can consider topological signals defined on the links of a simplicial complex (case $n = 1$) or alternatively one can consider signals defined on the triangles of a simplicial complex (case $n = 2$). The higher-order Kuramoto model is the most natural extension of the Kuramoto model to capture the synchronization of higher-order topological signals.

The standard Kuramoto dynamics describes the dynamics of the phases $\boldsymbol{\theta}$ associated to the nodes of the network. This dynamics, defined by Eq. (11.25), can be expressed in terms of the incidence matrix $\boldsymbol{B}_{[1]}$ (see Appendix) as

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega} - \sigma \boldsymbol{B}_{[1]} \sin \boldsymbol{B}_{[1]}^T \boldsymbol{\theta}. \tag{11.32}$$

The higher-order Kuramoto dynamics describes instead the dynamics of topological signals $\boldsymbol{\phi}$ with ϕ_α indicating the phase associated to the simplex α of dimension $n > 0$. Therefore, using the insights coming from algebraic topology, the natural definition of the *simple higher-order Kuramoto model* is

$$\dot{\boldsymbol{\phi}} = \hat{\boldsymbol{\omega}} - \sigma \boldsymbol{B}_{[n+1]} \sin \boldsymbol{B}_{[n+1]}^T \boldsymbol{\phi} - \sigma \boldsymbol{B}_{[n]}^T \sin \boldsymbol{B}_{[n]} \boldsymbol{\phi}, \tag{11.33}$$

where $\hat{\boldsymbol{\omega}}$ is the vector of intrinsic frequencies $\hat{\omega}_\alpha$ associated with each n-dimensional simplex α drawn independently from a normal distribution, i.e., $\hat{\omega}_\alpha \sim \mathcal{N}(\Omega_1, 1/\tau_1)$. As the standard Kuramoto model can be related to the graph Laplacian via linearization (see Eq. 11.27), the higher-order Kuramoto model can be related to the higher-order Laplacian upon linearization, leading to

$$\dot{\boldsymbol{\phi}} = \hat{\boldsymbol{\omega}} - \sigma \mathbf{L}_{[n]} \boldsymbol{\phi}. \tag{11.34}$$

The higher-order Kuramoto dynamics defined on topological signals $\boldsymbol{\phi}$ associated to n dimensional simplices can be projected on $(n + 1)$ and $(n - 1)$ dimensional simplices by applying to the signals the incidence matrices. For instance a dynamics defined on topological signals associated to links can be projected on nodes or on triangles. Specifically, we have that the projected dynamics $\boldsymbol{\phi}^{[-]}$ on $(n - 1)$-dimensional simplices and the projected dynamics $\boldsymbol{\phi}^{[+]}$ on $(n + 1)$-dimensional simplices is given by

$$\begin{aligned} \boldsymbol{\phi}^{[-]} &= \mathbf{B}_{[n]} \boldsymbol{\phi}, \\ \boldsymbol{\phi}^{[+]} &= \mathbf{B}_{[n+1]}^\top \boldsymbol{\phi}, \end{aligned} \tag{11.35}$$

where, for $n = 1$, $\mathbf{B}_{[n]}$ indicates the discrete divergence and $\mathbf{B}_{[n+1]}^\top$ indicates the discrete curl. Therefore $\boldsymbol{\phi}^{[-]}$ indicates the irrotational component of $\boldsymbol{\phi}$ while $\boldsymbol{\phi}^{[+]}$ indicates the solenoidal component of $\boldsymbol{\phi}$. For the simple higher-order Kuramoto model defined in Eq. (11.33), by recalling that $\mathbf{B}_{[n+1]}^\top \mathbf{B}_{[n]}^\top = \mathbf{0}$ and that $\mathbf{B}_{[n]} \mathbf{B}_{[n+1]} = \mathbf{0}$, it follows that the projected topological signals $\boldsymbol{\phi}^{[-]}$ and $\boldsymbol{\phi}^{[+]}$ obey the uncoupled

system of equations

$$\dot{\boldsymbol{\phi}}^{[-]} = \boldsymbol{B}_{[n]}\hat{\boldsymbol{\omega}} - \sigma \boldsymbol{L}^{[\text{up}]}_{[n-1]} \sin\left(\boldsymbol{\phi}^{[-]}\right),$$
$$\dot{\boldsymbol{\phi}}^{[+]} = \boldsymbol{B}^{T}_{[n+1]}\hat{\boldsymbol{\omega}} - \sigma \boldsymbol{L}^{[\text{down}]}_{[n+1]} \sin\left(\boldsymbol{\phi}^{[+]}\right). \tag{11.36}$$

Therefore, the solenoidal and the irrotational components of the topological signals are decoupled for the simple higher-order Kuramoto model.

The higher-order Kuramoto dynamics is remarkable from two perspectives:

(1) First of all it uses topology to naturally define the higher-order interactions between the topological signals. Indeed the incidence matrices define higher-order interactions with a clear prescription indicating the coupled variables and the sign of their interactions. For example, the higher-order Kuramoto dynamics for $n = 1$ dimensional simplices of the simplicial complex shown in Fig. 11.4 reads

$$\begin{aligned}
\dot{\theta}_{[1,2]} &= \hat{\omega}_{[1,2]} - \sigma \sin(\theta_{[2,3]} - \theta_{[1,3]} + \theta_{[1,2]}) - \sigma \left[\sin(\theta_{[1,2]} - \theta_{[2,3]}) + \sin(\theta_{[1,3]} + \theta_{[1,2]})\right], \\
\dot{\theta}_{[1,3]} &= \hat{\omega}_{[1,3]} + \sigma \sin(\theta_{[2,3]} - \theta_{[1,3]} + \theta_{[1,2]}) - \sigma \left[\sin(\theta_{[1,3]} + \theta_{[1,2]}) + \sin(\hat{\theta}_{[3]})\right], \\
\dot{\theta}_{[2,3]} &= \hat{\omega}_{[2,3]} - \sigma \sin(\theta_{[2,3]} - \theta_{[1,3]} + \theta_{[1,2]}) - \sigma \left[\sin(\theta_{[2,3]} - \theta_{[1,2]}) + \sin(\hat{\theta}_{[3]})\right], \\
\dot{\theta}_{[3,4]} &= \hat{\omega}_{[3,4]} - \sigma \left[\sin(\theta_{[3,4]}) - \sin(\hat{\theta}_{[3]})\right],
\end{aligned} \tag{11.37}$$

with $\hat{\theta}_{[3]}$ indicating the three-body interaction

$$\hat{\theta}_{[3]} = \theta_{[13]} + \theta_{[23]} - \theta_{[34]}. \tag{11.38}$$

Therefore, the choice of the higher-order interactions in the higher-order Kuramoto model is naturally dictated by topology.

(2) The synchronization of the higher-order Kuramoto model is only detectable if the right topological filtering of the data is performed. Indeed the naïve order parameter

$$R_n = \frac{1}{N_{[n]}} \left| \sum_{\alpha=1}^{N_{[n]}} e^{\mathrm{i}\phi_\alpha} \right| \tag{11.39}$$

associated to the unfiltered topological signal $\boldsymbol{\phi}$ does not detect any synchronization transition (see Fig. 11.8). Instead, the order parameter associated to the solenoidal and the irrotational components of the topological signal do detect the synchronization transition of the topological signals (see Fig. 11.8). These order parameters are given by

$$R_n^{[-]} = \frac{1}{N_{[n-1]}} \left| \sum_{\alpha=1}^{N_{[n-1]}} e^{\mathrm{i}\phi_\alpha^{[-]}} \right|, \quad R_n^{[+]} = \frac{1}{N_{[n+1]}} \left| \sum_{\alpha=1}^{N_{[n+1]}} e^{\mathrm{i}\phi_\alpha^{[+]}} \right|, \tag{11.40}$$

or alternatively by

$$R_n^{\downarrow} = \frac{1}{N_{[n]}} \left| \sum_{\alpha=1}^{N_{[n]}} e^{\mathrm{i}\phi_\alpha^{\downarrow}} \right|, \quad R_n^{\uparrow} = \frac{1}{N_{[n]}} \left| \sum_{\alpha=1}^{N_{[n]}} e^{\mathrm{i}\phi_\alpha^{\uparrow}} \right|, \tag{11.41}$$

where $\boldsymbol{\phi}^{\downarrow} = \mathbf{L}_{[n]}^{[\mathrm{down}]}\boldsymbol{\phi}$ and $\boldsymbol{\phi}^{\uparrow} = \mathbf{L}_{[n]}^{[\mathrm{up}]}\boldsymbol{\phi}$.

The synchronization transition described by the simple higher-order Kuramoto model leads to a continuous transition occurring at zero coupling, i.e., the synchronization threshold is $\sigma_c = 0$ as long as $n > 0$. However, the higher-order Kuramoto model admits a formulation called *explosive higher order Kuramoto model* that displays instead a discontinuous transition at a non zero coupling $\sigma_c > 0$.

The *explosive higher-order Kuramoto dynamics* [12] implements an adaptive coupling of the projected dynamics of $\boldsymbol{\phi}^{[+]}$ and $\boldsymbol{\phi}^{[-]}$ through their global order parameters. The adopted adaptive coupling is inspired by analogous couplings previously applied to multilayer and simple networks [49]. The explosive higher-order Kuramoto model [12] is defined by the system of equations

$$\dot{\boldsymbol{\phi}} = \hat{\boldsymbol{\omega}} - \sigma R^{[-]} \boldsymbol{B}_{[n+1]} \sin \boldsymbol{B}_{[n+1]}^T \phi - \sigma R^{[+]} \boldsymbol{B}_{[n]}^T \sin \boldsymbol{B}_{[n]} \phi. \tag{11.42}$$

It follows that the dynamics projected on the $(n+1)$ and $(n-1)$-dimensional simplices now obeys the coupled system of equations

$$\begin{aligned} \dot{\boldsymbol{\phi}}^{[+]} &= \boldsymbol{B}_{[n+1]}^T \hat{\boldsymbol{\omega}} - \sigma R^{[-]} \boldsymbol{L}_{[n+1]}^{[\mathrm{down}]} \sin\left(\boldsymbol{\phi}^{[+]}\right), \\ \dot{\boldsymbol{\phi}}^{[-]} &= \boldsymbol{B}_{[n]} \hat{\boldsymbol{\omega}} - \sigma R^{[+]} \boldsymbol{L}_{[n-1]}^{[\mathrm{up}]} \sin\left(\boldsymbol{\phi}^{[-]}\right). \end{aligned} \tag{11.43}$$

Numerical simulations on the configuration model of simplicial complexes with power-law distribution of generalized degrees reveal that the explosive higher-order Kuramoto model displays a discontinuous phase transition. The nature of the transition confirms the theoretical expectations obtained with an approximate phenomenological approach. This transition is clearly detected by a discontinuity in $R_n^{[+]}$ and $R_n^{[-]}$ and in $R_n^{\downarrow}$ and $R_n^{\uparrow}$ as well, but is not captured by the naïve order parameter R_n (see Fig. 11.8).

In Ref. [12] it has been shown that the nature of the phase transition does not change if the generalized degree distribution is more uniform or if the simplicial complex has a non trivial network geometry. Interestingly, simple and explosive higher-order Kuramoto models can be investigated on simplicial complexes constructed from real connectomes leading to continuous (for the simple higher-order Kuramoto model) and discontinuous (for the explosive higher-order Kuramoto model) synchronization transitions.

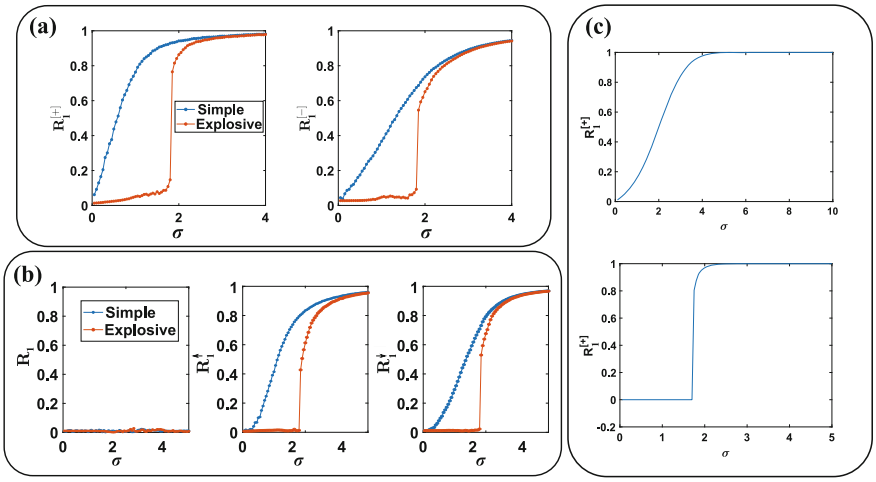


Fig. 11.8 Higher-order Kuramoto dynamics. The synchronization of higher-order topological signals is captured by the simple and explosive higher-order Kuramoto models. Panel **(a)** The order parameters $R_1^{[+]}$ and $R_1^{[-]}$ reveal the synchronization transition of topological signals defined on links ($n = 1$) of the configuration model of simplicial complexes. Panel **(b)** The naïve order parameter R_1 does not reveal the synchronization transition of topological signals defined on the links of the configuration model of simplicial complex, while the order parameters $R_1^{\uparrow}$ and $R_1^{\downarrow}$, sensible on the irrotational and solenoidal decomposition of the signal, to reveal the transition (as well as $R_1^{[+]}$ and $R_1^{[-]}$ shown in panel (a)). The underlying network is the same for both panels and has $N_{[0]} = 1000$ nodes, $N_{[1]} = 5299$ links and $N_{[2]} = 4147$ triangles. The generalized degree of the nodes is power-law distributed with power-law exponent $\gamma = 2.8$. Panel **(c)** shows the theoretical expectations provided by an effective phenomenological model treated in Ref. [12] for the simple (top) and explosive (bottom) higher-order Kuramoto models. Reprinted figure with permission from Ref. [12] ©Copyright (2020) by the American Physical Society

11.6.3 Coupled Topological Signals

So far we have considered the synchronization of topological signals defined on simplices of dimension $n > 0$. However, topological signals associated to simplices of different dimension can co-exist and co-evolve. For instance, phases associated to the nodes of a simplicial complex can be coupled with phases associated to its links. In this section we will show how different topological signals, i.e., phases defined on simplicial complexes of different dimensions, can be coupled to each other leading to simultaneous explosive synchronization transitions. For simplicity of presentation we will focus on phases defined on nodes and links, but we emphasize that our formalism allows one to consider the interaction of more general topological signals.

We start by considering Model 1, an explosive higher-order Kuramoto model of coupled signals of nodes and links. This model differs from the explosive higher-order Kuramoto model defined in the previous section as it includes an additional

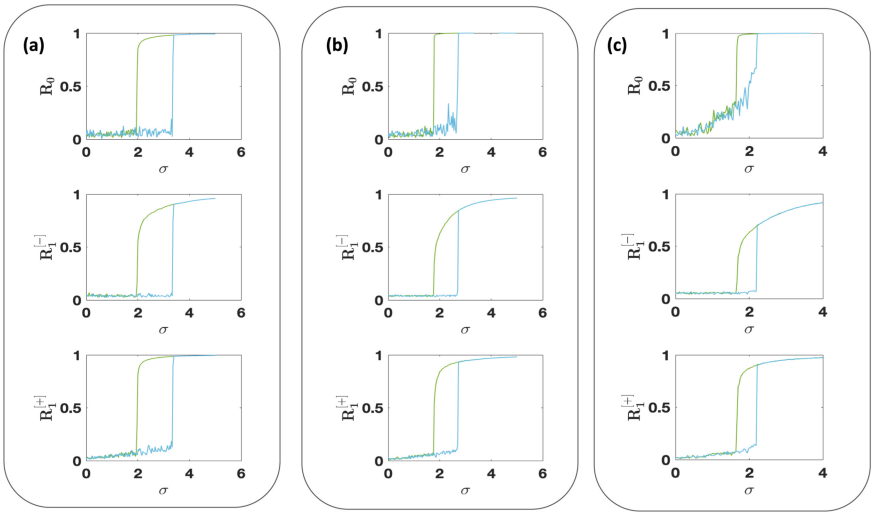


Fig. 11.9 Simultaneous explosive synchronization of coupled topological signals. Model 1 displays the simultaneous explosive (discontinuous) transition captured by the order parameters R_0, $R_1^{[-]}$ and $R_1^{[+]}$. Panels **(a), (b),** and (**c**) provide numerical evidence of this discontinuous transition on different simplicial network topologies: the NGF with flavor $s = -1$ and dimension d (panel **a**), the configuration model of pure $d = 3$ simplicial complex with power-law generalized degree distribution with power-law exponent $\gamma = 2.8$ (panel **b**) and the clique complex of the Caenorhabditis elegans (C. elegans) connectome coming from Ref. [50] (panel **c**). Figure adapted from [47]

adaptive coupling among the topological signals of nodes and links, and obeys the equations

$$\begin{aligned}\dot{\boldsymbol{\theta}} &= \boldsymbol{\omega} - \sigma R_1^{[-]} \mathbf{B}_{[1]} \sin(\mathbf{B}_{[1]}^{\top}\boldsymbol{\theta}),\\ \dot{\boldsymbol{\phi}} &= \hat{\boldsymbol{\omega}} - \sigma R_0 R_1^{[+]} \mathbf{B}_{[1]}^{\top} \sin(\mathbf{B}_{[1]}\boldsymbol{\phi}) - R_1^{[-]} \sigma \mathbf{B}_{[2]} \sin(\mathbf{B}_{[2]}^{\top}\boldsymbol{\phi}).\end{aligned} \tag{11.44}$$

This model simulated in a wide variety of simplicial complexes including the NGF, the configuration model of simplicial complexes, and the clique complex of real connectomes displays a simultaneous explosive (i.e., discontinuous) synchronization of the topological signals defined on nodes and of the soleinodal and irrotational component of the topological signals defined on links [47]. Indeed, at a critical threshold $\sigma = \sigma_c$ we observe a discontinuity in the three order parameters R_0, $R_1^{[-]}$ and $R_1^{[+]}$. In Fig. 11.9 we present numerical evidence of this discontinuous transition by displaying the corresponding hysteresis loop in the order parameters. In particular, instead of plotting the order parameters obtained at each value of σ starting from random initial conditions as in Fig. 11.8 here we display the order parameters along the forward and backward synchronization transitions obtained by first adiabatically increasing and then decreasing the coupling constant σ.

Interestingly, this model (Model 1) admits a variation (Model 2) that describes the simultaneous synchronization of the topological signals associated to the nodes and the irrotational component of the topological signal associated to the links. This simpler version of the explosive higher-order Kuramoto model of coupled topological signals can be also defined on pairwise networks and is amenable to an exact analytical treatment on fully connected networks and to an accurate annealed approximation solution on random networks with given degree distribution. More specifically, Model 2 only couples the signal of the nodes with the signal of the links projected to the nodes, as described by the differential equations

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega} - \sigma R_1^{[-]} \mathbf{B}_{[1]} \sin(\mathbf{B}_{[1]}^{\top} \boldsymbol{\theta}), \tag{11.45}$$

$$\dot{\boldsymbol{\phi}} = \hat{\boldsymbol{\omega}} - \sigma R_0 \mathbf{B}_{[1]}^{\top} \sin(\mathbf{B}_{[1]} \boldsymbol{\phi}) - \sigma \mathbf{B}_{[2]} \sin(\mathbf{B}_{[2]}^{\top} \boldsymbol{\phi}). \tag{11.46}$$

Projecting the dynamics of the link phases down to 0-simplices as in the previous section, we introduce $\boldsymbol{\psi}$ for simplicity of notation with

$$\boldsymbol{\psi} \equiv \boldsymbol{\phi}^{[-]} = \mathbf{B}_{[1]} \boldsymbol{\phi}. \tag{11.47}$$

By left multiplying Eq. (11.46) by $\mathbf{B}_{[1]}$, we obtain the closed system of equations for $\boldsymbol{\theta}$ and $\boldsymbol{\psi}$

$$\begin{aligned} \dot{\boldsymbol{\theta}} &= \boldsymbol{\omega} - \sigma R_1^{[-]} \mathbf{B}_{[1]} \sin(\mathbf{B}_{[1]}^{\top} \boldsymbol{\theta}), \\ \dot{\boldsymbol{\psi}} &= \tilde{\boldsymbol{\omega}} - \sigma R_0 \mathbf{L}_{[0]} \sin(\boldsymbol{\psi}), \end{aligned} \tag{11.48}$$

where $\tilde{\boldsymbol{\omega}} = \mathbf{B}_{[1]} \hat{\boldsymbol{\omega}}$. Here we assume $\omega \sim \mathcal{N}(0, 1)$ and $\hat{\omega}_\alpha \sim \mathcal{N}(\Omega_1, 1/\tau_1)$. With these hypotheses the internal frequencies of the links projected on the nodes $\{\tilde{\omega}_i\}_{i=1,2,\ldots,N_{[0]}}$ are Gaussian correlated variables with average

$$\langle \tilde{\omega}_i \rangle = \left[\sum_{j<i} a_{ij} - \sum_{j>i} a_{ij} \right] \Omega_1 \tag{11.49}$$

and with correlation matrix $\mathbf{C}$ of elements $C_{ij} = \langle \tilde{\omega}_i \tilde{\omega}_j \rangle - \langle \tilde{\omega}_i \rangle \langle \tilde{\omega}_j \rangle$ given by

$$\mathbf{C} = \mathbf{L}_{[0]} \frac{1}{\tau_1^2}. \tag{11.50}$$

To understand the nature of the synchronization transition analytically when Model 2 is defined on an uncorrelated random graph, in the following we discuss the solution of the model in the *annealed approximation*. The annealed approximation is a widely used approximation to study dynamical processes on random uncorrelated networks which consists in substituting the adjacency matrix entries of the network a_{ij} by their average values in an uncorrelated random network with given degree sequence, i.e.,

$$a_{ij} \to \frac{k_i k_j}{\langle k \rangle N}. \tag{11.51}$$

Using this approximation, Eq. (11.48) can be recast into the differential equations

$$\begin{aligned} \dot{\boldsymbol{\theta}} &= \boldsymbol{\omega} - \sigma R_1^{[-]} \hat{R}_0 \mathbf{k} \cdot \sin(\boldsymbol{\theta} - \hat{\Theta}), \\ \dot{\boldsymbol{\psi}} &= \tilde{\boldsymbol{\omega}} + \sigma R_0 \hat{R}_1^{[-]} \mathbf{k} \sin \hat{\Psi} - \sigma R_0 \mathbf{k} \odot \sin \boldsymbol{\psi}, \end{aligned} \tag{11.52}$$

where $\odot$ indicates the Hadamard product (element by element multiplication) and where two auxiliary complex order parameters are defined as

$$\hat{R}_0 e^{\mathrm{i}\hat{\Theta}} = \sum_{i=1}^{N_{[0]}} \frac{k_i}{\langle k \rangle N_{[0]}} e^{\mathrm{i}\theta_i}, \quad \hat{R}_1^{[-]} e^{\mathrm{i}\hat{\Psi}} = \sum_{i=1}^{N_{[0]}} \frac{k_i}{\langle k \rangle N_{[0]}} e^{\mathrm{i}\psi_i}, \tag{11.53}$$

with $\hat{R}_0$, $\hat{\Theta}$, $\hat{R}_1^{[-]}$ and $\hat{\Psi}$ being real. Let us indicate with $g(\omega)$ the probability distribution of the internal frequencies of the nodes and with $G_i(\tilde{\omega})$ the marginal probability distribution of the internal frequencies of the links projected on node i, i.e., the marginal probability that $\tilde{\omega}_i = \tilde{\omega}$. With this notation it is possible to derive the analytic solution of Eq. (11.52) which gives the following expression for the order parameters

$$\begin{aligned} R_0 &= \frac{1}{N_{[0]}} \sum_{i=1}^{N_{[0]}} r_0(i), \qquad \hat{R}_n = \sum_{i=1}^{N_{[0]}} \frac{k_i}{\langle k \rangle N_{[0]}} r_0(i), \\ R_1^{[-]} &= \frac{1}{N_{[0]}} \sum_{i=1}^{N_{[0]}} r_1^{[-]}(i), \qquad \hat{R}_1^{[-]} = \sum_{i=1}^{N_{[0]}} \frac{k_i}{\langle k \rangle N_{[0]}} r_1^{[-]}(i), \end{aligned} \tag{11.54}$$

with $r_0(i)$ and $r_1^{[-]}(i)$ given by

$$\begin{aligned} r_0(i) &= \int_{|\hat{c}_i|<1} d\omega g(\omega) \sqrt{1 - \left(\frac{\omega - \Omega_0}{\sigma k_i \hat{R}_0 R_1^{[-]}} \right)^2}, \\ r_1^{[-]}(i) &= \int_{|\hat{d}_i|<1} d\tilde{\omega} G_i(\tilde{\omega}) \sqrt{1 - \left(\frac{\tilde{\omega}}{\sigma k_i \hat{R}_0} \right)^2}, \end{aligned} \tag{11.55}$$

and $\hat{c}_i$ and $\hat{d}_i$ indicating

$$\hat{c}_i = \frac{\omega - \Omega_0}{\sigma k_i \hat{R}_0 R_1^{[-]}}, \quad \hat{d}_i = \frac{\tilde{\omega}}{\sigma k_i \hat{R}_0}. \tag{11.56}$$

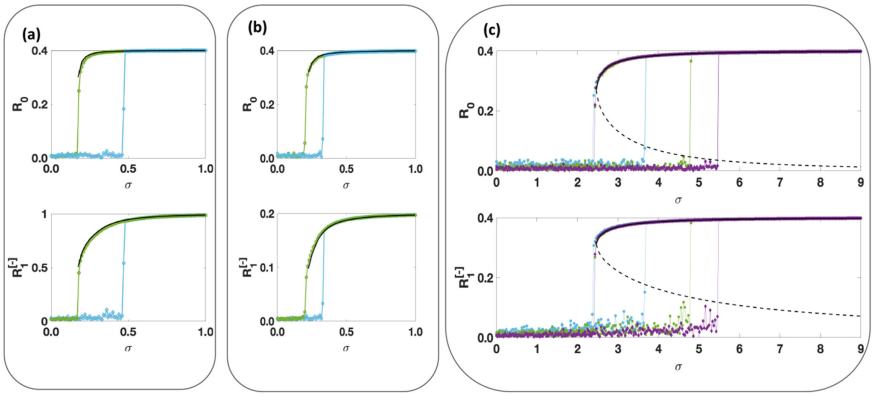


Fig. 11.10 Theoretical prediction for Model 2 of explosive higher-order Kuramoto model of coupled dynamical signals applied to random and fully connected networks. The hysteresis loop for Model 2 is shown for different types of networks. The backward transition is fully captured by the theoretical expectations (solid black lines). Panels **(a), (b)** and **(c)** show the order parameters R_1 and $R_1^{[-]}$ versus the coupling constant σ for: A Poisson network with average degree $c = 12$ and $N_{[0]} = 1600$ nodes (panel a), a scale-free network with minimum degree $m = 6$, power-law exponent $\gamma = 2.5$ and $N_{[0]} = 1600$ nodes (panel b) and a fully connected network of different network sizes $N_{[0]} = 500$ (cyan symbols), $N_{[0]} = 1000$ (green symbols)and $N_{[0]} = 2000$ (purple symbols) (panel **c**). From panel **(c)** it is evident that the forward transition occurs at higher values of σ for larger network size, confirming the theoretical prediction indicating that the transition is driven by finite size effects and it is absent in the infinite network limit. Figures adapted from[47]

Figure 11.10 shows excellent agreement between the simulation results of Model 2 and the analytical prediction obtained in the annealed approximation for a Poisson network with average degree $c = 12$ (panel (a)) and for an uncorrelated scale-free network with minimum degree $m = 6$ and power-law exponent $\gamma = 2.5$ (panel **b**).

Model 2 of explosive higher-order synchronization of coupled topological signals of nodes and links can be also solved exactly on a fully connected network. Before discussing these theoretical results let us highlight that when treating Model 2 on a fully connected network the model parameters need to be rescaled appropriately to give a well defined transition in the large network limit. In particular the coupling constant σ and τ_1 are rescaled to

$$\begin{aligned} \sigma &\to \frac{\sigma}{N}, \\ \tau_1 &\to \tau_1 \sqrt{N}. \end{aligned} \tag{11.57}$$

Moreover, for simplicity we set $\Omega_1 = 0$. With these hypotheses the marginal distribution $G_i(\tilde{\omega}) = G(\tilde{\omega})$ for every node i of the network can be derived to be equal to

$$G(\tilde{\omega}) = \frac{\tau_1}{\sqrt{2\pi/\tilde{c}}} \exp\left[-\tau_1^2 \tilde{c} \frac{\tilde{\omega}^2}{2}\right], \tag{11.58}$$

with $\bar{c} = N/(N-1)$. The self-consistent equations for the order parameters $R_0 = \hat{R}_0$ and $R_1^{[-]} = \hat{R}_1^{[-]}$ are given by

$$\begin{aligned} 1 &= \sigma R_1^{[-]} h\left(\sigma^2 R_0^2 (R_1^{[-]})^2\right), \\ R_1^{[-]} &= \sigma R_0 \tau_1 \sqrt{\bar{c}} h\left(\sigma^2 \tau_1^2 R_0^2\right), \end{aligned} \tag{11.59}$$

where the scaling function $h(x)$ is given by

$$h(x) = \sqrt{\frac{\pi}{2}} e^{-x/4} \left[I_0\left(\frac{x}{4}\right) + I_1\left(\frac{x}{4}\right) \right], \tag{11.60}$$

with I_0 and I_1 indicating the modified Bessel functions.

These equations agree perfectly well with direct simulation of Model 2 dynamical Eq. (11.46) on a fully connected network as it can be appreciated from Fig. 11.10c. Moreover a closer look to these equations reveals an important aspect of these transitions. While the backward transition has a well defined limit as $N \to \infty$, the forward transition occurs at larger value of the coupling constant for large network size N, and is only determined by finite size fluctuations, therefore the transition disappears in the limit $N \to \infty$. Interestingly, this lack of a proper hysteresis loop can be also predicted for Model 2 defined on uncorrelated random graphs with finite second moment of the degree distribution, starting from their annealed approximation solution.

11.7 Conclusions

In this chapter our goal has been to provide evidence that the interplay between simplicial complex structure and dynamics is mediated by simplicial geometry and topology. The spectral properties of the graph Laplacian and the higher-order Laplacian have been used here to reveal how simplicial synchronization is shaped by topological and geometry of the simplicial complex. In particular, we investigated how simplicial network geometry changes the phase diagram of the Kuramoto model defined on the network skeleton of simplicial complexes with notable geometrical properties and characterized by a finite spectral dimension. We have shown that a spectral dimension smaller or equal than four but larger than two can lead to a regime of frustrated synchronization characterized by large spatio-temporal fluctuations of the order parameter, while a spectral dimension smaller or equal than two leads to a Kuramoto model in the incoherent state for every finite value of the coupling constant. These theoretical results have been shown to apply to the simplicial complexes generated by the modelling framework called Network Geometry with Flavor (NGF) that is able to generate simplicial complexes with tunable spectral dimension of the graph Laplacian. Interestingly, the NGF are characterized also by displaying higher-order spectral dimension of the higher-order up Laplacian that describe higher-order diffusion processes.

This chapter introduces also a set of models for capturing synchronization of topological signals, i.e., phases not only associated to the nodes of a simplicial complex but also to the higher-order simplices such as links, triangles, and so on. This higher-order synchronization reveals itself in the order parameter of the irrotational and solenoidal projection of the topological signals. In the simple higher-order Kuramoto model the irrotational and solenoidal projection of the topological signal are uncoupled and undergo a sychronization transition at $\sigma_c = 0$. However, when these two projections are coupled to each other by an adaptive global coupling the synchronization becomes explosive, i.e., discontinuous, and occurs at a non-zero value of coupling constant.

The higher-order Kuramoto model can be further extended to capture coupled topological signals of different dimension, for instance coupling phases associated to nodes and to links of a network or of a simplicial complex. This generalized higher-order Kuramoto model can lead to an explosive phase transition affecting simultaneously the phases associated to the nodes and the irrotational and solenoidal projection of the phases associated to the links. Interestingly, the higher-order Kuramoto model of coupled topological signals defined on nodes and links can be treated analytically using the annealed approximation when it is defined on a random uncorrelated network and can be solved exactly on a fully connected network. This solution confirms the discontinuous nature of the transition of the explosive higher-order Kuramoto model and sheds light on the stability of the hysteresis loop associated to the transition on finite networks. The mathematical framework that we have proposed here can be explored and modified in different directions and we believe that an in-depth analysis of the model and its variations will provide important insights on the interplay between topology and dynamics. For instance, we note that the higher-order Kuramoto model has been recently modified [51] to investigate also the properties of a consensus model finding interesting results.

In conclusion, this chapter aims to provide an overview of the relation between network geometry topology and dynamics. We believe this topic will flourish in the incoming years and will transform our understanding of the relation between structure and dynamics of higher-order networks. Therefore our expectation is that this research line will play a relevant role for providing new insights in a variety of applications including brain dynamics and biological transportation networks.

Appendix: Kuramoto dynamics expressed in terms of the incidence matrix

In this Appendix our aim is to show that Eq. (11.25) that we rewrite here for convenience,

$$\dot{\theta}_i = \omega_i + \sigma \sum_{j=1}^{N_{[0]}} a_{ij} \sin\left(\theta_j - \theta_i\right), \tag{11.61}$$

is equivalent to Eq. (11.32) given by

$$\dot{\boldsymbol{\theta}} = \boldsymbol{\omega} - \sigma \boldsymbol{B}_{[1]} \sin \boldsymbol{B}_{[1]}^{T} \boldsymbol{\theta}. \tag{11.62}$$

In order to show this let us observe that the incidence matrix $\mathbf{B}_{[1]}$ has elements given by

$$[B_{[1]}]_{i\ell} = \begin{cases} -1 \text{ if } \ell = [i, j], \\ 1 \quad \text{if } \ell = [j, i], \\ 0 \quad \text{otherwise.} \end{cases} \tag{11.63}$$

To show the equivalence between Eq. (11.61) and Eq. (11.62) let us start by rewriting Eq. (11.62) element by element, getting

$$\dot{\theta}_i = \omega_i - \sigma \sum_{\ell \in S_1} [B_{[1]}]_{i\ell} \sin \left(\sum_{j \in S_0} [B_{[1]}^{\top}]_{\ell j} \theta_j \right), \tag{11.64}$$

where we indicate with S_1 the set of all links present in the simplicial complex or network under consideration.

Let us consider the particular link $\ell = [i, j]$ in this case we have

$$[B_{[1]}]_{i\ell} \sin \left(\sum_{j \in S_0} [B_{[1]}^{\top}]_{\ell j} \theta_j \right) = -a_{ij} \sin(\theta_j - \theta_i). \tag{11.65}$$

Equivalently, if we consider the same link with opposite orientation $\ell = [j, i]$ we get

$$[B_{[1]}]_{i\ell} \sin \left(\sum_{j \in S_0} [B_{[1]}^{\top}]_{\ell j} \theta_j \right) = a_{ij} \sin(\theta_i - \theta_j) = -a_{ij} \sin(\theta_j - \theta_i). \tag{11.66}$$

Since the incidence matrix $\mathbf{B}_{[1]}$ has non-zero elements only among nodes and links incident to each other, it follows that Eq. (11.61) is equivalent to Eq. (11.62).

References

1. A.L. Barabási, *Network Science* (Cambridge University Press, 2016)
2. S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, Critical phenomena in complex networks. Rev. Mod. Phys. **80**(4), 1275 (2008)
3. A. Barrat, M. Barthelemy, A. Vespignani. *Dynamical Processes on Complex Networks*. (Cambridge University Press, 2008)

4. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, G. Petri, *Networks beyond pairwise interactions: structure and dynamics* (Phys, Rep, 2020)
5. G. Bianconi, Higher-Order Networks: An Introduction to Simplicial Complexes. (Cambridge University Press, 2021)
6. G. Bianconi, Interdisciplinary and physics challenges of network theory. EPL (Europhys. Lett.) **111**(5), 56001 (2015)
7. L. Torres, A.S. Blevins, D.S. Bassett, T. Eliassi-Rad, The why, how, and when of representations for complex systems. ArXiv preprint arXiv:2006.02870 (2020)
8. S.H. Strogatz, From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. Phys. D: Nonlinear Phenomena **143**(1–4), 1–20 (2000)
9. A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Synchronization in complex networks. Phys. Rep. **469**(3), 93–153 (2008)
10. S. Boccaletti, A.N Pisarchik, C.I. Del Genio, A. Amann, *Synchronization: From Coupled Systems to Complex Networks* (Cambridge University Press, 2018)
11. Y. Kuramoto, Self-entrainment of a population of coupled non-linear oscillators, in *International Symposium on Mathematical Problems in Theoretical Physics*. ed. by H. Araki (Springer, Berlin Heidelberg, Berlin, Heidelberg, 1975), pp. 420–422
12. A.P. Millán, J.J. Torres, G. Bianconi, Explosive higher-order kuramoto dynamics on simplicial complexes. Phys. Rev. Lett. **124**(21), 218301 (2020)
13. J.J. Torres, G. Bianconi, Simplicial complexes: higher-order spectral dimension and dynamics. J. Phys.: Complex. **1**(1), 015002 (2020)
14. S. Barbarossa, S. Sardellitti, Topological signal processing over simplicial complexes. IEEE Trans. Sig. Process. **68**, 2992–3007 (2020)
15. O.T. Courtney, G. Bianconi, Generalized network structures: the configuration model and the canonical ensemble of simplicial complexes. Phys. Rev. E **93**(6), 062311 (2016)
16. G. Bianconi, C. Rahmede, Network geometry with flavor: from complexity to quantum geometry. Phys. Rev. E **93**(3), 032315 (2016)
17. G. Bianconi, C. Rahmede, Emergent hyperbolic network geometry. Sci. Rep. **7**, 41974 (2017)
18. D. Mulder, G. Bianconi, Network geometry and complexity. J. Stat. Phys. **173**(3–4), 783–805 (2018)
19. O.T. Courtney, G. Bianconi, Weighted growing simplicial complexes. Phys. Rev. E **95**(6), 062301 (2017)
20. P.S. Skardal, A. Arenas, Abrupt desynchronization and extensive multistability in globally coupled oscillator simplexes. Physical review letters **122**(24), 248301 (2019)
21. P.S. Skardal, A. Arenas, Higher-order interactions in complex networks of phase oscillators promote abrupt synchronization switching. arXiv preprint arXiv:1909.08057 (2019)
22. P.S. Skardal, A. Arenas, Memory selection and information switching in oscillator networks with higher-order interactions. J. Phys.: Compl. (2020)
23. Repository for higher-order network codes
24. M. Gromov, Hyperbolic groups, in *Essays in Group Theory*. (Springer, 1987), pp. 75–263
25. A.P. Millán, R. Ghorbanchian, N. Defenu, F. Battiston, G. Bianconi, Local topological moves determine global diffusion properties of hyperbolic higher-order networks. Phys. Rev. E **104**, 054302 (2021)
26. R. Burioni, D. Cassi, Random walks on graphs: ideas, techniques and results. J. Phys. A: Math General **38**(8), R45 (2005)
27. G. Rammal, Toulouse, Random walks on fractal structures and percolation clusters. J. de Phys. Lett. **44**(1), 13–22 (1983)
28. R. Burioni, D. Cassi, Universal properties of spectral dimension. P. Rev. Lett. **76**(7), 1091 (1996)
29. Z. Wu, G. Menichetti, C. Rahmede, G. Bianconi, Emergent complex network geometry. Sci. Rep. **5**, 10073 (2015)
30. A.P. Millán, G. Gori, F. Battiston, T. Enss, N. Defenu, Complex networks with tuneable dimensions as a universality playground. Phys. Rev. Res. **3**, 023015 (2021)

31. M.M. Dankulov, B. Tadić, R. Melnik, Spectral properties of hyperbolic nanonetworks with tunable aggregation of simplexes. Phys. Rev. E **100**(1), 012309 (2019)
32. A.P. Millán, J.J. Torres, G. Bianconi, Synchronization in network geometries with finite spectral dimension. Phys. Rev. E **99**(2), 022307 (2019)
33. A.P. Millán, J.J. Torres, G. Bianconi, Complex network geometry and frustrated synchronization. Sci. Rep. **8**(9910) (2018)
34. A. Muhammad, M. Egerstedt, Control using higher order laplacians in network topologies, in *Proceeding of 17th International Symposium on Mathematical Theory of Networks and Systems* (Citeseer, 2006), pp. 1024–1038
35. T.E. Goldberg, Combinatorial laplacians of simplicial complexes. *Senior Thesis, Bard College* (2002)
36. J.J. Torres, J. Marro, Brain performance versus phase transitions. Sci. Rep. **5**, 12216 (2015)
37. M. Reitz, G. Bianconi, The higher-order spectrum of simplicial complexes: a renormalization group approach. J. Phys. A: Math. Theor. (2020)
38. J.A. Acebrón, L.L. Bonilla, C.J. Pérez Vicente, F. Ritort, R. Spigler, The kuramoto model: a simple paradigm for synchronization phenomena. Rev. Mod. Phys. **77**(1), 137 (2005)
39. J.G. Restrepo, E. Ott, B.R. Hunt, Onset of synchronization in large networks of coupled oscillators. Phys. Rev. E **71**(3), 036151 (2005)
40. M. Chavez, D.-U. Hwang, A. Amann, H.G.E. Hentschel, S. Boccaletti, Synchronization is enhanced in weighted complex networks. Phys. Rev. Lett. **94**(21), 218701 (2005)
41. W. Huang, T.A.W. Bolton, J.D. Medaglia, D.S. Bassett, A. Ribeiro, D. Van De Ville, A graph signal processing perspective on functional brain imaging. Proc. IEEE **106**(5), 868–885 (2018)
42. M. Ruiz-Garcia, E. Katifori, Topologically controlled emergent dynamics in flow networks. arXiv preprint arXiv:2001.01811 (2020)
43. H. Hong, H. Park, M.Y. Choi, Collective synchronization in spatially extended systems of coupled oscillators with random frequencies. Phys. Rev. E **72**(3), 036217 (2005)
44. H. Hong, H. Chaté, H. Park, L.H. Tang, Entrainment transition in populations of random frequency oscillators. Phys. Rev. Lett. **99**(18), 184101 (2007)
45. P. Moretti, M.A. Muñoz, Griffiths phases and the stretching of criticality in brain networks. Nat. Commun. **4**, 2521 (2013)
46. P. Villegas, P. Moretti, M.A. Muñoz, Frustrated hierarchical synchronization and emergent complexity in the human connectome network. Sci. Rep. **4**, 5990 (2014)
47. R. Ghorbanchian, J.G. Restrepo, J.J. Torres, G. Bianconi, Higher-order simplicial synchronization of coupled topological signals. Commun. Phys. **4**, 1–13 (2021)
48. J.W. Rocks, A.J. Liu, E. Katifori, Revealing structure-function relationships in functional flow networks via persistent homology. Phys. Rev. Res. **2**(3), 033234 (2020)
49. X. Zhang, S. Boccaletti, S. Guan, Z. Liu, Explosive synchronization in adaptive and multilayer networks. Phys. Rev. Lett. **114**(3), 038701 (2015)
50. L.R. Varshney, B.L. Chen, E. Paniagua, D.H. Hall, D.B. Chklovskii, Structural properties of the Caenorhabditis elegans neuronal network. PLoS Comput. Biol. **7**(2), e1001066 (2011)
51. L. DeVille, Consensus on simplicial complexes, or: the nonlinear simplicial Laplacian. arXiv preprint arXiv:2010.07421 (2020)

Chapter 12
Signal Processing on Simplicial Complexes

Michael T. Schaub, Jean-Baptiste Seby, Florian Frantzen, T. Mitchell Roddenberry, Yu Zhu, and Santiago Segarra

Abstract Higher-order networks have so far been considered primarily in the context of studying the structure of complex systems, i.e., the higher-order or multi-way relations connecting the constituent entities. More recently, a number of studies have considered dynamical processes that explicitly account for such higher-order dependencies, e.g., in the context of epidemic spreading processes or opinion formation. In this chapter, we focus on a closely related, but distinct third perspective: how can we use higher-order relationships to process signals and data supported on higher-order network structures. In particular, we survey how ideas from signal processing of data supported on regular domains, such as time series or images, can be extended to graphs and simplicial complexes. We discuss Fourier analysis, signal denoising, signal interpolation, and nonlinear processing through neural networks based on simplicial complexes. Key to our developments is the Hodge Laplacian

This chapter is built on the exposition of [46], in particular with respect to the example applications considered. Our discussion here is however more geared towards a reader with a network science background who is less familiar with signal processing. In particular, we provide additional discussion on the relations between filtering and linear dynamical processes on networks.

M. T. Schaub (✉) · F. Frantzen
RWTH Aachen University, Aachen, Germany
e-mail: schaub@cs.rwth-aachen.de

F. Frantzen
e-mail: florian.frantzen@cs.rwth-aachen.de

J.-B. Seby
Université Paris-Sud, Paris, France
e-mail: jbseby@mit.edu

T. M. Roddenberry · Y. Zhu · S. Segarra
Rice University, Houston, USA
e-mail: mitch@rice.edu

Y. Zhu
e-mail: yz126@rice.edu

S. Segarra
e-mail: segarra@rice.edu

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_12

matrix, a multi-relational operator that leverages the special structure of simplicial complexes and generalizes desirable properties of the Laplacian matrix in graph signal processing.

12.1 Introduction

Graphs provide a powerful abstraction for a wide variety of complex systems. Underpinning this abstraction is the central idea of decomposing a system into its fundamental entities and their relations. Those entities are then modelled as nodes and pairwise interactions between those entities are encoded by edges in a graph. This highly flexible, yet conceptually simple framework has been employed in many fields [16, 36], with applications ranging from biology to social sciences.

However, the role that graphs play in such analysis can be substantially different, depending on the type of data or process under investigation [6]. At the risk of oversimplification, we distinguish between two perspectives: In one case, the data are the connectivity patterns of the network itself, in the other, we aim to comprehend high-dimensional signals supported on a network. This first viewpoint is central when modeling relational data, i.e., data corresponding to measured edges (connections, interactions) in a network, where we aim to learn about a system by finding patterns in these stochastic connections, e.g., via community detection, centrality analysis or a range of other tools. In contrast, when studying dynamical processes on networks, we often consider the network as an arbitrary but (essentially) fixed entity, and our goal is to leverage the graph structure to understand the dynamics associated with the nodes. This latter perspective, in which we aim to understand data *supported on a graph* is the point of view we will adopt in the following.

The problem of having to understand processes and data supported on the nodes of a graph does not only arise in the context of network dynamical systems. Another area focusing on data of this type is graph signal processing (GSP) [40, 43, 49], which is concerned with the analysis of general signals supported on the nodes of a network. In fact, while such signals may come in the form of a time series, e.g., from dynamical measurements at the nodes, many node signals consist of other types of attribute data supported on nodes—sometimes referred to as node-covariates, node meta-data, or node feature data. The goal of GSP is to extend concepts from signal processing such as the Fourier transform and the large set of filtering operations developed in signal processing to data supported on the nodes of a graph [40, 43, 49]. Drawing on the rich tradition of signal processing, GSP provides a range of methods to analyze and process graph supported data and, thus, has seen a steady increase in interest in the last decade.

Even though graph-based descriptions of systems have been the dominant paradigm to analyze heterogeneously structured systems and data so far (potentially augmented in terms of multiple layers or temporal dimensions), the utility of graphs for modeling certain aspects of complex systems has been scrutinized recently. Specifically, graphs do not encode interactions between more than two nodes, even though such multi-way interactions are widespread in complex systems [4, 6, 46,

51]: assemblies of neurons fire in unison [22], biochemical reactions typically include more than two chemical species [30], and group interactions are widespread in social context [28]. To represent such polyadic interactions, a number of modeling frameworks have been proposed in the literature to model such higher-order relations, including simplicial complexes [25], hypergraphs [5], and others [19]. Using these frameworks to analyze the organizational principles of the (higher-order edge) structure of polyadic relational data has garnered much attention in the literature lately.

In comparison to this line of work of representing and analyzing the *structure* of complex multi-relational systems, the literature on dynamical processes on such higher-order network structures is still relatively sparse. However, there is a fast-growing body of work considering epidemic spreading, diffusion and opinion formation, among other processes, on higher-order networks—see the recent review [4] and references therein. Similarly, the literature on *signal processing* on higher-order networks is still nascent and received so far comparably little attention [2, 41, 46]. Our goal in this manuscript is to present this area of signal processing on higher-order network in an accessible and coherent manner, focusing on simplicial complexes as modeling framework for higher-order network interactions. Along the way we point out connections to the study of certain dynamical processes on (higher-order) networks and highlight avenues for future research.

The remainder of this document is structured as follows. We assume most of our readers will be accustomed to graphs and networks, but perhaps less familiar with some of the ideas of signal processing. Hence, we review in Sect. 12.2 some central tenets from discrete signal processing that are relevant for our purposes, and highlight the close connection between discrete signal processing and linear dynamical systems. In Sect. 12.3, we use the interpretation of signal processing in terms of dynamical systems to explain how one can naturally generalize from the processing of time series to signals supported on graphs. In this context we introduce three example applications: (i) signal smoothing and denoising for graph signals, (ii) signal interpolation on graphs, and (iii)nonlinear signal processing via graph neural networks. In Sect. 12.4, we then extend the ideas from graph signal processing to signals supported on higher-order networks, and outline central ideas for signal processing on simplicial complexes (SC). We revisit our example applications and show how the Hodge Laplacian, a generalization of the graph Laplacian for SCs, plays a central role in the processing of signals supported on SCs. We finally point out some further connections to the study of certain dynamical systems on graphs and simplicial complexes and then conclude with a brief discussion on future research avenues and directions.

12.2 From Signal Processing to Dynamical Systems and Graphs

Before embarking on our study of signal processing on simplicial complexes, we revisit key concepts from signal processing that will be instrumental for our explanations in subsequent sections. In particular, our exposition highlights how linear signal processing and linear dynamical processes may be seen as two sides of the same coin.

12.2.1 Linear Signal Processing in a Nutshell

Discrete signal processing (DSP) [39] is concerned with the extraction of information from observed data. Arguably one of the most elementary scenarios is that we observe some n dimensional data vector $\mathbf{y} = \mathbf{s} + \boldsymbol{\eta} \in \mathbb{R}^n$, which is simply an addition of some signal $\mathbf{s}$ and a distortion term $\boldsymbol{\eta}$. Simply put, the data consists of signal plus noise, and our goal is to extract the signal component from the observed data.

To make this task well-defined, we have to characterize more precisely what properties the signal and the noise have, i.e., we have to provide some modelling assumptions that specify characteristics of the signal versus the noise. A typical assumption here is that the signal $\mathbf{s}$ is concentrated in a particular linear subspace $\mathcal{S} \subset \mathbb{R}^n$, whereas the noise $\boldsymbol{\eta}$ is not localized in $\mathbb{R}^n$. For instance, in the context of time series, we typically assume that the signal $\mathbf{s}$ is varying smoothly, i.e., can be well-approximated by a linear combination of smooth basis functions. By finding an expressive set of basis vectors such that the subspace $\mathcal{S}$ of "interesting signals" can be spanned via a (sparse) subset of basis vectors, we can filter out the noise by projecting the observed signal $\mathbf{y}$ onto $\mathcal{S}$. The classical choice for such a basis is the collection of discrete Fourier modes [39]. Through their inherent characterization in terms of frequencies, Fourier modes enable us to distinguish between slowly varying, smooth signals and more rapidly oscillating signals.

The discrete Fourier transform of a signal $\mathbf{y}$ is defined as $\tilde{\mathbf{y}} = \mathbf{F}\mathbf{y}$, where $\mathbf{F}$ is the discrete Fourier transform matrix:

$$\mathbf{F} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \omega^6 & \cdots & \omega^{2(n-1)} \\ 1 & \omega^3 & \omega^6 & \omega^9 & \cdots & \omega^{3(n-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \omega^{3(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{bmatrix} \in \mathbb{C}^{n \times n}. \tag{12.1}$$

Here $\omega = \exp(-2i\pi/n)$ is the n-th primitive root of unity, and i denotes the imaginary unit with $i^2 = -1$. Since $\mathbf{F}$ is a unitary matrix ($\mathbf{F}\mathbf{F}^* = \mathbf{I}$), we see that the original signal can be synthesized via $\mathbf{y} = \mathbf{F}^*\tilde{\mathbf{y}}$, where $\mathbf{F}^*$ is the conjugate transpose of $\mathbf{F}$. Hence, the vector of Fourier coefficients $\tilde{\mathbf{y}}$ is simply the representation of the signal $\mathbf{y}$ in the new Fourier basis given by the columns of $\mathbf{F}^*$, which are simply the complex conjugates of the rows of $\mathbf{F}$. Note that the Fourier basis functions are ordered such that the first Fourier mode, the constant vector, is associated with the smallest frequency possible (frequency 0).

Based on the Fourier transform, we can now define the concept of a linear time-invariant filter as follows.[1] Let us consider a signal vector $\mathbf{s}_{\text{in}}$ defined via a scalar time series $s_{\text{in}}(t)$ at time-steps $t = 0, 1, \ldots, n-1$, i.e., $\mathbf{s}_{\text{in}} = [s_{\text{in}}(0), \ldots, s_{\text{in}}(t), \ldots,$

[1] More precisely, as often encountered in DSP, we are dealing with a cyclic time-shift invariant filter.

$s_{\text{in}}(n-1)]$. A linear time-invariant filter now consists in a transformation of the input signal $\mathbf{s}_{\text{in}}$ via a linear operator $\mathbf{H}$ to an output signal $\mathbf{s}_{\text{out}}$:

$$\mathbf{s}_{\text{out}} = \mathbf{H}\mathbf{s}_{\text{in}}, \tag{12.2}$$

such that the filter $\mathbf{H} = \mathbf{F}^* \mathbf{\Lambda}_{\mathbf{H}} \mathbf{F}$ can be diagonalized by discrete Fourier modes, i.e., the discrete Fourier modes are eigenvectors of $\mathbf{H}$. Here $\mathbf{\Lambda}_{\mathbf{H}}$ is the diagonal matrix of eigenvalues of the filter $\mathbf{H}$, which is the so called frequency response of the filter. Accordingly, we can interpret the action of filter $\mathbf{H}$ in (12.2) in terms of three consecutive operations. We first express the initial signal $\mathbf{s}_{\text{in}}$ in the Fourier basis by applying the Fourier transformation $\mathbf{F}$. We then modulate (amplify or attenuate) the coefficients of the signal representation in this new basis representation in a desired way by multiplying with $\mathbf{\Lambda}_{\mathbf{H}}$. Finally, we project back the output signal onto the initial basis by applying the inverse Fourier transformation $\mathbf{F}^*$.

Via direct calculation, it can be shown that the matrix representation of any such time-invariant filter $\mathbf{H}$ is a circulant matrix of the form

$$\mathbf{H} = \begin{bmatrix} c_0 & c_{n-1} & c_{n-2} & \cdots & c_1 \\ c_1 & c_0 & c_{n-1} & \cdots & c_2 \\ c_2 & c_1 & c_0 & & c_3 \\ \vdots & & \ddots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix}. \tag{12.3}$$

Note that the vector $\boldsymbol{\lambda}(\mathbf{H})$ of eigenvalues of $\mathbf{H}$, i.e., the frequency response of the filter can be calculated as:

$$\boldsymbol{\lambda}(\mathbf{H}) = \sqrt{n} \cdot \mathbf{F}\mathbf{c} \quad \text{with} \quad \mathbf{c} = [c_0, \ldots, c_{n-1}]^\top, \tag{12.4}$$

which means that the eigenvalues are simply a Fourier transform of the coefficient vector $\mathbf{c}$ that defines the circulant matrix $\mathbf{H}$.

The above description of signal processing in terms of a change of basis of the original (time) signal to a frequency domain representation highlights that the choice of signal *representation* in terms of a basis can be crucial for processing signals—this is a scheme we will see reoccur in the context of graphs and simplicial complexes. However, based on this representation, the close connections between such linear filtering operations and linear dynamical systems are less apparent. In the following subsection we thus focus on an equivalent interpretation of such filtering processes in terms of linear dynamical processes defined on (certain) graphs.

12.2.2 Signal Processing via Linear Dynamical Systems on Graphs

We now concentrate on a formulation of the above signal processing procedures in terms of linear dynamical systems in discrete time. To this end, let us define a signal $c(t)$ based on the vector $\mathbf{c} = [c(0), \dots, c(t), \dots, c(n-1)]^\top$, analogously as we defined $\mathbf{s}_{\text{in}}$. Moreover, consider the periodic extensions of both of these signals defined as $s^\circ_{\text{in}}(t) = s_{\text{in}}(t - \ell n)$ and $c^\circ(t) = c(t - \ell n)$ for $\ell \in \mathbb{Z}$.

From (12.3), we observe that the filtering operation (12.2) may equivalently be written in terms of the linear convolution[2] of the (periodically extended) impulse response $c(t)$ and the input signal $s_{\text{in}}(t)$

$$s_{\text{out}}(t) = (c * s_{\text{in}})(t) = \sum_{i=-\infty}^{\infty} c(i) s^\circ_{\text{in}}(t-i) = \sum_{i=-\infty}^{\infty} c^\circ(t-i) s_{\text{in}}(i). \tag{12.5}$$

Clearly, the above formula defines a linear dynamical system in which $c(t)$ plays the role of an impulse response. The system is however not *memoryless* as inputs $s_{\text{in}}(t')$ at times $t' \neq t$ are important for the output $s_{\text{out}}(t)$ of the system at time t. To implement (or realize) the system we thus introduce a state vector $\mathbf{x} \in \mathbb{R}^n$ which keeps track of the inputs to the system $s_{\text{in}}(t)$ at previous times. Accordingly, we may express the filtering operation (12.2) as:

$$s_{\text{out}}(t) = \mathbf{c}^\top \mathbf{x}(t) \tag{12.6a}$$

$$\mathbf{x}(t+1) = \mathbf{S}\mathbf{x}(t) \quad \text{with} \quad \mathbf{x}(0) = \mathbf{s}_{\text{in}}, \tag{12.6b}$$

where the state transition matrix $\mathbf{S}$ takes the special form of a (circular) shift

$$\mathbf{S} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{bmatrix}. \tag{12.7}$$

Note that the above realization of the system may not be minimal and we may be able to implement our system with fewer states. However, the key idea of representing a linear system in terms of a set of internal states which are coupled by a shift operator is the aspect that is central to our further developments. In particular, using the shift operator we can compactly summarize the above filtering operation in vector form:

[2] Equivalently,this may also be interpreted as the cyclic convolution of the vectors $\mathbf{c}$ and$\mathbf{s}_{\text{in}}$.

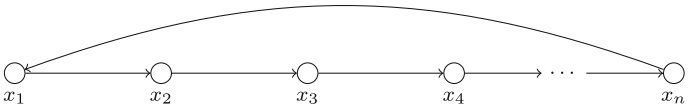


Fig. 12.1 Interpretation of the cyclic shift operator as a linear dynamical system on a graph. The cyclic shift operator $\mathbf{S}$ may be interpreted as a linear diffusion $\mathbf{x}(t+1) = \mathbf{S}\mathbf{x}(t)$ on a directed cycle graph of size n

$$\mathbf{s}_{\text{out}} = \sum_{k=0}^{n-1} c_k \mathbf{S}^k \mathbf{s}_{\text{in}} = \mathbf{H}\mathbf{s}_{\text{in}}, \tag{12.8}$$

which provides us with yet another way to express the output of our dynamical system. Note that this implies, in particular, that the filter $\mathbf{H}$ can be constructed from linear combinations of the simpler shift operator $\mathbf{S}$. Further, the Fourier basis $\mathbf{F}^*$ diagonalizes also the shift operator $\mathbf{S}$.

The rewriting (12.8) suggests an interpretation of the filtering operation as a linear process on a graph as illustrated in Fig. 12.1. Specifically, we may interpret the shift operator $\mathbf{S}$ as a cyclic graph $\mathcal{G}_C$ and associate each node with a state (time) of our dynamical system (see Fig. 12.1). Each iteration of the dynamical system $\mathbf{x}(t+1) = \mathbf{S}\mathbf{x}(t)$ can be interpreted as a linear filtering operation. Since a combination of linear filters is linear, the output $\mathbf{y} = \sum_t c_t \mathbf{x}(t)$ is linear and can thus also be interpreted in terms of a dynamical system on a cyclic graph. To summarize, a linear time-shift invariant filter can be interpreted as a weighted sum of the states of a linear dynamical system on a cyclic graph.

12.3 Signal Processing on Graphs

In this section, we provide a short introduction to graph signal processing (GSP), which extends the ideas of signal processing for time series or images to the processing of signals supported on general graphs. A key insight underpinning GSP is that the filtering operation (12.8) can be generalized from a cyclic graph to any graph, by defining an appropriate shift operator compatible with the structure of the graph on which the signals are supported.

12.3.1 Graph Signals, Fourier Transforms, and Filters

To introduce the ideas of graph signal processing mathematically, we will first recall some preliminary definitions and notation for graphs and signals supported on graphs. For simplicity, we will concentrate on undirected graphs in the ensuing discussion,

as this setup can be directly generalized to simplicial complexes. However, graph signal processing may also be considered in the context of directed graphs [20, 32].

We define an undirected graph $\mathcal{G}$ by a set of nodes $\mathcal{V} = \{v_1, \cdots, v_n\}$ with cardinality $|\mathcal{V}| = n$ and a set of edges $\mathcal{E}$, i.e., a collection of unordered pairs of nodes, with cardinality $|\mathcal{E}| = m$. A graph signal $s : \mathcal{V} \to \mathbb{R}$ is a mapping that assigns to each node $i \in \mathcal{V}$ a real-valued scalar. Such a graph signal may thus be suitably represented by a vector $\mathbf{s} \in \mathbb{R}^n$.

For computational purposes, we encode the structure of the graph $\mathcal{G}$ by an adjacency matrix $\mathbf{A}$, whose entry A_{ij} is 1 if there is an edge between nodes i and j, and 0 otherwise. The graph Laplacian of the graph $\mathcal{G}$ is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \mathrm{diag}(\mathbf{A1})$ is the diagonal matrix of (weighted) node degrees, i.e., D_{ii} is the degree of node i. Given an arbitrary orientation of the edges, an alternative description of the structure of $\mathcal{G}$ is the so-called *incidence* matrix $\mathbf{B} \in \mathbb{Z}^{n \times m}$, such that $B_{ie} = 1$ if node i is the tail of edge e, $B_{ie} = -1$ if i is the head of the edge e, and 0 otherwise. Using the operator $\mathbf{B}$, another expression for the graph Laplacian is $\mathbf{L} = \mathbf{B}\mathbf{B}^\top$. Notice that we think of the edge $e = (i, j)$ as oriented from the tail i to the head j, although the choice of the orientation is arbitrary and has nothing to do with a directed graph. For notational simplicity, we focus on unweighted graphs, although the presented ideas can be generalized. For instance, the entry A_{ij} of the adjacency matrix then simply becomes the weight of the edge from i to j if it exists.

To extend the idea of filtering operations to graphs, equation (12.8) provides a natural starting point. In particular, to define a linear filtering operation for general graphs, we may replace the cyclic shift operator $\mathbf{S}$ by any linear operator encoding the structure of the graph, e.g., the adjacency matrix $\mathbf{A}$ or the graph Laplacian $\mathbf{L}$. These matrices are accordingly referred to as graph shift operators in the context of GSP [48]. Since linear filtering in the time-domain is defined in terms of convolution, replacing the cyclic shift $\mathbf{S}$ in the filter (12.8) by a graph shift operator $\mathbf{G}$ gives rise to a graph convolutional filter:

$$\mathbf{s}_{\text{out}} = \sum_{k=0}^{n-1} c_k \mathbf{G}^k \mathbf{s}_{\text{in}} = \mathbf{H}_{\mathbf{G}} \mathbf{s}_{\text{in}}. \tag{12.9}$$

We remark that this definition also implies that any filtering operation can be implemented via localized computations in the graph, e.g., by exchanging node signals in a scheme akin to (12.6). In particular, if we choose a normalized adjacency matrix as the graph shift operator $\mathbf{G}$, we can see that this filtering scheme is interpretable as a weighted diffusion process on a graph. More generally, the graph filtering operation can always be understood in terms of a suitably defined linear dynamical process on the network. This highlights again the close similarities between signal processing and dynamical systems. There is a noteworthy difference in terms of the goals we typically have in mind in this context, however. In the study of (linear) dynamics on graphs, e.g., diffusion processes, we are often concerned with a fixed dynamics and aim to understand how the properties of an arbitrary graph may influence the behavior of this dynamics. In contrast, in the context of signal processing, we typically

consider the graph as a fixed entity and our goal is to design a filter (or equivalently a dynamical process) that achieves a desired behavior in terms of filtering.

At this point, two natural questions are: (i) What is the influence of the choice of the graph shift operator? (ii) Is there an advantage of choosing one over another graph shift operator? Let us concentrate on the first question for now. Since we are concerned with graph filters that can be expressed as a polynomial (or more generally as a power series) of the graph shift operator, the choice of the shift operator will induce an orthogonal basis in which we represent the graph signal $\mathbf{s}$. More precisely, given the spectral decomposition of the symmetric shift operator $\mathbf{G} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$, we say that the eigenvectors of the graph shift operator (the columns of $\mathbf{U}$) define a *Graph Fourier Transform* (GFT), i.e., they form a natural basis in which the signal $\mathbf{s}$ can be expressed [40, 43, 49]. The GFT of a signal is defined as

$$\tilde{\mathbf{s}} = \mathbf{U}^\top \mathbf{s}, \tag{12.10}$$

and the inverse Fourier Transform operation is

$$\mathbf{s} = \mathbf{U}\tilde{\mathbf{s}}. \tag{12.11}$$

Given a weight function $h : \mathbb{R} \to \mathbb{R}$, any shift-invariant graph filter can thus also be written as

$$\mathbf{H}_\mathbf{G} = \sum_{k=1}^{N} h(\lambda_k)\mathbf{u}_k\mathbf{u}_k^\top = \mathbf{U}h(\mathbf{\Lambda})\mathbf{U}^\top, \tag{12.12}$$

where we used the shorthand notation $h(\mathbf{\Lambda}) = \text{diag}(h(\lambda_1), h(\lambda_2), \cdots, h(\lambda_n))$. Note that this is exactly equivalent to the case of a time-invariant filter, apart from the choice of a different set of basis functions in which the signal is expressed. Indeed, $h(\mathbf{\Lambda})$ can again be interpreted as the frequency response of the filter $\mathbf{H}_\mathbf{G}$, and the filtering operation $\mathbf{s}_{\text{out}} = \mathbf{H}_\mathbf{G}\mathbf{s}_{\text{in}}$ can be decomposed into three steps: (i) express the signal in the Graph Fourier domain via multiplication by $\mathbf{U}^\top$; (ii) filter the signal by multiplication with $h(\mathbf{\Lambda})$, and (iii) project the filtered signal back into the graph domain via $\mathbf{U}$.

Based on this discussion, are there any advantages of choosing a particular graph shift operator over another? Potentially, yes. As the spectral decompositions of the various possible choices for a graph shift $\mathbf{G}$ can vary significantly, our choice can have a significant effect. However, whether the basis functions induced by the graph shift have favorable characteristics for the task at hand is largely dependent on the application context, and there is no choice that will be generally superior in all contexts. Nonetheless, we will focus primarily on the (combinatorial) graph Laplacian as our graph shift operator. This choice is on one hand motivated by the mathematical properties of the Laplacian. On the other hand, as we will see later, the Laplacian can be generalized in a natural way to simplicial complexes, which allows us to draw parallels to the processing of signals on graphs.

Importantly, the graph Laplacian $\mathbf{L}$ is positive semidefinite, and thus all its eigenvalues are real and non-negative, which enables us to interpret them as frequencies. In particular, we can order the GFT basis vectors (eigenvectors) according to these frequencies. This ordering indeed captures the amount of signal variation along the edges of the graph as we can see by considering the eigenvalues in terms of the Rayleigh quotient

$$r(\mathbf{s}) = \frac{\mathbf{s}^\top \mathbf{L}\mathbf{s}}{\mathbf{s}^\top \mathbf{s}} = \frac{\sum_{ij} A_{ij}(s_i - s_j)^2}{2\|\mathbf{s}\|^2}.$$

It follows, that eigenvectors associated with small eigenvalues have small variation along the graph edges (i.e., low frequency) and eigenvectors associated with large eigenvalues show large variation along edges (i.e., high frequency). Specifically, eigenvectors with eigenvalue 0 are constant over connected components, akin to a constant signal in the time domain.

12.3.2 Illustrative Applications of GSP

To make our discussion of signal processing on graphs more concrete, we consider three application scenarios. Later, we will use these applications as guiding examples to illustrate how the ideas of signal processing on graphs can be translated to simplicial complexes.

12.3.2.1 Signal Smoothing and Denoising

Consider a true signal $\mathbf{y}_0$ of interest supported on the set of nodes $\mathcal{V}$. In many settings, we only observe a noisy version $\mathbf{y}$ of it, i.e., $\mathbf{y} = \mathbf{y}_0 + \boldsymbol{\epsilon} \in \mathbb{R}^N$, where $\boldsymbol{\epsilon}$ is a vector of zero-mean white Gaussian noise. For instance, we may consider that $\mathbf{y}_0$ corresponds to a measurement of a sensor in a network, or an opinion of a person in a social network. Our goal is now to recover the true signal $\mathbf{y}_0$, a procedure that is called *denoising* or *smoothing* in GSP [10, 11, 38].

To make this problem well posed, we assume that the signal is *smooth* with respect to the graph structure, i.e., nodes that are connected should have a similar signal [15, 27]. This assumption translates into the optimization problem

$$\min_{\hat{\mathbf{y}}} \{\|\hat{\mathbf{y}} - \mathbf{y}\|_2^2 + \alpha \hat{\mathbf{y}}^\top \mathbf{L}\hat{\mathbf{y}}\}, \tag{12.13}$$

where $\hat{\mathbf{y}}$ is the estimate of the true signal $\mathbf{y}_0$. The coefficient $\alpha > 0$ can be interpreted as a regularization parameter that trades off the smoothness promoted by minimizing the quadratic form $\hat{\mathbf{y}}^\top \mathbf{L}\hat{\mathbf{y}} = \sum_{ij} A_{ij}(\hat{y}_i - \hat{y}_j)^2/2$ and the fit to the observed signal in terms of the squared 2-norm. The optimal solution for (12.13) is given by

$$\hat{\mathbf{y}} = (\mathbf{I} + \alpha\mathbf{L})^{-1}\mathbf{y} = \mathbf{H}_1\mathbf{y}. \tag{12.14}$$

Note, in particular, that $\mathbf{H}_1$ is a linear graph filter.

We can also obtain an estimate of the signal using the iterative smoothing operation

$$\hat{\mathbf{y}} = (\mathbf{I} - \mu\mathbf{L})^{k}\mathbf{y} = \mathbf{H}_2\mathbf{y}, \tag{12.15}$$

for a certain fixed number of iterations k and a suitably chosen update parameter μ. This may be interpreted in terms of k gradient descent steps, i.e., discretized gradient flow dynamics, for the potential function $\hat{\mathbf{y}}^\top\mathbf{L}\hat{\mathbf{y}}$ defining the regularization cost.

Both the denoising operator $\mathbf{H}_1$ and the smoothing operator $\mathbf{H}_2$ defined in (12.14) and (12.15) are instances of *low-pass filters*, i.e., the frequency responses $h(\boldsymbol{\lambda}) = \mathrm{diag}(\mathbf{U}^\top\mathbf{H}\mathbf{U})$ are vectors of non-increasing (decreasing) values [48]. Since eigenvectors with small eigenvalues show smaller variation along the graph edges, the low-pass filtering operation guarantees that variations over neighboring nodes are smoothed out. This is precisely in line with the intuition underpinning the optimization problem (12.13).

12.3.2.2 Graph Signal Interpolation

Given signal values, called *labels*, for a subset of the nodes $\mathcal{V}^L \subset \mathcal{V}$ of a graph, another common task in GSP is to interpolate the signal on unlabeled nodes $\mathcal{V}^U = \mathcal{V} \setminus \mathcal{V}^L$ in the graph [35, 47]. Similar to the signal denoising problem, we assume that connected nodes have similar labels, which translates into the following optimization problem:

$$\min_{\hat{\mathbf{y}}} \left\|\mathbf{B}^\top\hat{\mathbf{y}}\right\|_2^2, \text{ subject to } \hat{y}_i = y_i \text{ for all } v_i \in \mathcal{V}^L. \tag{12.16}$$

Thus, we again aim to minimize the sum-of-squares label difference between connected nodes, but now under the constraint that all observed node labels y_i should be kept in the optimal solution. Importantly, we assume here that these measurements are fully accurate. Further, notice that the objective function in (12.16) can again be written in terms of the quadratic form of the graph Laplacian $\left\|\mathbf{B}^\top\hat{\mathbf{y}}\right\|_2^2 = \hat{\mathbf{y}}^\top\mathbf{L}\hat{\mathbf{y}}$, highlighting again the inherent low-pass modeling assumption.

12.3.2.3 Graph Neural Networks

A recent extension to the domain of GSP is the introduction of graph neural networks [7] (GNNs) into the processing pipeline. In contrast to standard graph filters, GNN architectures include nonlinearities and learnable weights. GNNs can be built for a range of different tasks including node classification [12, 29], graph classification [21], and link prediction [54]. In intuitive terms, one may think of a graph neural network as a procedure to *automatically* find a nonlinear filter (or dynamical

process) that fits the desired behaviour given by a training set of graph signals and the desired outputs of such a filter on those signals.

One popular architecture is the graph convolutional network (GCN) [29], a generalization of the well-known convolutional neural network architecture for Euclidean data such as time series or images. A GCN can be understood as a form of iterative smoothing (12.15) with interleaved element-wise nonlinearities—usually called activation functions in this context—and learnable weights that perform linear transformations in the feature space

$$\mathbf{Y}_{k+1} = \sigma\left(\mathbf{H}\mathbf{Y}_k\mathbf{W}_{k+1}\right) \tag{12.17}$$

We run the network for K iterations and define $\mathbf{Y}_K$ as the output of the graph convolution. The input features F_0 are collected in the columns of $\mathbf{Y}_0 \in \mathbb{R}^{N\times|F_0|}$. Here, $\mathbf{H}$ is a shift-invariant graph filter based on the graph's adjacency or Laplacian matrix, which is adapted to the task at hand via a set $\left\{\mathbf{W}_k \in \mathbb{R}^{|F_{k-1}|\times|F_k|}\right\}_{k=1}^{K}$ of learnable weight matrices.

Note that we can interpret the GNN in terms of a dynamical system, if we treat each layer in the GNN as corresponding to one time-step. For simplicity, let us first consider the case where $\sigma(\cdot)$ is the identity mapping. Then, (12.17) can be expressed as a linear graph filter that is applied to each feature individually and whose outputs are linearly combined at each node using the matrices $\{\mathbf{W}_k\}$. That is,

$$\mathbf{Y}_K = \mathbf{H}^K\mathbf{Y}_0\mathbf{W}_1\cdots\mathbf{W}_K, \tag{12.18}$$

where $\mathbf{H}^K$ is itself a shift-invariant graph filter. From here, the iterative smoothing (12.15) can easily be recovered by restricting $\mathbf{Y}_k$ to only one feature and setting $\mathbf{H} := \mathbf{I} - \mu\mathbf{L}$. The key benefit of GCNs, however, lies in the interleaved nonlinearities and the linear combination of different features, which enables the network to learn more sophisticated relationships between nodes based on their neighborhoods and node features.

Recurrent graph neural networks take this idea of a nonlinear aggregation of information across neighborhoods to the extreme. In the basic case, their state evolution can be described as

$$\mathbf{Y}_{k+1} = \sigma\left(\mathbf{H}\mathbf{Y}_k\mathbf{W}\right)$$

and iterations continue until a stable equilibrium for all node states is reached [53], or some other predefined stopping criterion is fulfilled. Notice that, unlike in (12.17), the same (learnable) weights are shared across all layers, whose number does not have to be fixed a priori.

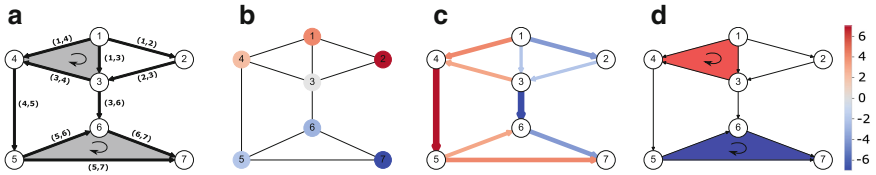


Fig. 12.2 Signals on simplicial complexes of different order. a Structure of the simplicial complexes used as a running example in the text. Arrows represent the chosen reference orientation. Shaded areas correspond to the 2-simplices {1, 3, 4} and {5, 6, 7}. **b** Signal on 0-simplices (nodes). **c** Signal on 1-simplices (edges). **d** Signal on 2-simplices (triangles). Reproduced from [46]

12.4 Signal Processing on Simplicial Complexes

In this section, we extend our analysis of graph signal processing to the case where signals are not only supported on nodes, but also on higher-order structures such as edges, triangles, and so on. One way to analyze such type of data is to model structures via *simplicial complexes* (SC). We first briefly review the formalism of simplicial complexes. We then discuss how the Hodge Laplacian provides an extension of the graph Laplacian for higher-order networks that can serve as a shift operation for SCs, and enables us to extend denoising and interpolating methods for signals on higher-order networks.

12.4.1 Brief Recap of Simplicial Complexes

Given a finite set of vertices $\mathcal{V}$, a k-simplex (or simplex of order k) $\mathcal{S}^k$ is a subset of $\mathcal{V}$ with cardinality $k+1$. A simplicial complex $\mathcal{X}$ is a set of simplices such that for any k-simplex $\mathcal{S}^k$ in $\mathcal{X}$, any subset of $\mathcal{S}^k$ must also be in $\mathcal{X}$.

Example 12.1 Consider the simplicial complex given in Fig. 12.2a. Simplices of order 0 and 1 can be understood as nodes and as edges, respectively. Simplices of order 2 correspond to filled triangular faces.

We can define a relation between k-simplices and simplices of order $(k+1)$ as follows: We call a k-simplex $\mathcal{S}^k$ a *face* of $\mathcal{S}^{k+1}$ if $\mathcal{S}^k$ is a subset of $\mathcal{S}^{k+1}$. Likewise, $\mathcal{S}^{k+1}$ is a *co-face* of $\mathcal{S}^k$ if $\mathcal{S}^{k+1}$ has exactly one additional element than $\mathcal{S}^k$.

Example 12.2 Consider again Fig. 12.2a. The edges {1, 3}, {1, 4} and {3, 4} are all faces of the 2-simplex {1, 23, 4}. The 2-simplex {5, 6, 7} is a co-face of each of the edges {5, 6}, {5, 7} and of {6, 7}.

To enable algebraic computations with simplicial complexes, we fix an arbitrary ordering of the nodes in the graph. This ordering induces an orientation for each simplex via the increasing order of the vertex labels. Note that this orientation is distinct from the notion of a direction found in a directed graph. Having defined an

orientation for each simplex, we can now keep track of the relationships between simplices of different orders by using linear maps called *boundary operators*. In our context, these boundary operators are nothing but matrices $\mathbf{B}_k$, whose rows are indexed by $(k-1)$-simplices and whose columns are indexed by k-simplices. In particular, the ijth entry of $\mathbf{B}_k$ is $+1$ (or -1) if the ith $(k-1)$-simplex is included in the jth k-simplex and their orientation is aligned (or anti-aligned), and 0 otherwise.

Example 12.3 In Fig. 12.2a, we fixed an arbitrary orientation by numbering the nodes from 1 to 7. The orientation of 2-simplices is chosen such that nodes appear in increasing order. For this simplicial complex, the boundary operators can be represented as follows:

$$\mathbf{B}_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{c} \begin{array}{cccccccccc} (1,2) & (1,3) & (1,4) & (2,3) & (3,4) & (3,6) & (4,5) & (5,6) & (5,7) & (6,7) \end{array} \\ \begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$\mathbf{B}_2 = \begin{array}{c} \\ (1,2) \\ (1,3) \\ (1,4) \\ (2,3) \\ (3,4) \\ (3,6) \\ (4,5) \\ (5,6) \\ (5,7) \\ (6,7) \end{array} \begin{array}{c} \begin{array}{cc} (1,3,4) & (5,6,7) \end{array} \\ \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ -1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} \end{array}$$

Note that $\mathbf{B}_1$ is nothing more than the incidence matrix of the graph and $\mathbf{B}_2$ is the edge-to-triangle incidence matrix.

12.4.2 The Hodge Laplacian as a Shift Operator for Simplicial Complexes

Given signals supported on the k-simplices of an SC, we need to define an appropriate shift operator in order to translate the results from the GSP setting to SCs. To this end, using the boundary operators $\mathbf{B}_k$ described above, we extend the definition of the

graph Laplacian for simplicial complexes. The *kth combinatorial Hodge Laplacian* is given by [18, 31]:

$$\mathbf{L}_k = \mathbf{B}_k^\top \mathbf{B}_k + \mathbf{B}_{k+1}\mathbf{B}_{k+1}^\top. \tag{12.19}$$

Note in particular that the graph Laplacian corresponds to $\mathbf{L}_0 = \mathbf{B}_1\mathbf{B}_1^\top$, with $\mathbf{B}_0 := 0$. Similar to the graph Laplacian, weighted versions of the Hodge Laplacian can be defined as well, but we stick to the unweighted versions here for simplicity.

While we have discussed general SCs so far, to make our discussion more concrete we will focus on signals supported on 1-simplices, which may be interpreted as edge flows. This choice is motivated from a practical point of view, as in many application scenarios we are confronted with flows supported on the edges of a network, e.g., in transportation and supply networks, or networks defined via information flows or human mobility. We argue that in those contexts the $\mathbf{L}_1$-Hodge Laplacian is a natural choice for a shift operator for signals supported on edges (1-simplices). The ensuing discussions are still applicable to signals supported on higher-order components of simplicial complexes, which may come up in domains such as electromagnetics [13]. Indeed, there are close connections between signals on simplicial complexes and differential forms on manifolds.

Specifically, similar to the graph Laplacian, the Hodge Laplacian is positive semi-definite, which ensures that we can interpret its eigenvalues in terms of non-negative frequencies. Moreover, these frequencies are again aligned with a notion of signal-smoothness displayed by the eigenvectors of the Hodge Laplacian. This notion of smoothness can be understood by means of the so-called *Hodge decomposition* [24, 31, 44], which states that the space of k-simplex signals can be decomposed into three orthogonal subspaces

$$\mathbb{R}^{n_k} = \mathrm{im}(\mathbf{B}_{k+1}) \oplus \mathrm{im}(\mathbf{B}_k^\top) \oplus \ker(\mathbf{L}_k), \tag{12.20}$$

where $\mathrm{im}(\cdot)$ and $\ker(\cdot)$ are shorthand for the *image* and *kernel* spaces of the respective matrices, $\oplus$ represents the union of orthogonal subspaces, and n_k is the cardinality of the space of signals on k-simplices (i.e., $n_0 = n$ for the node signals, and $n_1 = |\mathcal{E}|$ for edge signals). Here we have (i) made use of the fact that a signal on a finite dimensional set of n_k simplices is isomorphic to $\mathbb{R}^{n_k}$; and (ii) implicitly assumed that we are only interested in real-valued signals and thus a Hodge decomposition for a real valued vector space (see [31] for a more detailed discussion).

For the edge-space Hodge Laplacian $\mathbf{L}_1$ of an SC, this Hodge decomposition is the discrete analogue of the well-known vector calculus result that any vector field can be decomposed into gradient, curl, and harmonic components (see Fig. 12.3 for an illustration). First, the space $\mathrm{im}(\mathbf{B}_1^\top) = \{\mathbf{f} = \mathbf{B}_1^\top \mathbf{v},\ \text{for some } \mathbf{v} \in \mathbb{R}^n\}$ can be considered as the space of gradient flows (or potential flows). Specifically, we can create any such gradient flow by (i) assigning a scalar potential to all of the n nodes, and (ii) inducing a flow along the edges according to the difference of the potentials on the endpoints. Clearly, any such flow cannot have a positive net sum along any closed path in the graph and the space orthogonal to $\mathrm{im}(\mathbf{B}_1^\top)$ is thus called the cycle space. As indicated by (12.20), the cycle space is spanned by two types of cyclic flows. The

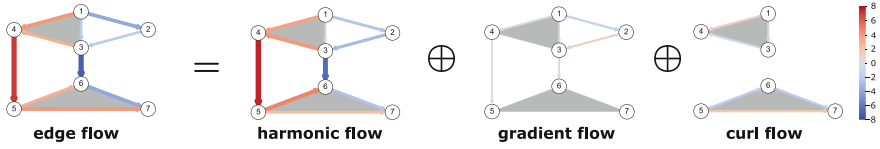


Fig. 12.3 Hodge decomposition of the edge flow in the example from Fig. 12.2. Any edge flow (left) can be decomposed into a harmonic flow, a gradient flow, and a curl flow. Reproduced from [46]

space $\mathrm{im}(\mathbf{B}_2)$ consists of curl flows that can be composed of linear combinations of local circulations along any 2-simplex (triangular face). Specifically, we may assign a scalar potential to each oriented 2-simplex, and consider the induced flows $\mathbf{f} = \mathbf{B}_2\mathbf{t}$, where $\mathbf{t}$ is the vector of 2-simplex potentials. Finally $\ker(\mathbf{L}_1)$ is the harmonic space, whose elements correspond to global circulations that cannot be represented as linear combinations of curl flows.

Importantly, these three subspaces are spanned by certain subsets of eigenvectors of $\mathbf{L}_1$ as described by the following result, which can be verified by direct computation [2, 44].

Theorem 12.1 *Let $\mathbf{L}_1 = \mathbf{B}_1^\top\mathbf{B}_1 + \mathbf{B}_2\mathbf{B}_2^\top$ be the Hodge 1-Laplacian of a simplicial complex. The eigenvectors with nonzero eigenvalues of $\mathbf{L}_1$ consist of two groups that span the gradient space and the curl space, respectively:*

- *Consider an eigenvector $\mathbf{v}_i$ of the graph Laplacian $\mathbf{L}_0$ with nonzero eigenvalue λ_i. Then, $\mathbf{u}_{grad}^{(i)} = \mathbf{B}_1^\top\mathbf{v}_i$ is an eigenvector of $\mathbf{L}_1$ with the same eigenvalue λ_i. Moreover, $\mathbf{U}_{grad} = [\mathbf{u}_{grad}^{(1)}, \mathbf{u}_{grad}^{(2)}, \ldots]$ spans the space of all gradient flows.*
- *Consider an eigenvector $\mathbf{t}_i$ of the matrix $\mathbf{T} = \mathbf{B}_2^\top\mathbf{B}_2$ with nonzero eigenvalue θ_i. Then, $\mathbf{u}_{curl}^{(i)} = \mathbf{B}_2\mathbf{t}_i$ is an eigenvector of $\mathbf{L}_1$ with the same eigenvalue θ_i. Moreover $\mathbf{U}_{curl} = [\mathbf{u}_{curl}^{(1)}, \mathbf{u}_{curl}^{(2)}, \ldots]$ spans the space of all curl flows.*

The above result shows that, unlike in the case of node signals, edge-flow signals can have a high frequency contribution due to two different types of (orthogonal) basis components being present in the signal: a high frequency may arise both due to a curl component as well as a strong gradient component present in the edge-flow. This has certain consequences for the filtering of edge signals that we will discuss in more detail in the following section.

12.4.3 Illustrative Applications

In the following subsections, we revisit the three application scenarios outlined in the context of graphs, but this time focusing on edge flows supported on general SCs. As we will see, in this context the Hodge Laplacian becomes a natural substitute for the graph Laplacian. While most of the mathematical formulations can be carried out in essentially the same way when using this substitution, it is important to consider how the interpretation of smoothing and denoising changes when using the Hodge Laplacian in the edge space as a shift operator.

12.4.3.1 Flow Smoothing and Denoising

Let us reconsider the problem of smoothing and denoising for oriented edge-signals $\mathbf{f}^0 \in \mathbb{R}^E$ (flows) supported on a simplicial complex $\mathcal{X}$. Let us assume again that we cannot observe these flows directly, but we get to see a noisy signal $\mathbf{f} = \mathbf{f}^0 + \boldsymbol{\epsilon}$, where

$\boldsymbol{\epsilon}$ is a zero-mean white Gaussian noise vector. As in the case of node signals, our objective is to recover the true underlying signal $\mathbf{f}^0$. By analogy with the estimation problem on graphs, we consider solving the following optimization program

$$\min_{\hat{\mathbf{f}}} \left\{ \left\| \hat{\mathbf{f}} - \mathbf{f} \right\|_2^2 + \alpha \hat{\mathbf{f}}^\top \mathbf{Q} \hat{\mathbf{f}} \right\}, \tag{12.21}$$

with optimal solution $\hat{\mathbf{f}} = \mathbf{H}_Q \mathbf{f} := (\mathbf{I} + \alpha \mathbf{Q})^{-1} \mathbf{f}$. Like before, the quadratic form $\hat{\mathbf{f}}^\top \mathbf{Q} \hat{\mathbf{f}}$ acts again as regularizer. Since the filter $\mathbf{H}_Q$ will inherit the eigenvectors of the matrix $\mathbf{Q}$, the eigenvectors will form a canonical basis for the filtered signal. A natural choice for a regularizer is thus an appropriate (simplicial) shift operator.

Here we discuss three possible choices for the regularizer (shift operator) $\mathbf{Q}$: (i) the graph Laplacian $\mathbf{L}_{\mathrm{LG}}$ of the line-graph [45] of the underlying graph skeleton of the complex $\mathcal{X}$, i.e., the line-graph of the graph induced by the 0-simplices (nodes) and 1-simplices (edges) of $\mathcal{X}$; (ii) the edge Laplacian $\mathbf{L}_e = \mathbf{B}_1^\top \mathbf{B}_1$, i.e., a form of the Hodge Laplacian that ignores all 2-simplices in the complex $\mathcal{X}$ such that $\mathbf{B}_2 = 0$; (iii) the Hodge Laplacian $\mathbf{L}_1 = \mathbf{B}_1^\top \mathbf{B}_1 + \mathbf{B}_2 \mathbf{B}_2^\top$ that takes into account all the triangles of $\mathcal{X}$ as well. To gain some intuition, let us illustrate the effects of these choices by means of an example.

Example 12.4 Figure 12.4a displays a conservative cyclic flow on the edges of an SC, i.e., all of the flow entering a node exits the node again. This flow is then distorted by a Gaussian noise vector $\boldsymbol{\epsilon}$ in Fig. 12.4b. The estimation error produced by the filter based on the line-graph (Fig. 12.4c) is comparatively worse than the estimation performance of the edge Laplacian (Fig. 12.4d) and the Hodge Laplacian (Fig. 12.4e) filters: Specifically the 2-norm of the error is 36.54 (line graph) versus 1.95 (Edge Laplacian), and 1.02 (Hodge Laplacian) respectively.

To explain the results obtained from the individual filters in the above example, it is essential to realize that the eigenvectors of the regularizer $\mathbf{Q}$ associated with small eigenvalues will incur a small regularization penalty. In the case of the line-graph Laplacian, these eigenvectors are determined by the edge connectivity: the low-frequency eigenvectors correspond to signals in which adjacent edges in the simplicial complex have a small difference. This is equivalent to the notion that low-frequency modes in the node space do not vary a lot on tightly connected nodes on a graph. However, for flow signals this notion of smoothness induced by the line-graph as shift operator is often inappropriate. Specifically, many real-world flow signals are approximately conservative, in that most of the flow signal entering a node exits the node again, but the relative allocation of the flow to the edges does not have to be similar. Accordingly, as can be seen from Fig. 12.4c, the line graph filtering operation leads to an increased error compared to the noisy input signal [45]. Another reason for this behavior is that the line-graph Laplacian does not reflect the arbitrary orientation of the edges, so that the performance is not invariant to the chosen sign of the flow.

Unlike the line-graph Laplacian, the Edge Laplacian captures a notion of flow conservation, and its zero frequency eigenvectors correspond to cyclic flows [45]. To

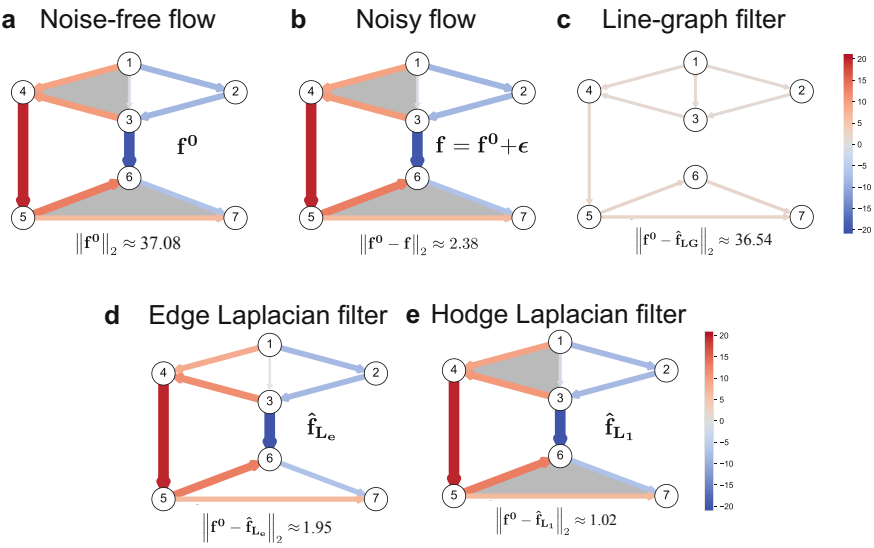


Fig. 12.4 Flow smoothing on an SC. a An SC with a pre-defined and oriented flow $\mathbf{f}^0$. **b** The observed flow is a noisy version of the flow $\mathbf{f}^0$, i.e., $\mathbf{f}^0$ is distorted by a Gaussian white noise vector $\boldsymbol{\epsilon}$. **c** We denoise the flow by applying a Laplacian filter based on the line-graph. This filter performs worse compared to the edge space filters in (**d**) and (**e**) that account for flow conservation. **d** Denoised flow obtained after applying the filter based on the edge Laplacian. **e** Denoised flow obtained after applying the filter based on the Hodge Laplacian. The estimation error is lower than in the edge Laplacian case as the filter accounts for filled faces in the graph. Reproduced from [46]

see this, it is insightful to inspect the quadratic regularizer $\mathbf{L}_e = \mathbf{B}_1^\top \mathbf{B}_1$. Note that this quadratic form can be written as $\mathbf{f}^\top \mathbf{L}_e \mathbf{f} = \|\mathbf{B}_1 \mathbf{f}\|_2^2$. This is precisely the (summed) squared divergence of the flow signal $\mathbf{f}$, as each entry $(\mathbf{B}_1\mathbf{f})_i$ corresponds to the difference of the inflow and outflow at node i. As a consequence, all cyclic flows will induce zero cost for the regularizer $\mathbf{f}^\top \mathbf{L}_e \mathbf{f}$. Stated differently, any flow that is *not* divergence free, i.e., not cyclic, will be penalized by the quadratic form. Since by the fundamental theorem of linear algebra $\ker(\mathbf{B}_1) \perp \mathrm{im}(\mathbf{B}_1^\top)$, any such non-cyclic flow can be written as a gradient flow $\mathbf{f}_{\mathrm{grad}} = \mathbf{B}_1^\top \mathbf{v}$ for some vector $\mathbf{v}$ of scalar node potentials—in line with the Hodge decomposition discussed in (12.20).

In contrast to the Edge Laplacian, the full Hodge Laplacian $\mathbf{L}_1$ includes the additional regularization term $\mathbf{f}^\top \mathbf{B}_2 \mathbf{B}_2^\top \mathbf{f} = \|\mathbf{B}_2^\top \mathbf{f}\|_2^2$, which may induce a non-zero cost even for certain cyclic flows. More precisely, any curl flow $\mathbf{f}_{\mathrm{curl}} = \mathbf{B}_2 \mathbf{t}$, for some vector $\mathbf{t}$ will have a non-zero penalty. This penalty is incurred despite the fact that $\mathbf{f}_{\mathrm{curl}}$ is a cyclic flow by construction: since $\mathbf{B}_1 \mathbf{f}_{\mathrm{curl}} = \mathbf{B}_1 \mathbf{B}_2 \mathbf{c} = 0$, the vector $\mathbf{f}_{\mathrm{curl}}$ is clearly in the cycle space; see also discussion in Sect. 12.4.2. The additional regularization term $\|\mathbf{B}_2^\top \mathbf{f}\|_2^2$ may thus be interpreted as a squared curl flow penalty.

From a signal processing perspective, the $\mathbf{L}_1$ based filter thus allows for a more refined notion of a smooth signal. Unlike in the Edge Laplacian filter, not all cyclic flows can be constructed from frequency (eigenvalue) 0 basis signals. Instead a signal can have a high-frequency even if it is cyclic, when it has a high curl compo-

nent. Hence, by constructing simplicial complexes with appropriate (triangular) 2-simplices, we have additional modeling flexibility for shaping the frequency response of an edge-flow filter. In our example above, this is precisely what leads to an improvement in the filtering performance. Indeed the displayed "ground truth" signal is a harmonic function with respect to the simplicial complex and thus does not contain any curl components. We remark that the eigenvector basis of $\mathbf{L}_e$ can always be chosen to be identical to the eigenvectors of $\mathbf{L}_1$; thus, we may represent any signal in exactly the same way in a basis of $\mathbf{L}_e$ or $\mathbf{L}_1$. Thus the difference in filtering performance is not due to the chosen eigenvector basis, but only due to the eigenvalues: the frequencies associated with all cyclic vectors will be 0 for the Edge Laplacian, while there will be cyclic flows with nonzero frequencies for $\mathbf{L}_1$, in general. This emphasizes that the construction of faces is an important modeling choice when defining the appropriate notion of a smooth signal.

12.4.3.2 Interpolation and Semi-supervised Learning

We now focus on the interpolation problem for data supported on the edges of a simplicial complex. Let us suppose that we measure signals on a subset of the edges in $\mathcal{X}$, i.e., we are given a set of labeled edges $\mathcal{E}^L \subset \mathcal{E}$, with cardinality $|\mathcal{E}^L| = E_L$. The objective is to estimate the labels of unobserved or unmeasured edges in the set $\mathcal{E}^U \equiv \mathcal{E} \backslash \mathcal{E}^L$, whose cardinality we will denote by $|\mathcal{E}^U| = E_U$. Following [26], we will again start by considering the problem setup with no 2-simplices ($\mathbf{B}_2 = 0$), before we consider the general case in which 2-simplices are present.

To derive a well-defined problem for imputing the remaining edge-flows, we need to assume that the true signal has some low-dimensional structure. Following our above discussions, we will again assume that the true signal has a low-pass characteristic in the sense of the Hodge 1-Laplacian, i.e., that the edge flows are mostly conserved. Let $\hat{\mathbf{f}}$ denote the vector of the true (partly measured) edge-flow. A convenient loss function to promote flow conservation is then again the sum-of-squares vertex divergence

$$\left\| \mathbf{B}_1 \hat{\mathbf{f}} \right\|_2^2 = \hat{\mathbf{f}}^\top \mathbf{B}_1^\top \mathbf{B}_1 \hat{\mathbf{f}} = \hat{\mathbf{f}}^\top \mathbf{L}_e \hat{\mathbf{f}}. \tag{12.22}$$

Accordingly, we formalize the flow interpolation problem

$$\min_{\hat{\mathbf{f}}} \left\| \mathbf{B}_1 \hat{\mathbf{f}} \right\|_2^2 + \alpha^2 \cdot \left\| \hat{\mathbf{f}} \right\|_2^2 \text{ s.t. } \hat{f}_r = f_r, \text{ for all measured edges } r \in \mathcal{E}^L, \tag{12.23}$$

Note that, in contrast to the node signal interpolation problem, we have to add an additional regularization term $\|\hat{\mathbf{f}}\|_2^2$ to guarantee the uniqueness of the optimal solution. If there is any independent cycle in the network for which we have no measurement available, we may otherwise add any flow along such a cycle while not changing the divergence in (12.22). To remedy this aspect, we simply add a 2-norm regularization

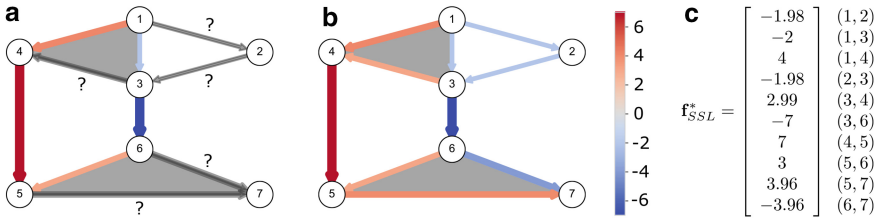


Fig. 12.5 Semi-supervised learning for edge flow. a Synthetic flow. 50% of the edges are labeled. Labeled edges are colored based on the value of their flow. The remaining edges in grey are inferred from the procedure explained in the text. **b** Edge flow obtained after applying the semi-supervised algorithm in (12.24). **c** Numerical value of the inferred signal. Reproduced from [46]

that promotes small edge-flow magnitudes by default. While other regularization terms are possible, with this formulation we can rewrite the above problem in a least squares form as described next.

To arrive at a least-squares formulation, we consider a trivial feasible solution $\hat{\mathbf{f}}^0$ for (12.23) that satisfies $\hat{f}^0_r = f_r$ if $r \in \mathcal{E}^L$ and $\hat{f}^0_r = 0$ otherwise. Let us now define the expansion operator $\mathbf{\Phi}$ as the linear map from $\mathbb{R}^{E_U}$ to $\mathbb{R}^E$ such that the true flow $\mathbf{f}$ can be written as $\mathbf{f} = \hat{\mathbf{f}}^0 + \mathbf{\Phi}\mathbf{f}^U$, where $\mathbf{f}^U \in \mathbb{R}^{E_U}$ is the vector of the unmeasured true edge-flows. Reducing the number of variables considered in this way, we can convert the constrained optimization problem (12.23) into the following equivalent unconstrained least-squares estimation problem for the *unmeasured* edges $\hat{\mathbf{f}}^U$:

$$\hat{\mathbf{f}}^{U*} = \arg\min_{\hat{\mathbf{f}}^U} \left\| \begin{bmatrix} \mathbf{B}_1\mathbf{\Phi} \\ \alpha\mathbf{I} \end{bmatrix} \hat{\mathbf{f}}^U - \begin{bmatrix} -\mathbf{B}_1\mathbf{f}^0 \\ 0 \end{bmatrix} \right\|_2^2 . \tag{12.24}$$

We illustrate the above procedure by the following example.

Example 12.5 We consider the network structure in Fig. 12.2a. The ground truth signal is $\mathbf{f} = [-2, -2, 4, -2, 3, -7, 7, 3, 4, -4]^\top$. We pick five labeled edges at random (colored in Fig. 12.5a). The goal is to predict the labels of the unlabeled edges (in grey with a question mark in Fig. 12.5a). The set of labeled edges is $\mathcal{E}^L = \{(1, 3), (1, 4), (3, 6), (4, 5), (5, 6)\}$. The set of unlabeled edges is $\mathcal{E}^U = \{(1, 2), (2, 3), (3, 4), (5, 7), (6, 7)\}$. Solving the optimization program (12.24), we obtain the predicted signal $\mathbf{f}^*_{SSL}$ in Fig. 12.5b. Numerical values are given in Fig. 12.5c. The Pearson correlation coefficient between $\mathbf{f}$ and $\mathbf{f}^*_{SSL}$ is 0.99. The 2-norm of the error is 0.064.

Analogous to our discussion above, it may be relevant to include 2-simplices for the signal interpolation problem. We interpret such an inclusion of 2-simplices in two ways. From the point of view of the cost function, it implies that instead of penalizing primarily gradient flows (which have nonzero divergence), we in addition penalize certain cyclic flows, namely those that have a nonzero curl component. From a signal processing point of view, it means that we are changing what we consider a smooth (low-pass) signal, by adjusting the frequency representation of certain flows. Accordingly, one possible formulation of the signal interpolation problem, including

information about 2 simplices is

$$\hat{\mathbf{f}}^{\star} = \arg\min_{\hat{\mathbf{f}}} \left\| \mathbf{B}_1 \hat{\mathbf{f}} \right\|_2^2 + \left\| \mathbf{B}_2^{\top} \hat{\mathbf{f}} \right\|_2^2 + \alpha^2 \left\| \hat{\mathbf{f}} \right\|_2^2, \tag{12.25}$$

subject to the constraint that the components of $\hat{\mathbf{f}}$ corresponding to measured flows are identical to those measurements. As in (12.24), we can convert this program into the following least-squares problem

$$\hat{\mathbf{f}}^{U\star} = \arg\min_{\hat{\mathbf{f}}^U} \left\| \begin{bmatrix} \mathbf{B}_1 \boldsymbol{\Phi} \\ \alpha \mathbf{I} \\ \mathbf{B}_2^{\top} \boldsymbol{\Phi} \end{bmatrix} \hat{\mathbf{f}}^U - \begin{bmatrix} -\mathbf{B}_1 \mathbf{f}^0 \\ 0 \\ -\mathbf{B}_2^{\top} \mathbf{f}^0 \end{bmatrix} \right\|_2^2. \tag{12.26}$$

Remark 12.1 Note that the problem of flow interpolation is tightly coupled to the issue of signal reconstruction from sampled measurements. Indeed, if we knew that the edge signal to be recovered was exactly bandlimited [2], then we could reconstruct the edge-signal if we had chosen the edges to be sampled appropriately. Just like the interpolation problem considered here may be seen as a semi-supervised learning problem for edge labels, finding and choosing such optimal edges to be sampled may be seen as an active learning problem in the context of machine learning. While we do not expand further on the choice of edges to be sampled here, we point the interested reader to two heuristic active learning algorithms for edge flows presented in [26]. We also refer the reader to [2, 3] for a theory of sampling and reconstruction of bandlimited signals on simplicial complexes, and to [1] for a similar overview that includes an approach for *topology inference* based on signals supported on simplicial complexes.

12.4.3.3 Simplicial Neural Networks

With a foundation in linear filtering based on the Hodge Laplacian for k-simplex signals, a natural next step is the interleaving of nonlinearities to form convolutional neural networks for data on simplicial complexes. In particular, convolutional layers analogous to (12.17) can be constructed by composing filters, boundary maps, and activation functions. Letting $\mathbf{Y}_0 \in \mathbb{R}^{n_k \times |F_0|}$ gather k-simplex signals in its columns, the layers of a simple convolutional neural network, as done by [17], can be written recursively as

$$\mathbf{Y}_{k+1} = \sigma\left(\mathbf{H}\mathbf{Y}_k\mathbf{W}_{k+1}\right), \tag{12.27}$$

where $\mathbf{H}$ is some suitable polynomial of the Hodge Laplacian $\mathbf{L}_k$, and each $\mathbf{W}_{k+1} \in \mathbb{R}^{|F_k| \times |F_{k+1}|}$ is a (learnable) weight matrix. Variants of this basic architecture allow weights for the upper ($\mathbf{B}_k^{\top}\mathbf{B}$) and lower ($\mathbf{B}_{k+1}\mathbf{B}_{k+1}^{\top}$) components to be learned independently [23], or for signals to be computed over all dimensions of the simplicial complex [8].

When representing a k-simplex signal for $k \geq 1$, the first step is to fix an arbitrary orientation of the simplices, as discussed in Sect. 12.4.1. In doing so, one fixes the signs of the entries in the incidence matrices $\mathbf{B}_k$, $\mathbf{B}_{k+1}$, as well as the signs of a given k-simplex signal. This is a distinct choice compared to architectures for signals on the nodes of a graph, such as graph neural networks, since nodes do not have a notion of orientation associated to them. Fortunately, the signs of the incidence matrices and the signs of the signal written as a real vector are compatible, so that linear filters based on the Hodge Laplacian commute with the choice of orientation: that is, such filters are *equivariant* to the chosen orientation of the k-simplices. This is an important property, since the orientation is indeed arbitrary: the choice of orientation for k-simplex signals is analogous to choosing a coordinate basis in a vector space. Although a choice of basis for a vector space makes computation intuitive, the "true" object is the vector itself, and not the list of coordinates in that basis.

This introduces a new problem for the design of neural networks, since convolutional layers such as (12.27) are not necessarily equivariant to the choice of orientation. This was considered by [23, 42], both advocating for the use of *odd, elementwise, and continuous* activation functions in neural networks for signals on simplicial complexes. Such functions are readily shown to yield architectures equivariant to the arbitrary choice of orientation.

12.5 Further Relations to Dynamical Systems on Graphs and Simplicial Complexes

As we have seen in Sect. 12.2, the graph Laplacian is intimately connected to linear dynamical systems on networks, such as diffusion processes or consensus processes.

In the context of consensus processes [37, 55], we consider networks composed of so called "agents" (entities, devices, people) that are connected by an edge if the corresponding agents interact. At each time step t, each agent is in a certain *state*. For instance, in the context of social networks, the *state* of node i at time t can be understood as the *opinion* of the individual i at time t. Let us denote by $s_i(t)$ the state of node i at time t and store the states of all agents at time t in a vector of states (or *opinions* in the social networks context) $\mathbf{s}(t) = [s_1(t), s_2(t), \cdots, s_n(t)]^\top$. The dynamics of s_i over time can be expressed with the averaging law

$$\dot{s}_i(t) = -\sum_{v_i \sim v_j} (s_i(t) - s_j(t)), \tag{12.28}$$

which we can rewrite as

$$\frac{d\mathbf{s}(t)}{dt} = -\mathbf{L}\mathbf{s}(t), \tag{12.29}$$

where $\mathbf{L}$ is the graph Laplacian.

Given an initial condition $\mathbf{s}(0) = \mathbf{s}_0$, the solution of the above linear dynamics is simply $\mathbf{s}(t) = \exp(-\mathbf{L}t))\mathbf{s}_0$, where $\exp(\cdot)$ denotes the matrix exponential. Note we may interpret this vector equivalently as a filtered graph signal. Stated differently, for every $t > 0$, we may interpret the current state vector of the system as a (low-pass) filtered version of the initial condition $\mathbf{s}_0$.

Following [33], we can also study a generalization of the dynamical system in (12.29) for higher-order Laplacians. Let us consider a discrete time-varying signal $\boldsymbol{\omega}(t)$ of order k. For instance, in the case $k = 1$, $\boldsymbol{\omega}(t)$ is an edge-flow, analog to $\mathbf{f}$ in Sect. 12.4.3, i.e., the i-th component of the vector $\boldsymbol{\omega}(t)$ is the state of the i-th edge, assuming that an ordering of the edges was initially defined. We consider the dynamical system

$$\frac{d\boldsymbol{\omega}(t)}{dt} = -\mathbf{L}_k\boldsymbol{\omega}(t), \qquad \boldsymbol{\omega}(0) = \boldsymbol{\omega}_0, \tag{12.30}$$

where $\mathbf{L}_k$ is the *k-th combinatorial Hodge Laplacian* as defined in (12.19). We can interpret (12.30) as the higher-order analog of the discretized heat equation given in (12.29) (note, however, that this is not a diffusion process). The equilibrium points of this dynamical system are given by the set [33]

$$\{\boldsymbol{\omega} \mid \mathbf{L}_k\boldsymbol{\omega} = 0\} = \ker(\mathbf{L}_k). \tag{12.31}$$

Since $\mathbf{L}_k$ is a positive semi-definite matrix and the system (12.30) is thus (semi-)stable, the dynamics in (12.30) can be seen as a means to compute a k-th order harmonic signal on a generic simplicial complex, starting from any arbitrary k-th order signal. If the simplicial complex has non-empty homology (many *holes*), then the system will converge to an element in the basis of $\ker(\mathbf{L}_k)$ depending on the initial condition. In other words, the above dynamical system acts as a perfect low-pass filter which projects the initial condition (asympotically) into the space of harmonic signals.

In particular, if the simplicial complex has exactly one hole, then for any initial condition, the system will converges to a vector that spans the harmonic subspace. The dynamical system in (12.30) thus offers a decentralized method to compute the *homology classes* in the simplicial complex [34]. This method finds a useful application in real sensor networks [34, 50], in particular for detecting coverage holes in a decentralized manner and without location information [50].

Generalizations of (12.30) have been discussed and analyzed in [14, 51]. In particular, [14] analyzes a nonlinear version of (12.30). For an unweighted SC, this model takes the form:

$$\frac{d\boldsymbol{\omega}(t)}{dt} = -\left(\mathbf{B}_{k+1} f(\mathbf{B}_{k+1}^\top \boldsymbol{\omega}(t)) + \mathbf{B}_k^\top f(\mathbf{B}_k \boldsymbol{\omega}(t))\right), \tag{12.32}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$ is a function that acts componentwise. It is shown in [14] that (12.32) under certain conditions be formulated as the gradient flow for an energy

functional defined by the simplicial complex. Based on these results, the stability of certain steady states in the nonlinear case can be deduced [14]. Note again that, apart from a lack of adjustable weights, there is again a remarkable similarity between the nonlinear equation (12.32) and the nonlinear signal processing approach of the neural network formulation for SCs in (12.27). In particular, both systems are based on alternating applications of a linear transformation and a point-wise nonlinearity. Indeed this close similarity of particular neural network architectures and ordinary differential equations has been the focal point for the development of the so-called neural ODE formulations [9].

12.6 Discussion

Simplicial complexes have emerged as a key modeling framework for abstracting complex systems with higher-order interactions [4, 6, 51]. The majority of previous works have focused on studying the structural properties of such complexes, in particular in the context of topological data analysis [52]. More recently, the study of dynamical processes acting on top of such complexes has gained attention. Here we have discussed signal processing for data supported on simplicial complex as a closely aligned, but different perspective to both of these viewpoints. Rather than trying to understand a dynamical behavior on a (arbitrary but fixed) simplicial complex, in the context of signal processing we aim to obtain a desired filtering output, based on a given input signal—which may be interpreted as aiming to *design* a particular dynamics that achieves a desired target specification as close as possible. We have centered our discussion on the Hodge Laplacian [18, 31] as a key operator whose spectral decomposition provides a unitary basis for signals supported on simplicial complexes, which is tightly coupled to the structural properties of the underlying SC due to the Hodge decomposition. Specifically, focusing on edge-flows, we discussed how the Hodge decomposition can be interpreted as a discrete analog of the well-known decomposition of a continuous vector field into gradient, curl, and harmonic components.

Our discussion opens up a number of avenues for future research. For instance, one pertinent question concerns the "optimal construction" of simplicial complexes from data and how this affects signal processing supported on SCs. This problem is only magnified when considering weighted SCs and corresponding weighted Hodge Laplacians, to which most of the theory discussed here can be readily applied as well [31, 44]. More generally, most of the developed theory can be readily extended to cell complexes, in which the atomic building blocks are not simplices but cells containing any number of nodes. How to choose an appropriate cell complex representation for a given system is a completely open topic that remains to be explored in more detail.

Acknowledgements This work was partially supported by USA NSF under award CCF-2008555. MTS and FF acknowledge partial funding from the Ministry of Culture and Science (MKW) of the

German State of North Rhine-Westphalia (NRW Rückkehrprogramm). We thank Lucille Calmon for carefully checking and providing feedback on the manuscript.

References

1. S. Barbarossa, S. Sardellitti, Topological signal processing: making sense of data building on multiway relations. IEEE Signal Process. Mag. **37**(6), 174–183 (2020)
2. S. Barbarossa, S. Sardellitti, Topological signal processing over simplicial complexes. IEEE Trans. Signal Process. **68**, 2992–3007 (2020)
3. S. Barbarossa, S. Sardellitti, E. Ceci, Learning from signals defined over simplicial complexes, in *IEEE Data Scientific Wrokshop (DSW)* (IEEE, 2018), pp. 51–55
4. F. Battiston, G. Centetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, Networks beyond pairwise interaction: structure and dynamics. Phys. Rep. **874**, 1–92 (2020)
5. C. Berge, *Hypergraphs* (Elsevier, 1989)
6. C. Bick, E. Gross, H.A. Harrington, M.T. Schaub, *What are Higher-Order Networks*? arXiv preprint arXiv:2104.11329 (2021)
7. M.M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst, Geometric deep learning: going beyond Euclidean data. IEEE Signal Process. Mag. **34**(4), 18–42 (2017)
8. E. Bunch, Q. You, G. Fung, V. Singh, Simplicial 2-complex convolutional neural networks, in *NeurIPS Workshop on Topological Data Analysis and Beyond* (2020)
9. R.T. Chen, Y. Rubanova, J. Bettencourt, D. Duvenaud, Neural ordinary differential equations. Adv. Neural Info. Process. Syst. (NeurIPS), 6572–6583 (2018)
10. S. Chen, A. Sandryhaila, J.M. Moura, J. Kovacevic, Signal denoising on graphs via graph filtering, in *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)* (2014), pp. 872–876
11. S. Chen, A. Sandryhaila, J.M. Moura, J. Kovačević, Signal recovery on graphs: variation minimization. IEEE Trans. Signal Process. **63**(17), 4609–4624 (2015)
12. M. Defferrard, X. Bresson, P. Vandergheynst, Convolutional neural networks on graphs with fast localized spectral filtering. Adv. Neural Info. Process. Syst. (NeurIPS), 3844–3852 (2016)
13. G. Deschamps, Electromagnetics and differential forms. Proc. IEEE **69**(6), 676–696 (1981). https://doi.org/10.1109/PROC.1981.12048
14. L. DeVille, *Consensus on Simplicial Complexes, or: The Nonlinear Simplicial Laplacian*. arXiv Preprint (2020)
15. X. Dong, D. Thanou, P. Frossard, P. Vandergheynst, Learning Laplacian matrix in smooth graph signal representations. IEEE Trans. Signal Process. **64**(23), 6160–6173 (2016)
16. D. Easley, J. Kleinberg, *Networks, Crowds, and Markets*, vol. 8 (Cambridge University Press, 2010)
17. S. Ebli, M. Defferrard, G. Spreemann, *Simplicial Neural Networks*. arXiv Preprint (2020)
18. B. Eckmann, Harmonische funktionen und randwertaufgaben in einem komplex. Commentarii Mathematici Helvetici **17**(1), 240–255 (1944)
19. P. Frankl, Extremal set systems, in *Handbook of Combinatorics* (1995)
20. S. Furutani, T. Shibahara, M. Akiyama, K. Hato, M. Aida, Graph signal processing for directed graphs based on the hermitian laplacian, in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (Springer, Heidelberg, 2019), pp. 447–463
21. F. Gama, A.G. Marques, G. Leus, A. Ribeiro, Convolutional neural network architectures for signals supported on graphs. IEEE Trans. Signal Process. **67**(4), 1034–1049 (2018)
22. C. Giusti, R. Ghrist, D.S. Bassett, Two's company, three (or more) is a simplex. J. Comput. Neurosci. **41**(1), 1–14 (2016)
23. N. Glaze, T.M. Roddenberry, S. Segarra, Principled simplicial neural networks for trajectory prediction. Intl. Conf. Mach. Learn. (ICML) (to appear) (2021)
24. L.J. Grady, J.R. Polimeni, *Discrete Calculus: Applied Analysis on Graphs for Computational Science* (Springer Science & Business Media, 2010)

25. A. Hatcher, *Algebraic Topology* (Cambridge University Press, 2002)
26. J. Jia, M.T. Schaub, S. Segarra, A.R. Benson, Graph-based semi-supervised & active learning for edge flows. ACM Intl. Conf. Know. Disc. Data Mining (SIGKDD), 761–771 (2019)
27. V. Kalofolias, How to learn a graph from smooth signals, in *Artificial Intelligence and Statistics*, PMLR (2016), pp. 920–929
28. K.F. Kee, L. Sparks, D.C. Struppa, M. Mannucci, Social groups, social media, and higher dimensional social structures: a simplicial model of social aggregation for computational communication research. Commun. Quar. **61**(1), 35–58 (2013)
29. T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks. Intl. Conf. Learn. Repres. (ICLR) (2017)
30. S. Klamt, U.U. Haus, F. Theis, Hypergraphs and cellular networks. PLoS Comput Biol **5**(5), e1000, 385 (2009)
31. L.H. Lim, Hodge laplacians on graphs. SIAM Rev. **62**(3), 685–715 (2020)
32. A.G. Marques, S. Segarra, G. Mateos, Signal processing on directed graphs: the role of edge directionality when processing and learning from network data. IEEE Signal Process. Mag. **37**(6), 99–116 (2020)
33. A. Muhammad, M. Egerstedt, Control using higher order laplacians in network topologies, in *Proceedings of 17th International Symposium on Mathematical Theory of Networks and Systems* (2006), pp. 1024–1038
34. A. Muhammad, A. Jadbabaie, Decentralized computation of homology groups in networks by gossip, in *American Control Conference* (IEEE, 2007), pp. 3438–3443
35. S.K. Narang, A. Gadde, A. Ortega, Signal processing techniques for interpolation in graph structured data, in *IEEE Intl. Conf. Acoust., Speech and Signal Process. (ICASSP)*, IEEE (2013), pp. 5445–5449
36. M. Newman, *Networks* (Oxford University Press, 2018)
37. R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems. Proc. IEEE **95**(1), 215–233 (2007)
38. M. Onuki, S. Ono, M. Yamagishi, Y. Tanaka, Graph signal denoising via trilateral filter on graph spectral domain. IEEE Trans. Sig. Inf. Process. Over Netw. **2**(2), 137–148 (2016)
39. A.V. Oppenheim, R.W. Schafer, *Discrete-Time Signal Processing*, 3rd edn. (Prentice Hall Press, 2009)
40. A. Ortega, P. Frossard, J. Kovacevic, P. Vandergheynst, Graph signal processing: overview, challenges and applications. Proc. IEEE **106**(5), 808–828 (2018)
41. M. Robinson, *Topological Signal Processing*, vol. 81 (Springer, Heidelberg, 2014)
42. T.M. Roddenberry, S. Segarra, Hodgenet: Graph neural networks for edge data, in *Asilomar Conf. Signals, Systems, and Computers*, IEEE (2019), pp. 220–224
43. A. Sandryhaila, J.M. Moura, Discrete signal processing on graphs. IEEE Trans. Signal Process. **61**(7), 1644–1656 (2013)
44. M.T. Schaub, A.R. Benson, P. Horn, G. Lippner, A. Jadbabaie, Random walks on simplicial complexes and the normalized hodge 1-laplacian. SIAM Rev. **62**(2), 353–391 (2020)
45. M.T. Schaub, S. Segarra, Flow smoothing and denoising: graph signal processing in the edge-space, in *IEEE Global Conf. Signal and Info. Process. (GlobalSIP)*, IEEE (2018), pp. 735–739
46. M.T. Schaub, Y. Zhu, J.B. Seby, T.M. Roddenberry, S. Segarra, Signal processing on higher-order networks: Livin' on the edge... and beyond, in *Signal Processing* p. 108149 (2021). https://doi.org/10.1016/j.sigpro.2021.108149
47. S. Segarra, A.G. Marques, G. Leus, A. Ribeiro, Reconstruction of graph signals through percolation from seeding nodes. IEEE Trans. Signal Process. **64**(16), 4363–4378 (2016)
48. S. Segarra, A.G. Marques, A. Ribeiro, Optimal graph-filter design and applications to distributed linear network operators. IEEE Trans. Signal Process. **65**(15), 4117–4131 (2017)
49. D. Shuman, S. Narang, P. Frossard, A. Ortega, P. Vandergheynst, The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process. Mag. **30**(7), 83–98 (2013)
50. A. Tahbaz-Salehi, A. Jadbabaie, Distributed coverage verification in sensor networks without location information. IEEE Trans. Auto. Control **55**(8), 1837–1849 (2010)

51. J.J. Torres, G. Bianconi, Simplicial complexes: higher-order spectral dimension and dynamics. J. Phys.: Complexity **1**(1), 015,002 (2020)
52. L. Wasserman, Topological data analysis. Ann. Rev. Statistics Appl. **5**, 501–532 (2018)
53. Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, S.Y. Philip, A comprehensive survey on graph neural networks, in *IEEE Trans. Neural Netw. and Learn, Syst* (2020)
54. M. Zhang, Y. Chen, Link prediction based on graph neural networks. Adv. Neural Info. Process. Syst. (NeurIPS), 5171–5181 (2018)
55. Y. Zhu, M.T. Schaub, A. Jadbabaie, S. Segarra, Network inference from consensus dynamics with unknown parameters. IEEE Trans. Sig. Inf. Process. Over Netw. **6**, 300–315 (2020)

Chapter 13
Social Contagion on Higher-Order Structures

Alain Barrat, Guilherme Ferraz de Arruda, Iacopo Iacopini, and Yamir Moreno

Abstract In this Chapter, we discuss the effects of higher-order structures on SIS-like processes of social contagion. After a brief motivational introduction where we illustrate the standard SIS process on networks and the difference between simple and complex contagions, we introduce spreading processes on higher-order structures starting from the most general formulation on hypergraphs and then moving to several mean-field and heterogeneous mean-field approaches. The results highlight the rich phenomenology brought by taking into account higher-order contagion effects: both continuous and discontinuous transitions are observed, and critical mass effects emerge. We conclude with a short discussion on the theoretical results regarding the nature of the epidemic transition and the general need for data to validate these models.

13.1 Introduction

The standard modeling and study of social or biological contagion processes in populations is based on two types of ingredients. First, the evolution of the process within each individual is often described through compartmental models [1, 2], such that each individual is at any time in one of several possible compartments or states. For instance, in the description of many infectious diseases, the considered states include susceptible (S, healthy), infectious (I, having the disease and able to transmit it to others), or recovered (R, cured from the disease and immunized). This type of modeling gives a simplified description of the disease course, abstracting the

A. Barrat (✉) · I. Iacopini
Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France
e-mail: alain.barrat@cpt.univ-mrs.fr

G. Ferraz de Arruda · Y. Moreno
ISI Foundation, Torino, Italy

Y. Moreno
Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, 50018 Zaragoza, Spain

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_13

continuous growth and decrease of the viral load and viral shedding of an individual. The modeling also defines the possible transitions between states: in the SIR model, an S individual can become I upon interaction with I individual(s), and an I individual becomes R upon recovery. In the SIS model instead, an I individual becomes again susceptible upon recovery.

The second type of modeling hypothesis concerns the definition and representation of the interactions between individuals. This representation is crucial as it describes the way in which the process spreads between individuals. Numerous results have been obtained under the simplest homogeneous mixing hypothesis, in which any individual can interact with any other, and contagion occurs with a certain probability per unit time upon each contact [1, 2]. Even within this simplistic picture, the SIS and SIR models exhibit an interesting phenomenology, with a continuous phase transition at the so-called epidemic threshold: when the ratio of the contagion to the recovery rate is smaller than the epidemic threshold, the spread dies out, while it reaches a finite fraction of the population above the threshold. In the SIS case, a steady state is then reached, in which the epidemic is sustained by a non-zero number of individuals.

One of the most successful impacts of network science has been to go beyond the homogeneous mixing hypothesis and study how more realistic structures of interactions between individuals affect the dynamics of compartmental models of contagion processes, and in particular the epidemic threshold [3–6]. Indeed, network-based representations are conveniently used to describe many systems of various nature, including the social structures on which many dynamical processes occur, such as the spread of diseases and of information, the formation of opinions and the diffusion of innovations [4, 7, 8]. In the resulting modeling, the transmission process is assumed to occur through pairwise interactions and through a single exposure: in other words, an infectious individual can transmit the disease to a susceptible one upon a single interaction (along one of the links of the network representation).

While such "simple contagion" frameworks are still widely used in the modeling of infectious diseases, the situation is more complex when dealing with social contagion phenomena, such as the adoption of norms or new products, or the diffusion of rumors. Indeed, empirical evidence has shown that simple epidemic-like contagion processes do not provide a satisfactory description of the complex dynamics occurring when peer influence and reinforcement mechanisms are at work [9–15]. Complex contagion mechanisms have been proposed to account for these effects: broadly speaking, they are defined as any process in which exposure to multiple sources presenting the same stimulus is needed for the contagion to occur [9]. Modeling of complex contagion has been developed in two main directions. On the one hand, threshold models consider that an individual can be convinced to adopt e.g. a new behaviour if and only if a fraction of their contacts larger than a given threshold is already convinced (have already adopted the behaviour) [9, 13, 16–19]. On the other hand, epidemic-like processes have been generalized, with contagion rates that depend on the number of sources of exposure to which an individual is linked [12, 19–24].

From the homogeneous mixing simple contagion models to the complex contagion occurring on complex networks, the assumption of transmission processes occurring along pairwise interactions has remained an ubiquitous and most often undiscussed

norm. It fits well with the representation of social groups as networks, since links of the networks are pairwise associations of nodes (the individuals of the population). However, a number of social phenomena occur as the result of group interactions. Let us consider for instance the adoption of a product or a norm. An individual might be convinced by a single interaction with an adopter (simple contagion), or by successive interactions with two distinct adopters (complex contagion), along the links of their social networks. However, a qualitatively different process is at work if the individual gets convinced as part of a social group of three individuals, the other two being adopters. It might occur because the individual wants to be similar to the rest of the group, or, in a group discussion, the two adopters' arguments might reinforce each other in a way that would be impossible in separate pairwise discussions.

To account for such interactions between individuals occurring in groups of various sizes, it is thus necessary to expand the representation of the social structure from networks, which can only encode pairwise interactions, to higher-order structures, namely hypergraphs [25]: the building blocks of hypergraphs are indeed hyperedges that can join an arbitrary number of nodes. Clearly, the modeling of spreading processes on hypergraphs also implies to generalize contagion processes from pairwise to group processes: one needs for instance to define which contagion events can take place on a hyperedge joining n nodes among which m are infectious. A number of recent works have focused on the definition and study of such models [26–31], and we review in this chapter some of the corresponding approaches and results, highlighting in particular how the obtained behaviour is richer than in the usual (network-based) contagion models. The emerging phenomenology indeed includes both continuous and discontinuous transitions, hysteresis phenomena and critical mass phenomena reminiscent of the recently observed minimal size of committed minorities required to initiate social changes [32].

13.2 Spreading Processes on Higher-Order Structures

Group interactions can be encoded as hyperedges of an hypergraph, where each hyperedge is thus a set $[i_0, i_1, \ldots, i_{k-1}]$ that involves k elements. In this language, pairwise interactions are called 1-hyperedges, 3-body interactions are called 2-hyperedges, etc. In the broadest definition, there are no limitations to the size and relative inclusions of hyperedges. In some cases, it can be convenient to represent a social structure using the more restricted framework of simplicial complexes: such a representation assumes that in any group interaction all the sub-interactions among the group members should be considered as well [33]. While this hypothesis has been used in Ref. [26], further developments have shown that similar dynamical outcomes for contagion processes can be found even under the more general framework of hypergraphs [27]. Thus, in this chapter the latter setup will be used.

As the interactions are not necessarily pairwise anymore, but can occur in groups of more than two individuals, this implies moreover that the models used to describe the contagion processes need to be redefined. In this section, we present a rather

general mathematical formulation of such possible contagion models on higher-order structures, defining it in terms of Bernoulli random variables and Poisson processes. Obtaining results directly from these definitions is, however, very hard, so that we mainly restrict this subsection to the definition of the models and of the quantities of interest, leaving to the following subsections the development of analytical approximations and the numerical simulations.

Mathematically, in the social contagion process the states of the nodes are modeled as Bernoulli random variables, $Y_i = 1$ (with its complementary $X_i = 0$) if the node is active and $Y_i = 0$ otherwise (and then $X_i = 1$). Individual states change either spontaneously or as a consequence of their interactions. Formally, this is a collection of independent Poisson processes. First, we associate to each active node i a Poisson process with parameter δ_i, modeling its spontaneous deactivation, $\{Y_i = 1\} \xrightarrow{\delta_i} \{X_i = 1\}$. This transition is similar to the healing in disease spreading dynamics. On the other hand, spreading processes occur along the hyperedges as follows. For each hyperedge e_j we define a random variable $T_j = \sum_{k \in e_j} Y_k$: T_j is by definition the number of active nodes in the hyperedge. If T_j is equal to or above a given threshold Θ_j, we model the contagion by a Poisson process with parameter λ_j. In other words, if $T_j \geq \Theta_j$, then $\{X_k = 1\} \xrightarrow{\lambda_j} \{Y_k = 1\}$, $\forall k \in e_j$. This corresponds to a threshold process that becomes active only above a critical mass of active nodes. Finally, if $|e_j| = 2$, we assume directed Poisson processes, recovering a traditional SIS contagion process. For the sake of simplicity, we assume that $\delta_i = \delta$ and $\lambda_j = \lambda \times \lambda^*(|e_j|)$, where λ is the control parameter and $\lambda^*(|e_j|)$ is an arbitrary function of the cardinality of the hyperedge. The first assumption considers that every individual deactivates at the same rate. The second condition assumes that a hyperedge that is above its critical-mass threshold activates its nodes with a rate that depends only of its cardinality (scaled by a global control parameter λ). The exact equation describing the resulting dynamics can be written as

$$\frac{d\mathbb{E}(Y_i)}{dt} = \mathbb{E}\left(-\delta Y_i + \lambda (1 - Y_i) \sum_{e_j | i \in e_j} \lambda^*(|e_j|) \sum_{B} \mathbb{1}_{\{Y_i = 0, T_j \geq \Theta_j\}}\right), \tag{13.1}$$

where the first summation is over all hyperedges containing node i and the second over the set B of all possible dynamical micro-states inside the hyperedge e_j. Furthermore, $\mathbb{1}_{\{Y_i = 0, T_j \geq \Theta_j\}}$ is an indicator function depending on both the specific node and the hyperedge, taking the value 1 if $Y_i = 0$ and the critical mass in the hyperedge is reached (i.e., if node i is inactive and can potentially become active), and 0 otherwise. We also use for convenience a global threshold ratio Θ^*, with $\Theta_j = \lceil \Theta^* |e_j| \rceil$.

The order parameter is defined as the expected fraction of active nodes, i.e., $\rho = \frac{1}{N} \sum_i \mathbb{E}(Y_i)$. Although a formal proof is yet lacking, we observed, through simulations and numerical solutions of several analytical approaches, a rather general phenomenology when λ is varied at fixed Θ^*, as illustrated in Fig. 13.1. Two solutions for ρ as a function of λ are generically obtained, here called ρ^{Lower} and ρ^{Upper} ($\rho^{\text{Upper}} > \rho^{\text{Lower}}$). Moreover, under certain conditions, the ρ^{Lower} solution presents

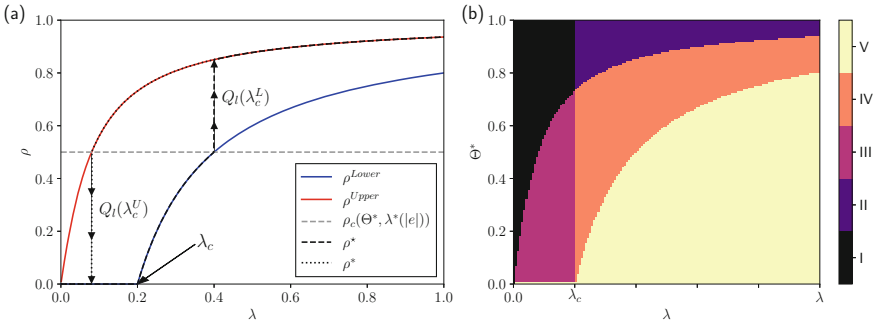


Fig. 13.1 Example of a phase diagram and parameter space for the hyperblob (See Sect. 13.3.2 for details). Panel **a** shows the solutions for a fixed $\Theta^* = 0.5$. The red and blue curves show respectively ρ^{Upper} and ρ^{Lower}. The lower solution presents a second-order phase transition at $\lambda_c = 0.2$. When increasing λ from 0 to 1 the transition from the lower to the upper solution occurs at the intersection of the lower solution with a value of ρ in which the upper solution becomes the only stable one, ρ_c. The jump between the two solutions is $Q_l(\lambda_c^L)$. Similarly, when decreasing back λ, the jump from ρ^{Upper} to ρ^{Lower} takes place when ρ^{Upper} crosses the value ρ_c and becomes unstable, and the density jump is $Q_l(\lambda_c^U)$. Panel **b** shows a sketch of the parameter space. Region I: the system reaches the absorbing state, $\rho = 0$; Region II: only the lower solution is stable; Region III: ρ^{Upper} is stable and $\rho^{\text{Lower}} = 0$ (bistable region below the critical point); Region IV: $\rho^{Upper} > \rho^{Lower} > 0$ and both are stable (bi-stable); Region V: only the upper solution is stable

a continuous phase transition between the absorbing state, where all the individuals are deactivated ($\rho^{\text{Lower}} = 0$), and an active state ($\rho^{\text{Lower}} > 0$). This transition occurs at a critical value of the parameter λ denoted λ_c. A bistable region can also exist, in which the final state depends on the initial condition $\rho(t = 0)$ being below or above a so-called global critical-mass denoted ρ_c. Let us denote by $\rho^\star$ the solution that is obtained if $\rho(t = 0) < \rho_c$ and by ρ^* solution obtained if $\rho(t = 0) \geq \rho_c$. In the bistable region $\lambda_c^{\text{U}} < \lambda < \lambda_c^{\text{L}}$, $\rho^\star = \rho^{\text{Upper}}$ and $\rho^* = \rho^{\text{Lower}}$, while for $\lambda < \lambda_c^{\text{U}}$ we have $\rho^\star = \rho^* = \rho^{\text{Lower}}$, and for $\lambda > \lambda_c^{\text{L}}$, $\rho^\star = \rho^* = \rho^{\text{Upper}}$.

When increasing λ from 0 (forward phase diagram), the system thus first follows ρ^{Lower} and jumps to ρ^{Upper} at λ_c^{L} when ρ^{Lower} becomes unstable. When decreasing back λ (backward phase diagram), the system follows ρ^{Upper} and jumps back to ρ^{Lower} at λ_c^{U} where ρ^{Upper} becomes unstable. The length of these two jumps are defined as

$$Q_l(\lambda_c^X) = \left(\rho^{\text{Upper}} - \rho^{\text{Lower}}\right)_{\lambda=\lambda_c^X}, \tag{13.2}$$

where $Q_l(\lambda_c^X)$ can be $Q_l(\lambda_c^{\text{L}})$ or $Q_l(\lambda_c^{\text{U}})$. These quantities give the sudden change in the fraction of active nodes at these jumps. These concepts are exemplified in Fig. 13.1a, where we show an example obtained for a homogeneous hypergraph composed of a random regular network and a hyperedge containing all the nodes. This structure's symmetries allow us to analytically explore their solutions following a first-order approximation and serve as a didactic example of the behaviors present in our model. In Fig. 13.1b, we show a sketch of the (λ, Θ^*) parameter space for the same structure. We present the analytical aspects of this solution in Sect. 13.3.2.

13.3 Individual-Based or Quenched Mean-Field Approach

As mentioned above, the exact formulation provides only a conceptual understanding of our model but fails to provide a quantitative characterization. Here we consider the individual-based or also called quenched mean-field approximation. This approach neglects dynamical correlations but takes into account the structural correlations of the interactions of the nodes. It is possible to solve the resulting equations numerically (without resorting to stochastic numerical simulations), obtaining a better understanding of the model's behaviour. We first derive the general dynamical equations of this approximation in Sect. 13.3.1; we then consider a toy example and solve numerically the corresponding equations in Sect. 13.3.2 in order to exemplify the variety of behaviors present in our model. Finally, in Sect. 13.3.3 we consider a hypergraph with power-law distributed cardinalities of hyperedges, which has a more complex and heterogeneous structure than the toy example of Sect. 13.3.2.

13.3.1 The General Formulation

Since Eq.(13.1) cannot be numerically solved, here we assume that the random variables are independent, allowing us to significantly reduce the complexity of our model. Denoting $y_i = \mathbb{E}(Y_i)$, this first-order approximation is given by

$$\frac{dy_i}{dt} = -\delta y_i + \lambda (1 - y_i) \sum_{e_j | i \in e_j} \sum_{k=\Theta_j}^{|e_j|} \lambda^*(|e_j|) \mathbb{P}_{e_j}(K = k), \tag{13.3}$$

where $\mathbb{P}_{e_j}(K = k)$ is the probability that the hyperedge e_j has k active nodes. In this formulation, we have used that the expectation of the indicator function in Eq. (13.1) follows a Poisson binomial distribution, which can be formally expressed as

$$\mathbb{E}\left(\mathbb{1}_{\{(T_j - Y_k) \geq \Theta_j\}}\right) \approx \sum_{m=\Theta_j}^{|e_j|} \mathbb{P}_{e_j}(K = m) \tag{13.4}$$

$$\mathbb{P}_{e_j}(K = m) = \sum_{A \in F_m} \prod_{i \in A} y_i \prod_{i' \in A^c} (1 - y_{i'}), \tag{13.5}$$

where F_m is the set of all subsets of m integers in $\{1, 2, ..., |e_j|\}$ and A^c is the complementary of A. The summation in Eq. (13.5) considers all possible microconfigurations in a given hyperedge, with A accounting for the active nodes and A^c for the inactive ones. Using it directly for numerical computations can introduce numerical stability problems for large hyperedges [34]. Fortunately, this issue can

be solved by considering the discrete Fourier transform, obtaining the following numerically stable solution [34]:

$$\mathbb{P}_{e_j}(K=k) = \frac{1}{n+1}\sum_{l=0}^{n} C^{-lk}\prod_{m=1}^{n}\left(1+(C^l-1)y_m\right), \tag{13.6}$$

where $C = \exp\left(\frac{2i\pi}{n+1}\right)$. This expression allows to compute the solution for arbitrarily large hyperedges. Although the whole argument is quite intricate, Eq. (13.6) is simple and robust enough, allowing the numerical evaluation of Eq. (13.3) for arbitrary hypergraphs and parameters.

13.3.2 The Hyperblob

For the sake of simplicity, let us focus here on a very particular and homogeneous structure: the hyperblob. The hyperblob is a hypergraph constructed as a homogeneous set of pairwise interactions with average degree $\langle k\rangle$, to which a single additional hyperedge containing all nodes is added. This structural simplicity allows us to solve the model analytically. Indeed, given the symmetry of the system, all y_i are equal ($y_i = \rho\ \forall i$) and their evolution can be expressed by the following single equation:

$$\frac{d\rho}{dt} = -\delta\rho + \lambda(1-\rho)\left[\langle k\rangle\rho + \lambda^* F\left(\Theta^*, \rho\right)\right]. \tag{13.7}$$

Here λ^* stands for $\lambda^*(|e_j|)$ and

$$F\left(\Theta^*, \rho\right) = 1 - \sum_{l=0}^{\Theta-1}\mathbb{P}_{N-1}(K=l) \approx \begin{cases} 1, & \text{if } \rho \geq \Theta^* \\ 0, & \text{otherwise} \end{cases}, \tag{13.8}$$

where the approximation on the right-most part of the equation assumes that the hypergraph is sufficiently large (for more on this approximation, we refer to the supplemental material of [27]).

The approximation in Eq. (13.8) suggests the possibility of having two solutions, one such that $F(\Theta^*, \rho) = 0$, and another one such that $F(\Theta^*, \rho) = 1$. Note that the first solution represents the case in which the largest hyperedge is inactive, while it is active in the second case. We remark that, as the only higher-order structure contains all the nodes, the global critical-mass is here $\rho_c = \Theta^*$. In other words, the activation of this hyperedge determines which solution the system is in, and the jumps between upper and lower solutions happen when they "cross" the value $\rho = \Theta^*$. From the approximation in Eqs. (13.8) and (13.7), we can analytically obtain the model's parameter space, obtaining the two solutions (see [27]).

$$\rho^{\text{Lower}} = \begin{cases} 1 - \frac{\delta}{\langle k\rangle\lambda}, & \text{if } \frac{\lambda}{\delta} \geq \frac{1}{\langle k\rangle} \\ 0, & \text{otherwise} \end{cases} \tag{13.9}$$

$$\rho^{\text{Upper}} = \frac{-\delta + \langle k\rangle\lambda - \lambda^*\lambda + \sqrt{4\langle k\rangle\lambda^*\lambda^2 + (\delta + (-\langle k\rangle + \lambda^*)\lambda)^2}}{(2\langle k\rangle\lambda)}. \tag{13.10}$$

As anticipated in Sect. 13.2, a second-order phase transition is obtained for ρ^{Lower} as the feasibility condition $\frac{\lambda}{\delta} \geq \frac{1}{\langle k\rangle}$ [35]. We remark that this lower solution is here simply the solution of a mean-field approach for an homogeneous structure with average degree $\langle k\rangle$. The next quantities of interest are the limits of the bistable region, which can be calculated as

$$\lambda_c^{\text{L}} = \frac{\delta}{\langle k\rangle - \Theta^*\langle k\rangle} \tag{13.11}$$

$$\lambda_c^{\text{U}} = -\frac{\delta\Theta^*}{\lambda^*\Theta^* - \lambda^* + (\Theta^*)^2\langle k\rangle - \Theta^*\langle k\rangle}. \tag{13.12}$$

Finally, the jump length is expressed as

$$Q_l(\lambda_c^X) = \left(\frac{\delta - \lambda(\lambda^* + \langle k\rangle) + \sqrt{(\delta + \lambda(\lambda^* - \langle k\rangle))^2 + 4\lambda^*\langle k\rangle\lambda^2}}{2\langle k\rangle\lambda} \right)_{\lambda=\lambda_c^X}, \tag{13.13}$$

where λ_c^X can be λ_c^{L} or λ_c^{U}. Although these equations are reasonably simple, the upper solution depends on a quadratic equation, where only one of the solutions is physical. The details for the complete derivation of these results can be found in the supplemental material of [27]. In the same reference, the interested reader can also find similar results when the considered lower-order structure is a star graph.

Figure 13.2 shows the phase diagram for a hyperblob with $N = 10^3$, $\langle k\rangle = 5$, $\delta = 1$ and $\lambda^*(|e_j|) = \log_2(|e_j|)$. This result complements Fig. 13.1 and exemplifies the five regions of the diagram in Fig. 13.1b. As predicted by our solutions, in (a), we

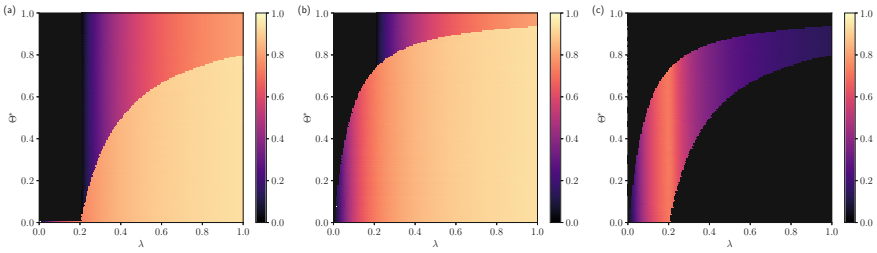


Fig. 13.2 Phase diagram for the Hyperblob with $\langle k\rangle = 5$, $\delta = 1$ and $\lambda^*(|e_j|) = \log_2(|e_j|)$. In **a**–**c** the colormaps are obtained changing λ and Θ^*. In **a** the solution of forward phase diagram, in **b** the solution of the backward one and in **c** the jump length (i.e., difference between **b** and **a**), emphasizing the bi-stability region

observe that the absorbing state plays a major role in the forward diagram, $\rho^{\bigstar}$, as there is a region of the parameter space that is not active, which is a consequence of the second-order phase transition present in ρ^{Lower}. This is also depicted as Regions I and III in Fig. 13.1b. Conversely, for the backward phase diagram, ρ^{*}, the set of parameters in which the system can reach the absorbing state is rather reduced, being restricted to Region I in Fig. 13.1b. We highlight that substituting the random regular network by a star would slightly change the parameter space as the second-order phase transition of ρ^{Lower} vanishes in the limit $N \to \infty$, thus implying that the Regions I and III in Fig. 13.1b vanish as well.

13.3.3 Example of a Hypergraph with a Power-Law Distribution of Cardinalities

In order to consider more complex and heterogeneous structures, we show in Fig. 13.3 an example of the solutions of the system of equations (13.3) for a hypergraph with $N = 10^4$, power-law distributed cardinalities, $P(|e_j|) \sim |e_j|^{-\gamma}$ with $\gamma = 2.25$, and $\min\{|e_j|\} = 2$. Here we use spreading rates $\lambda_j = \lambda \times \log_2(|e_j|)$ and we fix the deactivation parameter as $\delta = 1$. Figure 13.3a shows the phase diagram, while Fig. 13.3b displays the temporal behavior for $\lambda = 0.25$. The agreement is qualitatively good, with the upper solution being well captured. Both λ_c^{U} and λ_c^{L} seem to be underestimated by the ODE. We however remark that an accurate determination of the transition points from numerical simulations is not an easy task and requires more sophisticated algorithms (see Ref. [27] and its supplemental material for more details). The temporal behavior shown in Fig. 13.3b suggests that the upper solution is better

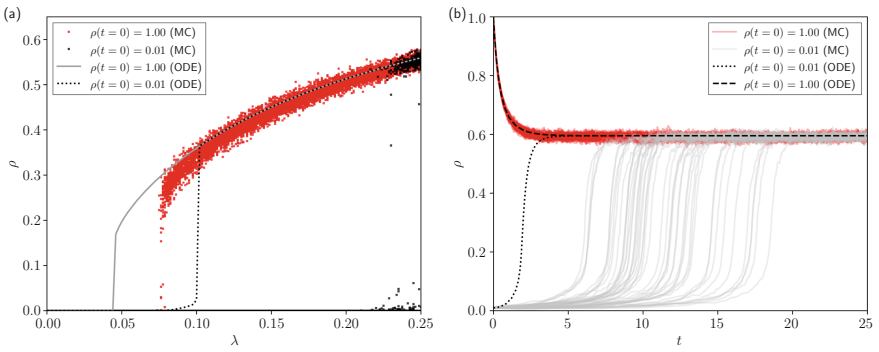


Fig. 13.3 Comparison of the numerical solution of the ODE system Eq. (13.3) with Monte Carlo (MC) simulations for a hypergraph with $N = 10^4$, power-law distributed cardinalities, $P(|e_j|) \sim |e_j|^{-\gamma}$ with $\gamma = 2.25$, and $\min\{|e_j|\} = 2$. The spreading rates are taken as $\lambda_j = \lambda \times \log_2(|e_j|)$ and the deactivation parameter is fixed as $\delta = 1$. We show in panel **a** shows the phase diagram for both the upper and lower solutions, and in panel **b** the temporal behavior for $\lambda = 0.25$

captured by our approach even at the dynamical level. For the lower solution, the steady-state value is very well captured, but the duration of the transient is longer in Monte Carlo simulations.

In summary, these examples highlight that the first-order approximation can provide a qualitative picture of the phenomenology at work, but that its limitations still need to be further evaluated. For instance, the accuracy of the estimated discontinuities might be related to specificities of the considered structure (e.g., low average degree or hyperedge intersections). The strong interest of this approximation lies in the relatively easy numerical implementation, as the system of equations (13.3) can be solved, e.g., by using Runge-Kutta methods. The qualitative picture obtained also suggests that the first-order approximation might be a good starting point for further analytical explorations of this type of models.

13.4 Annealed Mean-Field Approach

13.4.1 Homogeneous Mean-Field

We now focus on the simplest analytical framework, the mean-field (MF) approach, in which we assume that the population is fully mixed such that nodes are statistically equivalent, their states are independent, and all interactions can happen with identical probabilities. This is indeed the simplest scenario, which completely neglects the underlying structure. The mean-field form of Eq. (13.1) is given by

$$\frac{d\rho}{dt} = -\delta\rho + P^{\nu}(\rho) = -\delta\rho + \sum_{m=2}^{\nu} \lambda_m c(m)(1-\rho)\rho^{m-1}, \tag{13.14}$$

where $\nu = \max_j\{|e_j|\}$ is the maximum cardinality and $c(m)$ is the ratio between the average number of hyperedges with cardinality m and the average number of pairwise interactions incident on a node i, which characterizes the structure of the hypergraph. In the steady state this is a polynomial equation whose solutions are the fixed points of the process.

We now restrict our attention to a tractable case in which we can find a solution to the MF approximation. To do this, we consider a hypergraph formed solely by 1-hyperedges (standard pairwise links) and 2-hyperedges (3-body interactions). In this case, the maximum cardinality is $\nu = 3$, and Eq. (13.14) simplifies to:

$$\frac{d\rho}{dt} = -\delta\rho + \lambda_2 c(2)(1-\rho)\rho + \lambda_3 c(3)(1-\rho)\rho^2. \tag{13.15}$$

Notice that $c(2) = 1$ by definition, while $c(3)$ is given by the ratio between the average number of 2-hyperedges and the average number of 1-hyperedges adjacent to a node, so that $c(3) = \langle k_3\rangle/\langle k_2\rangle$. We can thus rewrite Eq. (13.15) as

$$\frac{d\rho}{dt} = -\delta\rho + \lambda_2(1-\rho)\rho + \frac{\langle k_3\rangle}{\langle k_2\rangle}\lambda_3(1-\rho)\rho^2. \tag{13.16}$$

After defining $\beta_2 = \lambda_2\langle k_2\rangle/\delta$ and $\beta_3 = \lambda_3\langle k_3\rangle/\delta$, we can rewrite Eq. (13.16) as:

$$\frac{d\rho}{dt} = -\rho + \beta_2(1-\rho)\rho + \beta_3(1-\rho)\rho^2. \tag{13.17}$$

From Eq. (13.17) it is evident that we can recover the standard MF equation for the SIS model by setting $\beta_3 = 0$. In this case, we get back the two standard stationary solutions which correspond to the absorbing state with no infected nodes $\rho_1^{*[\beta_3=0]} = 0$ and the endemic state $\rho_2^{*[\beta_3=0]} = 1 - 1/\beta_2$. When $\beta_2 < 1$, $\rho_1^{*[\beta_3=0]}$ is the only (stable) solution; it becomes unstable when $\beta_2 > 1$ and $\rho_2^{*[\beta_3=0]}$ appears (stable). The standard epidemic threshold $\beta_2 = 1$ represents the points at which the system undergoes a continuous transition between the two regimes.

Let us now consider the more interesting case in which there are contributions coming from the higher-order interactions (2-hyperedges), i.e., $\beta_3 > 0$. In this case, there are up to three stationary solutions of the steady state equation $d_t\rho = 0$ that fall within the range $\rho \in [0, 1]$. One is the trivial solution $\rho_1^* = 0$, which corresponds to the usual absorbing state where the epidemics dies out. The other two non-trivial solutions are given by

$$\rho_{2\pm}^* = \frac{\beta_3 - \beta_2 \pm \sqrt{(\beta_2-\beta_3)^2 - 4\beta_3(1-\beta_2)}}{2\beta_3}. \tag{13.18}$$

These correspond to the lower (ρ_{2-}^*) and upper (ρ_{2+}^*) branch that have been previously discussed. This simple mean-field description allows to go further and study the stability of the system (see [26] for details), confirming that:

- When $\beta_3 \leq 1$, if $\beta_2 < 1$ there is only one acceptable solution, that is the trivial absorbing state $\rho_1^* = 0$. If instead $\beta_2 > 1$, the non-trivial solution ρ_{2+}^* is positive and stable, while ρ_1^* becomes unstable. Thus, when moving–using the standard control parameter–from $\beta_2 < 1$ to $\beta_2 > 1$, it is possible to show that the system undergoes a continuous transition at the standard epidemic threshold $\beta_2 = 1$. While this is similar to what happens when $\beta_3 = 0$ (standard SIS model), if $0 < \beta_3 \leq 1$ there is a higher density of infected nodes in the endemic state.
- When $\beta_3 > 1$, algebraic manipulations of Eq. (13.18) show that if $\beta_2 < \beta_c = 2\sqrt{\beta_3} - \beta_3$, $\rho_{2\pm}^*$ are not in the acceptable domain and the only (stable) solution is, again, the trivial one $\rho_1^* = 0$. Contrarily, if $\beta_2 > \beta_c$, the system presents two different regimes. If $\beta_2 > 1$, we have a scenario similar to the one above, where ρ_1^* is unstable and the stable state is $\rho_{2+}^* > 0$. If instead $\beta_c < \beta_2 < 1$, both solutions $\rho_{2\pm}^*$ are positive ($0 < \rho_{2-}^* < \rho_{2+}^*$). More precisely, ρ_{2-}^* is an unstable solution that splits the phase space into two regions and determines–according to the initial conditions–in which one of the other two stable solutions ρ_1^* and ρ_{2+}^* the system will end up. We can thus confirm what we had previously observed, that is the presence of a discontinuous transition at β_c and of a bistable region in which the

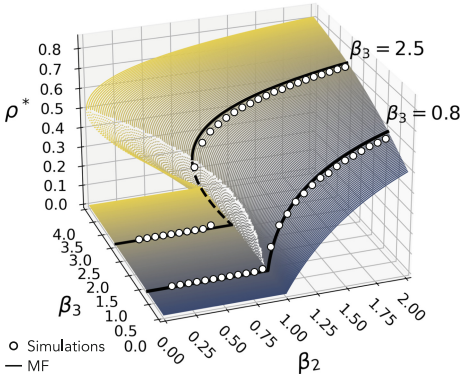


Fig. 13.4 Analytical solution in the mean-field approximation. Three-dimensional phase diagram, where the density of infected nodes in the stationary state ρ^* is plotted as a function of the 1-hyperedges rescaled infectivity $\beta_2 = \lambda_2 \langle k_2 \rangle / \delta$ and the 2-hyperedges rescaled infectivity $\beta_3 = \lambda_3 \langle k_3 \rangle / \delta$. When $\beta_3 = 0$ the dynamics obeys the one of the standard SIS model on networked systems with no higher-order interactions (links only). Two example curves are shown (at constant values of $\beta_3 = 0.8$ and $\beta_3 = 2.5$), where MF results (black lines) are compared to results of stochastic simulations on random simplicial complexes (white circles) [26]

system reaches ρ^*_{2+} only if the initial seed of infected nodes is above a critical mass ($\rho(t = 0) > \rho^*_{2-}$).

These results are also illustrated in Fig. 13.4, which gives a three-dimensional representation of the phase diagram associated to the system. These are the solutions of Eq. (13.17) just described, representing the density of infected nodes in the large-time limit as a function of the rescaled infectivity parameters β_2 and β_3. For visualization purposes only stable solutions are shown when $\beta_2 > 1$. We also plot two representative curves (black lines) that highlight the possible types of transitions. For $\beta_3 = 0.8$ the system still presents the standard continuous transition at $\beta_2 = 1$, while for greater values ($\beta_3 = 2.5$ shown in the figure) the transition becomes discontinuous. The presence of a bistable region is evident from the "folding" of the surface, in which a line parallel to the vertical axes can cross the surface in two distinct points.

We also compare the MF results with average stationary values extracted from multiple runs of stochastic simulations (white circles). Notice how, despite the oversimplified MF approach, the analytical predictions–on the position of the epidemic threshold and the nature of the transition–are in good agreement with the simulations when higher-order structures with homogeneous degree distributions are considered, such as the random simplicial complex structure used in this case ($N = 2000$, $\langle k_2 \rangle = 20$, $\langle k_3 \rangle = 6$). More details on the construction of this random structure are given in Ref. [26].

13.4.2 Heterogeneous Mean-Field

The MF approach can be improved by relaxing the assumption that all nodes are equivalent, and considering instead that nodes within the same hyperdegree class behave similarly [3]. Let us thus call $\mathbf{k}_i$ the vector containing all the generalized degrees associated to node i up to the maximum cardinality ν, such that $\mathbf{k}_i = [k_{2,i}, k_{3,i}, \ldots, k_{\nu,i}]$ [36]. By doing that, we are effectively removing the actual structure and describing it in terms of the probabilities of nodes to share a hyperedge. The equation for the heterogeneous mean-field (HMF) approach, as introduced in Ref. [31], reads:

$$\frac{d\rho_{\mathbf{k}}}{dt} = -\delta\rho_{\mathbf{k}} + (1-\rho_{\mathbf{k}}) \times \sum_{m=2}^{\nu} \frac{\lambda_m}{(m-1)!} \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_m} \prod_{\ell}^{m-1} P(\mathbf{k}_\ell) f_m(\mathbf{k}, \mathbf{k}_1, \ldots, \mathbf{k}_{m-1}) G(\rho_{\mathbf{k}_1}, \ldots, \rho_{\mathbf{k}_{m-1}}) \quad (13.19)$$

where $\rho_{\mathbf{k}}$ denotes the density of active nodes having hyperdegree $\mathbf{k}$, and $P(\mathbf{k})$ the number of nodes with hyperdegree $\mathbf{k}$. In the second term of the r.h.s. of Eq. (13.19), the first summation runs over all hyperedges of size m that can infect a node having hyperdegree $\mathbf{k}$. This means that for each hyperedge there are $m-1$ other nodes that could be infected, and their combinations are counted by the second summation. The ability to actually transmit the infection depends on the fraction of hyperedges (among all their possible combinations) that include the given node, given by $f_m(\mathbf{k}, \mathbf{k}_1, \ldots, \mathbf{k}_{m-1})$, and the probability $G(\rho_{\mathbf{k}_1}, \ldots, \rho_{\mathbf{k}_{m-1}})$ that the given hyperedge can transmit the infection. If we assume that a hyperedge can infect a node only if all the remaining nodes composing it are infected, this reads $G(\rho_{\mathbf{k}_1}, \ldots, \rho_{\mathbf{k}_{m-1}}) = \prod_{\ell=1}^{m-1} \rho_{\mathbf{k}_\ell}$. In addition, if we consider as before a hypergraph containing 1- and 2-hyperedges only ($\nu = 3$), and we assume that the connection probabilities are only determined by the links, i.e., $f_m(\mathbf{k}, \mathbf{k}_1, \ldots, \mathbf{k}_{m-1}) = f_m(k, k_1, \ldots, k_{m-1})$, Eq. (13.19) simplifies to

$$\begin{aligned} \frac{d\rho_k}{dt} &= -\delta\rho_k + (1-\rho_k)\lambda_2 \textstyle\sum_{k_1} P(k_1) f_2(k, k_1) \rho_{k_1} \\ &\quad + (1-\rho_k)\tfrac{\lambda_3}{2} \textstyle\sum_{k_1,k_2} P(k_1) P(k_2) f_3(k, k_1, k_2) \rho_{k_1} \rho_{k_2} \end{aligned} \quad (13.20)$$

where it is now possible to explicitly distinguish the contributions coming from links and "triangles", respectively the second and third term of the r.h.s. of Eq. (13.20).

The process described by Eq. (13.20) can be analyzed using linear stability analysis. Although an analytical solution for the fixed points of Eq. (13.20) is not possible, we can restrict our analysis to the inactive state, i.e., $\rho_k = 0$ for all k. And as it turns out [31], the inactive state becomes unstable for

$$\frac{\lambda_2}{\delta} > \frac{\langle k_2 \rangle}{\langle k_2^2 \rangle}, \quad (13.21)$$

where k_2 is the pairwise degree. Interestingly, the take-home message from this analysis is that only pairwise interactions are responsible for the inactive state's stability. In this case, the parameter λ_3 is responsible for the presence or absence of bi-stable solutions. As λ_3 increases, the dynamics allow for a discontinuity, bi-stability, and hysteretic behavior. These results are in agreement with the approach of Ref. [30], where the authors arrived at a similar conclusion using a quenched formalism. Furthermore, in Ref. [31], the authors used the HMF formalism to investigate the effect of heterogeneity in hypergraph contagion models. They showed that in the extreme case where a hyperedge can transmit infection if there is at least one infectious node (as opposed to $m - 1$ discussed here), the bi-stability disappears, and the critical point depends on both λ_2 and λ_3. Interestingly, they also showed that the explosive transition could disappear for specific heterogeneous structures, e.g., when power-law distributions of pairwise interactions are used as a starting structure to construct the hypergraph. This can also happen when 2-hyperedges are placed at random, as opposed to degree-correlated structures where higher-order interactions are more likely to involve nodes that have a high pairwise degree (more details on the effects of heterogeneity and the HMF formalism can be found in Ref. [31]).

13.5 Simulations on Real-World Structures

While the analytical approximations developed in the above sections correspond to simplified structures of interactions between nodes, real-world interactions are expected to involve complex and intricate structural correlations at various scales that are not easily reproduced by models. Therefore, we now briefly investigate the dynamics of the higher-order social contagion model on empirical higher-order structures. We focus in particular on the simplicial contagion model in its original formulation, where the social structure is modeled as a simplicial complex and each simplex of size k can transmit the infection (at its order-dependent rate) to a susceptible node incident on it only if the remaining $k - 1$ nodes are infectious [26].

To this aim, we construct empirical simplicial complexes from temporally resolved interactions data. In fact, data already encoded into graphs are intrinsically ill-suited for the task–since they have already been "projected" into pairwise relations (the links of the graph). Although recovering the hidden higher-order interactions from pairwise networks surely represent a challenging task, recent efforts have addressed this problem with a Bayesian approach [37]. Here, leveraging high-resolution proximity contact data provided by the SocioPatterns collaboration,[1] we consider simplicial complexes representing interactions in four different social contexts: a workplace (InVS15 [38]), a conference (SFHH [39]), a hospital (LH10 [40]) and a high school (Thiers13 [41]). More precisely, as described in Ref. [26], we first aggregate temporally the recorded (temporal) interactions into windows of 5 min. Maximal cliques within each temporal window are then "promoted" to simplices (with associated

[1] http://www.sociopatterns.org/datasets/.

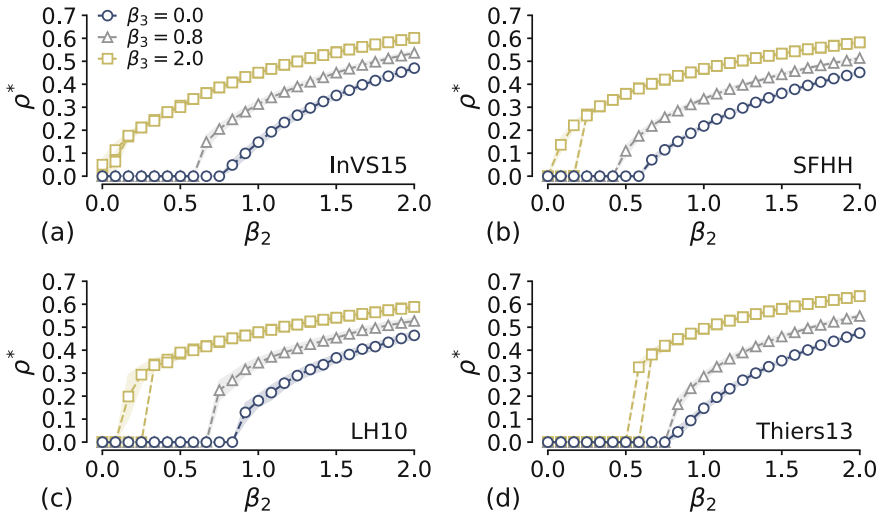


Fig. 13.5 Simulations on real-world social structures. Density of infected nodes ρ^* in the stationary state as a function of the 1-hyperedges rescaled infectivity $\beta_2 = \lambda_2\langle k_2\rangle/\delta$ and for three different values of the 2-hyperedges rescaled infectivity $\beta_3 = \lambda_3\langle k_3\rangle/\delta$. When $\beta_3 = 0$ the dynamics obeys the one of the standard SIS model on networked systems with no higher-order interactions. Points and shaded areas correspond to median values and standard deviations as extracted from stochastic simulations on top of four different empirical simplicial complexes constructed from the SocioPatterns data sets: a workplace (**a**), a conference (**b**), a hospital (**c**) and a high school (**d**). See [26] for details

frequency of appearance) and the final simplicial complex is formed by retaining the 20% most frequent simplices (up to 2-simplices). More detailed information can be found in Ref. [26]. The results of the stochastic simulations run on each structure are displayed in Fig. 13.5, where the density of infected nodes in the stationary state is plotted as a function of β_2 for different values of β_3. Despite the very different nature of these datasets and their different generalized degree distributions, we encounter a similar phenomenology to the one described in the previous sections. Namely, when contributions from the higher-order interactions are stronger (higher values of β_3) we observe a lower (almost vanishing in some cases) epidemic threshold. Moreover, the bi-stability is present for the highest value of β_3, confirming the overall phenomenology obtained by the analytical approaches.

13.6 Conclusions

In this chapter we have reviewed some recent conceptual advances in the modeling of social contagion processes, based on the idea to consider group interactions as such, and not simply as a superposition of dyadic ones. To this aim, the substrate

of the contagion models has to be changed, moving from a network picture to representations by hypergraphs or simplicial complexes, which are able to represent interactions involving an arbitrary number of individuals [42]. Notably, the models of interactions themselves have to be redefined, as contagion models are traditionally defined with dyadic interactions in mind. While we have not covered all the relevant literature,[2] we have highlighted the main approaches and results, and in particular the rich phenomenology emerging from hyperedge interactions, with co-existence of continuous and discontinuous phase transitions, bi-stability regions and critical mass phenomena.

Moreover, while the discovery of this rich phenomenology has already prompted a wealth of studies and brought both analytical and numerical insights, a number of interesting points remain open.

First, few analytical or mathematical results are available regarding the nature of the phase transitions: these results have been obtained under specific approximations or for specific structures. It would be of clear interest to have more general results on the conditions (either on the structure or on the dynamical model's rules) for the emergence of discontinuous transitions.

Another important point regards the availability of empirical data to feed models defined on hypergraphs. Indeed, given the popularity and convenience of the network representation, relational datasets are traditionally represented as sets of dyadic interactions and often fail to include higher-order interactions (with some exceptions, e.g. for scientific collaboration data that is easily represented as group interactions [46]). While temporally resolved data can help understand whether cliques in an aggregated network actually correspond to group meetings or not, as discussed in Sect. 13.5, using dyadic data to reconstruct the actual higher-order interactions is in general far from trivial [37] and it seems crucial to develop new methods to this aim.

Empirical validation of the rich phenomenology uncovered in the models remains also very challenging. On the one hand, it has been shown that complex contagion processes might become indistinguishable from simple contagion at the population level when multiple contagion processes interact [47]. For simple contagions taking place along networks, it is possible to infer the structure on which the process unfolds and the process' parameters [48], but the generalization to higher order processes remains an open challenge.

Validation could also come from specifically designed experiments in which the structure of the groups in which individuals interact is controlled. In the case of networks, controlled experiments have indeed helped discuss the role of the interaction network structure on the emergence of conventions or on the outcome of game theoretical models [49, 50]. For higher order structures, such experiments would also need to be carefully crafted and performed, a difficult yet promising challenge ahead.

[2] For instance, the authors of [31] study the case in which $\lambda_3 < 0$, i.e., an individual is less likely to adopt a trend if this trend is popular in the group, and call this ingredient the "hipster effect"; this effect also can lead to a region of bi-stability in the phase diagram [31]. Note that heterogeneous recovery rates [43, 44] and "complex recovery" rates depending on the state of the surrounding individuals have also been considered in the literature [45].

References

1. R.M. Anderson, R.M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1992)
2. M.J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2011)
3. R. Pastor-Satorras, A. Vespignani, Phys. Rev. Lett. **86**(14), 3200 (2001)
4. A. Barrat, M. Barthelemy, A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008)
5. R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, Rev. Mod. Phys. **87**(3), 925 (2015)
6. I.Z. Kiss, J.C. Miller, P.L. Simon, et al., (Springer, Cham, 2017)
7. R. Albert, A.L. Barabási, Rev. Mod. Phys. **74**(1), 47 (2002)
8. M.E.J. Newman, SIAM Rev. **45**, 167 (2003)
9. D. Centola, M. Macy, Am. J. Sociol. **113**(3), 702 (2007)
10. D. Centola, Science **329**(5996), 1194 (2010)
11. J. Ugander, L. Backstrom, C. Marlow, J. Kleinberg, Pro. Natl. Acad Sci. U.S.A. 201116502 (2012)
12. L. Weng, A. Flammini, A. Vespignani, F. Menczer, Sci. Rep. **2**, 335 (2012)
13. M. Karsai, G. Iniguez, K. Kaski, J. Kertész, J.R. Soc, Interface **11**(101), 20140694 (2014)
14. B. Mønsted, P. Sapieżyński, E. Ferrara, S. Lehmann, PLoS One **12**(9), e0184148 (2017)
15. D. Guilbeault, J. Becker, D. Centola, in *Complex Spreading Phenomena in Social Systems* (Springer, 2018), pp. 3–25
16. D.J. Watts, Proc. Natl. Acad. Sci. U.S.A. **99**(9), 5766 (2002)
17. S. Melnik, J.A. Ward, J.P. Gleeson, M.A. Porter, Chaos **23**(1), 013124 (2013)
18. Z. Ruan, G. Iniguez, M. Karsai, J. Kertész, Phys. Rev. Lett. **115**(21), 218702 (2015)
19. A. Czaplicka, R. Toral, M. San Miguel, Phys. Rev. E **94**(6), 062301 (2016)
20. E. Cozzo, R.A. Banos, S. Meloni, Y. Moreno, Phys. Rev. E **88**(5), 050801 (2013)
21. N.O. Hodas, K. Lerman, Sci. Rep. **4**, 4343 (2014)
22. M. Herrera, G. Armelini, E. Salvaj, PLoS One **10**(10), e0140891 (2015)
23. D.J. O'Sullivan, G.J. O'Keeffe, P.G. Fennell, J.P. Gleeson, Front. Phys. **3**, 71 (2015)
24. P. Tuzón, J. Fernández-Gracia, V.M. Eguíluz, Front. Phys. **6**, 21 (2018)
25. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, Phys. Rep. **874**, 1 (2020)
26. I. Iacopini, G. Petri, A. Barrat, V. Latora, Nat. Commun. **10**(1), 2485 (2019)
27. G.F. de Arruda, G. Petri, Y. Moreno, Phys. Rev. Res. **2**(2), 023032 (2020)
28. B. Jhun, M. Jo, B. Kahng, J. Stat. Mech. Theory Exp. **2019**(12), 123207 (2019)
29. J.T. Matamalas, S. Gómez, A. Arenas, Phys. Rev. Res. **2**(1), 012049 (2020)
30. G.F. de Arruda, M. Tizzani, Y. Moreno, Commun. Phys. **4**, 24 (2021)
31. N.W. Landry, J.G. Restrepo, Chaos **30**(10), 103117 (2020)
32. D. Centola, J. Becker, D. Brackbill, A. Baronchelli, Science **360**(6393), 1116 (2018)
33. A. Hatcherr, (2002)
34. M. Fernández, S. Williams, IEEE Trans. Aerosp. Electron. Syst. **46**(2), 803 (2010)
35. G.F. de Arruda, F.A. Rodrigues, Y. Moreno, Phys. Rep. **756**, 1 (2018)
36. O.T. Courtney, G. Bianconi, Phys. Rev. E **93**(6), 062311 (2016)
37. J.G. Young, G. Petri, T.P. Peixoto, Commun. Phys. **4**(1), 1 (2021)
38. M. Génois, C.L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, A. Barrat, Netw. Sci. **3**(3), 326 (2015)
39. L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.F. Pinton, W. Van den Broeck, J. Theor. Biol. **271**(1), 166 (2011)
40. P. Vanhems, A. Barrat, C. Cattuto, J.F. Pinton, N. Khanafer, C. Régis, B.a. Kim, B. Comte, N. Voirin, PLoS One **8**(9), e73970 (2013)
41. R. Mastrandrea, J. Fournet, A. Barrat, PLoS One **10**(9), e0136497 (2015)

42. S. Battiston, J.B. Glattfelder, D. Garlaschelli, F. Lillo, G. Caldarelli, in *Network Science* (Springer, 2010), pp. 131–163
43. A. Darbon, D. Colombi, E. Valdano, L. Savini, A. Giovannini, V. Colizza, R. Soc. Open Sci. **6**, 181404 (2019)
44. G.F. de Arruda, G. Petri, F.A. Rodrigues, Y. Moreno, Phys. Rev. Res. **2**(1), 013046 (2020)
45. I. Iacopini, B. Schäfer, E. Arcaute, C. Beck, V. Latora, Chaos **30**(1), 013153 (2020)
46. G. Petri, A. Barrat, Phys. Rev. Lett. **121**(22), 228301 (2018)
47. L. Hébert-Dufresne, S.V. Scarpino, J.G. Young, Nat. Phys. **16**, 426 (2020)
48. P. Sah, M. Otterstatter, S.T. Leu, S. Leviyang, S. Bansal, BioRxiv p. 169573 (2018)
49. D. Centola, A. Baronchelli, Proc. Natl. Acad. Sci. U.S.A. **112**(7), 1989 (2015). https://doi.org/10.1073/pnas.1418838112
50. C. Gracia-Lázaro, A. Ferrer, G. Ruiz, A. Tarancón, J.A. Cuesta, A. Sánchez, Y. Moreno, Proc. Natl. Acad. Sci. U.S.A. **109**(32), 12922 (2012)

Chapter 14
Consensus Dynamics and Opinion Formation on Hypergraphs

Leonie Neuhäuser, Renaud Lambiotte, and Michael T. Schaub

Abstract In this chapter, we derive and analyse models for consensus dynamics on hypergraphs. As we discuss, unless there are *nonlinear* node interaction functions, it is always possible to rewrite the system in terms of a new network of effective pairwise node interactions, regardless of the initially underlying multi-way interaction structure. We thus focus on dynamics based on a certain class of non-linear interaction functions, which can model different sociological phenomena such as peer pressure and stubbornness. Unlike for linear consensus dynamics on networks, we show how our nonlinear model dynamics can cause shifts away from the average system state. We examine how these shifts are influenced by the distribution of the initial states, the underlying hypergraph structure and different forms of non-linear scaling of the node interaction function.

14.1 Background: Modelling Group Interactions

Group interactions are present in various areas in nature [34], society [11] and technology [26]. Examples range from collaborations of authors [29] to neuronal activity [10, 31]. In sociology, for instance, it is well known that the dynamics in a social clique is determined not just by the pairwise relationships of its members, but often by complex mechanisms of peer influence and reinforcement [30]. This is illustrated by the example of joint parental discipline shown in Fig. 14.1. In Fig. 14.1a, a strong link between the parents reinforces the influences dynamics resulting in a stronger

L. Neuhäuser (✉) · M. T. Schaub
RWTH Aachen University, Aachen, Germany
e-mail: neuhaeuser@cs.rwth-aachen.de

M. T. Schaub
e-mail: schaub@cs.rwth-aachen.de

R. Lambiotte
University of Oxford, Oxford, UK
e-mail: lambiotte@maths.ox.ac.uk

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_14

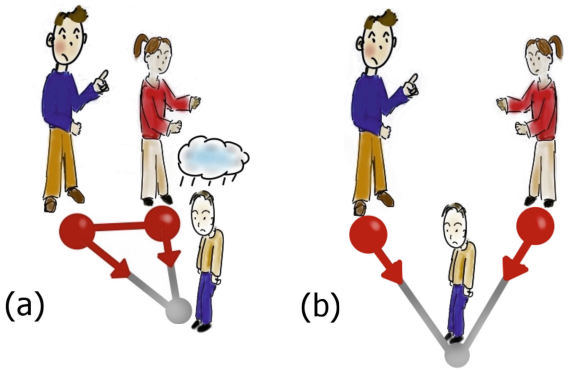


Fig. 14.1 Higher-order group interactions in social context. Higher-order group interactions **a** can result in greater influence on target nodes than **b** pairwise interactions

effect on the child. This is not captured by the independent pairwise influences in Fig. 14.1b.

In the context of modelling such multi-way dynamics, it is thus important to distinguish between the pairwise interactions between individuals and higher-order interactions, which cannot be decomposed further into pairwise interactions. Specifically, if the influence on an agent can be fully explained by its pairwise relationships to other group members, then the system can be abstracted by an (effective, derived) pairwise network representation. In contrast, higher-order interactions account for the effect of the group as a whole, and thus different frameworks than graphs are required to encode the interactions between agents.

Especially complex social processes such as the adoption of norms or opinion spreading might not be explainable by a simple exchange of the states of neighbouring nodes, as simple models for, e.g., epidemic spreading would suggest. For instance, experiments in social psychology such as the conformity experiment [2] indicate that multiple exposures might be necessary for an agent to adopt a certain state. This type of behaviour is also at the core of threshold models on networks, which model adoption processes (e.g. opinion spreading) in social systems. A threshold model posits that each node in a network has an associated binary state, and the (binary) state of agents only switches if a certain fraction (or a certain number) of their neighbours agrees on the same opinion [14, 39]. More generally, such a nonlinear dependence of a node on all its neighbors may be captured via a generalized linear model, in which each node is influenced according to a nonlinear map applied to a linear transformation of the states of its neighbors, i.e., we first linearly accumulate pairwise influences and then transform the result in a nonlinear way [23]. For instance, in the case of the threshold model this nonlinear function is a threshold or Heaviside function, which is applied to the (linear) mean of the neighbors' opinions. While generalised linear models can capture certain aspects of a group dynamics, these models may nonetheless provide an over-simplified view of the system. Consider again the previous example of parental discipline. If we were to model this situation with a threshold model, then any pair of adults could influence the child in the same way, irrespective of the relation between the adults. However, it is not difficult to

imagine that if there is strong relation between the parents, the influence on the child may be stronger. We are therefore interested in models that can capture *multi-way* relations that cannot be encoded with a network of pairwise influences.

There are different ways to encode such multi-way group relations in a network, including set systems, general hypergraphs [6] and simplicial complexes [16, 21, 22, 28, 30, 35]. Note that we are using the term "network" in the sense of system and not as a synonym for a graph. In this chapter, we explore dynamical models based on hypergraphs, where each node has an associated state and the evolution of those states depends on the values of all the nodes inside each hyperedge.

14.2 From Pairwise to Multi-way Interactions

Pairwise dynamical systems can describe a wide range of nonlinear dynamics on graphs [38]. We can formalize these dynamical systems as follows. Let $\mathscr{G}$ be a (directed) graph consisting of a set $V(\mathscr{G}) = \{1, \ldots, N\}$ of N nodes connected by a set of edges $E(\mathscr{G}) = \{(i, j) : i, j \in V(\mathscr{G})\}$, described by ordered tuples of nodes. The structure of the network can be represented by the adjacency matrix $A \in \mathbb{R}^{N\times N}$ with entries

$$A_{ij} = \begin{cases} 1 & (i, j) \in E(\mathscr{G}) \\ 0 & \text{otherwise.} \end{cases} \tag{14.1}$$

For simplicity we will consider only undirected networks in this chapter, in which case the adjacency matrix A is symmetric.

We endow each node $i \in V(\mathscr{G})$ with a dynamical variable, $x_i \in \mathbb{R}$. For a pairwise dynamical system as we consider here, the evolution of these variables is mediated by the underlying graph $\mathscr{G}$, whose edges constrain which nodes can interact with each other, and a set of node interaction functions

$$\mathcal{F} = \left\{ f_{ij} \;\middle|\; f_{ij} : \mathbb{R}^2 \to \mathbb{R}, (i, j) \in E(\mathscr{G}) \right\}, \tag{14.2}$$

that quantify how the states of neighbouring nodes affect each other. Given the information about $\mathscr{G}$ and $\mathcal{F}$, the time-evolution of the pairwise system is then defined as

$$\dot{x}_i = \sum_{j=1}^{N} A_{ij} f_{ij}\left(x_i, x_j\right), \tag{14.3}$$

where each node is affected by the sum of possibly nonlinear interactions with its neighbours. Particular examples of (14.3) include the Kuramoto model [1], continuous-time random walks and linear consensus on networks [15].[1]

Multi-way Interactions on Hypergraphs

To encode possible multi-way interactions in a dynamical system we use a hypergraph $\mathscr{H}$. A hypergraph $\mathscr{H}$ consist of a set $V(\mathscr{H}) = \{1, 2, \ldots, N\}$ of N nodes, and a set $E(\mathscr{H}) = \{E_1, E_2, \ldots, E_M\}$ of M hyperedges. Each hyperedge E_α is a subset of the nodes, i.e. $E_\alpha \subseteq V(\mathscr{H})$ for all $\alpha = 1, 2, \ldots, M$, where each hyperedge may have a different cardinality $n_\alpha = |E_\alpha|$. A graph is thus simply a hypergraph constrained to contain only 2-edges. We use $E^n(\mathscr{H})$ to denote the set of all hyperedges of cardinality n, which are henceforth referred to as n-edges. In the following we will concentrate on hypegraphs without self-loops, i.e., $n_\alpha \geq 2$ for all hyperedges E_α.

We can describe the structure of a hypergraph $\mathscr{H}$ by a set of adjacency tensors $\{A^{(n)}, n = 2, 3 \ldots, N\}$, where each tensor $A^{(n)}$ represents the connections made by n-edges.

$$A^{(n)}_{ij\ldots} = \begin{cases} 1 & \{i, j \ldots\} \in E^n(\mathscr{H}) \\ 0 & \text{otherwise} \end{cases} \tag{14.4}$$

Thus the adjacency tensor $A^{(n)}$ is symmetric with respect to any permutation of its indices, and as we do not allow for self-loops its entries $A^{(n)}_{ij\ldots}$ can only be nonzero if all indices are distinct.

Generalising Eq. (14.3), a multi-way dynamical system is now defined by a hypergraph $\mathcal{H}$ encoding the structure of the interactions between the nodes, and by a set of node interaction functions

$$F = \left\{ f^{E_\alpha} \;\middle|\; f^{E_\alpha} : \mathbb{R}^{N\times N\times\ldots\times N} \to \mathbb{R},\ E_\alpha \in E(\mathscr{H}) \right\}. \tag{14.5}$$

Analogous to the definition of $E^n(\mathscr{H})$, we can also partition the function set F into subsets F^n based on the cardinality of the hyperedges, which then can be matched to the corresponding adjacency tensors $A^{(n)}$:

$$F^n = \left\{ f^{\{i,j\ldots\}} \;\middle|\; \{i, j \ldots\} \in E^n(\mathscr{H}) \right\}. \tag{14.6}$$

The set of all node interaction functions is then given by the union $F = \bigcup_{n=2}^{N} F^n$. Finally, we assume that the time-evolution of system is governed via a linear combination of the effects of each hyperedge in which a node is involved (possibly in a

[1] Note that the case of generalized linear models mentioned in the introduction does *not* fit the above description. Instead of consisting of a linear sum of possibly nonlinear functions, a generalized linear model would be of the form $\dot{x}_i = f(\sum_j A_{ij} x_j)$, a nonlinear function applied to a (linear) sum.

nonlinear way):

$$\dot{x}_i = \sum_{n=2}^{N} \sum_{j,k,\ldots} A^{(n)}_{ijk\ldots} f^{\{i,j\ldots\}}(x_i, x_j, \ldots) \tag{14.7}$$

14.3 Higher-Order Effects and Nonlinearity

In this section our main goal is to examine under what conditions a multi-way dynamical system as described by (14.5) cannot be rewritten as an appropriately defined pairwise dynamical system (14.2), that is when a multi-body formalism is truly necessary to capture the complexity of a system. We tackle this problem for classes of node interaction functions satisfying desirable symmetries, described as follows.

14.3.1 Symmetries and Quasilinearity

As is often the case for models of non-linear consensus or synchronisation on standard networks, we would like our model to be invariant to translation and rotation. This is a reasonable assumption for physical and sociological interaction processes ensuring independence on the global reference frame. A function is rotational and translational invariant if it is invariant under application of elements from the special Euclidean group $\mathrm{SE}(N)$, which is defined as the symmetry group of all translations and rotations around the origin. As we restrict the scope to scalar values x_i on nodes, and do not consider vectors here, the rotational invariance simply means invariance under a change of signs of the values. In the case of two-body dynamical systems, it is known that a necessary and sufficient condition for these symmetries to be satisfied is the quasi-linearity of the interaction [38], that is

$$\dot{x}_i = \sum_j A_{ij}\, k_{ij}(|x_j - x_i|)\,(x_j - x_i), \tag{14.8}$$

where k_{ij} is an arbitrary function from $\mathbb{R}$ to $\mathbb{R}$. This form implies that the node interaction function is, for each edge, an odd function of $(x_j - x_i)$, which is a popular choice in the study of non-linear consensus [37]. Within the language of non-linear consensus, this model belongs to the family of relative non-linear flow. While we cannot generally transfer these results to multi-body dynamical systems, they will provide us a guide on how to define a 'minimal non-linear' model in that case.

14.3.2 Linear Dynamics and Motif Matrices

We first investigate the relations between pairwise and multi-way dynamical systems in the case of linear node interaction functions. Linear dynamics are crucial for modeling a range of different phenomena and serve as a first approximation for many nonlinear systems. With pairwise dynamical systems, the node interaction function is given by $f_{ij}(x_i, x_j) = c(x_j - x_i)$ where $c \in \mathbb{R}$ is a scaling constant, and the resulting dynamics reads

$$\dot{x}_i = \sum_j A_{ij} c(x_j - x_i) = -c \sum_j L_{ij} x_j, \tag{14.9}$$

where $L_{ij} = D_{ij} - A_{ij}$ is the network Laplacian. Here the degree matrix $D_{ij} = \delta_{ij} d_i$ is a diagonal matrix of the degrees $d_i = \sum_j A_{ij}$. Equation (14.9) naturally arises when modelling continuous-time random walks on networks [15], but also appears in the context of opinion-formation and decentralized consensus, as in the continuous-time DeGroot model [27]. For undirected, connected networks, the dynamics asymptotically converges to an average consensus $\lim_{t\to\infty} x(t) = \mathbf{1}\alpha$ for some $\alpha \in \mathbb{R}$, with a convergence rate determined by the second dominant eigenvalue of the Laplacian.

With multi-way interaction systems, the linear node interaction function is given by

$$\begin{aligned} \dot{x}_i &= \textstyle\sum_{n=2}^{N} \frac{1}{(n-1)!} \sum_{jk\ldots} A^{(n)}_{ijk\ldots} c \left(x_j - x_i + x_k - x_i + \ldots\right) \\ &= \textstyle\sum_{n=2}^{N} \frac{c}{(n-2)!} \sum_{jk\ldots} A^{(n)}_{ijk\ldots} \left(x_j - x_i\right) \\ &= -c \textstyle\sum_{n=2}^{N} \sum_j L^{(n)}_{ij} x_j. \end{aligned} \tag{14.10}$$

Here, we generalise (14.3.2), scaled according to the symmetry of the linear multi-way interactions on the n-edges. We have defined the *motif Laplacian* for fully connected n-cliques as:

$$L^{(n)} = D^{(n)} - W^{(n)} \tag{14.11}$$

which is simply the standard Laplacian for a graph with adjacency matrix

$$W^{(n)}_{ij} = \frac{1}{(n-2)!} \sum_{kl\ldots} A^{(n)}_{ijk\ldots}. \tag{14.12}$$

This rescaled network is thus obtained by weighting each edge by the number of n-edges to which it belongs.[2] Equation (14.3.2) can now be written as:

$$\dot{x}_i = -c \sum_j \mathscr{L}_{ij} x_j \tag{14.13}$$

[2] This is a standard procedure to project a hypergraph on a weighted network, but other possibilities exist, for instance based on the dynamics of biased random walkers [7].

where the Laplacians for all hyperedge cardinalities are summed up to one Laplacian $\mathscr{L}_{ij} = \sum_{n=2}^{N} L_{ij}^{(n)}$.

In other words, a multi-way dynamical system can be rewritten as a pairwise dynamical system in the case of linear dynamics, after a proper rescaling of the adjacency matrix. This observation reveals that a genuine multi-way dynamics on hypergraphs requires a non-linear node interaction function. Hence, linear multi-way interactions are not sufficient to produce dynamics that cannot be reduced to pairwise dynamical systems. It is therefore essential to not simply study the hypergraph structure in isolation, but also consider the dynamical process evolving on top of this hypergraph.

Remark Before exploring the dynamics of multi-way systems further, let us remark upon a connection between the operation (14.12) and the 'motif matrix' used to uncover communities in higher-order networks [4]. A *motif* on k nodes is defined by a tuple (B, P), where B is a $k \times k$ binary matrix that encodes the edge pattern between the k nodes, and $P \subset \{1, 2, \ldots, k\}$ is a set of anchor nodes. The study of such motifs is an important objective in network science [19]. In Ref. [4], the authors define a generalisation of conductance and cut, based on motifs rather than edges. In this context they define the *motif adjacency matrix*

$$(W_M)_{ij} = \text{number of instances of motifs in } M \text{ containing } i \text{ and } j,$$

from which a motif Laplacian could be defined. Equation (14.12) provides a dynamical interpretation of this quantity for n-edges. An interesting path for future research could be to employ such motif Laplacians to extend random-walk based community detection techniques such as the Map equation [32] and Markov stability [9, 13, 36] to higher-order networks.

14.4 Non-linear Consensus Dynamics on Hypergraphs

In this section we explore how a non-linear node interaction function f can lead to higher-order effects that do not exist in a pairwise setting. In particular, we investigate non-linear consensus dynamics on hypergraphs.

14.4.1 Non-linear Consensus Dynamics in Three-Way Interactions (3CM)

As previously mentioned, group effects that cannot be reduced to pairwise interactions appear in various contexts. In the area of sociology, reinforcing group effects such as peer pressure are a long-standing area of study, for instance in social psychology [2]. It is thus important to develop models that capture these multi-way

mechanisms to better understand phenomena such as hate communities [12], echo chambers and polarisation [3] in society.

Motivated by these observations, and inspired by Eq. (14.3), we first consider the case when all hyperedges have the same size 3, and we introduce a three-way consensus model (3CM)[24] with a non-linear node interaction function of a specific form, which is akin to a consensus dynamics with group reinforcement:

$$f_i^{\{j,k\}}(x_i, x_j, x_k) = s\left(|x_j - x_k|\right)\left((x_j - x_i) + (x_k - x_i)\right). \tag{14.14}$$

where we assume the function on each 3-edge is the same, for the sake of simplicity.[3] This expression models, for each 3-edge $\{i, j, k\}$, the multi-way influence of nodes j and k on node i by the standard linear term $\left((x_j - x_i) + (x_k - x_i)\right)$ modulated by a scaling function $s\left(|x_j - x_k|\right)$ of their state differences. If the scaling function $s(x)$ is monotonically decreasing, the influence of j and k on i is increased if j and k have similar states. In such a situation we will say that j and k *reinforce* each other's influence. In contrast, the joint influence is diminished if j and k have very different states which will be called *inhibiting*. This property is reminiscent of non-linear voter models in the case of discrete dynamics [14, 18, 20], where the voters change opinion with a probability p that depends non-linearly on the fraction of disagreeing neighbours. Note that other choices of non-linear node interaction functions f, akin to Watts threshold model [39], have also been considered recently for information spreading [8].

For the dynamics (14.14) the resulting dynamics for each node i are then given by

$$\dot{x}_i = \sum_{j,k=1}^{N} \frac{1}{2} A_{ijk}^{(3)}\, s\left(|x_j - x_k|\right)\left((x_j - x_i) + (x_k - x_i)\right). \tag{14.15}$$

In Eq. (14.14), the node interaction function is non-linear for non-constant scaling functions $s(x)$ and captures multi-way effects, as the interactions on a triangle can no longer be split into pairwise node interaction functions. If the scaling function $s(x)$ is constant, we recover the linear dynamics discussed in Sect. 14.3.2.

The functional form of our model has some further symmetries. In particular, we remark that (14.14) is invariant to translation ($x_i \to x_i + a$ for $a \in \mathbb{R}$) and equivariant to reflections through the origin ($x_i \to -x_i$) of all node states. This is a desirable property for many opinion formation process, as it ensures that the opinion formation is only influenced by the relative position of the node states x_i and independent of a specific global reference frame. This is still true for the more general case of vector valued states: any rotation of the node states is norm preserving, and thus $s(\|x_j - x_k\|)$ is rotational and translational invariant. Since the term $\left((x_j - x_i) + (x_k - x_i)\right)$ is translation invariant and linear, any translation and rotation applied to all states will

[3] Note that we adapted the notation of a multi-way node interaction as given in Eq. (14.6) to emphasise that this function is symmetric in j and k and influencing node i.

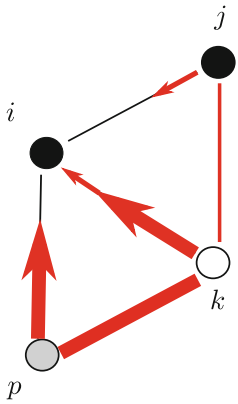


Fig. 14.2 Reinforcing group dynamics on 3-edges (3CM). The influence on node i due to the interactions on 3-edge $\{i, j, k\}$ and $\{i, k, p\}$. The value $x_i \in [0, 1]$ of the nodes is visualised by a colour gradient between white and black. We consider a monotonically decreasing scaling function (e.g. $s(x) = \exp(\lambda x)$ for $\lambda < 0$). As a result, nodes p and k reinforce each other as they have similar values, and hence a large $s(|x_p - x_k|)$. Nodes k and j have instead distinct values, which leads to a smaller scaling function $s(|x_j - x_k|)$. Reproduced and adapted from [24]

leave (14.14) equivariant, both in the case of scalar and vector valued states. Note that this 'quasi-linearity' of the three-way node interaction function is in close correspondence to the necessary and sufficient conditions for translation and rotational invariance for pairwise interaction systems [38]. In the following, we restrict our scope to scalar states x_i.

As we want to be able to model a reinforcement effect for nodes with a similar opinion, a natural choice for the scaling function $s(x)$ is

$$s\left(|x_j - x_k|\right) = \exp\left(\lambda\,|x_j - x_k|\right), \tag{14.16}$$

where the sign of the parameter λ determines if the function monotonically decreases or increases. Specifically, if $\lambda < 0$, then similar node states x_j and x_k will lead to a stronger influence on node i. If $\lambda > 0$, then dissimilar node states x_j and x_k will lead to a stronger influence on node i. Finally, if $\lambda = 0$, then we recover the linear dynamics discussed above (since the scaling function $s(x) = 1$ is constant). Figure 14.2 shows the influences on node i for $\lambda < 0$, i.e. where similar node states reinforce each other.

14.4.2 Derivation of a Weighted, Time-Dependent Laplacian

In Sect. 14.3.2, we showed that, in the case of linear node interactions, a multiway, and thus a three-way dynamical system can be rewritten as a pairwise dynamical system, defined on a network where the weight of an edge is the number of 3-nodes to which the edge belongs. Let us explore how this result extends to 3CM. Recall that we assumed that the adjacency tensor $A^{(3)}_{ijk}$ is symmetric. We define $\mathcal{I}_{ij}$ as the index-set containing all nodes k that are part of 3-edge containing node i and j. Note

that $\mathcal{I}_{ij} = \emptyset$ if the nodes i, j are not part of any 3-edge. We now define the weighted adjacency matrix $\mathfrak{W}$ as

$$(\mathfrak{W}^{(3)})_{ij} = \sum_k A^{(3)}_{ijk} s\left(|x_j - x_k|\right) = \sum_{k \in \mathcal{I}_{ij}} s\left(|x_j - x_k|\right). \tag{14.17}$$

Since $\mathcal{I}_{ii} = \emptyset$, because a node cannot appear more than once in a hyperedge, the weighted adjacency matrix has zero diagonal ($\mathrm{diag}(\mathfrak{W}^{(3)}) = 0$). The corresponding degree matrix, that measures the total three-way influence on node i is defined as

$$(\mathcal{D}^{(3)})_{ii} = \sum_{jk} A^{(3)}_{ijk} s\left(|x_j - x_k|\right) = \sum_j \mathfrak{W}^{(3)}_{ij} \tag{14.18}$$

and the corresponding Laplacian is given by

$$\mathcal{L}^{(3)} = \mathcal{D}^{(3)} - \mathfrak{W}^{(3)}. \tag{14.19}$$

Using (14.19), we can rewrite the dynamics in (14.15) as

$$\begin{aligned} \dot{x}_i &= \sum_{jk} \frac{1}{2} A^{(3)}_{ijk}\, s\left(|x_j - x_k|\right)\, ((x_j - x_i) + (x_k - x_i)) \\ &= \sum_{jk} A^{(3)}_{ijk}\, s\left(|x_j - x_k|\right)\, (x_j - x_i) \\ &= \sum_j \mathfrak{W}^{(3)}_{ij} (x_j - x_i) = -\sum_j \mathcal{L}^{(3)}_{ij} x_j. \end{aligned} \tag{14.20}$$

The 3CM can thus also be rewritten in terms of the Laplacian of a network represented by $\mathfrak{W}^{(3)}$, whereas the entries $(\mathfrak{W}^{(3)})_{ij}$ measure the three-way influence on node i over edge $\{i, j\}$. However, as the adjacency matrix $\mathfrak{W}^{(3)} = \mathfrak{W}^{(3)}(t)$ depends on the dynamical node states $x_i = x_i(t)$ in (14.17) and this projection is thus *time-dependent* and *node state-dependent*, which implies that we cannot write this dynamics in terms of pairwise interactions. We drop these dependencies in our notation for simplicity, and simply use $\mathfrak{W}$ from now on. The weighted time-dependent Laplacian $\mathcal{L}^{(3)}$ is the matrix representation of the non-linear dynamics and therefore the analogue of the motif Laplacian, introduced in Sect. 14.3.2 for linear dynamics.

14.4.3 Diffusive Processes on Hypergraphs

In order to generalise the 3CM for hypergraphs with arbitrary edge cardinality, we introduce a general model of diffusive processes on hypergraphs. Here, we consider $(x_j - x_i)$ as a diffusive coupling of nodes i and j. The influence of the nodes in a

hyperedge E_α on a node i is given by the node interaction function

$$f_i^{E_\alpha}(x_i, x_j, x_k, \cdots) = \begin{cases} \sum_{j \in E_\alpha} s_i^j(x_i, x_j, x_k, \ldots)\left(x_j - x_i\right) & i \in E_\alpha \\ 0 & i \notin E_\alpha. \end{cases} \tag{14.21}$$

The joint effect of all nodes within a hyperedge is given by the diffusive couplings of each node pair of influenced node (i) and acting node (j), modulated by the scaling function $s_i^j(E_\alpha)$ which captures the influence of the hyperedge E_α as a whole. The scaling function $s_i^j(E_\alpha)$ is invariant with respect to any permutation of the indices $k \in E_\alpha$, where $k \neq i$ and $k \neq j$. For the special case where it is symmetric in all node indices $k \neq i$, we can write $s_i(E_\alpha)$.

The overall dynamics of node i is then obtained by linearly combining the effect of each hyperedge node i is part of:

$$\dot{x}_i = \sum_\alpha f_i^{E_\alpha}(x_i, x_j, x_k, \cdots). \tag{14.22}$$

For the special case of a scaling function s_i which is independent of acting node (j), we can derive analytical results for the behavior of this opinion formation process. We can then write the effect of all n-edges on i as:

$$\dot{x}_i = \sum_{n=2}^{N} \sum_{jk\ldots} A^{(n)}_{ijk\ldots} s_i(x_i, x_j, x_k, \ldots)\left(x_j - x_i\right) \times \frac{1}{(n-2)!} \tag{14.23}$$

Analogously to the 3CM, we define weight matrices $\mathfrak{W}^{(n)}$and corresponding degree matrices $\mathfrak{D}^n$ as follows:

$$\begin{array}{l} \mathfrak{W}^{(n)}_{ij} = \sum_{kl\ldots} A^{(n)}_{ijk\ldots} s_i(x_i, x_j, \ldots) \frac{1}{(n-2)!} \\ \mathfrak{D}^{(n)}_{ii} = \sum_j \mathfrak{W}^{(n)}_{ij} \end{array} \tag{14.24}$$

Here, $\mathfrak{D}^{(n)}_{ij} = 0$ for all $i \neq j$. Equation (14.23) can now be written as:

$$\dot{x}_i = -\sum_{n=2}^{N} \sum_j \mathfrak{L}^{(n)}_{ij} x_j = -\sum_j \mathscr{L}_{ij} x_j \tag{14.25}$$

where $\mathscr{L}_{ij} = \sum_{n=2}^{N} \mathfrak{L}^{(n)}_{ij}$ and $\mathfrak{L}^{(n)}_{ij} = \mathfrak{D}^{(n)}_{ij} - \mathfrak{W}^{(n)}_{ij}$.

We see again that when the scaling function s_i is constant, the dynamics reduces to a linear dynamics on an (effective) static weighted network as shown in Sect. 14.2. However, when the interactions are non-linear, the corresponding network is time-dependent and multi-way effects are created.

14.4.4 Multi-way Consensus Model (MCM)

There are multiple possibilities for specific forms of the nonlinear scaling functions $s_i^j(E_\alpha)$ in the general diffusion process in (14.21). They can have different sociological motivations and result in distinct mathematical properties, which we will show for two specific node interaction functions. They define two submodels, MCM I and MCM II, of a general Multi-way Consensus Model (MCM) [33].[4]

14.4.4.1 MCM I Models Homophily

Homophily is a central concept in sociology describing the tendency of like-minded individuals to interact [17]. The topology of social interactions is often influenced heavily by homophily. Here, we consider the underlying topology fixed, but interpret homophily instead as a force modulating the effect of a hyperedge, depending on the proximity of the opinions inside it.

This is captured by the following node interaction function (MCM I)

$$f_i^{E_\alpha}(x_i, x_j, x_k, \cdots) = s_i^I(g_i^{E_\alpha}(x)) \sum_{j \in E_\alpha} (x_j - x_i), \tag{14.26}$$

where the argument function $g_i^{E_\alpha}$ of the scaling function s_i^I measures the distance of the state of **node** i to the mean state of the hyperedge **including** i:

$$g_i^{E_\alpha}(x_i, x_j, x_k, \cdots) = \left| \frac{\sum_{k \in E_\alpha} x_k}{\mid E_\alpha \mid} - x_i \right| = |\langle x \rangle_{E_\alpha} - x_i| \tag{14.27}$$

In sociological terms, the argument function $g_i^{E_\alpha}$ quantifies the difference between the opinion of individual i and the average opinion of group E_α that i belongs to. The influence of a group on a node is thus determined by the proximity of its average state to the state of the node, modulated by the scaling function s_i^I. For instance, a monotonically decreasing function s_i^I represents an individual i who is less influenced by groups with opinions very different from its own than by groups with similar opinions.

Mathematically speaking, $g_i^{E_\alpha}$ is independent of acting node j within the hyperedge, and $s_i^I(g_i^{E_\alpha}(x))$ thus modulates the competing effect of different hyperedges on the state of an incident node. In other words, the scaling function determines the rate at which a certain hyperedge influences the state of a node.

[4] Note that 3CM (MCM) was originally named three-body (multi-body) consensus model in [24, 33]. Here, we prefer the more descriptive name three-way consensus model, which emphasises more clearly that it is not the number of entities involved but the type of interaction between those entities (pairwise vs multi-way) which is different to classical models.

14.4.4.2 MCM II Models Conformity

Conformity is used to describe the tendency of an individual to align its beliefs to those of its peers, and is usually affected by the reinforcing nature of shared opinions (peer pressure). Modelling peer pressure has already been the motivation for the node interaction function of 3CM in (14.14) before. We generalise this for hyperedges of arbitrary size (MCM II):

$$f_i^{E_\alpha}(x_i, x_j, x_k, \cdots) = \sum_{j \in E_\alpha} s_i^{II}(g_{i \leftarrow j}^{E_\alpha}(x))(x_j - x_i). \tag{14.28}$$

where the argument function $g_{i \leftarrow j}^{E_\alpha}$ of s_i^{II} now measures the distance of the state of a **participating node** j from the mean state of the hyperedge **excluding** i.

$$g_{i \leftarrow j}^{E_\alpha}(x_i, x_j, x_k, \cdots) = \left| \frac{\sum_{k \in E_\alpha, k \neq i} x_k}{|E_\alpha| - 1} - x_j \right| = |\langle x \rangle_{E_\alpha \setminus i} - x_j| \tag{14.29}$$

When all the hyperedges have size 3, we then have that

$$g_{i \leftarrow j}^{\{i,j,k\}}(x_i, x_j, x_k) = \left| \frac{x_j + x_k}{2} - x_j \right|.$$

and the node interaction function is thus given by

$$f_i^{\{i,j,k\}}(x_i, x_j, x_k) = s_i^{II}\left(\frac{|x_k - x_j|}{2} \right) \left((x_j - x_i) + (x_k - x_i) \right). \tag{14.30}$$

which indeed recovers the node interaction function (14.14) of 3CM, with a constant that can be absorbed in the time scale.

In sociological terms, the argument function $g_{i \leftarrow j}^{E_\alpha}$ of s_i^{II} captures the difference between the opinion of individual j to the average opinion of the group except individual i. Thus, in MCM II the influence exerted by j inside a hyperedge depends on the proximity of its opinion to those of the rest of the group.

Mathematically, $g_{i \leftarrow j}^{E_\alpha}$ is dependent on acting node (j) and $s_i^{II}(g_{i \leftarrow j}^{E_\alpha}(x))$ thus determines which nodes inside a single hyperedge are the most influential.

Like in 3CM, for a monotonically decreasing function an individual i tends to be more influenced by individuals who agree with the rest of the group. For an increasing function, individuals are more attracted to the outliers of a group (anti-conformists or contrarians).

It becomes clear here that even if we choose $s(x) = s_i^I(x) = s_i^{II}(x)$ to be of the same form, the behaviour of the two facets will differ due to their arguments.

14.4.4.3 General Multi-way Consensus Model

We can combine both of these submodels in a more general Multi-way Consensus Model (MCM) [33], which can capture both homophily and conformity. The overall effect of the hyperedge E_α on node $i \in E_\alpha$, is then given by

$$f_i^{E_\alpha}(x_i, x_j, x_k, \cdots) = s_i^I(g_i^{E_\alpha}(x)) \sum_{j \in E_\alpha} s_i^{II}(g_{i \leftarrow j}^{E_\alpha}(x))(x_j - x_i) \tag{14.31}$$

We note that, as in the case of 3CM, both node interaction functions are invariant under translations ($x_i \mapsto x_i + a$ for $a \in \mathbb{R}$) and reflection through the origin ($x_i \mapsto -x_i$). Thus, the dynamics are independent of the frame of reference in $\mathbb{R}$.

14.5 Opinion Drifts: Higher-Order Effects of Non-linearity

In the previous section we saw that non-linearity of the dynamics is important to make genuine higher-order effects appear, which can not be explained by pairwise interactions. This emphasises that in the non-linear cases, we have to pay extra attention to how the interaction of dynamics and higher-order topology affects the overall dynamics of the system. Therefore, our primary objective in this section is to determine if, and how, non-linear consensus models on hypergraphs asymptotically reach consensus and which aspects can influence the dynamics. We can do that by identifying and analysing conserved quantities, which is usually an essential step to understand the properties of dynamical systems.

14.5.1 Conservation and Shifts of the Average Node State

In particular, we will investigate the average node state of the system. For consensus dynamics on graphs, it is well-known that the average state at time t,

$$\bar{x}(t) = \frac{1}{N} \sum_{i=1}^{N} x_i(t), \tag{14.32}$$

is conserved under general conditions. Specifically, consider a pairwise dynamical system described by

$$\dot{x}_i(t) = \sum_{j=1}^{N} A_{ij} f_{ij}(x_i(t), x_j(t)) = \sum_{j=1}^{N} A_{ij} h(x_j(t) - x_i(t)).$$

The initial average $\bar{x}(0)$ is conserved if the derivative $\dot{\bar{x}}(t) = \frac{1}{N}\sum_{i,j=1}^{N} A_{ij}h(x_j(t) - x_i(t))$ is zero for all times. This is true if the adjacency matrix A_{ij} of the graph is symmetric and the node interaction function $h(x)$ is odd, which is fulfilled by quasi-linear dynamics (14.8).

We now investigate how these conditions for conservation of the average state change for a general multi-way interaction system. We consider a general node interaction function $f(x_1(t), \ldots, x_n(t)) = h(\sum_{j\neq i}(x_j(t) - x_i(t)))$, where the form of the node interaction function h ensures that the dynamics influence node i. To determine conditions for a conservation of the average state we write

$$\dot{x}_i(t) = \sum_{jk\ldots} A_{ijk\ldots} h\left(\sum_{j\neq i}(x_j(t) - x_i(t))\right). \tag{14.33}$$

Let $\Pi(i, j, k, \ldots)$ be the set of all permutations of the n indices. Using this notation, we can conclude that the derivative

$$\dot{\bar{x}}(t) = \frac{1}{N}\sum_{i,j,k,\cdots=1}^{N} A_{ijk\ldots} h\left(\sum_{j\neq i}(x_j(t) - x_i(t))\right) \tag{14.34}$$

is zero for all times, if we have $A_\pi = A_\tau$ for all permutations $\pi, \tau \in \Pi(i, j, k, \ldots)$ and moreover $\sum_{i=1}^{n} h(\sum_{j\neq i} x_j(t) - (n-1)x_i(t)) = 0$. This is the case for an undirected multi-way interaction (for which A is symmetric in all indices), provided $h(x)$ is a linear function. We can thus conclude that, in line with our previous discussion, for a multi-way dynamical systems of the form (14.33), a linear dynamics conserves the average state of the system. For non-linear dynamics, however, we can not generally guarantee a conservation of the average state.

14.5.2 *Factors Influencing the Consensus Process for Non-linear Dynamics*

For concreteness let us consider a simple but illustrative case for three-way interactions to gain some intuition for the possible effects we can observe in a multi-way interaction system. For 3CM, where the node interaction function takes the form $f(x_i, x_j, x_k) = s\left(|x_j - x_k|\right)((x_j - x_i) + (x_k - x_i))$, the change in the average state can be written as

$$\dot{\bar{x}}(t) = \frac{1}{N}\sum_{i=1}^{N}\dot{x}_i(t) = \frac{1}{N}\sum_{i,j=1}^{N}\mathfrak{W}(t)_{ij}(x_j(t) - x_i(t)). \tag{14.35}$$

It is important to note that when the dynamics is non-linear, $\mathfrak{W}(t)$ is time-dependent. Hence, the average state is only conserved if $\mathfrak{W}(t)$ is symmetric for all times. In particular, this means that for all i, j,

$$\mathfrak{W}(t)_{ij} = \mathfrak{W}(t)_{ji} \Leftrightarrow \sum_{k \in \mathcal{I}_{ij}} s\left(\left|x_j(t) - x_k(t)\right|\right) = \sum_{k \in \mathcal{I}_{ij}} s\left(\left|x_i(t) - x_k(t)\right|\right),$$

where $\mathcal{I}_{ij}$ is the index-set containing all nodes k that are part of 3-edge containing node i and j. This is only true for all times if the scaling function is constant ($s(x) = c$), i.e., when the dynamics is linear. In this case, the weighted matrix $\mathfrak{W}_{ij} = c(W^{(3)})_{ij}$ simply equals the motif adjacency matrix scaled by the constant c (cf. Sect. 14.3.2).

Otherwise, the weighted matrix $\mathfrak{W}(t)$ may be asymmetric at some time point t, which implies that the average state can shift, i.e., that we can get a drift of opinions. These possible shifts are influenced by an interplay between

1. the initial node states,
2. the scaling function $s(g(x))$, and
3. the hypergraph topology

as all of these aspects are encoded in the weighted matrix $\mathfrak{W}^{(3)}$. This is also true for the more general case of non-linear diffusive dynamics on general hypergraphs, represented by the weighted matrix $\mathfrak{W}^{(n)}$(14.24).

14.5.3 *Influence of the Initial Node States on the Final Consensus Value in a Fully Connected Hypergraph*

We first consider a system which eliminates all topological effects, given by a fully connected hypergraph $\mathcal{H}$ with nodes $V(\mathcal{G}) = \{1, \ldots, N\}$, in which each n-tuple of distinct nodes is connected by a hyperedge. This is to investigate the isolated effect of the initial node states.

Let us first stay with the model example of the 3CM to get some intuition of the higher-order effects that appear in a fully connected hypergraph of 3-edges. As all possible hyperedges exist, the node set $\mathcal{I}_{ij} = V(\mathcal{G})/\{i, j\}$ is given by all nodes except i and j and the symmetry condition

$$\sum_{k \in \mathcal{I}_{ij}} s\left(\left|x_j(t) - x_k(t)\right|\right) = \sum_{k \in \mathcal{I}_{ij}} s\left(\left|x_i(t) - x_k(t)\right|\right)$$

thus implies that the multi-way effects have to balance out for all nodes in the network in order for the average state to be conserved. The equality here only depends on the initial node states and the scaling function $s(x)$, as the topology of the hypergraph captured by the node set $\mathcal{I}_{ij}$ does not have an influence in a fully connected system.

As an illustration, consider a situation where the number of nodes is even and when the initial values $x_i(0)$ on the nodes is binary, that is either zero or one. The symmetry

condition (14.36) will be satisfied only if the initial configuration $x(0)$ is balanced, that is when $\bar{x}(0) = 0.5$, in which case this average state is conserved in time. In contrast, if the initial configuration is unbalanced, (14.36) will not hold in general, and the average state will evolve in time. If $s\left(\left|x_i - x_j\right|\right)$ is given by a decreasing function, that is when similar node states reinforce each other, the deviation from 0.5 is expected to grow in time, with a drift towards the majority. In contrast, if the scaling function is such that dissimilar node states reinforce each other, one will observe a drift to the balanced state 0.5.

To validate these findings, we performed numerical simulations of 3CM on a fully connected 3-edge hypergraph of 100 nodes. We used the exponential scaling function $s(x) = \exp(\lambda x)$ with the parameter λ set to $\lambda = -1$ to obtain an example for a decreasing function, $\lambda = 1$ as an example for a growing function, and $\lambda = 0$ to obtain a constant function (i.e., a linear dynamics).

Considering different initial distributions of the node states, we compared

1. a deterministic, symmetric distribution $\mathcal{B}(0.5)$, where 50% of the initial node states have value 0 and 50% have value 1 ($\bar{x}(0) = 0.5$)
2. a deterministic, asymmetric distribution $\mathcal{B}(0.2)$, where 80% of the node states have value 0 and 20% have value 1 ($\bar{x}(0) = 0.2$)
3. a random initialisation according to a uniform distribution $\mathcal{U}([0, 1])$ ($\mathrm{E}(\bar{x}(0))$=0.5)

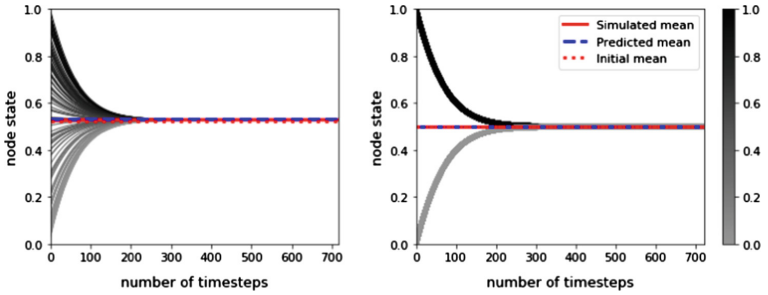
In the first two cases we do not observe any shift in the average state, as expected. We can observe this conservation in Fig. 14.3, independently of the non-linear dynamics that are applied. However, we do see a shift for the asymmetric initialisation, as shown in Fig. 14.4. The simulations confirm that the average state is conserved for linear dynamics ($\lambda = 0$) and multi-way effects only occur for non-linear node interaction functions with $\lambda \neq 0$. For $\lambda < 0$ we observe a shift of $x(t)$ towards the majority, resulting in an asymptotic average $\lim_{t\to\infty} \bar{x}(t)$ smaller than the initial value of $\bar{x} = 0.2$. For $\lambda > 0$ we see the opposite phenomenon, with a shift of the average opinion towards 0.5.

In order to approximate the asymptotic state from the initial configuration of the system, we can use a simple method which estimates the dynamical importance w_i of a node i based on the initial configuration as

$$w_i(t) = \frac{\text{influence of node } i}{\text{total weight in the system}} = \frac{\sum_{j,k=1}^{N} A^{(3)}_{ijk} s\left(\left|x_i(t) - x_j(t)\right|\right)}{\sum_{i,j,k=1}^{N} A^{(3)}_{ijk} s\left(\left|x_i(t) - x_j(t)\right|\right)}. \tag{14.36}$$

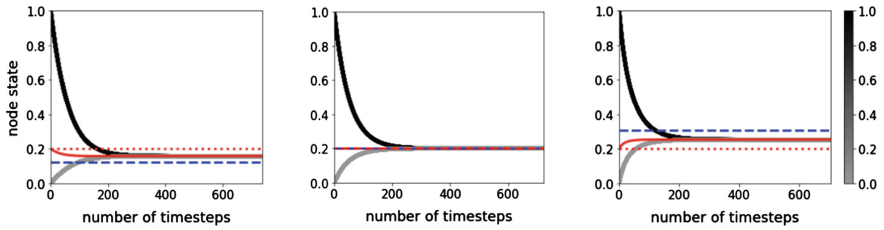
The asymptotic value of $\bar{x}$ is then obtained by one explicit Euler step of the dynamics from the initial configuration $\bar{x}(0)$

$$\bar{x}^p = \bar{x}(0) + \sum_{i=1}^{N} w_j(0)(x_j(0) - x_i(0)). \tag{14.37}$$



(a) $\mathcal{U}([0, 1])$ (b) $\mathcal{B}(0.5)$

Fig. 14.3 Symmetric initialisation: Conservation of average state. For uniform random **a** and binary **b** symmetric initialisation the average is preserved for dynamics with a scaling function $s(x) = \exp(\lambda x)$ for all $\lambda = \{1, 0, -1\}$, so for both linear and non-linear dynamics. Shown are the results for $\lambda = -1$ and the line colors represent the initial states of the nodes. The predicted average agrees with the simulated average in this case



(a) $\lambda = -1$ (reinforcing) (b) $\lambda = 0$ (linear) (c) $\lambda = 1$ (inhibiting)

Fig. 14.4 Asymmetric initialisation: Conservation only for linear dynamics. An asymmetric initialisation, with $\bar{x}(0) = 0.2$, may shift the average node state in 3CM for fully-connected hypergraphs. The scaling function is $s(x) = \exp(\lambda x)$. For $\lambda < 0$ (left), the dynamics exhibits a drift towards the majority as similar node states reinforce each other. The opposite effect occurs for $\lambda > 0$ (right), as the dynamics exhibits a drift towards balance. The average state is conserved for $\lambda = 0$ (center), as expected for linear dynamics. Again, the line colors represent the initial states of the nodes. Dotted red lines indicate the initial value of the average node state. Black (grey) solid lines represent the evolution of the state of nodes whose initial configuration is one (or zero). Dashed blue lines are the final state approximation, $\bar{x}^p$. Reproduced and adapted from [24]

The simulations in Figs. 14.3 and 14.4 also display the predicted value (14.37), which correctly identifies the direction of the shift.

14.5.3.1 Analytical Examination of MCM I

Let us now investigate how the shifts in the average node state of the system generalise to the case of non-linear consensus on hypergraphs modeled by the MCM. We will look at the two submodels MCM I and MCM II separately. Similar to the 3CM, the analytical examination of the MCM II is difficult as the argument of the scaling

function s_i^{II}, given by the argument function $g_{i\leftarrow j}^{E_\alpha}$, depends on both node i and node j. In contrast, in the case of the MCM I, the argument function $g_i^{E_\alpha}$ for the scaling function s_i^I is independent of acting node j and symmetric in all node indices $k \in E_\alpha, k \neq i$. We can thus quantify the shifts of the average opinion in a symmetric system analytically.

We assume that the scaling function is the same for all nodes and we call it s^I. We assume a homogeneous mixing, where the nodes have equal probability of being part of a hyperedge. We consider a hypergraph $\mathscr{H}$ with m_k hyperedges of cardinality k for $k = 2, 3, \ldots, N$. Then, each node participates in $\frac{km_k}{N}$ hyperedges of cardinality k and that the mean of every hyperedge is the global mean $\bar{x}$. For an arbitrary node in some hyperedge $i \in E_\alpha$, we thus have

$$\dot{x}_i = \sum_{k=2}^{N} \frac{k^2 m_k}{N} s^I\left(|\bar{x} - x_i|\right)(\bar{x} - x_i) \tag{14.38}$$

The mean state evolves as

$$\dot{\bar{x}} = \frac{1}{N^2}\left(\sum_{k=2}^{N} k^2 m_k\right)\left(\sum_{i=1}^{N} s^I\left(|\bar{x} - x_i|\right)(\bar{x} - x_i)\right) \tag{14.39}$$

and we observe that in a homogeneously mixed system, the mean does not shift if the distribution of x_i about the mean is symmetric. This result is equivalent to the results for 3CM which we examined in Fig. 14.3.

We can now investigate the effect of an unbalanced initial distribution of the states analytically. Consider a situation where the initial states are binary (either 1 or 0). Suppose at $t = 0$, f_0 fraction of the nodes have state 0, and the rest ($f_1 = 1 - f_0$) have state 1. From Eq. 14.39, we can write

$$\begin{aligned}\dot{\bar{x}} &= \frac{1}{N^2}\left(\sum_{k=2}^{N} k^2 m_k\right)\left(\sum_{i=1}^{N} s^I\left(|f_1 - x_i|\right)(f_1 - x_i)\right) \\ &= \frac{1}{N}\left(\sum_{k=2}^{N} k^2 m_k\right) f_0 f_1 (s^I(f_1) - s^I(f_0))\end{aligned}$$

If s^I is monotonically increasing, $f_1 > f_0$ implies that $\dot{\bar{x}} > 0$ and $f_1 < f_0$ that $\dot{\bar{x}} < 0$, i.e. $\bar{x}$ shifts towards the majority. Similarly, $\bar{x}$ shifts towards the minority for monotonically decreasing s^I. This is fundamentally different to the case of 3CM, where the same shift appears, but in the opposite direction. We therefore also expect the generalised MCM II to behave in the opposite way to MCM I. In the next section, we thus investigate these contrasting effects, which arise from the different form of non-linear argument.

14.5.3.2 Fundamental Differences Between MCM I and MCM II

To compare the outcomes of the two models, we run numerical simulations on identical topologies (a fully connected hypergraph with $N = 10$ nodes) with the same choice of scaling function with different parameters:

$$s_i^I(x) = s_i^{II}(x) = e^{\lambda x} \quad \forall i \in V(\mathscr{H}) \tag{14.40}$$

As before, the initial node states have binary values (0 or 1), with n_0 nodes of state 0.

Numerical results in Fig. 14.5 show that the two submodels MCM I and MCM II evolve in opposite ways. Further, the results for MCM I validate the analytical results in Sect. 14.5.3.1. For a monotonically increasing scaling function $s(x) = e^{\lambda x}$ with $\lambda > 0$, we see that the average state of the nodes shifts towards the opinion of the initial majority. Similarly, the average state shifts towards the opinion of the initial minority for a monotonically decreasing scaling function s^I ($\lambda < 0$). While MCM II is a direct generalisation of 3CM and the results of the simulations in Fig. 14.5 (right) align with the numerical results of the previous section, MCM I shows an opposite behaviour to MCM II despite the same scaling function $s^I(x) = s^{II}(x)$. These drastic differences underline the huge effect of the argument functions, i.e. $g_i^{E_\alpha}$ in the case of MCM I and $g_{i\leftarrow j}^{E_\alpha}$ in the cas MCM II, on the long term behavior of the dynamical system.

14.5.4 Influence of Node Specific Function Parameters

Up until here, we have focused on a dynamics of the form (14.26) (MCM I) or (14.28) (MCM II/3CM) with a scaling functions $s(x)$ which was the same for all

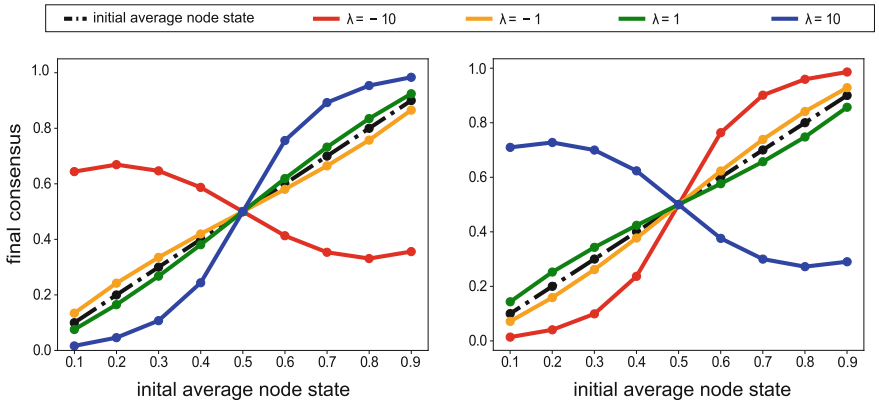


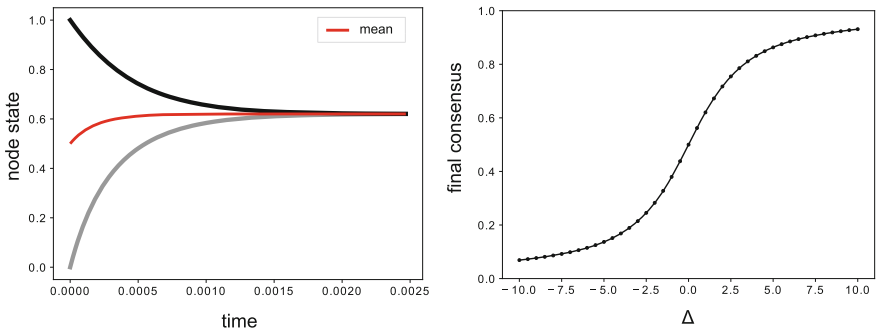
Fig. 14.5 MCM I and MCM II show opposite effects. Numerical simulations to compare the evolution of MCM I (left) and II (right) on a fully connected hypergraph. Reproduced and adapted from [33]

nodes, e.g. $s_i(x) = s(x) = e^{\lambda x}$ for all $i \in V(\mathcal{G})$. However, we can also choose node specific parameters for the scaling functions $s_i(x)$. These different scaling functions s_i may, e.g., be motivated by sociological aspects such as character traits of different individuals. In order to explore the higher-order effects of these individual node traits, consider again for each node $i \in V(\mathcal{G})$ the submodel MCM I (14.26) with exponential function:

$$s_i^I(x) = e^{\lambda_i x}. \tag{14.41}$$

For $\lambda_i < 0$ ($\lambda_i > 0$), the function is monotonically decreasing (increasing). Following the sociological motivation of MCM I in Sect. 14.4.4.1, a decreasing function s_i can model an individual i that is comparably less influenced by groups E_α with an average opinion $\langle x_{E_\alpha} \rangle$ that is very different from its own opinion x_i, than by groups that have an average opinion similar to its own opinion x_i. This can be thought of as individual i resisting change, or some form of 'stubbornness'. On the other hand, an increasing function s_i can be thought of as representing 'gullibility', i.e. a rather susceptible individual i. In Fig. 14.6, we present the temporal evolution of the dynamics of the MCM I with $s_i^I(x) = e^{\lambda_i x}$ on a fully connected hypergraph ($N = 10$) with binary, symmetric initialisation. The nodes whose states were initialised to $x_i = 1$ (or to initial state $x_i = 0$) have a scaling function with parameter $\lambda_i = -\Delta$ (or parameter $\lambda_i = \Delta$), respectively. The numerical results in 14.6 show that the final consensus of opinion values shifts towards the initial opinion of stubborn individuals.

Note that this is an important aspect of the MCM I, as it enables an individual to heavily influence other members of a group while being resistant to their influence. Hence our models allows certain individuals to be 'trendsetters', that can pull entire groups towards their opinion. Specifically, stubborn individuals within in a group of people with different opinions will become trendsetters.



(a) Evolution with $\Delta = 1$

(b) Consensus vs Δ

Fig. 14.6 Influence of node specific parameters. Evolution of MCM I on a fully connected hypergraph of 10 nodes initialised with 5 nodes each of opinions 0 (with $\lambda_i = \Delta$) and 1 (with $\lambda_i = -\Delta$). Reproduced and adapted from [33]

14.5.5 *Influence of Clustered Hypergraph Topologies*

In the previous experiments we have focused on fully connected hypergraphs, which eliminated the influence of specific hypergraph topologies on the dynamics. We could thus investigate, how asymmetric distributions of the initial node states cause shifts in the average opinion of the system, which would not be present in a pairwise or a linear higher-order setting. The direction of these shifts were additionally dependent on the form of the scaling function s_i. We now want to investigate the additional influence of the hypergraph topology, if it is not symmetric as in the fully connected case.

If we consider a hypergraph with two clusters A and B, we can define two types of hyperedges: If all the nodes of a hyperedge E_α are contained in either cluster A or in cluster B, we call E_α a *cluster* hyperedge. However, if E_α contains nodes from both clusters, it will be referred to as a *connecting* hyperedge. In the latter case, the connecting hyperedge is called *oriented* towards one of the cluster, if the minority of the hyperedge nodes is part of that cluster. Otherwise, if the nodes are part of cluster A or B in equal numbers, the hyperedge is *unoriented*.

As an illustrative example we consider a 3-edge hypergraph consisting of two equally sized fully connected clusters. In addition, we assume that these clusters are connected by a (small) set of 3-edges. This setting is illustrated in Fig. 14.7.

The dynamical effect of this construction becomes clear if we consider how the initial node states will influence the future dynamics. We consider binary initial node states, whereas all the nodes in cluster A have the initial state $x_A(0) = 0$ and the nodes in cluster B the initial state $x_B(0) = 1$. We consider the 3CM with a positive-definite,

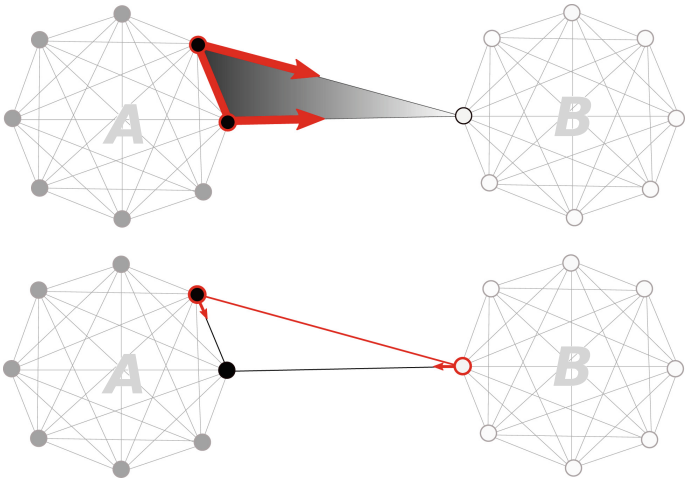


Fig. 14.7 Cluster dynamics. If we consider a binary initialisation of the two clusters, here in black and white, and a 3-edge oriented towards cluster B (top), the consensus in cluster A accelerates the rate of change of the neighbour in B. In contrast, the node-state difference between the clusters is maximal, which slows down the effect of cluster B on A. Reproduced and adapted from [24]

decreasing scaling function $s(x)$, such that the influence of nodes with similar states is reinforced within a hyperedge. Moreover, in this example, we chose to add only a single 3-edge between the cluster A and B, that is oriented towards cluster B. Due to the consensus in cluster A and the fact that the connecting hyperedge is oriented towards cluster B, the diffusion dynamics is accelerated towards cluster B, as the majority of the nodes in the connecting hyperedge have the initial opinion of cluster A and are thus reinforcing each other. On the contrary, the influence of cluster B is inhibited in the opposite direction, as the node couplings of nodes belonging to different clusters damp the diffusion because of their large state difference. For this reason, one expects the average initial value in A to dominate that in B and thus to dominate the asymptotic consensus value. Note that we thus achieve directed dynamics (or an asymmetric flow) from one cluster to the other.

In order to quantitatively analyse this mechanism, we perform numerical simulations on two fully connected clusters, each consisting of 10 nodes, with the binary initialisation specified above. We then connect the clusters with 80 randomly placed 3-edges, such that a fraction $p \in [0, 1]$ of 3-edges are oriented towards cluster A and the rest towards cluster B.

14.5.5.1 Cluster Dominance Through Directedness of Cluster connection

We first examine the influence of the orientation parameter p. For that purpose, we take the scaling function $s(x) = \exp(\lambda x)$ with $\lambda = -100$, so that pairs of similar nodes exert a strong influence on other nodes. We show the results of our model simulations averaged over 20 random instances in Fig. 14.9. In Fig. 14.9 (left), we observe a shift in the final consensus value towards the initial value in cluster A (or cluster B, respectively). The direction of this shift depends on the orientation of the connecting 3-edges, quantified by p. For $p = 0$, all the connecting 3-edges are oriented towards cluster B and the initial opinion of cluster A thus dominates the dynamics. On the contrary, for $p = 1$ all connecting traingles are oriented towards cluster A, therefore we observe a maximal influence of the initial value of cluster B.

The asymmetry of the dynamics which results from the orientation of the connecting hyperedges also influences the rate of convergence towards consensus, as shown in Fig. 14.9 (right). More asymmetric configurations lead to a faster rate of convergence. The simulations also reveal higher fluctuations in the asymptotic state for values close to $p = 0.5$. This result indicates that the process is sensitive to even small deviations from balance in the initial topology, which can lead to large differences in the consensus value.

Note that the effect of the orientation reverses if we consider an increasing scaling function such that dissimilar node states reinforce each other. In Sect. 14.8, we examine this effect by changing λ for the scaling function $s(x) = \exp(\lambda x)$. We observe a transition from the initial value in cluster A to that in cluster B, as expected.

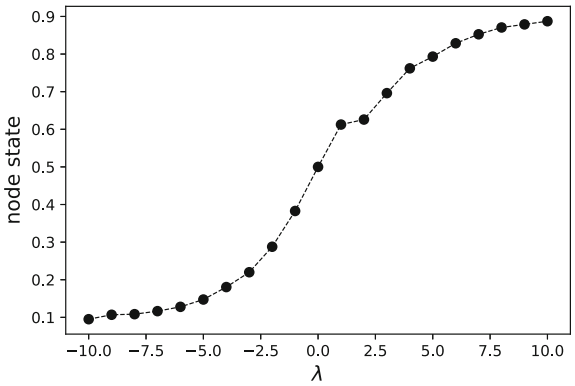


Fig. 14.8 Cluster dominance for reinforcing versus inhibiting dynamics. The final consensus value of the two cluster system (with $p = 0$), dependent on the parameter λ. As the connecting 3-edges are all oriented towards cluster B, it depends on λ if the nodes in cluster A are reinforcing each others' influence which leads to directed dynamics towards cluster B ($\lambda < 0$) or inhibit each other ($\lambda > 0$) which leads to the contrary effect. Therefore, the consensus value shifts towards the mean of cluster B with growing λ. For $\lambda = 0$ we have linear dynamics and the initial average 0.5 is conserved. Reproduced and adapted from [24]

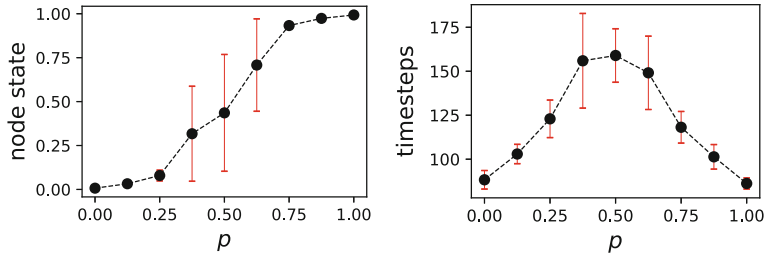


Fig. 14.9 Cluster dominance depends on directedness. Simulations of 3CM on two interconnected clusters of 10 nodes, with the scaling function $s(x) = \exp(-100x)$ (see main text for a complete description). We compute the final consensus value, averaged over 20 simulations, where the error bars display one standard deviation (left). As the fraction of 3-edges directed from cluster A to cluster B increases, so does the consensus value towards the initial state in cluster A. The rate of convergence is significantly faster when the initial configuration is very asymmetric, that is extreme values of p (right). Reproduced and adapted from [24]

14.5.5.2 Minority Influence

Instead of considering equally sized groups as in the last section, we can also consider settings in which one cluster forms a "minority" and is comparably smaller than the other cluster (the majority) [25]. As shown in Fig. 14.10, even in this case the opinion of the global minority may have a stronger influence on the final consensus value than the majority cluster, depending on the relative number of 3-edges oriented towards the majority. In the context of opinion dynamics, this type of behavior is akin to a "minority influence", where small groups can dominate the formation of an opinion.

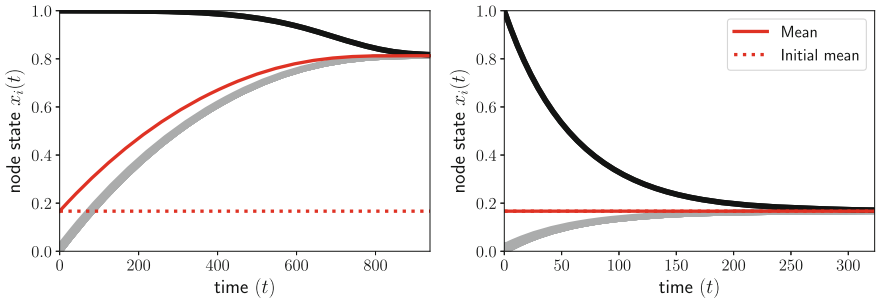


Fig. 14.10 Global minority influence through reinforcing opinion. We display simulations for a scaling function $s(x) = \exp(\lambda x)$ for $\lambda = -10$ and two clusters of different sizes, which are connected with a single 3-edge oriented towards cluster A. Cluster A comprises the majority of nodes (10 nodes) whereas cluster B consists only of 5 nodes. While intuition may suggest a final consensus that is leaning towards the initial opinion 0 of the majority cluster A, we observe the opposite behavior due to opinion reinforcing effect of the nonlinear coupling, which leads to an (effectively) directed dynamics between B and A (left). In contrast, if the dynamics are linear (right), the initial average opinion is conserved and therefore the majority opinion dominates the final consensus value. Reproduced and adapted from [25]

Note that this happens not because of the size of the minority group, but due to the internal cohesion of opinions within the minority and because the minority nodes form the local majority in the connecting subgroups. Accordingly, if the minority does not agree on the same opinion or the connecting subgroups are not oriented towards the majority, the minority influence is diminished. Likewise, if we remove the nonlinear effect of opinion reinforcement via the scaling function s and consider simply a linear coupling, then the initial opinion of the majority will have the strongest effect on the final consensus state.

14.5.5.3 Heaviside Function: Bounded Confidence Models on hypergraphs

Up until now, the scaling function s has always been an exponential function, $s(x) = \exp(\lambda x)$, to demonstrate general properties of the model. However, we can choose other functions as scaling functions. One interesting option is the Heaviside function, given by

$$s(|x_j - x_k|) = H(|x_j - x_k| - \phi) = \begin{cases} 0 & \text{if } |x_j - x_k| < \phi \\ 1 & \text{otherwise,} \end{cases} \tag{14.42}$$

which switches between a zero interaction and linear diffusion when the difference of the neighbouring nodes becomes smaller than a threshold $\phi \in (0, 1)$. This property is reminiscent of the bounded confidence model [5]. Note that the Heaviside function is not positive-definite, so that nonlinear consensus dynamics (MCM I and II) with

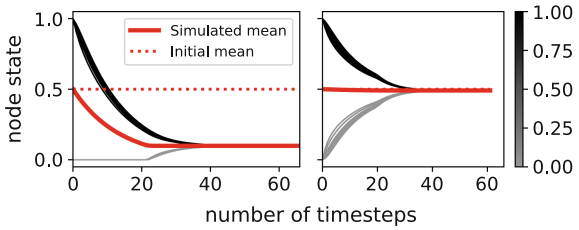


Fig. 14.11 Bounded confidence on hypergraphs. Time evolution of the node states for a Heaviside function with $\phi = 0.2$. For $p = 0$ (left), only diffusion from cluster A towards B is enabled, until the threshold of the Heaviside threshold $\phi = 0.2$ is reached. The dynamics then become linear and the average state becomes conserved. For $p = 0.5$ (right), the dynamics are initially symmetric, as the orientation of the connecting 3-edges are balanced, and the dynamics are simultaneously switched on. Reproduced and adapted from [24]

this scaling function do not necessarily converge to a consensus state asymptotically, in which all nodes have the same value.

In Fig. 14.11, we show the simulation results for $\phi = 0.2$ and a cluster scenario with $p = 0$, i.e., the 3-edges are all oriented towards cluster B. As the difference of the node states of the two clusters is initially larger than ϕ, only the nodes of cluster A in the connecting 3-edges are close enough such that linear diffusion takes place towards cluster B. Therefore, only nodes of cluster B change their value initially as shown in Fig. 14.11 (left). As soon as the difference of the node states of the two clusters is smaller than ϕ, the opinion dynamics is switched on for all node couplings and the dynamics becomes linear. Therefore, the asymmetry of the dynamics disappears. This mimics a situation of two polarised clusters where one side makes unilateral concessions until the other side starts to participate in the consensus formation again. For $p = 0.5$ the dynamics are symmetric as the orientation of the 3-edges is balanced.

14.5.5.4 Time-Scale Separation

Finally, we investigate our multi-way interaction dynamics concerning the interplay between the topology in a clustered hypergraph and initial conditions that is not bimodal. In particular, we are interested in examining different time scales in the dynamics induced by the clustered topology with two groups. The different time-scales are here associated to a fast convergence of states inside the clusters, followed by a slower convergence towards global consensus.

Specifically, we reconsider the clustered 3-edge hypergraph with $p = 1$ and with a dynamics governed by the 3CM. This time, however, the nodes in each cluster have different states initially. For our experiments, we initialise nodes in different clusters uniformly at random over disjoint intervals, such that nodes of cluster A have random initial states in the interval $I_A = [0, 0.5]$ and nodes in cluster B have random initial states in $I_B = [0.5, 1]$ (see Fig. 14.12). The initial cluster averages of the node states are thus separated.

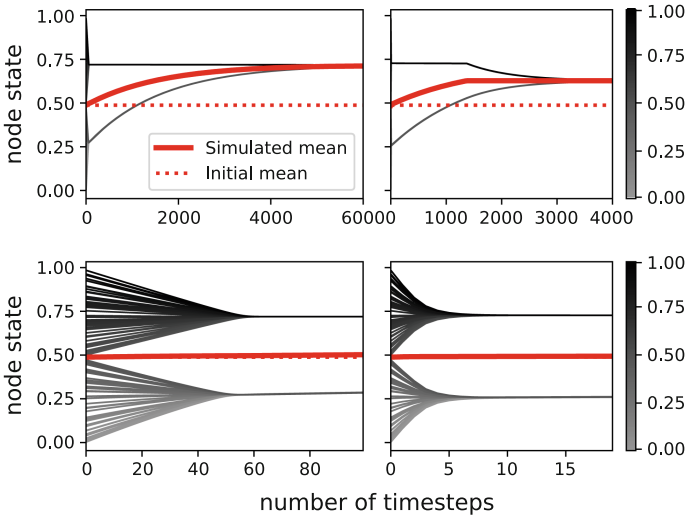


Fig. 14.12 Time-scale separation. Dynamics of two clusters A and B connected with $p = 1$, e.g. with 3-edges oriented towards nodes in A, and initialised with uniform distributions over separate intervals I_A, I_B with $I_A \cap I_B = \emptyset$. The left figures correspond to the exponential scaling function $s(x) = \exp(\lambda x)$ for $\lambda = -100$, and the right to a Heaviside function with threshold $\phi = 0.2$. We observe a timescale separation with a fast, symmetric dynamics inside the clusters, followed by a slow, asymmetric dynamics between the clusters. The fast dynamics is shown in the bottom figures, with qualitatively similar results for both scaling functions. The top figures show a shift towards cluster B for the slow dynamics. For the Heaviside function, the process becomes linear when the node states in the two clusters are less separated than the Heaviside-threshold. Reproduced and adapted from [24]

Two effects lead now to a fast multi-way consensus inside each cluster. First, each of the clusters are internally fully-connected. Second, the inter-cluster-dynamics will have a weaker effect, since the difference in the distribution of the initial conditions implies that $s\left(\left|x_i - x_j\right|\right)$ will be small if nodes i and j are in different clusters. As a result, we first observe a fast dynamics within the clusters in which nodes approach the cluster-average state (Fig. 14.12, bottom) and then a slower dynamics between the two clusters (Fig. 14.12, top).

The observed effect of the dynamics additionally depends on the scaling function. For an scaling function $s(x) = \exp(\lambda x)$, with $\lambda = -100$, we observe an asymmetric shift of the average node state as shown in Fig. 14.12 (left). If we consider the Heaviside function as a scaling function $s(x)$ instead, the dynamics show a similar asymmetry as in the exponential case, until the two cluster means are less separated than the Heaviside threshold $\phi = 0.2$. As shown in Fig. 14.12 (right), the dynamics then become linear and symmetric.

14.6 Conclusion

In this chapter, we have emphasised the importance of non-linearity for a dynamical process to exhibit higher-order effects. Specifically, our results show that it is important to distinguish between the model of the multi-body structure of a system (here: a hypergraph) and the model of its multi-way dynamics (here: non-linear multi-way consensus). The interplay of both aspects is important for genuine higher-order effects to emerge. This is particularly apparent for linear consensus models, whose dynamics can always be reduced to a pairwise dynamical system even when defined on a hypergraph. In other words, it is a necessary (but not sufficient) condition that the node interaction function is non-linear for genuine, non-reducible multi-way dynamical phenomena to emerge. In that case, adjacency matrices are not adequate to encode that sub-groups of nodes interact together, and higher-order objects like hypergraphs are required.

We have then analysed possible higher-order dynamical effects by looking at specific non-linear interaction functions, which are inspired by models in opinion dynamics. We introduced a general Multi-way consensus model (MCM) in which the adoption of an opinion by a node within a group is scaled non-linearly by the similarity of the group members, either including or excluding the affected node. This leads to submodels, the MCM I and the MCM II, which have different mathematical properties due to the dependencies of the argument of the non-linear scaling. In sociological terms, the two submodels can represent consensus dynamics that are either driven by homophily (MCM I) or by conformity or peer-pressure (MCM II).

The resulting dynamics lead to shifts of the average opinion state in the system, which would not be present in the case of pairwise or linear multi-way interactions.

In a fully connected system, we find that the shift in the final consensus value only depends on the interplay between (i) the distribution of the initial states of the nodes (no shifts if $\bar{x}(0) = 0.5$ versus shifts for $\bar{x}(0) \neq 0.5$) and (ii) the form of the non-linearity of the dynamics, i.e., the scaling function $s(g(x))$ (reinforcing (inhibiting) dynamics for monotonically decreasing (increasing) s in the case of 3CM, opposite effects of MCM I and MCM II due to the different form of their argument function g). If we additionally consider a scaling function s_i with node specific parameters, which classify certain nodes to be more stubborn than others (in the case of MCM I) we can observe shifts even in the case of $\bar{x}(0) = 0.5$.

If we go beyond fully connected systems and thus additionally consider the influence of the hypergraph topology, we observe that the influence of certain node subsets can dominate the final consensus value in clustered hypergraphs. This depends on the orientation of the hyperedges connecting the cluster. In the case of bounded-confidence dynamics, this mechanisms can even lead to a situation in an opinion dynamics where only one subgroup makes concessions initially. Moreover, in the case of initial state distributions which are not bimodal, a combination of symmetric and asymmetric dynamics is possible: we observe a timescale separation with a fast, symmetric dynamics inside the clusters, followed by a slow, asymmetric dynamics between the clusters.

Generally, we conclude that non-linearity is needed for higher-order dynamical effects to appear on hypergraphs. The effects that appear depend on a complex interplay between the type of the non-linear dynamics, the topology of the hypergraph and the initial node states. We have explored this interplay for a family of models for consensus dynamics, and the rich phenomenology that we observed motivates the study of these questions for other families of dynamical models in the future.

Acknowledgements We acknowledge partial funding from the Ministry of Culture and Science (MKW) of the German State of North Rhine-Westphalia (NRW Ruckkehrprogramm).

References

1. A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, C. Zhou, Synchronization in complex networks. Phys. Rep. **469**(3), 93–153 (2008)
2. S.E. Asch, Effects of group pressure on the modification and distortion of judgments. Groups, Leadership Men, 177–190 (1951)
3. C.A. Bail, L.P. Argyle, T.W. Brown, J.P. Bumpus, H. Chen, M.B.F. Hunzaker, J. Lee, M. Mann, F. Merhout, A. Volfovsky, Exposure to opposing views on social media can increase political polarization. Proc. Nat. Acad. Sci. **115**(37), 9216–9221 (2018)
4. A.R. Benson, D.F. Gleich, J. Leskovec, Higher-order organization of complex networks. Science **353**(6295), 163–166 (2016)
5. V.D. Blondel, J.M. Hendrickx, J.N. Tsitsiklis, On Krause's multi-agent consensus model with state-dependent connectivity. IEEE Trans. Autom. Cont. **54**(11), 2586–2597 (2009)
6. B. Bollobas, *Combinatorics, Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability*. Number 101. (Cambridge University Press, July 1987)
7. T. Carletti, D. Fanelli, R. Lambiotte, Random walks and community detection in hypergraphs. arXiv preprint arXiv:2010.14355 (2020)
8. G.F. de Arruda, G. Petri, Y. Moreno, Social contagion models on hypergraphs (2019)
9. J.-C. Delvenne, S.N. Yaliraki, M. Barahona, Stability of graph communities across time scales. Proc. Nat. Acad. Sci. **107**(29), 12755–12760 (2010)
10. C. Giusti, E. Pastalkova, C. Curto, V. Itskov, Clique topology reveals intrinsic geometric structure in neural correlations. Proc. Nat. Acad. Sci. **112**(44), 13455–13460 (2015)
11. I. Iacopini, G. Petri, A. Barrat, V. Latora, Simplicial models of social contagion. Nat. Commun. **10**(1), 1–9 (2019)
12. N.F. Johnson, R.M. Leahy, N.J. Restrepo, N. Velasquez, M. Zheng, P. Viña Manrique. Hidden resilience and adaptive dynamics of the global online hate ecology. Nature (2019)
13. R. Lambiotte, J.-C. Delvenne, M. Barahona, R. Walks, Markov processes and the multiscale modular organization of complex networks. IEEE Trans. Netw. Sci. Eng. **1**(2), 76–90 (2014)
14. R. Lambiotte, S. Redner, Dynamics of non-conservative voters. EPL (Europhysics Letters) **82**(1), 18007 (2008)
15. N. Masuda, M.A. Porter, R. Lambiotte, Random walks and diffusion on networks. Phys. Rep. **716–717**, 1–58 (2017)
16. P. Masulli, A. Villa. Algebro-topological invariants in network theory, **09** 2015
17. M. McPherson, L. Smith-Lovin, J.M. Cook, Birds of a feather: Homophily in social networks. Ann. Rev. Sociol. **27**(1), 415–444 (2001)
18. A. Mellor, M. Mobilia, R.K.P. Zia, Heterogeneous out-of-equilibrium nonlinear q-voter model with zealotry. Phys. Rev. E **95**(1), 012104 (2017)
19. R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks. Science **298**(5594), 824–827 (2002)

20. J. Molofsky, R. Durrett, J. Dushoff, D. Griffeath, S. Levin, Local frequency dependence and global coexistence. Theoret. Population Biol. **55**(3), 270–282 (1999)
21. A. Muhammad, M. Egerstedt, Control using higher order laplacians in network topologies, in *Proceeding of 17th International Symposium on Mathematical Theory of Networks and Systems, Kyoto*, pp. 1024–1038 (2006)
22. S. Mukherjee, J. Steenbergen, Random walks on simplicial complexes and harmonics. Random Struct. Algor. **49**(2), 379–405 (2016)
23. J.A. Nelder, R.W.M. Wedderburn. Generalized linear models. J. Royal Statis. Soc. Series A (General), **135**(3), 370 (1972)
24. L. Neuhäuser, A. Mellor, R. Lambiotte, Multibody interactions and nonlinear consensus dynamics on networked systems. Phys. Rev. E, **101**(3), 032310 (2020). Publisher: American Physical Society
25. L. Neuhäuser, M.T. Schaub, A. Mellor, R. Lambiotte, Opinion dynamics with multi-body interactions, in *Network Games, Control and Optimization, Communications in Computer and Information Science*, ed. by S. Lasaulce, P. Mertikopoulos, and A. Orda (Springer International Publishing, Cham, 2021), pp. 261–271
26. R. Olfati-Saber, J. Alex Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems. Proc. IEEE **95**(1), 215–233 (2007)
27. R. Olfati-Saber, J.A. Fax, R.M. Murray, Consensus and cooperation in networked multi-agent systems. Proc. IEEE **95**(1), 215–233 (2007)
28. O. Parzanchevski, R. Rosenthal, Simplicial complexes: spectrum, homology and random walks. Random Struct. Algor. **50**(2), 225–261 (2017). ((arXiv: 1211.6775))
29. A. Patania, G. Petri, F. Vaccarino, The shape of collaborations. EPJ Data Sci. **6**(1), 18 (2017)
30. G. Petri, A. Barrat, Simplicial activity driven model. Phys. Rev. Lett. **121**(22), 228301 (2018)
31. M.W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, H. Markram, Cliques of neurons bound into cavities provide a missing link between structure and function. Front. Comput. Neurosc. **11**, 48 (2017)
32. M. Rosvall, D. Axelsson, C.T. Bergstrom, The map equation. Eur. Phys. J. Spec. Top. **178**(1), 13–23 (2009)
33. R. Sahasrabuddhe, L. Neuhäuser, R. Lambiotte, Modelling non-linear consensus dynamics on hypergraphs. J. Phys.: Complex. (2020)
34. F.A.N. Santos, E.P. Raposo, M.D. Coutinho-Filho, M. Copelli, C.J. Stam, L. Douw, Topological phase transitions in functional brain networks. Phys. Rev. E, **100**(3), 032414 (September 2019). Publisher: American Physical Society
35. M.T. Schaub, A.R. Benson, P. Horn, G. Lippner, A. Jadbabaie, Random walks on simplicial complexes and the normalized hodge laplacian. CoRR, abs/1807.05044 (2018)
36. M.T. Schaub, J.-C. Delvenne, S.N. Yaliraki, M. Barahona, Markov dynamics as a zooming lens for multiscale community detection: non clique-like communities and the field-of-view limit. PLoS ONE **7**(2), e32210 (2012)
37. V. Srivastava, J. Moehlis, F. Bullo, On bifurcations in nonlinear consensus networks. J. Nonlinear Sci. **21**(6), 875–895 (2011)
38. C.-I. Vasile, M. Schwager, C. Belta, *SE(N) invariance in networked systems, in 2015 European Control Conference (ECC)* (Linz, Austria, IEEE, 2015), pp. 186–191
39. D.J. Watts, A simple model of global cascades on random networks. Proc. Nat. Acad. Sci. **99**(9), 5766–5771 (2002)

Chapter 15
Collective Games on Hypergraphs

Unai Alvarez-Rodriguez, Federico Battiston, Guilherme Ferraz de Arruda, Yamir Moreno, Matjaž Perc, and Vito Latora

Abstract Human activities often require simultaneous decision-making of individuals in groups. These processes cannot be coherently addressed by means of networks, as networks only allow for pairwise interactions. Here, we propose a general implementation for collective games in which higher-order interactions are encoded on hypergraphs. We employ it for the study of the public goods game by first validating the analytical expression of the replicator dynamics in uniform and heterogeneous populations, and then by introducing a procedure for retrieving empirical synergistic effects of group interactions from real datasets.

U. Alvarez-Rodriguez (✉)
Data Analytics Group, University of Zurich, Zurich, Switzerland

F. Battiston
Department of Network and Data Science, Central European University, Vienna, Austria

Department of Anthropology, University of Zurich, Zurich, Switzerland

G. Ferraz de Arruda · Y. Moreno
ISI Foundation, Turin, Italy

Y. Moreno
Institute for Biocomputation and Physics of Complex Systems, University of Zaragoza, Zaragoza 50008, Spain

Department of Theoretical Physics, University of Zaragoza, Zaragoza 50009, Spain

M. Perc
Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia

Department of Medical Research, China Medical University Hospital, China Medical University, Taichung, Taiwan

Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria

V. Latora
School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK

Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, 95123 Catania, Italy

The Alan Turing Institute, The British Library, London NW1 2DB, UK

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_15

15.1 Introduction

Cooperation has been detected to be a key element in the explanation of the evolutionary success of our species [1, 2]. The problem of understanding the emergence of cooperation lies at the boundary of a plethora of scientific disciplines [3–9]. Previous works have been able to explain how cooperation in a population is sustained by the structure of the social network encoding the interactions amongst individuals [10, 11]. Evolutionary game theory is the branch of applied mathematics providing a theoretical framework to address these questions, enabling quantitative statements about the conditions that give rise to stable cooperation [12–14]. In this context, a social dilemma is a scenario in which defection results in a higher payoff for individuals while cooperation entails a higher payoff for the collective of players [15]. The challenge presented by social dilemmas is solved by network reciprocity [16], a property by which groups of nodes are strongly connected between them and weakly connected to nodes outside the group, and thus protected from defector invasion. This feature may be observed in structures of different nature, such as networks with a heterogeneous degree distribution [17–19], networks with community structures [20, 21], and even in multilayered systems [22–29].

These advances in evolutionary game theory have been restricted to the realm of two-player games, as standard networks are not suitable for encoding group interactions. In order to bypass this limitation, it was proposed to infer higher-order interactions from the dyadic structure, by assuming that every node would act as a vehicle for a group-game involving all of his neighbours [30, 31]. However, such approach is structurally ambiguous, and thus incompatible with the well known mechanisms favouring cooperation [32–39]. To overcome the constraints of traditional graphs, higher-order interactions have been suggested as the natural way to encode non-diadic relationships [40, 41]. In particular, it has recently been shown that hypergraphs are a natural solution to formalise n-player games [42]. In the following we elaborate on this idea by explaining how to implement n-player games on hypergraphs and applying this to the study of the public goods game (PGG) [43, 44] for uniform and heterogeneous structures.

The PGG is an n-player game of two strategies where at each round of the game participants are requested to contribute to a common pool with a token of value c. We shall call cooperators C those players who do contribute and defectors D those that do not contribute. The collected amount is multiplied by the synergy factor R and the outcome is equally split between all the participants [8]. It is standard to assign a fixed value of $c = 1$ to the token, and to describe the payoffs in terms of the reduced synergy factor $r = R/g$ where g is the number of players. In a round with w_C cooperators the payoff for the defectors is $\pi_D = rw_C$, while the payoff of the cooperators is given by $\pi_C = rw_C - 1$.

15.2 Collective Games on Higher-Order Networks

The goal of evolutionary dynamics is to predict the number of cooperators and defectors in a population undergoing a continuous iteration of the game combined with the adaptation of strategies. A game implementation is a set of rules that determines how this process occurs. In the hypergraph implementation (HI) each round of the PGG is hosted by a hyperlink $l \in \mathcal{L}$ of the hypergraph $H(\mathcal{N}, \mathcal{L})$ representing the system. The evolutionary process is a concatenation of micro-steps: At each micro-step a hyperlink $l \in \mathcal{L}$ and one of its nodes $n \in l$ are randomly selected. All the nodes present in the hyperlink play a round of the game for each hyperlink they are part of and accumulate their payoffs. The payoffs are then normalised and compared to select the node with the highest payoff per game ratio. Only then node n copies the winning strategy with a probability proportional to the payoff difference $\frac{1}{\Delta}[(\max_l \pi_j) - \pi_i]$. Here Δ is the normalisation factor that accounts for the maximal payoff difference between two nodes. For a system with $|\mathcal{N}| = N$ nodes, N micro-steps add up to a step, in which every node has the opportunity to change its strategy at least once. In Fig. 15.1 we graphically explain HI, and we compare it with the original network implementation [30] or graph implementation (GI).

In the next subsections we analyse in depth the hypergraph implementation on different families of hypergraphs. Thorough this chapter we adopt the fixed cost per game perspective, where players contribute with a token to every round they play.

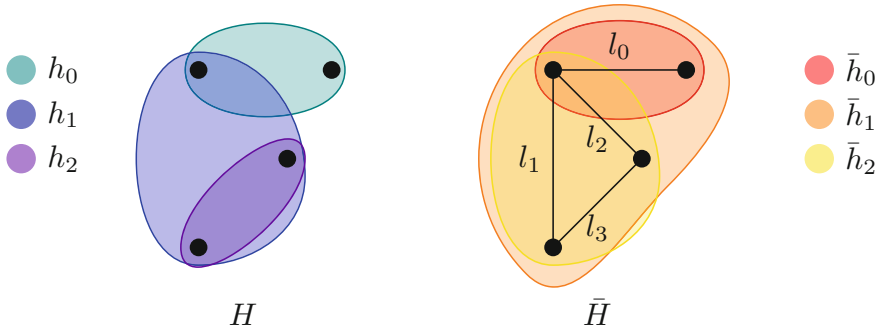


Fig. 15.1 Collective games on Hypergraphs. We graphically argument that the hypergraph implementation provides a reliable alternative to the ambiguous graph implementation for collective games. The hypergraph on the left H accounts for the real higher-order connections represented by hyperlinks h_i. The network with links l_i is inferred by linking all the nodes that are part of a common hyperlink in the original structure. The hypergraph on the right $\bar{H}$ with hyperlinks $\bar{h}_i$ is obtained as a product of the graph implementation, which imposes a group interaction between a node and its first neighbours. The difference between the real groups h_i and the ones imposed by the graph implementation $\bar{h}_i$ shows the inconsistency of the dyadic approach. Adapter from Ref. [42]

15.2.1 Uniform Hypergraphs

Uniform Hypergraphs are a subset of all the possible hypergraphs in which all the hyperlinks have the same cardinality. This means in our context that all the rounds of the PGG will have the same number of players g. The system may be described by using the replicators approximation, in which the relevant properties of the evolutionary dynamics are captured by the fraction of population selecting each strategy. We use x_C and x_D for the fraction of cooperators and defectors respectively. We first compute the average payoffs for cooperators π_C and defectors π_D for a generic order g by counting all the possible configurations of strategies

$$\pi_D = \sum_{i=0}^{g-1} \binom{g-1}{i} x_D^{g-1-i} x_C^i ir$$
$$\pi_C = \sum_{i=0}^{g-1} \binom{g-1}{i} x_D^{g-1-i} x_C^i ((i+1)r - 1)$$

and obtain the average payoff difference as $\pi_D - \pi_C = 1 - r$. We also compute Δ, the maximal payoff difference

$$\Delta = \begin{cases} r(g-2)+1 & \text{if } r < 1 \\ gr - 1 & \text{if } r > 1 \end{cases}$$

The time evolution equation can be derived by counting all the possible combinations leading to a strategy change: for every group in which at least two strategies are present one has to consider the probabilities that a cooperator defector pair is involved in a potential strategy change combined with the probability of the strategy change actually occurring given the payoff difference. Although the equation for cooperators is not presented here, its formulation is analogue to that of defectors.

$$\dot{x}_D = \sum_{i=0}^{g-2} \binom{g}{1+i} x_D^{g-1-i} x_C^{1+i} \frac{(g-1-i)(1+i)}{g(g-1)} Q\Big[\theta(\pi_C - \pi_D) + \theta(\pi_D - \pi_C)\Big]$$
$$= Q x_D x_C$$

where Q is the normalized payoff difference, $Q \equiv (\pi_D - \pi_C)/\Delta$. From this equation we observe that the dynamics has two absorbing states $x_D = 0, x_C = 1$ and $x_D = 1, x_C = 0$ and a phase transition at $r = 1$. Therefore one should expect cooperators emerging in uniform hypergraphs only if $R > g$ holds.

We test the replicators prediction on synthetically designed uniform random hypergraphs (URH). For a fixed order g, these hypergraphs are composed of L independent hyperlinks created by randomly selecting L g-tuples of different nodes. We run $T = 10^4$ steps of the game in a system with $N = 1000$ nodes for $g = 2, 3, 4, 5$. The number of hyperlinks L is tuned to exceed the critical threshold guaranteeing a

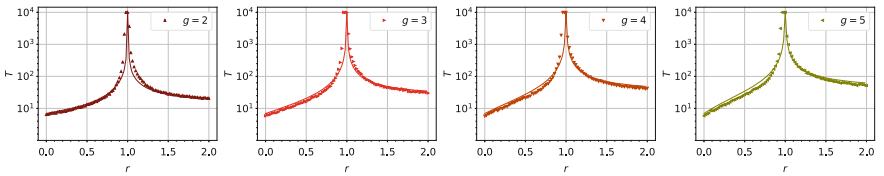


Fig. 15.2 Uniform Hypergraphs. Relaxation time as a function of the reduced synergy factor on uniform random hypergraphs with $N = 1000$ nodes and $L = 5L_c$ hyperlinks for groups of different sizes g. The triangles report the numerical simulations and the continuous line accounts for the replicators prediction. Adapted from Ref. [42]

connected hypergraph L_c, $L = 5L_c$, with $L_c = \frac{N}{g} \ln N$. The plot in Fig. 15.2 reports the relaxation time for an initial population in which cooperators and defectors are evenly distributed. A good agreement between the theory and the numerical simulations is observed.

The replicators equation is derived by assuming that all nodes are indistinguishable and potentially connected to each other, and therefore an increasingly alike behaviour to that of the replicator is expected when density is increased in URHs. However, real-world scenarios seldom provide these conditions, as structures emerging from optimisation processes tend to display strong heterogeneities and low densities. Hence, further analysis of the limits of the replicators approximation is needed to understand its applicability range.

Let us start by introducing the hyperdegree of a node as an additional degree of freedom in our model. We have $p(k)$ for the probability of a node having hyperdegree k, and $p(D|k)$ or $p(C|k)$ for the probabilities that a node with hyperdegree k is either a defector or a cooperator. We can recover the fraction of defectors and cooperators by adding the contributions from all the possible hyperdegrees $\mathcal{K}$.

$$\begin{aligned} x_D &= \sum_{k \in \mathcal{K}} p(k) p(D|k) \\ x_C &= \sum_{k \in \mathcal{K}} p(k) p(C|k) \end{aligned} \tag{15.1}$$

From the rest of this section we will use p_{Dk} and p_{Ck} to lighten the notation. Following the procedure explained in Eq. (15.1) one may derive the time evolution of the hyperdegree restricted variables by adding all the possible channels that lead to a strategy change. In essence, the population of defectors with degree k can only increase if a cooperator of degree k changes its strategy, or if a defector of degree k becomes a cooperator. Mathematically this can be expressed as

$$\dot{x}_{Dk} = -x_{Dk} \sum_{k' \in \mathcal{K}} p(k'|k) p(C|k') \sum_{k'' \in \mathcal{K}^{g-2}} p(k''|kk') \sum_{x \in \mathcal{X}^{g-2}} p(x|k'') \times \frac{(\pi_{Ck'} - \pi_{Dk})\theta(\pi_{Ck'} - \pi_{Dk})}{\Delta} + x_{Ck} \sum_{k' \in \mathcal{K}} p(k'|k) p(D|k') \sum_{k'' \in \mathcal{K}^{g-2}} p(k''|kk') \sum_{x \in \mathcal{X}^{g-2}} p(x|k'') \times \frac{(\pi_{Dk'} - \pi_{Ck})\theta(\pi_{Dk'} - \pi_{Ck})}{\Delta} \tag{15.2}$$

While different, the terms associated to each of the channels respect the same principle: In the first summation we consider all the possibilities of a neighbouring node being a cooperator with a given degree k'. In the second summation we account for the hyperdegrees of the rest of $g-2$ nodes in the group, conditioned to the hyperdegrees of the defector-cooperator pair. In the third summation we include all the possible strategy selections by these nodes. The last element in the product includes the normalized average payoff difference between a defector with degree k and a cooperator with degree k'. We shall compute these to move forward. We have

$$\pi_{Dk} = \sum_{k'' \in \mathcal{K}^{g-1}} p(k''|k) \sum_{x'' \in \mathcal{X}^{g-1}} p(x''|k'')(nr)$$
$$\pi_{Ck'} = \sum_{k'' \in \mathcal{K}^{g-1}} p(k''|k') \sum_{x'' \in \mathcal{X}^{g-1}} p(x''|k'')((n+1)r - 1)$$

where n is a function of x'' accounting for the number of cooperators. The first term in each expression, $p(k''|k)$ and $p(k''|k')$ represent the hyperdegree-hyperdegree correlations, i.e., how likely is that a node with hyperdegree k or k' is part of a group of $g-1$ with a given combination of hyperdegrees. We notice that if we make these probabilities node independent, i.e., $p(k''|k) = p(k''|k') = p(k'')$ we can simplify the expression for the average payoff difference, and recover the result of the original replicators approach $\pi_{Dk} - \pi_{Ck'} = 1 - r$. By introducing this result on Eq. (15.2) one obtains the more simplified

$$\dot{x}_{Dk} = Q\theta(r-1)x_{Dk}x_C + Q\theta(1-r)x_{Ck}x_D \tag{15.3}$$

where Q is the normalized average payoff difference, and the absence of hyperdegree-hyperdegree correlations is used again. This expression is then combined with Eq. (15.1) to yield a final formula for the time evolution of cooperators and defectors that exactly coincides with the one derived above without considering the hyperdegrees.

$$\dot{x}_D = Q\theta(r-1)x_Dx_C + Q\theta(1-r)x_Cx_D = Qx_Dx_C$$

This result establishes a precise boundary between the hypergraphs that are suited to be described with a replicators approach and those that are not. We point out that these derivations corroborate the intuition behind the indistinguishability of the

nodes, as the absence of hyperdegree-hyperdegree correlations implies that not only the nodes but also their neighbourhoods are equal.

15.2.2 Heterogeneous Hypergraphs

Heterogeneous Hypergraphs are the next step in the path towards the application of HI to more realistic scenarios. In these, hyperlink orders are not fixed, and therefore effects arising from the mixture of group sizes are expected. A particularly interesting question is to understand the phenomenology driven by a group-size dependent synergy factor. We start by providing a description for heterogeneous hypergraphs in terms of the abundance of hyperlinks of order g. We say that $\mathbf{p} = \{p^g\}_{g=g^-}^{g^+}$ of elements $p^g = k^g/k$ contains the fraction of hyperlinks of order g one node is part of, where g^- and g^+ are the minimal and maximal orders respectively, and $g \in \mathcal{G} = \{g^-, g^- + 1, ..., g^+ - 1, g^+\}$. We are interested in describing the dynamics for synergy factors modelled by non-linear functions of the group size, $R(g) = \alpha g^\beta$ with $\alpha, \beta \geq 0$. Given that g takes only discrete values, we find more convenient to work with $\mathbf{r} = \{r^g\}_{g=g^-}^{g^+}$ of elements $r^g = \alpha g^{\beta-1}$. With $\beta = 1$, we would factor out the dependence of g in the reduced synergy factor, and therefore recover the uniform case. The average payoff difference can be computed by adding the contribution of all the group sizes

$$\pi_D - \pi_C = \sum_{g \in \mathcal{G}} p^g (1 - r^g) = 1 - \alpha \sum_{g \in \mathcal{G}} p^g g^{\beta-1}$$

We use Q again as $Q = (\pi_D - \pi_C)/\Delta$ to represent the normalized payoff difference, in terms of which Eq. (15.1) is obtained for the dynamics. Therefore, the condition $Q = 0$ yields the critical value α_c

$$\alpha_c = \frac{1}{\sum_{g \in \mathcal{G}} p^g g^{\beta-1}}$$

We validate our predictions on a series on numerical experiments with synthetic heterogeneous hypergraphs. We consider hypergraphs whose $\mathbf{p}$ is restricted to $p^g = n^g/4$ with $n^g \in \{0, 1, 2, 3, 4\} \forall g \in \mathcal{G}$ with $g^- = 2$ and $g^+ = 5$. We construct a hypergraph with $N = 1000$ and $L = 2L_c$ for each of the possible values of $\mathbf{p}$ fulfilling the aforementioned condition, and collect all of them in an hypergraph ensemble $\mathcal{H}$ with $|\mathcal{H}| = 35$. We then run $T = 10^4$ steps of the evolutionary dynamics and obtain the asymptotic fraction of cooperators as a function of α for $\beta \in \{0, 1, 2, 3\}$ for all the hypergraphs in $\mathcal{H}$. The simulations in Fig. 15.3 (from [42]) display a good agreement between our prediction of the critical point and the empirical transition between cooperators and defectors.

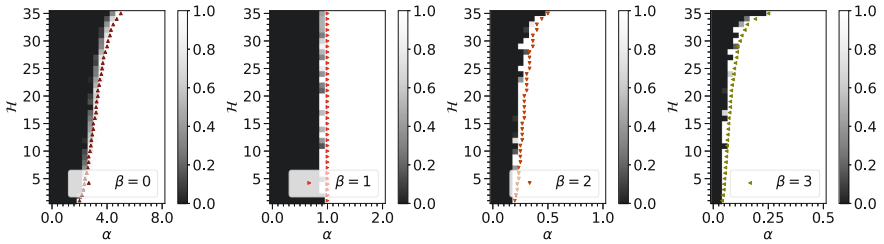


Fig. 15.3 Heterogeneous Hypergraphs. Fraction of cooperators x_C after $T = 10^4$ time steps as a function of α for each hypergraph in $\mathcal{H}$ with $N = 1000$ nodes. The panels correspond to values of $\beta = 0, 1, 2, 3$ from left to right. The coloured triangles mark the analytical phase transition yield by the replicators approximation. Adapted from Ref. [42]

15.2.3 Synergy Factor on Real Games

The advances in uniform and heterogeneous hypergraphs pave the way towards the application of evolutionary dynamics to explain real-life systems. In particular, a challenging question is to understand how group size influences the performance of teams of cooperating individuals. In this section we introduce a procedure based on a series of assumptions to extract the synergy factor from datasets of interacting individuals and apply it to study the bibliographic dataset of the American Physical Society (APS).

Our technique is grounded on the hypothesis that the structure of the hypergraph is the outcome of an optimisation process, in which the players have selected their connections to maximise their payoff. Therefore, one should expect a one to one correlation between the hyperdegree distribution $\mathbf{p}$ and the group size dependent synergy factor r^g. Based on this idea, we argue that the system has to be constrained by two conditions: $\mathbf{r}$ has to be proportional to $\mathbf{p}$ and the system has to be at equilibrium, and thus defectors and cooperators have the same average payoff. The combination of $r^g = zp^g$ and $\sum_{g\in\mathcal{G}} p^g(1 - r^g) = 0$ yields

$$r^g = \frac{p^g}{\sum_{g\in\mathcal{G}}(p^g)^2} \tag{15.4}$$

In a random graph with no hyperdegree heterogeneities the average fraction of hyperdegrees $p^g = k^g/k$ can be obtained from the total number of hyperlinks at each order as $k^g = gL^g/N$ as long as hyperlinks are uniformly distributed. Under these conditions the synergy factor can be extracted as a function of the total number of groups at each order L^g. We employ this technique to retrieve the synergy factor of 13 different APS journals with a total of 577886 papers. Authors and papers are respectively represented as nodes and hyperlinks of 13 different hypergraphs from where L^g is measured as the total number of papers produced by a given number of authors.

We have now an algorithm for detecting the synergy factor, but we are also interested in explaining its origin. With that goal in mind we propose to model the group

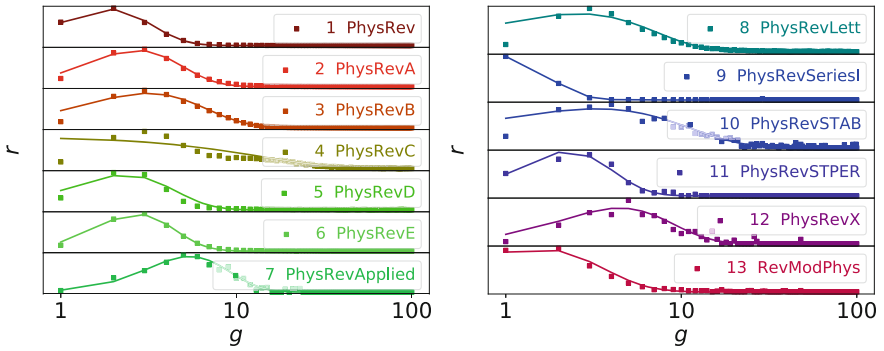


Fig. 15.4 Synergy factor on real games. We show the reduced synergy factor as a function of the group size extracted from the bibliographic data of the APS for a selection of 13 journals. The dots represent the experimental values while the continuous line corresponds to the model at Eq. (15.5) explaining the synergy factor as the combination of positive and negative group effects. Adapted from Ref. [42]

size dependence of the synergy factor as the product of two opposite contributions, a first one increasing with g, and a second one decreasing with g

$$f(g, \alpha, \beta, \gamma) = \alpha g^{\beta} e^{-\gamma(g-1)} \tag{15.5}$$

From these three parameters α has a fixed value given by normalisation, and therefore we are left with β and γ, both larger than zero. Due to their functional dependence on g, β accounts for the positive aspects of group interactions while γ represents an exponential dumping and therefore is associated to lower synergy factors in larger groups. For the particular dataset we are studying, these opposite terms have specific meanings in the production of scientific manuscripts: One could associate the benefits (β) with a multiplication in the amount and depth of ideas and discussions preceding the manuscript preparation when increasing the group size. Analogously, the costs (γ) may be associated with difficulties in coordination when recruiting additional authors for a paper. This interpretation is also compatible with the shape of r^g, whose maximum is predicted to be at β/γ.

In Fig. 15.4 (from [42]) the empirically derived profiles of the reduced synergy factor for the APS dataset are shown, as well as the curve that better fits such profile according to the model in Eq. (15.5). The parameters are reported in Table 15.1.

15.3 Discussion

In this chapter we have presented the formalism introduced in [42] for studying the evolutionary dynamics of systems with explicit higher-order interactions. We have derived the replicators approximation and showed that it successfully accounts

Table 15.1 Synergy factor on real games. Parameters of the study carried out for the APS bibliographic dataset. For each journal we indicate the total number of hyperlinks L, the average order $< g >$, the order with the highest synergy factor $g(\max r)$, the benefit and cost parameters β, γ and the distance between the approximation by Eq. (15.5) and the data d_r

Journal	L	$< g >$	$g(\max r)$	β	γ	d_r
PhysRev	47313	1.95	2	2.936	1.573	0.033
PhysRevA	70502	3.07	3	2.679	0.986	0.067
PhysRevB	171268	3.75	3	1.531	0.49	0.05
PhysRevC	36290	5.98	3	0.02	0.075	0.146
PhysRevD	74715	3.02	2	2.178	0.941	0.206
PhysRevE	50988	2.93	3	3.84	1.41	0.048
PhysRevApplied	327	5.39	5	3.356	0.62	0.09
PhysRevLett	113674	4.57	3	0.848	0.33	0.175
PhysRevSeriesI	1240	1.21	1	2.691	2.831	0.019
PhysRevSTAB	2393	5.52	4	0.566	0.173	0.127
PhysRevSTPER	484	2.42	3	2.75	1.21	0.078
PhysRevX	611	5.28	5	1.85	0.416	0.127
RevModPhys	3153	2.05	2	1.19	0.79	0.112

for the system's dynamics in uniform and heterogeneous hypergraphs as long as no hyperdegree-hyperdegree correlations are present. We have then discussed a proposal for extracting the synergy factor of real games and apply it to the analysis of the bibliographic dataset of the APS.

This new framework for higher-order interactions calls for a hypergraph adaptation of additional game characteristics that complement the PGG to facilitate the emergence of cooperation, such as image scoring [34–36], rewarding [45], and punishment [46–49]. Similarly, the hypergraph implementation motivates new research in the direction of understanding the influence of more complex structures, such us communities or multilayer organization, which were well characterized for games in standard networks [21–29]

All these assets will surely strengthen the already consistent and reliable hypergraph implementation of evolutionary dynamics for modelling the emergence of cooperation.

Acknowledgements U.A.-R. acknowledges support from the Swiss National Science Foundation through grant 176938. F.B. acknowledges partial support from the ERC Synergy Grant 810115 (DYNASNET). V. L. acknowledges support from the Leverhulme Trust Research Fellowship "CREATE: the network components of creativity and success". Y. M. acknowledges partial support from the Government of Aragón and FEDER funds, Spain through grant E36-20R to FENOL, by MINECO and FEDER funds (grant FIS2017-87519-P) and from Intesa Sanpaolo Innovation Center. M. P. was supported by the Slovenian Research Agency (Grant Nos. J1-2457 and P1-0403). G. F. A. acknowledges support from Intesa Sanpaolo Innovation Center. We thank M. Clarin from COSNET Lab for help and assistance with the figures. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. M.A. Nowak, R. Highfield, *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed* (Free Press, New York, 2011)
2. S.B. Hrdy, *Mothers and Others: The Evolutionary Origins of Mutual Understanding* (Harvard University Press, Cambridge, MA, 2011)
3. J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, Am. Econ. Rev. **91**, 73 (2001)
4. M.A. Nowak, Science **314**, 1560 (2006)
5. N. Henrich, J.P. Henrich, *Why Humans Cooperate: A Cultural and Evolutionary Explanation* (Oxford University Press, Oxford, UK, 2007)
6. D.G. Rand, M.A. Nowak, Trends Cogn. Sci. **17**, 413 (2013)
7. G. Kraft-Todd, E. Yoeli, S. Bhanot, D. Rand, Curr. Opinion Behav. Sci. **3**, 96 (2015)
8. M. Perc, J.J. Jordan, D.G. Rand, Z. Wang, S. Boccaletti, A. Szolnoki, Phys. Rep. **687**, 1 (2017)
9. M.O. Jackson, Y. Zenou, *Economic Analyses of Social Networks* (Edward Elgar Publishing, Cheltenham, 2013)
10. F.C. Santos, J.F. Rodrigues, J.M. Pacheco, Proc. R. Soc. B **273**, 51 (2006)
11. D.G. Rand, S. Arbesman, N.A. Christakis, Proc. Natl. Acad. Sci. USA **108**, 19193 (2011)
12. J.W. Weibull, *Evolutionary Game Theory* (MIT Press, Cambridge, MA, 1995)
13. J. Hofbauer, K. Sigmund, *Evolutionary Games and Population Dynamics* (Cambridge University Press, Cambridge, UK, 1998)
14. M.A. Nowak, *Evolutionary Dynamics* (Harvard University Press, Cambridge, MA, 2006)
15. R. Axelrod, *The Evolution of Cooperation* (Basic Books, New York, 1984)
16. M.A. Nowak, R.M. May, Nature **359**, 826 (1992)
17. F.C. Santos, J.M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005)
18. F.C. Santos, J.M. Pacheco, T. Lenaerts, Proc. Natl. Acad. Sci. USA **103**, 3490 (2006)
19. J. Gómez-Gardeñes, M. Campillo, L.M. Floría, Y. Moreno, Phys. Rev. Lett. **98**, 108103 (2007)
20. C.E. Tarnita, T. Antal, H. Ohtsuki, M.A. Nowak, Proc. Natl. Acad. Sci. USA **106**, 8601 (2009)
21. B. Fotouhi, N. Momeni, B. Allen, M.A. Nowak, J.R. Soc, Interface **16**, 20180677 (2019)
22. Z. Wang, A. Szolnoki, M. Perc, EPL **97**, 48001 (2012)
23. J. Gómez-Gardeñes, I. Reinares, A. Arenas, L.M. Floría, Sci. Rep. **2**, 620 (2012)
24. J. Gómez-Gardeñes, C. Gracia-Lázaro, L.M. Floría, Y. Moreno, Phys. Rev. E **86**, 056113 (2012)
25. Z. Wang, A. Szolnoki, M. Perc, Sci. Rep. **3**, 1183 (2013)
26. Z. Wang, L. Wang, M. Perc, Phys. Rev. E **89**, 052813 (2014)
27. F. Battiston, M. Perc, V. Latora, New J. Phys. **19**, 073017 (2017)
28. F. Fu, X. Chen, New J. Phys. **19**, 071002 (2017)
29. B. Fotouhi, N. Momeni, B. Allen, M.A. Nowak, Nat. Human Behav. **2**, 492 (2018)
30. F.C. Santos, M.D. Santos, J.M. Pacheco, Nature **454**, 213 (2008)
31. A. Szolnoki, M. Perc, G. Szabó, Phys. Rev. E **80**, 056109 (2009)
32. R.L. Trivers, Q. Rev. Biol. **46**, 35 (1971)
33. K. Sigmund, Trends Ecol. Evol. **22**, 593 (2007)
34. M.A. Nowak, K. Sigmund, Nature **393**, 573 (1998)
35. M. Milinski, D. Semmann, T.C.M. Bakker, H.J. Krambeck, Proc. R. Soc. Lond. B **268**, 2495 (2001)
36. H.H. Nax, M. Perc, A. Szolnoki, D. Helbing, Sci. Rep. **5**, 12145 (2015)
37. E. Fehr, Nature **432**, 449 (2004)
38. S. Gächter, Scand. J. Econ. **104**, 1 (2002)
39. F. Fu, C. Hauert, M.A. Nowak, L. Wang, Phys. Rev. E **78**, 026117 (2008)
40. F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.G. Young, G. Petri, Phys. Rep. **874**, 1 (2020)
41. F. Battiston, E. Amico, A. Barrat, G. Bianconi, G.F. de Arruda, B. Franceschiello, I. Iacopini, S. Kefi, V. Latora, Y. Moreno, M. Murray, T. Peixoto, F. Vaccarino, G. Petri, Nat. Phys. **17** (10), 1093–1098 (2021)

42. U. Alvarez-Rodriguez, F. Battiston, G.F. de Arruda, Y. Moreno, M. Perc, V. Latora, Nat. Human Behav. **5**(5), 586 (2021)
43. M. Archetti, I. Scheuring, J. Theor. Biol. **299**, 9 (2012)
44. M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L.M. Floría, Y. Moreno, J.R. Soc, Interface **10**, 20120997 (2013)
45. D.G. Rand, A. Dreber, T. Ellingsen, D. Fudenberg, M.A. Nowak, Science **325**, 1272 (2009)
46. J. Andreoni, W. Harbaugh, L. Vesterlund, Am. Econ. Rev. **93**, 893 (2003)
47. S. Gächter, E. Renner, M. Sefton, Science **322**, 1510 (2008)
48. R. Boyd, H. Gintis, S. Bowles, Science **328**, 617 (2010)
49. J.J. Jordan, M. Hoffman, P. Bloom, D.G. Rand, Nature **530**, 473 (2016)

Chapter 16
Topological Data Analysis of Spatial Systems

Michelle Feng, Abigail Hickok, and Mason A. Porter

Abstract In this chapter, we discuss applications of topological data analysis (TDA) to spatial systems. We briefly review a recently proposed level-set construction of filtered simplicial complexes, and we then examine persistent homology in two cases studies: street networks in Shanghai and anomalies in the spread of COVID-19 infections. We then summarize our results and provide an outlook on TDA in spatial systems.

16.1 Introduction

To improve our understanding of spatial systems, it is important to develop methods that directly probe the effects of space on their structure and dynamics. Many complex systems have a natural embedding in a low-dimensional space or are otherwise influenced by space [1, 2]. Spatial effects significantly influence both their structure and their dynamics.

One way to gain information about the global structure of spatial systems is by studying notions of 'connectedness', 'holes', and 'cavities'. Consequently, it is not surprising that many researchers have used topological data analysis (TDA), usually in the form of persistent homology (PH), to study a diverse variety of spatial systems. For example, TDA has been used to study granular and particulate systems [3, 4], neuronal networks [5], leaf-venation patterns [6], networks of blood vessels [7], aggregation models [8], spatial percolation [9], human migration [10], and voting patterns [11].

Analyzing PH allows one to quantify holes in data in a meaningful way and has made it possible to apply homological ideas to a wide variety of empirical data sets [12]. To study PH, one needs to construct a filtered simplicial complex. (See, e.g., Chap. 3 and [13].) In [11], Feng and Porter developed new types of filtered simpli-

M. Feng
California Institute of Technology, Pasadena, CA, USA

A. Hickok · M. A. Porter (✉)
University of California, Los Angeles, Los Angeles, CA, USA
e-mail: mason@math.ucla.edu

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_16

cial complexes that incorporate spatial information. In [14], they applied their new constructions to synthetic spatial networks, city street networks, spiderwebs, and snowflakes. Other studies have also incorporated spatial information into PH (see, e.g., [7, 15, 16]). Recently, researchers have also extended TDA methods other than PH—specifically, ones that use persistent cohomology [17] and the Euler characteristic [18]—to incorporate spatial information.

In the present chapter, we discuss two case studies of PH to spatial systems. We use a recently introduced level-set construction of simplicial complexes [11] to study (1) city street networks in Shanghai[1] and (2) anomalies in the spread of COVID-19 infections. Through these examples, we illustrate the importance of incorporating spatial information when studying spatial systems using TDA.

Our chapter proceeds as follows. In Sect. 16.2, we discuss the level-set construction of filtered simplicial complexes. We use the PH of these complexes to study city street networks in Shanghai in Sect. 16.3 and anomalies in the spread of COVID-19 infections in Sect. 16.4. In Sect. 16.5, we conclude and give a brief outlook on TDA in spatial systems.

16.2 Level-Set Complexes

We now briefly review the level-set construction of filtered simplicial complexes that was introduced recently in [11]. For discussions of other types of filtered simplicial complexes (which are often called simply 'filtrations'), see Chap. 3 and [12].

In a level-set filtration, one describes data as a manifold. Let M denote a two-dimensional (2D) manifold that is embedded in $\mathbb{R}^2$, such as data in an image format. We construct a sequence

$$M_0 \subseteq M_1 \subseteq \cdots \subseteq M_n$$

of manifolds (where M_0 is an approximation of M) as follows. At each time t, we evolve the boundary Γ_t of M_t outward according to the level-set equation of front propagation [19]. Specifically, for a 2D manifold M that is embedded in $\mathbb{R}^2$, we define a function $\phi(\mathbf{x}, t) : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$, where $\phi(\mathbf{x}, t)$ is the signed distance function from $\mathbf{x}$ to Γ_t at time $t \geq 0$. We propagate Γ_t outward at velocity v using the partial differential equation

$$\frac{\partial \phi}{\partial t} = v|\nabla \phi| \tag{16.1}$$

until all homological features die. The evolution (16.1) gives a signed distance function at each time t. We take M_t to be the set of points $\mathbf{x}$ such that $\phi(\mathbf{x}, t) > 0$. (This corresponds to points inside the boundary Γ_t.) In our examples in this chapter, we use $v = 1$.

[1] This case study is related to an example in [14].

By imposing $\{M_i\}$ over a triangular grid of points (see [11]), we obtain a corresponding simplicial complex X_i for each M_i. Because the level-set equation (16.1) evolves outward, we satisfy the condition that $X_i \subseteq X_{i+1}$ for all i, so $\{X_i\}$ is a filtered simplicial complex.

16.3 Case Study: Street Networks in Shanghai

This case study is closely related to one of the examples in [14]. In this case study, we use level-set complexes to examine patterns in city street networks. We focus on the city of Shanghai, which has a long history of urban development [20]. The discussion in [14] used the PH of level-set complexes to classify a variety of small street networks from different neighborhoods of Shanghai. In the present discussion, we closely examine the PH of street networks in several different neighborhoods of Shanghai. Computing PH (and, more generally, using TDA) allows us to detect both topological and geometric properties of city blocks in these neighborhoods. These properties may reflect differences in the development of city streets across time and cultural influences.

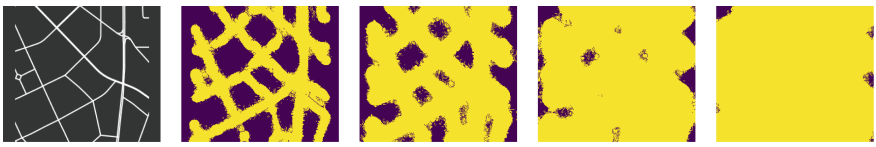
The city of Shanghai was first inhabited about 6000 years ago during China's Warring States period. Over the course of several millennia, Shanghai has experienced urban growth, with a variety of developmental styles, over many distinct time periods [20]. These different architectural and urban-planning styles reflect a diversity of different views by the various powers of Shanghai for how the city should be structured. In the following paragraphs, we use PH to highlight street networks in several distinct neighborhoods of Shanghai. We draw connections between the histories of these neighborhoods and the topological features that we observe in their PH.

We use networks from OSMnx [21] as input data. Our street networks are images of street maps; they consist of a 2 km by 2 km square that is centered at a given set of (latitude, longitude) coordinates. We show three such street maps in Fig. 16.1. In Fig. 16.1a, we show a street map from Laoximen ('Old West Gate'), a neighborhood that was built up around the western gate of Shanghai's original city walls. In Figure 16.1b, we show a street map from the former French Concession, which was a French colonial territory from 1849 to 1943. In Fig. 16.1c, we show a street map from Pudong New Area, which is a modern financial district that has developed mostly over the last few decades.

From the street maps in Fig. 16.1, we obtain level-set complexes following the approach in Sect. 16.2. In Fig. 16.2, we show the level-set complex that corresponds to the map in Fig. 16.1c. This level-set complex begins with line segments that represent the streets in the map. The streets expand outward as we add simplices to the complex. We thereby capture city blocks as homological features, whose death times increase as the sizes of the blocks increase. (Larger blocks take longer to be 'filled' by the expanding streets in the simplicial complex.)

(a) Laoximen (b) Former French Concession (c) Pudong New Area

Fig. 16.1 Street networks from three different neighborhoods of Shanghai. (We generated these maps using OSMnx [21].)

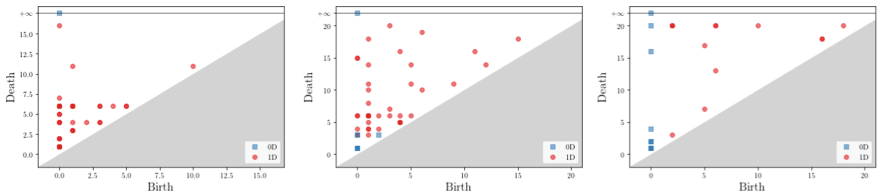


(a) (b) (c) (d) (e)

Fig. 16.2 Selected steps of the level-set evolution of the map of Pudong New Area from Fig. 16.1c. As the level-set complex evolves, the streets expand and fill in the blocks. Smaller blocks fill in faster

To visualize the features of the PH of the street-map images, we use *persistence diagrams* (PDs). PDs represent homological features as points on a scatter plot. We plot each feature at a point (b, d), where b is the birth time of the feature and d is its death time. We show zero-dimensional (0D) features as blue squares and one-dimensional (1D) features as red disks. Because features cannot die before they are born, all points must lie on or above the identity line $g(x) = x$. More persistent features lie farther above this line. See Chap. 3 and [12] for more information about PDs.

In Fig. 16.3, we show the PDs that correspond to the maps in Fig. 16.1. The PD of Laoximen (see Fig. 16.3a) reveals that most of the 1D features have death times of less than 10. This indicates that the city blocks in this area are relatively small. Additionally, although many of the features of the map of Laoximen are born at early times (such features are close to the vertical axis of a PD), there are also several points close to the diagonal that have later birth times. These points correspond to features that tend to occur when a street map has dead ends. As the level-set complex evolves, dead ends expand. This can result in the 'pinching' of a single block into multiple smaller blocks when a dead end connects to the streets that border that block. Similarly, blocks that are not rectangular because of winding roads can be 'pinched' into smaller blocks when narrower areas fill in faster than wider areas. In the street map of Laoximen, there are a large number of dead ends and winding streets. Street designs like these, which do not resemble rectangular grids, are less common in modern street layouts than in older ones [22]. We observe in Fig. 16.1 that the southern part of our Laoximen map seems to contain more of these features



(a) Laoximen (b) Former French Concession (c) Pudong New Area

Fig. 16.3 Persistence diagrams of street networks from three neighborhoods of Shanghai. We show the (birth, death) coordinates of 0D features as blue squares and those of 1D features as red disks; darker shapes indicate that more features have the same coordinates. By comparing the birth and death times of the features in the PDs, we observe differences in block-size distributions and block shapes in these neighborhoods. These differences reflect the different developmental histories of these three areas

than other parts of the map. This particular area of the map includes one of the oldest remaining neighborhoods of Laoximen.[2] Much of the area around it has been demolished and redeveloped.

The PD of the former French Concession (see Fig. 16.3b) has more features with death times between 10 and 20 than is the case for the PD of Laoximen. This indicates the existence of medium-size blocks, and we see in Fig. 16.1 that the blocks in the former French Concession are generally larger than those in Laoximen. We still observe many features with death times that are less than 10, so the street network of the former French Concession does have a variety of block sizes. Although it has fewer dead ends than Laoximen, many of the blocks in the former French Concession are not rectangular because of its curved roads. Like Laoximen, the former French Concession has experienced much redevelopment in the last several decades [25]. However, many of the original buildings and streets remain, and the former French Concession is a popular tourist destination because of its European-style buildings and streets. Its extra-settlement roads (which were built by the French colonial government), spacious residential lots, and wide and tree-lined streets are reflected in its street map.

The final district that we discuss is Pudong New Area, a financial hub that has developed rapidly in the last few decades. This area, which is located across the Huangpu River from European concession territories and the old city of Shanghai, was initially developed only modestly before the late 20th century. In the 1990s, the Chinese government set up a Special Economic Zone in Pudong New Area [26], and this district now has some of Shanghai's most famous skyscrapers. The PD of our street map of Pudong New Area (see Fig. 16.3c) has several 1D features with death times that are larger than 20, indicating the existence of large blocks. We also observe several features with early and intermediate death times; these correspond

[2] This part of Laoximen has been slated for redevelopment since 2017 [23]. When we obtained these street maps in 2019, residents were fighting redevelopment efforts and development had not yet begun [24]. It remains to be seen how this part of our Laoximen street map will change as a result of redevelopment.

to a few small blocks on the map. For example, there appears to be a small traffic circle towards the western part of the street map (see Fig. 16.1c). The large blocks are indicative of modern styles of urban planning, with large blocks laid out along grids. Although these blocks are much larger than those in the street maps of the other two regions, many of them are not rectangular, so we again observe several features with late birth times.

16.4 Case Study: Anomalies in the Spread of COVID-19 Infections

The spread of coronavirus disease 2019 (COVID-19), which is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has resulted in a global pandemic [27]. Modeling the spread of COVID-19 is an important and complicated task [28], in part because of the spatial heterogeneity in how it spreads.

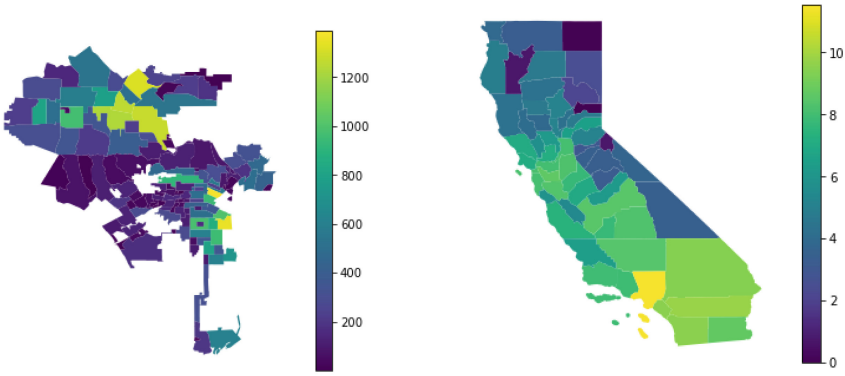
TDA can be useful for the analysis of spreading phenomena. For example, PH has been used previously in epidemiological applications to forecast the spread of Zika [29] and to analyze the Watts threshold model of a contagion on noisy geometric networks [30]. PH provides a different perspective than the many spatiotemporal forecasting models that have been developed for COVID-19 without TDA (see, e.g., [31, 32]).

In our case study,[3] we use PH to analyze the spatial properties of the spread of COVID-19 in neighborhoods in the city of Los Angeles (LA) and counties in California. In contrast to prior work, we use PH in a way that incorporates the underlying geographic structure and various spatial relationships. We consider two data sets. The first is a highly granular data set that consists of COVID-19 case counts in 136 LA neighborhoods on 30 June 2020. The second is a coarser data set that consists of case counts in the 58 counties of California on the same day [34]. For each data set, we also have geographic information in the form of a SHAPEFILE [35, 36]. We visualize these data sets in Fig. 16.4.

Let M^{LA} denote the 2D manifold that consists of the union of LA neighborhoods with fewer than 750 cumulative cases, and let M^{CA} denote the union of California counties with fewer than 5000 cumulative cases. We approximate these manifolds by rasterizing the associated SHAPEFILES to obtain manifolds M_0^{LA} and M_0^{CA}. We show M_0^{LA} and M_0^{CA} in Fig. 16.5. As we described in Sect. 16.2, we construct sequences of manifolds starting from M_0^{LA} and M_0^{CA} using level-set dynamics (16.1). We then construct a level-set filtration for each of these sequences by imposing the manifolds in them on a triangulation of the plane.

In Fig. 16.6, we show the PDs that we compute for the 1D PH of our level-set complexes for the two data sets. These PDs can help us identify COVID-19 anomalies. We define an 'anomaly' to be a collection of regions—a set of neighborhoods in the

[3] See [33] for a study of geographic patterns in COVID-19 case rates and COVID-19 vaccination rates that uses PH in a different way.

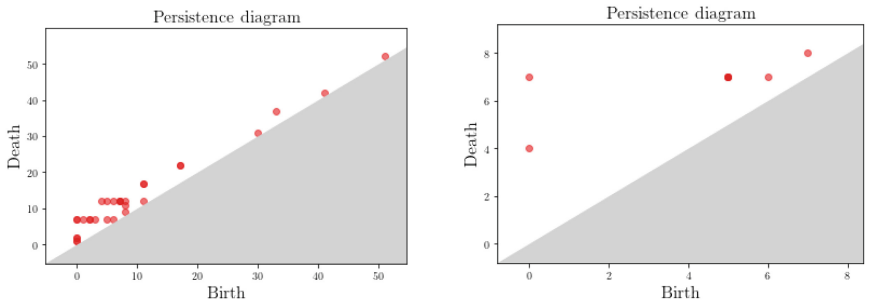


(a) LA neighborhoods (b) California counties

Fig. 16.4 Cumulative COVID-19 case counts on 30 June 2020 in (**a**) neighborhoods in the city of Los Angeles and (**b**) counties in California. We plot the LA case counts on a linear scale and the California county case counts on a (natural) logarithmic scale

(a) M_0^{LA} (b) M_0^{CA}

Fig. 16.5 Initial manifolds for the level-set filtrations that we construct from data of the spread of COVID-19. (**a**) The manifold M_0^{LA} is an approximation of the manifold M^{LA}, which consists of the union of LA neighborhoods with fewer than 750 cumulative cases on 30 June 2020. (**b**) The manifold M_0^{CA} is an approximation of the surface M^{CA}, which consists of the union of California counties with fewer than 5000 cumulative cases on 30 June 2020



(a) LA neighborhoods (b) CA counties

Fig. 16.6 The PDs for the 1D PH of the level-set filtrations for COVID-19 cases in (**a**) LA and (**b**) California

LA data and a set of counties in the California data—in which the case count is larger than in the surrounding area. This notion of an anomaly[4] is analogous to the political 'islands' that were studied using PH in [11]. Anomalies with case counts that are at least as large as the threshold (750 for LA neighborhoods and 5000 for California counties) appear as holes in M_0, unless the anomaly is adjacent to the boundary of the associated map (e.g., LA County in California). The anomalies that are not adjacent to the boundary correspond to homology classes that are born at time 0. There is not a one-to-one correspondence between anomalies and homology classes that are born at time 0. Some of the homology classes that are born at 0 are simply holes in the map (e.g., see Fig. 16.4a), and anomalies that are adjacent to a boundary do not necessarily correspond to any homology class. Homology classes that are born after time 0 usually reflect only the geography of the regions, although they sometimes correspond to anomalies on the boundary of the associated map (much like the homology classes that are created by city blocks with dead ends in the Shanghai street networks). The PDs reflect both the numbers of anomalies and the sizes of the anomalies.

16.5 Conclusions

In this chapter, we discussed the importance of incorporating spatial information into TDA when using it to study spatial systems. As case studies, we computed PH using a level-set construction of filtered simplicial complexes for two case studies: city street networks in Shanghai and anomalies in the spread of COVID-19 infections.

[4] We are concerned with local maxima. By contrast, the Centers for Disease Control and Prevention (CDC) defines COVID-19 'hotspots' using an absolute threshold for the number of cases and criteria that are related to the temporal increase in the number of cases [37].

In our case study of street networks in Shanghai, we illustrated that PH can capture both topological and geometric properties of the organization of city streets. We also observed that the PH of Shanghai's street networks reflect underlying differences in urban planning and organization. This suggests that topological tools can summarize information about how humans organize themselves in space, although further study is necessary to fully understand what types of spatial organization are amenable to TDA.

In our case study of the spread of COVID-19, we showed that one can use a level-set filtration to study the number and sizes of COVID-19 anomalies on both a granular level (by considering neighborhoods in the city of Los Angeles) and a coarse level (by considering counties in California). We used only case counts in our computations, but one can also construct level-set filtrations for death counts, hospitalization counts, or other quantities. The level-set filtration is flexible, but our approach has important limitations. For example, we only detected anomalies with case counts that are at least some fixed threshold. This restricts us to measuring the severity of an outbreak based on its geographic area. One way to address this issue is by applying the level-set filtration after constructing a cartogram [38], instead of directly from a SHAPEFILE. Additionally, the level-set filtration is unable to detect anomalies that occur on the boundary of a map (e.g., in Los Angeles County when considering counties in California). Addressing these limitations is part of ongoing work [33].

Many spatial systems are also social in nature, and there are major challenges to overcome when studying such systems using TDA. In this chapter, we studied spatial systems, but it is important to point out that many spatial systems (including the examples in this chapter) reflect complicated social dynamics. For example, the Shanghai street networks have been shaped by social processes like colonial occupation and displacement of historical neighborhoods. Additionally, COVID-19 disproportionately affects certain communities because of a confluence of social factors, including who is in prison [39] and where hospitals are located [40]. The interaction between social and spatial systems is complicated and inseparable, and intense work is necessary to connect approaches like TDA in spatial systems to the social factors at play.

Acknowledgements We thank the Los Angeles County Department of Public Health for providing the LA data on COVID-19, and we thank Federico Battiston and Giovanni Petri for the invitation to write this chapter. We thank Deanna Needell for helpful comments. We acknowledge support from the National Science Foundation (grant number 1922952) through the Algorithms for Threat Detection (ATD) program. MAP also acknowledges support from the National Science Foundation (grant number DMS-2027438) through the RAPID program.

References

1. M. Barthelemy, *Morphogenesis of Spatial Networks* (Springer International Publishing, Cham, Switzerland, 2018)

2. M.E.J. Newman, *Networks*, 2nd edn. (Oxford University Press, Oxford, UK, 2018)
3. M. Buchet, Y. Hiraoka, I. Obayashi, Persistent homology and materials informatics, in *Nanoinformatics* ed. by I. Tanaka (Springer-Verlag, Heidelberg, Germany, 2018), pp. 75–95
4. L. Papadopoulos, M.A. Porter, K.E. Daniels, D.S. Bassett, Network analysis of particles and grains. J. Complex Netw. **6**(4), 485–565 (2018)
5. A.E. Sizemore, J.E. Phillips-Cremins, R. Ghrist, D.S. Bassett, The importance of the whole: topological data analysis for the network neuroscientist. Netw. Neurosci. **3**(3), 656–673 (2019)
6. H. Ronellenfitsch, E. Katifori, Global optimization, local adaptation, and the role of growth in distribution networks. Phys. Rev. Lett. **117**(13), 138301 (2016)
7. H.M. Byrne, H.A. Harrington, R. Muschel, G. Reinert, B.J. Stolz, U. Tillmann, Topology characterises tumour vasculature. Math. Today **55**(5), 206–210 (2019)
8. C.M. Topaz, L. Ziegelmeier, T. Halverson, Topological data analysis of biological aggregation models. PLOS ONE, **10**(5), e0126383 (2015)
9. L. Speidel, H.A. Harrington, S.J. Chapman, M.A. Porter, Topological data analysis of continuum percolation with disks. Phys. Rev. E **98**(1), 012318 (2018)
10. P.S.P. Ignacio, I.K. Darcy, Tracing patterns and shapes in remittance and migration networks via persistent homology. Euro. Phys. J. Data Sci. **8**, 1 (2019)
11. M. Feng, M.A. Porter, Persistent homology of geospatial data: a case study with voting. SIAM Rev. **63**(1), 67–99 (2021)
12. N. Otter, M.A. Porter, U. Tillmann, P. Grindrod, H.A. Harrington, A roadmap for the computation of persistent homology. Euro. Phys. J. Data Sci. **6**, 17 (2017)
13. G. Carlsson, Topological methods for data modelling. Nat. Rev. Phys. **2**, 697–707 (2020)
14. M. Feng, M.A. Porter, Spatial applications of topological data analysis: cities, snowflakes, random structures, and spiders spinning under the influence. Phys. Rev. Res. **2**, 033426 (2020)
15. L. Kanari, P. Dłotko, M. Scolamiero, R. Levi, J. Shillcock, K. Hess, H. Markram, A topological representation of branching neuronal morphologies. Neuroinformatics **16**(1), 3–13 (2018)
16. H. Ronellenfitsch, J. Lasser, D.C. Daly, E. Katifori, Topological phenotypes constitute a new dimension in the phenotypic space of leaf venation networks. PLOS Comput. Biol. **11**(12), e1004680 (2015)
17. B.J. Stolz, J. Tanner, H.A. Harrington, V. Nanda, Geometric anomaly detection in data. Proc. Nat. Acad. Sci. U.S.A. **117**(33), 19664–19669 (2020)
18. A. Smith, V. Zavala, The Euler characteristic: a general topological descriptor for complex data (2021). arXiv:2103.03144
19. S.J. Osher, R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces* (Springer-Verlag, Heidelberg, Germany, 2003)
20. Y.M. Yeung, Sung Y.-w. (eds.), *Shanghai: Transformation and Modernization Under China's Open Policy* (Chinese University of Hong Kong Press, Hong Kong, 1996)
21. G. Boeing, OSMnx: new methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. **65**, 126–139 (2017)
22. G. Boeing, Urban spatial order: street network orientation, configuration, and entropy. Appl. Netw. Sci. **4**(1), 67 (2019)
23. T. Kanagaratnam, K. Knyazeva, Demolition of Laoximen: Shanghai's best link to its pre-colonial past may soon be gone. *SupChina*. https://supchina.com/2017/12/13/demolition-of-laoximen-shanghai/ (13 December 2017)
24. M. Walsh, In old Shanghai, a last spring festival before the bulldozers. https://www.sixthtone.com/news/1003537/in-old-shanghai%2C-a-last-spring-festival-before-the-bulldozers (4 February 2019)
25. Q. Guan, Lilong housing, a traditional settlement form. M. Arch. Thesis, McGill University (1996). https://www.mcgill.ca/mchg/student/lilong
26. B.X. Sang, Pudong: another special economic zone in China?—An analysis of the special regulations and policy for Shanghai's Pudong New Area. Northwest. J. Int. Law Bus. **14**(1), 130–160 (1993)
27. World Health Organization. Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019 (19 March 2021), 2021

28. A. Vespignani, H. Tian, C. Dye, J.O. Lloyd-Smith, R.M. Eggo, M. Shrestha, S.V. Scarpino, B. Gutierrez, M.U.G. Kraemer, J. Wu, K. Leung, G.M. Leung, Modelling COVID-19. Nat. Rev. Phys. **2**, 279–281 (2020)
29. M. Soliman, V. Lyubchich, Y.R. Gel, Ensemble forecasting of the Zika space-time spread with topological data analysis. Environmetrics **31**(7), e2629 (2020)
30. D. Taylor, F. Klimm, H. A. Harrington, M. Kramár, K. Mischaikow, M.A. Porter, P.J. Mucha, Topological data analysis of contagion maps for examining spreading processes on networks. Nat. Commun. **6**, 7723 (2015)
31. G. Bobashev, I. Segovia-Dominguez, Y.R. Gel, J. Rineer, S. Rhea, H. Sui, Geospatial forecasting of COVID-19 spread and risk of reaching hospital capacity. SIGSPATIAL Spec. **12**(2), 25–32 (2020)
32. S. Zhu, A. Bukharin, L. Xie, M. Santillana, S. Yang, Y. Xie. High-resolution spatio-temporal model for county-level COVID-19 activity in the U.S., ACM Trans. Manage. Inf. Syst. **12**(4), 33 (2021)
33. A. Hickok, D. Needell, M.A. Porter, Analysis of spatiotemporal anomalies using persistent homology: case studies with COVID-19 data (2021). arXiv:2107.09188
34. USA Facts. US COVID-19 cases and deaths by state. https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ (1 July 2020)
35. Los Angeles GeoHub. COVID19 by neighborhood. https://geohub.lacity.org/datasets/covid19-by-neighborhood/about (3 June 2020)
36. California Open Data Portal. California county boundaries. https://data.ca.gov/dataset/ca-geographic-boundaries/resource/b0007416-a325-4777-9295-368ea6b710e6 (10 September 2019), 2019
37. A.M. Oster et al., Trends in number and distribution of COVID-19 hotspot counties—United States, March 8–July 15, 2020. MMWR Morb. Mortal. Wkly. Rep. **69**, 1127–1132 (2020)
38. M.T. Gastner, M.E.J. Newman, Diffusion-based method for producing density-equalizing maps. Proc. Nat. Acad. Sci. U.S.A. **101**(20), 7499–7504 (2004)
39. A. Nellis, The color of justice: racial and ethnic disparity in state prisons. https://www.sentencingproject.org/publications/color-of-justice-racial-and-ethnic-disparity-in-state-prisons/ (14 June 2016)
40. C.A. Nguyen, M.E. Chernew, I. Ostrer, N.D. Beaulieu, Comparison of healthcare delivery systems in low- and high-income communities. Am. J. Accountable Care **7**(4), 11–18 (2019)

Chapter 17
Higher-Order Description of Brain Function

Paul Expert and Giovanni Petri

Abstract Higher-order interactions have long figured both at the microscopic and macroscopic level in neuronal and whole-brain descriptions, with the aim to capture structural, functional, and ultimately cognitive aspects. They are systematizing the paradigm shift that graph theory introduced by moving from studying neural and brain activation to co-activation patterns. Recently, topology has emerged as a central tool in this context due to its natural capacity to describe relations beyond pairwise interactions, and to recent advances in its computational applications. In this chapter, we summarize fundamental concepts and results of the application of higher-order descriptions to neuroscience. We start from the microscopic scale, describing how higher-order interactions have been introduced and measured in the context of neuronal populations activation patterns and in neural coding theory. We then move to the macroscopic scale, discussing recent applications of topological data analysis to whole-brain data, and finally highlight the challenges related to extracting higher-order signals from low-order ones.

The role of higher-order interactions in neuroscience has been been actively debated at both microscopic and macroscopic levels over the last decade. In both cases, there is evidence that higher-order terms are present, yet at the same time in many instances it is still unclear to what degree such interactions dominate, or are dominated by, pairwise interactions. To further complicate the matter, a certain confusion is present on what exactly is meant by higher-order interactions or effects, since they might be encoded as many-body coefficients in spin models of neuron firing, as hyper-

P. Expert (✉)
University College London, London WC1E 6BT, UK
e-mail: p.expert@ucl.ac.uk

Imperial College London, Global Digital Health Unit, London SW72AZ, UK

Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Yokohama 226-8503, Japan

G. Petri
ISI Foundation, via Chisola 5, 10126 Turin, Italy
e-mail: giovanni.petri@isi.it

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_17

graphical structures for population coding, or even homological properties of the functional spaces of whole-brain activations. The landscape of higher-order effects in neuroscience includes contributions in which higher-order interactions can feature as a fundamental dynamical unit, as a methodological tool, and, at times, as both. In this chapter, we discuss these different aspects moving from the microscopic to the macroscopic scale, while explicitly highlighting the role that higher-order interactions take in the different cases.

Specifically, we describe how higher-order interactions have been introduced and measured in the context of neuronal populations and in coding theory. We then discuss recent applications of topological data analysis to whole-brain data, and finally highlight the challenges related to promoting signals from low-order to higher-order interactions.

17.1 Higher-Order Interactions and Descriptions at the Neuronal Scale

Neuronal activation is the atomic unit of brain activity, and the firing patterns of groups of neurons underpin human society. These patterns are not unique and do reoccur, showing that neurons communicate and generate spatio-temporal correlations in their firing activity. Interestingly, even cortical slices in a Petri dish display non-trivial spatio-temporal correlation patterns [5] and can be used to show the delicate neurochemical balance underpinning neural activation [59]. Measuring neural activity is by definition difficult. The size and density of neurons makes it impossible to measure the activation of a single neuron, and electrodes array typically measure the firing activity of a group of neurons. Measurements taken from live subjects are invasive as they require implanting electrodes. Most experimental data therefore comes from animal studies, with the exception of measurement obtained from subjects suffering from certain forms of epilepsy or neurodegenerative disorders. Computational models are commonly used to generate data, but are—of course—short of the real thing. Despite these limitations on data size—i.e. number of electrodes, length of clean times series or the artificiality of data—, important work has emerged in the study of the role of higher-order interactions in neural coding.

Neuronal activity can be encoded as a two states variable, inactive and firing. A paradigmatic model to study binary variables interactions is the Ising spin model that can easily be extended to include any order of interactions, i.e. pairwise, threeway, and higher-order analogues. The probability of a given neuronal configuration for a population of N neurons $(\sigma_1, \ldots, \sigma_N)$ is given by the maximal entropy distribution, also known as the Gibbs distribution:

$$p(\sigma_1, \ldots, \sigma_N) \propto \exp\left(\sum_i \alpha_i \sigma_i + \sum_{i<j} \beta_{ij}\sigma_i\sigma_j + \sum_{i<j<k} \gamma_{ijk}\sigma_i\sigma_j\sigma_k + \cdots\right), \tag{17.1}$$

where the coupling parameters α_i, β_{ij}, γ_{ijk} are to be estimated from experimental data. In practice, the estimation of the coupling parameters and its reliability limits the order that can be considered [54], and early results found that considering pairwise interactions was in some setups sufficient to capture most of the firing patterns structure [58].

However, investigations of larger neuron ensembles (of the order of 100), showed that higher-order interactions can easily encode responses to stimuli [19, 70] and—additionally—that the structure of the interactions is hierarchical and modular, suggesting scalability [19]. Further work has shown that including higher-order terms to encode firing pattern elicited in response to stimuli improves the goodness of fit when introducing a state-space for patterns [7, 60]. Simultaneous silence, i.e. patterns of inactivity, are also characterisation of higher-order interactions and highlight the role of inhibitory neurons in creating spatio-temporal interaction patterns [61]. Moreover, the higher-order models reveal activity patterns closely related to the underlying structure of cortical columns [31], indicating a relationship between structure and function.

While useful to capture the statistics of neuron firing patterns in response to stimuli, these models suffer from several limitations. Their scalability is a problem, as obtaining good and reliable estimates of the model parameters requires long time-series, even for small systems [54]. They also remain "fitting" models, that make assumptions about the process generating the data that, albeit intuitively reasonable, are nevertheless without theoretical or empirical foundation. The last limitation is built in the model class, they inherently lack a temporal dimension and dynamics that is central to spatio-temporal neural coding.

To continue the study of higher-order driven neural activation patterns while alleviating some of the model-based limitations, we turn to model free, data-based methods from topological data analysis, and focus particularly on place cells [41]. Place and grid cells have complementary roles in encoding and memorising spatial information in the hyppocampus and the enthorinal cortex [24, 41]. They also display reliable and long lasting transient patterns [28], making them ideal candidates for detecting structure in neural activation patterns and understand the function of neural circuits. Although we argue for common neural mechanism across species [4, 30], the experimental data in the works we discuss come from rodents.

Remarkably, the firing patterns of hyppocampal place cells are shown to encode the topology of an animal's environment rather than its exact geometry, as well as its position within its environment [12–14]. Place cells' activations therefore reflect the environment an animal is moving in. The question of how the brain activates the appropriate "environmental" map is currently unknown, but research has shed light on possible mechanisms that allow maps to be consistent and robust over time [1]. Co-activation complexes are constructed by building simplices from coactive place cells. Over time and exploration of the environment, the coactivation simplices progressively become a better topological representation of the physical space (Fig. 17.1). It is not clear that this mechanism is enough to ensure the maps are committed to memory once the animal is removed from the test environment and can be reused in the future or in mental exploration, i.e. memory trip. A potential

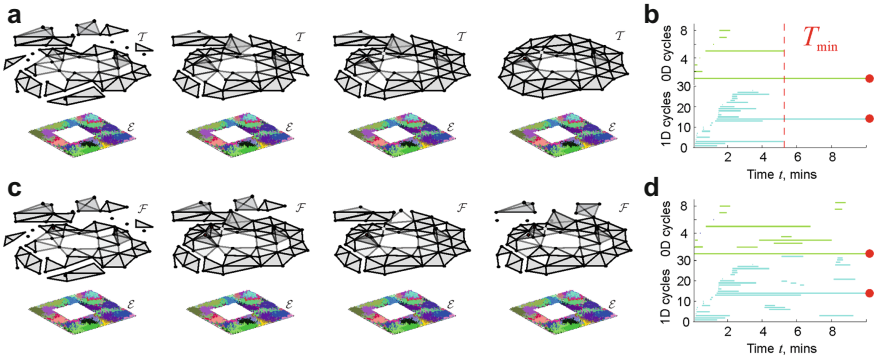


Fig. 17.1 Coactivation complexes can explain sustained and robust representation of spaces. **a** A simulated place cell field map $M(\epsilon)$ of a small planar environment ϵ with a hole in the center is shown together with temporal snapshots of the temporal dynamics of the coactivity complex, which evolved from a small and fragmented one, early during the exploration, to a stable representation of the underlying environment later on. **b** The timelines, encoded as persistent barcodes, of topological persistent H_0 and H_1 cycles for the coactivity simplicial complexes: 0-dimensional persistent generators are shown in light-blue lines, 1-dimensional ones in light-green. Most 1-dimensional cycles correspond to noise, while the persistent topological loops (red dots) encode true physical features of the environment. The time to eliminate the spurious cycles is a proxy for the estimation of the minimal time needed to learn the path connectivity of ϵ. **c** Since simplices can also disappear due to noise and unstable neuronal firing, the coactivity complex can flicker, resulting in **d** the timelines of the topological cycles to be interrupted by opening and closing topological gaps. Figures reproduced from Ref. [1]

such mechanism is proposed in a computational study [2] in the form of replays, where the cells regularly and autonomously reproduce firing sequences corresponding to specific maps, reinforcing existing patterns in a Hebbian learning way. One might conjecture that replays happen during sleep as part of a memory consolidation process [49].

Furthermore, [22] studied how the correlations of spike trains can be used to detect intrinsic structures in place cell activity, without recurring to external stimuli, and how they relate to the topology and geometry of the animal's space. Each correlation matrix was then transformed into an order complex, a filtration of simplicial complexes, obtained by adding a each filtration step a new edge corresponding to the next highest non-diagonal correlation matrix value. The clique complex corresponding to that filtration was then built. They found that the Betti curves that encode the homological properties of the cell activation patterns measured from the animal free-roaming have consistently lower values than from reshuffled version of the correlation matrices. These observations suggested that the correlation structure of hippocampal neurons intrinsically represent the low dimension of the ambient space.

While the geometry of place cells is constrained by the nature of the information they encode, [50] investigated the topology of excitation networks built from simulated activity on reconstructed generic cortical micro-circuitry. The homological

structure of such networks was strikingly complex, showing a surprisingly large number of high-dimensional cliques and a wide variety of high-dimensional homological holes. Further simulations on synthetic and null models found different organisations, suggesting that the topological properties of the activation patterns are not purely driven by the neuron interaction topology, but also by their particular function.

Due to the difficulty of obtaining precise measurement of physical connections for large collection of neurons, topologically based higher-order methods are currently limited to decoding the activity or neuron rather than their structure. At the single neuron level, there is however a correlation between a neuron's topology and its function [29], opening to door to investigating the topological properties of groups of neurons and their function.

So far we have only discussed the role of neural activity at the very small scale, focusing on small neuronal ensembles or very specific functions, e.g. spatial representation. However, one of the great challenge of neuroscience is to understand how behaviour emerges from neural activity, to bridge the scales at which brain activity can be measured [66]. A unified model of brain function remains elusive, despite progresses being made, such as models relating interneuron dysfunction to schizophrenia [65], which has found some experimental confirmation [8, 42], or gene co-expression maps being correlated with fMRI brain activity and neurotransmitter pathways [43, 51]. One is often focused on the difference between the micro scale, i.e. neuronal activity, and macro scale dynamics, i.e. brain activity measured with neuroimaging techniques such as fMRI, EEG, or MEG. There are however similarities, such as the spatio-temporal statistics of neuron [5] and voxels [64] activity. Moreover, the aim at both scales is to link spatio-temporal activity patterns to behaviour. The same set methods can therefore be applied in both cases as they are agnostic to the source of the data, and they can be used to bridge across scales.

For example, binarized fMRI signal [64] can also be used as an input for the extended Ising spin model. The structure of the energy landscape defined by the spin-voxels configuration (Eq. (17.1)) reveals transition dynamics between tasks [17, 68, 69]. However, the models fitted in these studies do not include higher-order terms, as the length of fMRI time series is typically too short for a reliable estimation of the model parameters. Methods relying on signal correlation analysis and topological data analysis are less sensitive to this limitation and have seen their popularity increase for the analysis of macroscopic brain function [16]. We discuss a selection of relevant results in the next section.

17.2 Higher-Order Topology in Whole-Brain Descriptions

At the whole-brain level [15], the question of the importance of higher-order interactions is faced with contrasting evidence. For example, it has been suggested that weak higher-order interactions exist in large-scale functional networks, but are dominated by the pairwise interactions, which would therefore be the main shapers of

brain function [25]. From this perspective, the higher-order terms could be neglected and functional connectivity descriptions based purely on network properties would be enough to fully characterize brain function.

On the other hand, however, higher-order observables were identified as important in multiple studies, e.g. test-retest analysis [72], aberrant connectivity in mental disorders [48] and mild cognitive impairment [71], as well as model inference for EEG signals [32]. Further evidence in this direction has recently come from the study of the shape of the functional spaces described by whole-brain structural and functional data, using tools borrowed from topological data analysis [16, 21]. In structural networks, typically obtained from DTI measurements, persistent homology was used to distinguish healthy and pathological states in developmental [34, 37, 57] and neurodegenerative diseases [35]. For example, considering white matter fibers between brain regions as a weighted network, it was possible to detect loops and cavities between regions that were coherent with biologically-inspired principles of parsimonious wiring (Fig. 17.2a) [62]. Such cavities act as obstructions for information flows and were surrounded by large cliques, which could be interpreted as local dense units able to perform rapid processing. The cavities were reproducible across subjects and connected regions belonging to different phases of brain evolutionary history (Fig. 17.2b).

At the functional level, topological differences were found in healthy and pathological subjects [34, 36]. Higher-dimensional topological features in these cases corresponded to the homological structure of the correlation spaces extracted from functional connectivity analysis, e.g. spaces with Pearson correlation metric. They were useful to discriminate between brain functional configurations in neuropsychiatric disorders and altered states of consciousness relative to controls [11], and to characterize intrinsic geometric structures in neural correlations [22, 55]. One of the more known examples of this type of analysis compared the topology of the functional connectivity of subjects under the effect of psilocybin, a psychedelic drug, with their own under placebo [44], finding that the topological structure was very different between the two conditions, and that the difference could be quantified at the level of persistence diagrams (Fig. 17.2c). To improve the interpretability of the H_1 topological summaries extracted from the data, homological scaffolds were produced to map the topological information back to the brain regions. Such scaffolds can be understood as topological backbones, built from approximated minimal homological generators (Fig. 17.2d), and showed that altered states of consciousness induced by psilocybin (and likely, other psychedelics) arise from different patterns of information integration and importance across brain regions [38] with respect to the normal state (Fig. 17.2e).

Other examples of the application of homology can be found in the literature. Lee et al. [34] have proposed methods to discriminate between cohorts of children with attention deficit hyperactivity disorder, autism spectrum disorder and pediatric control subjects on the basis of their functional topology. In following works [33], the topological substructure of brain networks was represented through the eigenvectors of the corresponding Hodge Laplacians and used it to discriminate between mild and progressive cognitive impairments, and Alzheimer's disease, used to describe the

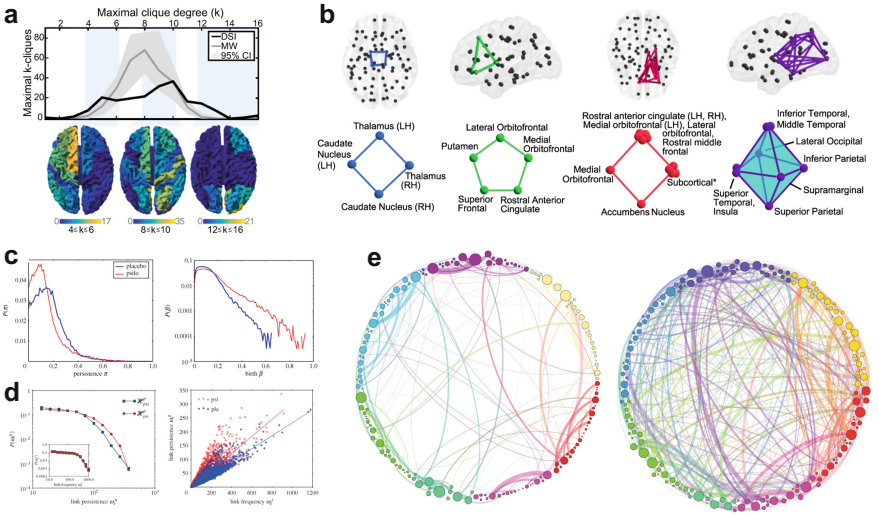


Fig. 17.2 Structural and functional brain topology. **a** Distribution of maximal cliques in the average DSI (black) and individual minimally wired (gray) networks, thresholded at an edge density of $\rho = 0.25$. Heat maps of node participation shown on the brain surfaces for a range of clique degrees equal to 4–6 (left), 8–10 (middle), and 12–16 (right). **b** Minimal cycles representing each persistent cavity at the density at birth represented in the brain (top) and as a schematic (bottom) (adapted from [62]). **c** Comparison of persistence p and birth b distributions. Left, H_1 generators' persistence distributions for the placebo group and psilocybin group. Right, distributions of homology cycles' births. **d** Statistical features of group level homological scaffolds. Left, probability distributions for the edge weights in the persistence homological scaffolds (main plot) and the frequency homological scaffolds (inset). Right, scatter plot of the scaffold edge frequency versus total persistence for both placebo and psilocybin scaffolds. **e** Simplified visualization of the persistence homological scaffolds for subjects injected with placebo (left) and with psilocybin (right). Colours represent communities obtained by modularity optimization on the placebo scaffold and display the departure of the psilocybin connectivity structure from the placebo baseline. Figures adapted from Ref. [44]

heritability of differences in whole-brain functional topology in a cohort of twins [10], and related to topological functional structure of EEG data during imagery to functional equivalence in a population of skilled versus unskilled imagers [26, 27].

17.3 Beyond Functional Connectivity

The analysis methods presented so far mostly focus on notions of functional connectivity, the prototypical example being the canonical Pearson correlation matrices. Using topological description, it is however possible to investigate different features of the spaces in which brain activity can be represented. An interesting example is a topological simplification analysis [56] which focuses on extracting new network representations from temporally resolved fMRI signals. This approach starts

by considering each instantaneous BOLD signal measurement as a point in a high-dimensional space. This space of activation is then filtered using a PCA-based function, which is then used to create a binning of the time points in overlapping bins. Inside each bin, points are then grouped using standard clustering algorithms. The resulting clusters constitute the node set of a new graph, typically called shape graph or Mapper graph [43]. Since the binning allows for overlap across bins, it is possible for the same time point to belong to multiple clusters in different bins. Whenever two clusters share a time point in this way, an edge is added linking the Mapper graph nodes corresponding to the two clusters. In such way, it is possible to build a Mapper graph for each subject, which captures the topology of a simplified representation of the landscape of an individual's activation space (Fig. 17.3a). Interestingly, Saggar and collaborators [56] found that the properties of the individual Mapper graphs were predictive of changes in performance over a multiple tasks: Mapper graphs with large modularity were linked to higher accuracy and smaller response times (Fig. 17.3b). This suggests that a brain activation space that has more diverse and specialised representations of tasks guarantees better multitasking performances [45], as opposed to representations shared across multiple tasks, which instead have been linked to generalization. Moreover, it also suggests that changes in function can be both localized, i.e. specific altered states that induce functional change, and global, i.e. they affect the whole dynamical landscape of brain function rather than only specific configurations. Results supporting this picture were also obtained using related embedding techniques, e.g. low-dimensional projections based on topological distances [6] or persistent homological features obtained from spatial activation patterns [52].

17.4 Higher-Order Signals and Reconstruction

An open and interesting question regarding higher-order interactions in neuroscience is how to measure and—in some cases—even define them. In the case of co-firing neurons, it is natural to identify their firing patterns as a higher-order interaction, as it is done for example in the co-activation simplicial complexes of [3]. In such cases, the interactions also have a natural downward closure—groups of 4 co-firing neurons, also co-fire in groups of 3 and in pairs—, making simplices and simplicial complexes natural descriptions for the system. Moreover, it is also straight-forward and natural to define binary activation signals for these higher-order interactions by considering when they are and are not present.

On the other side, when dealing with large-scale brain dynamics, signals for higher-order interactions are almost never directly available. In general, neuroimaging signals recorded from regions of interest, i.e. 0-th order signals, are encoded as metric spaces [11, 44] or weighted clique simplicial complexes [46], using their correlation properties. Filtrations of simplicial complexes are then extracted from these representations to compute persistent homological features [47]. While this allows to capture information that is not available from a network representation perspective, the dynamics of higher-order interactions strongly depends on the structure of

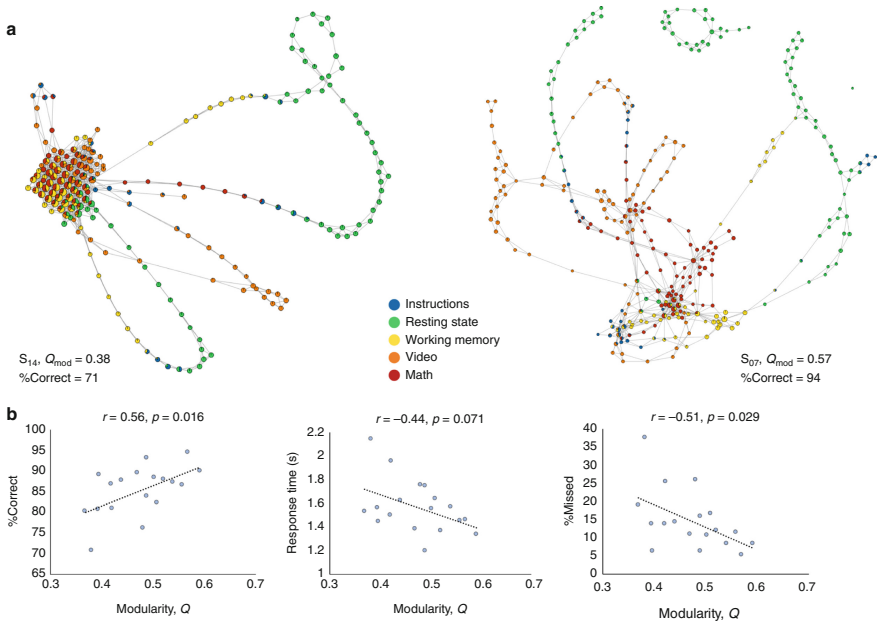


Fig. 17.3 Mesoscale properties of graph representations of brain activation predict individual task performances. Panel **a** shows the Mapper graphs obtained for two subjects [56] (labeled S14 and S07). The pie charts on the nodes show the fraction of timepoints corresponding to each task in the graph node. The Mapper graph for S14 has a low modularity score [40], while that of S07 shows a high degree of modularity structure, in particular showing nodes that are most often connected between to other nodes of same task type. Panel **b** reports the correlations found between the graphs' modularity scores (Qmod) and task performances. Adapted from [56]

pairwise interactions. Although this type of analyses has encountered large success already [16], it would be very valuable to be able to measure or—at least—construct higher-order signals from low-order ones in a principled and controlled way.

A first possibility in this direction is to explicitly leverage low-order signals to define higher-order ones. An example is the edge-level signals and the corresponding edge-centric connectivity introduced by Faskowitz et al. [18]. In standard functional connectivity studies, after z-scoring each time series, the correlation r_{ij} between regions (nodes) i and j is computed as

$$r_{ij} = \frac{1}{T-1} \sum_t \left[z_i(t) \cdot z_j(t) \right] \tag{17.2}$$

where $z_{i,j}$ are the z-scored timeseries. The correlation coefficient r_{ij} is by definition time independent, however, if one discards the sum over t and the normalization, then it is possible to consider its time evolution

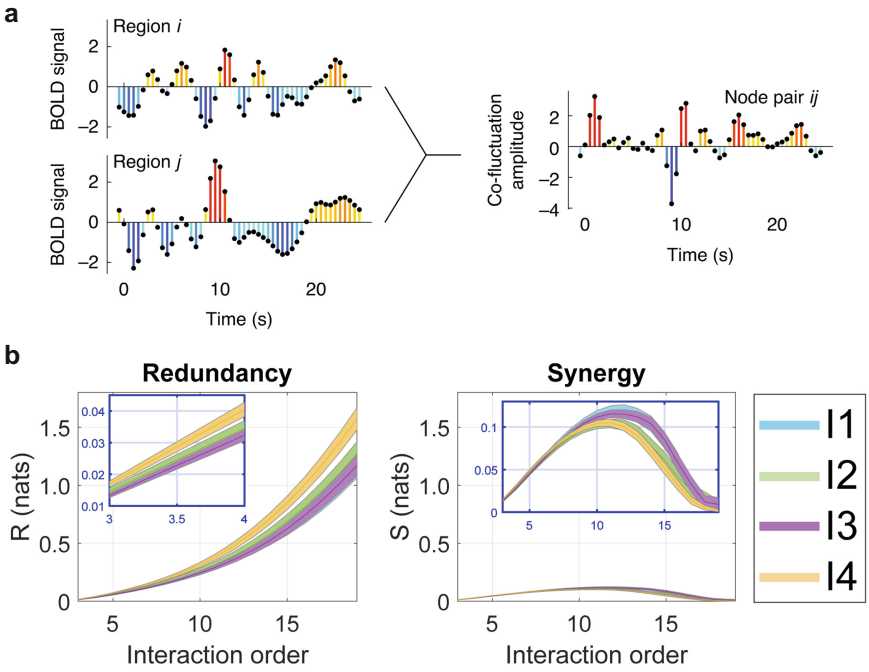


Fig. 17.4 Construction of higher-order timeseries from low-order ones. **a** the construction in the case of edge-centric connectivity. Adapted from Faskowitz et al. [18]. **b** Results for redundancy and synergy based on O-information for age groups: I_1 30 subjects, ages 10–20; I_2 46 subjects, ages 20–40; I_3 29 subjects, ages 40–60; I_4 59 subjects, ages 60–80. Adapted from Gatica and Cofré [20]

$$c_{ij}(t) = z_i(t) \cdot z_j(t) \tag{17.3}$$

as the timeseries describing the coherent fluctations of the functional edge ij and therefore as a genuine higher-order signal. In [18] the authors used this construction to define an edge-based functional connectivity $eFC_{ij,uv}$ among pairs of edges (Fig. 17.4a):

$$eFC_{ij,uv} = \frac{\sum_t c_{ij}(t) \cdot c_{uv}(t)}{\sqrt{\sum_t c_{ij}(t)^2}\sqrt{\sum_t c_{uv}(t)^2}} \tag{17.4}$$

and then studied it using conventional network-based observables. The construction could be generalized to arbitrary orders, which could provide a way to construct weighted and temporally-resolved higher-order representations of brain neuroimaging data. Note that, however, there is no concept of what type of information we are encoding, e.g. whether the co-fluctuations of a set of regions are due to the effect of yet another region, absent which they would be conditionally independent.

A second recent approach to the inference of higher-order interactions offers a possible solution to this problem by adopting an information-theoretic point of

view. O-information [53] is a (real-valued) observable that discriminates between redundant and synergistic components of the information in systems composed by multiple variables. Redundant information here means information that is present in the low-order marginals, e.g. at the node level, while synergistic refers to information that is absent at low orders and is only present at the group level. Formally, for a system composed by n discrete variables, $\mathbf{X}^n = (X_1, \ldots, X_n)$, the O-information $\Omega(\mathbf{X}_n)$ is defined as:

$$\Omega(\mathbf{X}^n) = TC(\mathbf{X}^n) - DTC(\mathbf{X}^n) = (n-2)H(\mathbf{X}^n) + \sum_{j=1}^{n} \left[H(X_j) - H(\mathbf{X}^n_{-j})\right] \tag{17.5}$$

where TC and DTC are respectively the total correlation [67] and dual total correlation [63], and $\mathbf{X}^n_{-j}$ is the vector $\mathbf{X}^n$ with the jth variable omitted. A positive value of $\Omega(\mathbf{X}^n)$ implies that the interdependence are mostly dominated by redundancy, while a negative value implies that synergistic effects are dominant. A further advantage of this quantity over previous multivariate measures of dependency is that it does not require a division between predictors and target variables, but rather provides a genuine measure of group synergy. As a proof of principle, O-information was used to quantify the changes in relevance of interactions of different orders in groups with different ages [20]. In particular, they found significant increases in redundancy in older participants for all interaction orders, and also that synergy and redundancy display different functional forms across all age groups and interaction orders (Fig. 17.4b).

Although promising, these two approaches still face challenges to be widely adopted in conjunction with the existing topological and network tools. For example, when generalizing edge-connectivity to higher-order interactions, the sign of the co-fluctuations, being the result of the multiplication of more than two terms can be misleading and could result in misinterpretations if not properly accounted for. On the other hand, information-theoretic observables typically require a discretization of the signals in states which is non-trivial already in simpler applications, e.g. multivariate mutual information [39]. Finally, in both cases, while it is possible to compute the strength of interactions at all orders, it is unclear how values for interactions at different orders could be compared directly, making the definition of a valid filtration on a weighted simplicial complex non-trivial.

17.5 Outlook for the Future

There is little doubt in our opinion that higher-order interactions play a central role in the brain dynamical organisation and in cognition. They might be weak and difficult to quantify in general at the moment, but in the context of complex systems, weak does not imply negligible [9, 23, 66]. They are likely to play a central role in multitasking [45], focus [27] and neural coding [12, 22] among other things.

They have also been shown to improve model fitting [54, 68], prediction [3, 13] and separation [56]. Hence, higher-order interactions and related analysis methods might be good candidate for biomarkers. It is however difficult at the moment to know whether they are simply good tools and representation of signals for analysis, or have a deeper, more fundamental role in brain theory. Information-based signal analysis [53] might be a good candidate for a first investigation of the precise role of higher-order interactions and structure in brain organisation. However, as future steps, we envision the inclusion of such interactions in testable theoretical brain models so that theory and experiments feed on each other.

References

1. A. Babichev, D. Morozov, Y. Dabaghian, R. spatial memory maps encoded by networks with transient connections. PLOS Comput. Biol. **14**(9), e1006433 (2018)
2. A. Babichev, D. Morozov, Y. Dabaghian, Replays of spatial memories suppress topological fluctuations in cognitive map. Network Neurosc. (Cambridge, Mass.), 3(3):707–724, 2019
3. A. Babichev, D. Morozov, Y. Dabaghian, Replays of spatial memories suppress topological fluctuations in cognitive map. Netw. Neurosci. **3**(3), 707–724 (2019)
4. H.C. Barron, R.B. Mars, D. Dupret, J.P. Lerch, Cassandra sampaio-baptista. Cross-species neuroscience: closing the explanatory gap. Philosoph. Trans. of the Royal Society B, **376**(1815), 20190633 (2021)
5. J.M. Beggs, D. Plenz, Neuronal avalanches in neocortical circuits. J. Neurosci. Official J. Soc. Neurosc. **23**(35), 11167–11177 (2003)
6. J. Billings, M. Saggar, J. Hlinka, S. Keilholz, G. Petri, Simplicial and topological descriptions of human brain dynamics. Network Neurosci. **5**(2), 549–568 (2021)
7. N.A. Cayco-Gajic, J. Zylberberg, E. Shea-Brown, Triplet correlations among similarly tuned cells impact population coding. Front. Comput. Neurosci. **9**, 57 (2015)
8. K.K.A. Cho, R. Hoch, A.T. Lee, T. Patel, J.L.R. Rubenstein, V.S. Sohal, Gamma rhythms link prefrontal interneuron dysfunction with cognitive inflexibility in Dlx5/6+/- Mice. Neuron. **85**(6), 1332–1343 (2015)
9. A. Choudhary, A. Saha, S. Krueger, C. Finke, Phys. Rev., *(A curious case of weak interactions* (Weak-winner phase synchronization, Physical Review Research, 2021), p. 2021
10. M.K. Chung, H. Lee, A. DiChristofano, H. Ombao, V. Solo, Exact topological inference of the resting-state brain networks in twins. Netw. Neurosci. **3**(3), 674–694 (2019)
11. M.K. Chung, V. Villalta-Gil, H. Lee, P.J. Rathouz, B.B. Lahey, D.H. Zald, Exact topological inference for paired brain networks via persistent homology. in *International Conference on Information Processing in Medical Imaging* (Springer, 2017), pp. 299–310
12. C. Curto, V. Itskov, Cell groups reveal structure of stimulus space. PLoS Comput. Biol. **4**(10), e1000205 (2008)
13. Y. Dabaghian, V.L. Brandt, L.M. Frank, Reconceiving the hippocampal map as a topological template. Elife **3**, e03476 (2014)
14. Y. Dabaghian, F. Mémoli, L. Frank, G. Carlsson, A topological paradigm for hippocampal spatial map formation using persistent homology. PLOS Comput. Biol. **8**(8), e1002581 (2012)
15. F. de Vico Fallani, J. Richiardi, M. Chavez, S. Achard, Graph analysis of functional brain networks: practical issues in translational neuroscience. Philosoph. Trans. Royal Soc. B: Biolog. Sciences, **369**(1653), 20130521 (2014)
16. P. Expert, L.-D. Lord, M.L. Kringelbach, G. Petri, Editorial: topological neuroscience. Netw. Neurosci. **3**(3), 653–655 (2019)
17. T. Ezaki, T. Watanabe, M. Ohzeki, N. Masuda, Energy landscape analysis of neuroimaging data. Philosoph. Trans. Royal Soc. A: Mathe. Phys. Eng. Sci. **375**(2096), 20160287 (2017)

18. J. Faskowitz, F.Z. Esfahlani, Y. Jo, O. Sporns, R.F. Betzel, Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. Nat. Neurosc. **23**(12), 1644–1654 (2020)
19. E. Ganmor, R. Segev, E. Schneidman, Sparse low-order interaction network underlies a highly correlated and learnable neural population code. Proc. Natl. Acad. Sci. U.S.A. **108**(23), 9679–9684 (2011)
20. M. Gatica, R. Cofré, P.A.M. Mediano, F.E. Rosas, P. Orio, I. Diez, S.P. Swinnen, J.M. Cortes (High-order interdependencies in the aging brain, Brain connectivity, 2021)
21. R. Ghrist, Barcodes: the persistent topology of data. Bull. Am. Math. Soc. **45**(1), 61–75 (2008)
22. C. Giusti, E. Pastalkova, C. Curto, V. Itskov, Clique topology reveals intrinsic geometric structure in neural correlations. Proc. Natl. Acad. Sci. U.S.A. **112**(44), 13455–13460 (2015)
23. M.S. Granovetter, The strength of weak ties. Am. J. Sociol. 1–22 (1973)
24. T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, E.I. Moser, Microstructure of a spatial map in the entorhinal cortex. Nature **436**(7052), 801–806 (2005)
25. X. Huang, X. Kaibin, C. Chu, T. Jiang, Y. Shan, Weak higher-order interactions in macroscopic functional networks of the resting brain. J. Neurosci. **37**(43), 10481–10497 (2017)
26. E. Ibáñez-Marcelo, L. Campioni, D. Manzoni, E.L. Santarcangelo, G. Petri, Spectral and topological analyses of the cortical representation of the head position: does hypnotizability matter? Brain Behav. **9**(6), e01277 (2019)
27. E. Ibáñez-Marcelo, L. Campioni, A. Phinyomark, G. Petri, E.L. Santarcangelo, Topology highlights mesoscopic functional equivalence between imagery and perception: the case of hypnotizability. NeuroImage **200**, 437–449 (2019)
28. V. Itskov, C. Curto, E. Pastalkova, G. Buzsáki, Cell assembly sequences arising from spike threshold adaptation keep track of time in the hippocampus. J. Neurosc?: Official J. Soc. Neurosci. **31**(8), 2828–2834 (2011)
29. L. Kanari, S. Ramaswamy, Y. Shi, S. Morand, J. Meystre, R. Perin, M. Abdellah, Y. Wang, K. Hess, H. Markram, Objective morphological classification of neocortical pyramidal cells. Cerebral Cortex (New York, N.Y. 1991), **29**(4), 1719–1735 (2019)
30. P.S. Katz, Neural mechanisms underlying the evolvability of behaviour. Philosoph. Trans. Royal Soc. B: Biolog. Sci. **366**(1574), 2086–2099 (2011)
31. U. Köster, J. Sohl-Dickstein, C.M. Gray, B.A. Olshausen, Modeling higher-order correlations within cortical microcolumns. PLOS Comput. Biol. **10**(7), e1003684 (2014)
32. B. Kralemann, A. Pikovsky, M. Rosenblum, Reconstructing effective phase connectivity of oscillator networks from observations. New J. Phys. **16**(8), 085013 (2014)
33. H. Lee, M.K. Chung, H. Choi, H. Kang, S. Ha, Y.K. Kim, D.S. Lee, Harmonic holes as the submodules of brain network and network dissimilarity, in *International Workshop on Computational Topology in Image Context* (Springer, 2019), pp. 110–122
34. H. Lee, M.K. Chung, H. Kang, B.-N. Kim, D.S. Lee. Discriminative persistent homology of brain networks, in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, 2011), , pp. 841–844
35. H. Lee, M.K. Chung, H. Kang, D.S. Lee, Hole detection in metabolic connectivity of Alzheimer's disease using k- Laplacian, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2014), pp. 297–304
36. H. Lee, H. Kang, M.K. Chung, B.-N. Kim, D.S. Lee, Persistent brain network homology from the perspective of dendrogram. IEEE Trans. Med. Imaging **31**(12), 2267–2277 (2012)
37. H. Lee, H. Kang, M.K. Chung, S. Lim, B.-N. Kim, D.S. Lee, Integrated multimodal network approach to PET and MRI based on multidimensional persistent homology. Hum. Brain Mapp. **38**(3), 1387–1402 (2017)
38. L.-D. Lord, P. Expert, H.M. Fernandes, G. Petri, T.J. Van Hartevelt, F. Vaccarino, G. Deco, F. Turkheimer, M.L. Kringelbach, Insights into brain architectures from the homological scaffolds of functional connectivity networks. Front. Syst. Neurosci. **10**, 85 (2016)
39. A. Montalto, L. Faes, D. Marinazzo, Mute: a matlab toolbox to compare established and novel estimators of the multivariate transfer entropy. PloS one **9**(10), e109462 (2014)

40. M.E.J. Newman, Modularity and community structure in networks. Proc. Natl. Acad. Sci. U.S.A. **103**(23), 8577–8582 (2006)
41. J. O'Keefe, J. Dostrovsky, The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res. **34**(1), 171–175 (1971)
42. F. Orhan, H. Fatouros-Bergman, M. Goiny, A. Malmqvist, F. Piehl, K. Schizophrenia Project (KaSP) Consortium, S. Cervenka, K. Collste, P. Victorsson, C.M. Sellgren, L. Flyckt, S. Erhardt, G. Engberg CSF GABA is reduced in first-episode psychosis and associates to symptom severity. Mol. Psych. **23**(5), 1244–1250 (2018)
43. A. Patania, P. Selvaggi, M. Veronese, O. Dipasquale, P. Expert, G. Petri, Topological gene expression networks recapitulate brain anatomy and function. Netw. Neurosci. (Cambridge, Mass.) **3**(3), 744–762 (2019)
44. G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer, F. Vaccarino, Homological scaffolds of brain functional networks. J. R. Soc. Interface **11**(101), 20140873 (2014)
45. G. Petri, S. Musslick, B. Dey, K. Özcimder, D. Turner, N.K. Ahmed, T.L. Willke, J.D. Cohen, Topological limits to the parallel processing capability of network architectures. Nat. Phys. **17**(5), 646–651 (2021)
46. G. Petri, M. Scolamiero, I. Donato, F. Vaccarino. Topological strata of weighted complex networks. PLOS ONE, **8**(6) (2013)
47. A. Phinyomark, E. Ibanez-Marcelo, G. Petri, Resting-state fMRI functional connectivity: big data preprocessing pipelines and topological data analysis. IEEE Trans. Big Data **3**(4), 415–428 (2017)
48. S.M. Plis, J. Sui, T. Lane, S. Roy, V.P. Clark, V.K. Potluru, R.J. Huster, A. Michael, S.R. Sponheim, M.P Weisend et al. High-order interactions observed in multi-task intrinsic networks are dominant indicators of aberrant brain function in schizophrenia. Neuroimage **102**, 35–48 (2014)
49. B. Rasch, J. Born, About sleep's role in memory. Physiolog. Rev. **93**(2), 681–766 (2013)
50. M.W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, H. Markram, Cliques of neurons bound into cavities provide a missing link between structure and function. Front. Comput. Neurosci. **11**, 48 (2017)
51. J. Richiardi, A. Altmann, A.-C. Milazzo, C. Chang, M.M. Chakravarty, T. Banaschewski, G.J. Barker, A.L.W. Bokde, U. Bromberg, C. Büchel, P. Conrod, M. Fauth-Bühler, H. Flor, V. Frouin, J. Gallinat, H. Garavan, P. Gowland, A. Heinz, H. Lemaître, K.F. Mann, J.-L. Martinot, F. Nees, T. Paus, Z. Pausova, M. Rietschel, T.W. Robbins, M.N. Smolka, R. Spanagel, A. Ströhle, G. Schumann, M. Hawrylycz, M.D. Greicius, IMAGEN consortium, L. Albrecht, C. Andrew, M. Arroyo, E. Artiges, S. Aydin, C. Bach, T. Banaschewski, A. Barbot, G. Barker, N. Boddaert, A. Bokde, Z. Bricaud, U. Bromberg, R. Bruehl, C. Büchel, A. Cachia, A. Cattrell, P. Conrod, P. Constant, J. Dalley, B. Decideur, S. Desrivieres, T. Fadai, H.Flor, J. Gallinat, H. Garavan, F.G. Briand, P. Gowland, B. Heinrichs, A. Heinz, N. Heym, T. Hübner, J. Ireland, B. Ittermann, T. Jia, M. Lathrop, D. Lanzerath, C. Lawrence, H. Lemaitre, K. Lüdemann, C. Macare, C. Mallik, J.-F. Mangin, K. Mann, J.-L. Martinot, E. Mennigen, F.M. de Carvahlo, X. Mignon, R. Miranda, K. Müller, F. Nees, C. Nymberg, M.-L. Paillere, Z. Pausova, J.-B. Poline, L. Poustka, M. Rapp, G. Robert, J. Reuter, M. Rietschel, S. Ripke, T. Robbins, S. Rodehacke, J. Rogers, A. Romanowski, B. Ruggeri, C. Schmäl, D. Schmidt, S. Schneider, M. Schumann, F. Schubert, Y. Schwartz, M. Smolka, W. Sommer, R. Spanagel, C. Speiser, T. Spranger, A. Stedman, S. Steiner, D. Stephens, N. Strache, A. Ströhle, M. Struve, N. Subramaniam, L. Topper, R. Whelan, S. Williams, J. Yacubian, M. Zilbovicius, C.P. Wong, S. Lubbe, L. Martinez-Medina, A. Fernandes, A. Tahmasebi, Correlated gene expression supports synchronous activity in brain networks. Science **348**(6240), 1241–1244 (2015)
52. B. Rieck, T. Yates, C. Bock, K. Borgwardt, G. Wolf, N. Turk-Browne, S. Krishnaswamy, Uncovering the topology of time-varying fmri data using cubical persistence (2020). ArXiv preprint arXiv:2006.07882
53. F.E. Rosas, P.A.M. Mediano, M. Gastpar, H.J. Jensen, Quantifying high-order interdependencies via multivariate extensions of the mutual information. Phys. Rev. E. **100**(3), 032305 (2019)

54. Y. Roudi, J. Tyrcha, J. Hertz, Ising model for neural data: model quality and approximate methods for extracting functional connectivity. Phys. Rev. E **79**(5), 051915–12 (2009)
55. E. Rybakken, N. Baas, B. Dunn, Decoding of neural data using cohomological feature extraction. Neural Comput. **31**(1), 68–93 (2019)
56. M. Saggar, O. Sporns, J. Gonzalez-Castillo, P.A. Bandettini, G. Carlsson, G. Glover, A.L. Reiss, Towards a new approach to reveal dynamical organization of the brain using topological data analysis. Nat. Commun. **9**(1), 1–14 (2018)
57. F.A.N. Santos, E.P. Raposo, M.D. Coutinho-Filho, M. Copelli, C.J. Stam, L. Douw, Topological phase transitions in functional brain networks. Phys. Rev. E **100**(3), 032414 (2019)
58. E. Schneidman, M.J. Berry II., R. Segev, W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population. Nature **440**(7087), 1007 (2006)
59. W.L. Shew, H. Yang, S. Yu, R. Roy, D. Plenz, Information capacity and transmission are maximized in balanced cortical networks with neuronal avalanches **31**(1), 55–63 (2011)
60. H. Shimazaki, S. Amari, E.N. Brown, S. Grün, State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data. PLOS Comput. Biol. **8**(3), e1002385 (2012)
61. H. Shimazaki, K. Sadeghi, T. Ishikawa, Y. Ikegaya, T. Toyoizumi, Simultaneous silence organizes structured higher-order interactions in neural populations. Sci. Rep. **5**, 9821 (2015)
62. A.E. Sizemore, C. Giusti, A. Kahn, J.M. Vettel, R.F. Betzel, D.S. Bassett, Cliques and cavities in the human connectome. J. Comput. Neurosci. **44**(1), 115–145 (2018)
63. T.H. Sun, Linear dependence structure of the entropy space. Inf. Control. **29**(4), 337–68 (1975)
64. E. Tagliazucchi, P. Balenzuela, D. Fraiman, D.R. Chialvo, Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. Front. Phys. 1–12 (2013)
65. F.E. Turkheimer, R. Leech, P. Expert, L.-D. Lord, A.C. Vernon, The brain's code and its canonical computational motifs. From sensory cortex to the default mode network: a multi-scale model of brain function in health and disease. Neurosc. Biobehav. Rev. **55**, 211–222 (2015)
66. F.E. Turkheimer, F.E. Rosas, O. Dipasquale, D. Martins, E.D. Fagerholm, P. Expert, F. Váša, L.-D. Lord, R. Leech, A Complex systems perspective on neuroimaging studies of behavior and its disorders. The Neuroscientist: a Rev J. Bring. Neurob. Neurol. Psychiatry pp. 1073858421994784 (2021)
67. S. Watanabe, Information theoretical analysis of multivariate correlation. IBM J. Res. Dev. **4**(1), 66–82 (1960)
68. T. Watanabe, S. Hirose, H. Wada, Y. Imai, T. Machida, I. Shirouzu, S. Konishi, Y. Miyashita, N. Masuda, Energy landscapes of resting-state brain networks. Front. Neuroinformatics **8**, 12 (2014)
69. T. Watanabe, N. Masuda, F. Megumi, R. Kanai, G. Rees, Energy landscape and dynamics of brain activity during human bistable perception. Nature **5**, 4765 (2014)
70. S. Yu, H. Yang, H. Nakahara, G.S. Santos, D. Nikolić, D. Plenz, Higher-order interactions characterized in cortical activity. J. Neurosci. **31**(48), 17514–17526 (2011)
71. H. Zhang, X. Chen, F. Shi, G. Li, M. Kim, P. Giannakopoulos, S. Haller, D. Shen, Topographical information-based high-order functional connectivity and its application in abnormality detection for mild cognitive impairment. J. Alzheimers Dis. **54**(3), 1095–1112 (2016)
72. H. Zhang, Y. Xiaobo Chen, Zhang, D. Shen, Test-retest reliability of "high-order" functional connectivity in young healthy adults. Front. Neurosci. **11**, 439 (2017)

Chapter 18
Higher-Order Interactions in Biology: The Curious Case of Epistasis

C. Brandon Ogbunugafor and Samuel V. Scarpino

Abstract As with other arenas of complex systems, the biological world is driven by interactions between actors, parcels, and forces of various kinds. Higher-order interactions between these elements defines the complexity underlying many biological systems, from species interactions, the microbiota, to biomechanics and others. Here we explore higher-order interactions through a discussion of epistasis, a cutting-edge concept in population and evolutionary genetics. We examine the concept's history and controversies, measure higher-order epistasis operating in a gene encoding an enzyme, and discuss the implications of higher-order interactions for contemporary conversations surrounding genetic modification and other technical challenges that require a more refined understanding of the relationship between genotype and phenotype.

18.1 Introduction

Biology occupies a unique niche among the domains where higher-order interactions exist and are the object of study. By some standards, the biological realm is home to many of the oldest examples of complex systems where higher-order interactions have been observed and cataloged. On the other, it is also among the arenas where higher-order actions remain the most intractable and elusive. But, as we argue in the

C. Brandon Ogbunugafor (✉)
Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA
e-mail: brandon.ogbunu@yale.edu

C. Brandon Ogbunugafor · S. V. Scarpino
Vermont Complex Systems Center, Burlington, VT 05401, USA
e-mail: sscarpino@rockfound.org

S. V. Scarpino
Pandemic Prevention Institute, The Rockefeller Foundation, Network Science Institute, Northeastern University, Boston, MA 02115, USA

Santa Fe Institute, Santa Fe, NM 87501, USA

F. Battiston and G. Petri (eds.), *Higher-Order Systems*, Understanding Complex Systems,
https://doi.org/10.1007/978-3-030-91374-8_18

coming sections, the biological realm is poised to make rapid progress in our scientific understanding of higher-order interactions and our ability to engineer desired outcomes in complex adaptive systems.

While dozens of studies have affirmed that biological systems have all the associated jazz that we associate with complex systems (e.g., emergence), there are fewer areas of biology where a consideration of higher-order interactions has improved our understanding enough to make accurate predictions. That is, while the biological realm recognizes the importance of higher-order interactions in its constitution—across scales of biological organization, from the molecular to the cellular to the ecosystem—the study of higher-order interaction has yet to fully reconcile nor solve any longstanding challenges in biology.

In this chapter, we discuss higher-order interactions within biology, using genetics—and specifically a concept called "epistasis"—as a model problem to highlight how higher-order interactions can be identified, measured, and interpreted. We will focus on higher-order epistasis, a phenomenon that captures both the caprice and the ubiquity of higher-order interactions in biological systems. We examine its multiple definitions within the fields of population, evolutionary, and quantitative genetics, and discuss some of its implications for fundamental theoretical questions and conflicts in the many subfields of genetics. We offer an analysis of a small dataset to demonstrate how higher-order interactions manifest within genes but focus our attention on large questions and conflicts that have undermined a more rigorous understanding of cutting-edge problems in modern genetics.

18.2 Higher-Order Interactions Across Biological Domains

As in all the foundational fields where higher-order interactions exist, higher-order epistasis is defined by nonlinear interactions between actors, entities, replicators, or parcels of information. Such interactions exist in the many subfields of biology, including ecology, where the study of high-dimensional species interactions has long been an important area of research. Indeed, in one of the foundational papers of both mathematical ecology and complex systems, May 1972 asked (and tried to answer) the question, "Will a large complex system be stable?" [1]. Not surprisingly, the answer was that it depends. Whether a large complex system—think about an ecosystem or even a genome—will be stable depends on the number of interacting entities, how they interact, the density of their connections, and—crucially—how we quantify stability. Nevertheless, examples of large, complex biological systems abound [2]. How then should we quantify the nature of the interactions and the degree of stability in these systems? Or said differently, how do we study higher-order interactions in biological systems?

Higher-order interactions have been examined in ecological systems of various kinds, ranging from statistical tests of interactions in ecology [3] to community

ecology [4–6], predator–prey interactions [7], food webs [8], and microbial systems [9, 10]. The importance of higher-order interactions between taxa of microbes has become the focus of study in the microbiota. The microbiota includes the complex community of microbes—trillions in number, across thousands of taxa—that live within other, often multicellular organisms. The composition of the microbiota is now understood to play a large role in many organismal-level phenotypes—disease states, behavior, and many other phenotypes. And higher-order interactions have been measured and are known to occur between the taxa of the microbiota, which may have consequences in the construction of microbial consortia [11–13].

Biomedical systems have highlighted the presence of higher-order interactions between drugs used at the clinical bedside, and especially to treat microbes [14–16]. These findings have reframed our understanding of the challenges of therapy, as drug-drug interactions can foster treatment environments that are challenging to predict from the effects of individual treatments. The future of biomedicine must, then, properly incorporate details of how drugs interact in a higher-order fashion to responsibly predict the effects of multiple drug environments. And, perhaps surprisingly, biomechanical and physiological systems also embody higher-order interactions. For example, recent studies have identified how interactions between anatomical traits may have influenced their evolution across taxa of flatfishes [17]. And, in the model organism *Arabidopsis thaliana*, physiological stress response to cold vs. drought treatments result in different patterns of higher-order interaction in gene expression [18].

In sum, the biological world is full of examples that fortify the importance of higher-order interactions between actors, entities, and parcels of biological information. And biology's many subfields utilize a wide breadth of methods to detect and quantify higher-order interactions.

However, across all the subfields of biology, genetics is an area where higher-order interactions remain most present in both theoretical and empirical research. The notion that genetic information interacts with other parcels of genetic information in crafting phenotypes is a defining feature of genotype–phenotype mapping. And it manifests in many of the most sophisticated aspects of modern genetics, including the search for the genetic underpinnings of complex traits (e.g., disease, behavior) and genetic-modification technology. But how did we end up here? Where did some of these concepts related to higher-order epistatic effects originate? While a full examination of the history of these ideas is beyond the scope of this chapter, we will offer a take, which we hope can provide some context for how higher-order effects ended up on the menu of central ideas in evolutionary and population genetics.

18.3 On Epistasis

18.3.1 History

Shortly after the rediscovery of Gregor Mendel's work on mechanisms of inheritance, biologists demonstrated that certain phenotypes appeared to violate the independence assumptions of gene interactions. Work by Bateson et al. 1910 and Weinberg 1910 [19, 20] both found that offspring phenotypes significantly deviated from expectation, which could not be accounted for by dominance effects nor differences in environment. In 1909, Bateson published an extensive description of what he referred to as *epistasis* [21], which is now the most commonly used term for gene interactions leading to deviation from independence. However, he first used the term in a 1907 correspondence to Muriel Wheldale who was conducting breeding experiments on snapdragons in Bateson's laboratory [22]. In Wheldale's paper [23], she ended up using a different term to describe the effect of the expected phenotype being repressed by another site in the genome. Because Bateson 1907 was trying to describe this repression effect, he settled on the term *epistasis*, which comes from Greek and most closely means "standing upon" [24]. Today, epistasis is used quite broadly to mean any deviation from non-additive (or in some cases even multiplicative) interactions between regions in the genome.

Deviations from the expected phenotype due to genetic interaction were first rigorously demonstrated in plant and animal breeding experiments in the early 1900s. Concurrently, there was a debate raging between geneticists broadly referred to as "Mendelians" and those referred to as "Biometricians" [25]. The debate was between a continuous view of variation between individuals for the same phenotype (Biometricians) and the discrete view that was empirically demonstrable for many phenotypes in breeding experiments (Mendelians) [24]. Reminiscent of the emergence of quantum mechanics, the seeming incongruence of the continuous and discrete views of variation was resolved by R.A. Fisher in 1918, who showed analytically that the two views were entirely compatible [26]. Although, when discussing statistical deviations from non-independence, Fisher used the term *epistacy*, which has fallen out of favor and is no longer used [24]. Instead, geneticists now use the term epistasis to refer to any genetic interaction that leads to non-independence (regardless of whether its suppression—as the term was originally defined—or enhancement and whether the effect is demonstrated using continuous or discrete models [24]).

We now know that the myriad factors interact to determine an individual's trait value for any specific phenotype. These factors include genetic effects, environmental effects, and all possible interactions. Broadly, we use the term *genetic architecture* to represent the joint effects of all determinants of a phenotype. For some phenotypes, e.g., the smooth vs. wrinkly peas from Mendel's experiments, there may be exclusively independent genetic effects. For others, such as schizophrenia, there will be higher-order interactions between genes, environments, and other modifiers. Nearly 100 years of genetic studies have concluded that the vast majority of genetic architectures are complex [27, 28]. Despite this complexity, animal and plant breeders are

able to direct the evolutionary trajectory of numerous traits, from milk production to cold tolerance, using equations and frameworks from quantitative genetics [29].

18.3.2 A Contemporary Example: Biological Networks

Today we know that the structure of molecular interactions clearly affects evolutionary processes across biological levels of organization: from cells, to species, to populations and ecosystems. However, the precise role of molecular network evolution in the process of species formation and adaptation to novel environments, i.e. innovation, remains contentious, with prominent researchers claiming regulatory network changes are critical for macroevolutionary processes [30] and others concluding that they rarely matter [31]. These differences in opinion manifest in various settings.

For example, a recent comparative study of biological regulatory networks found that such networks exist at the edge of criticality, straddling the border of chaotic and ordered states [32]. That biological regulatory networks should exhibit the kind of dynamic stability associated with near-critical networks has been theorized as adaptive, both from the perspective of functional robustness [33] and their ability to effectively process information [34]. However, there is also empirical and theoretical evidence for the importance of change in these networks, e.g., if species must evolve to meet shifting environmental or ecological selection pressures [35]. This tradeoff between robustness and evolvability is hypothesized as an explanation for the commonality of "small-world" networks in biology [36]. Nevertheless, foundational work on self-organized criticality and 1/f noise demonstrated that dynamical systems embedded in a spatial dimension, e.g., biological regulatory networks, might naturally evolve to near-critical states [37, 38].

What then is the role of networks in evolution? For adaptation occurring along a single phenotypic axis, e.g., temperature, RA Fisher's geometric model (1930) of adaptive substitutions provides a great deal of insight [39]. Simply put, the farther away a population is from its fitness optimum, the larger the expected effect-size of a mutational substitution. Recently, a number of experimental evolution studies have provided empirical support for this model (e.g., Hietpas et al. 2013) [40]. However, it seems unlikely that most historical episodes of adaptation occurred along single phenotypic or genotypic axes. And, as Haldane (1957) concluded [41], for many species experiencing selection along multiple axes of variation (e.g., temperature and salinity), that importantly are under independent genetic control, natural selection is expected to drive them rapidly to extinction. As we now know, the genetic architectures underlying even diverse phenotypes are rarely independent: genes underlying all traits exist in complex webs termed molecular networks [42]. Therefore, the role of individual genes in the process of adaptation is affected by their position in these networks and their very interconnectedness may explain why species are able to simultaneously adapt to multiple axes of selection.

18.4 Epistasis: Definitions and a Brief Survey of Methods

One simple dichotomy that captures many of the different uses of the term "epistasis" is the difference between statistical and physiological epistasis. This difference was summarized by Sackton and Hartl (2016) as: "*any situation in which the genotype at one locus modifies the phenotypic expression of the genotype at another* [43]." This is compatible with another very useful definition used by Weinreich et al. (2013), who offered that epistasis is the "*the surprise at the phenotype when mutations are combined, given the constituent mutations' individual effects* [44]". Both are highly mechanistic definitions of epistasis and differ from the more statistical pictures offered that focus on the role of epistasis on genetic variance in populations. Though each form of epistasis—physiological and statistical—are referring to the same sort of phenomena, the means through which epistasis is measured differs greatly depending on which form we are discussing. In general, questions surrounding which methods one should use to study epistasis are analogous to statistical debates regarding the most defensible ways to measure nonlinear effects in complex systems. Quite often, the specific question and context dictate which methods that one should use, and some methods specialize in capturing a particular feature of epistasis. For example, some methods address how noise can affect our understanding of epistasis [45, 46]. Others consider the limits of regression in detecting epistatic effects [47, 48], or propose ways to measure epistasis in incomplete data sets [49].

One method for quantifying epistasis considers the marginal effect of non-independence across large sets of mutations in genomes [50]. This test, called "MArginal ePIstasis Test" or MAPIT, which is a linear mixed modeling strategy for detecting genetic variants (e.g., single nucleotide polymorphisms, SNPs) that are involved in the study of epistasis in genomic mapping studies. With respect to Genome-Wide Association Studies (GWAS), MAPIT estimates and tests the marginal epistatic effect or the combined epistatic effect between SNPs of interest and all other SNPs in the data. By inferring the marginal epistatic effects of SNPS, MAPIT can identify variants that exhibit epistatic interactions with any other variant without the need to identify the specific combinations that drive the epistatic association. Therefore, MAPIT represents an important alternative to standard methods for measuring epistasis [50]. Although somewhat intriguingly, this class of models may perform differentially well at reconstructing epistatic interactions depending on whether the species is outcrossing or self-fertilizing [51]. Despite the documented effectiveness of methods like MAPIT, such approaches are unable to detect epistatic interactions that are greater than second order (pairwise). However, recent applications of approximate Bayesian inference and neural networks, e.g., Biologically Annotated Neural Networks (BANNs), show promise for reconstructing higher-order interactions leading to epistasis [52].

In sum, the methods to diagnose and measure epistasis are as diverse as the settings and varied definitions of epistasis, which provides a challenge to whomever wants to tell a singular narrative about higher-order interactions in genetics. But there are smaller problem cases in genetics where the conceptualization and study

of higher-order epistasis is more tractable. These can serve as useful model problems to discuss how interactions between mutations in a single genetic locus may manifest. And for these purposes, often the physiological definition of epistasis, one marked by clear ways of measuring the "surprise" at the phenotypic effect of combinations of mutations, is the simpler way of demonstrating larger implications of the phenomenon.

18.5 A Demonstration of Epistasis Operating in a Gene Encoding an Enzyme

To demonstrate how epistatic interactions for even a relatively "simple" genetic system can lead to complexity, we examine a data set corresponding to a protein that carries multiple mutations that are associated with resistance to an antibiotic (trimethoprim). The example that follows is based on real-world data, but is simplified as a model to illustrate how epistasis manifests in a biological system. We will measure how interactions between mutations are computed from values for individual variants of a protein containing a different suite of mutations. The dataset itself might be called a "fitness landscape," whereby scientists ask questions about how evolution might be expected to occur across a small discrete portion of sequence space [53, 54].

Specifically, we utilized a suite of three mutations (P21L, A26T, and L28R; single amino acid abbreviations) in bacterial dihydrofolate reductase (DHFR, an essential enzyme target of many antimicrobial drugs) constructed in combination ($2^3 = 8$ alleles) across 3 different genetic backgrounds [55, 56]. The three different genomic backgrounds are as follows: (i) wildtype, (ii) a GroEL$^+$ strain, where a protein chaperone is overexpressed, (ii) ΔLon protease, where an important protease (Lon) has been deleted. Both GroEL and Lon have been demonstrated to regulate DHFR activity in bacterial cells [57].

18.5.1 Data Structure and Methods to Detect Epistasis

The data used here are a subset of those taken from a 2019 study on higher-order epistasis [56]. Importantly, the growth rate and IC_{50} are different, but related traits, part of a canonical tradeoff that has often been observed between growth and resistance in microbial systems. In this examination, we measure epistasis across both traits.

Because this analysis is invoked to be a general examination of higher-order effects in biological systems, rather than a focused study of proteins (or a protein of a certain kind), we will rename the elements in our discussion:

- Dihydrofolate reductase will be referred to as "an enzyme."

- The mutations corresponding to P21L, A26T, and L28R will be referred to with regards to their combinatorial arrangement. For example, "PAL" corresponds to the enzyme variant with amino acids Proline (P), Alanine (A), and Leucine (L) at the three loci of interest.
- The genomic backgrounds corresponding to the wildtype GroEL^{+} and ΔLon protease will be referred to as "environment A" (wildtype) "environment B" (GroEL^{+}) and "environment C" (ΔLon).
- Growth rate will be referred to as "trait 1," and IC_{50} (resistance) will be referred to as "trait 2"

We will make use of a method pioneered in theoretical computer science called the Walsh-Hadamard transform, which computes a coefficient corresponding to the magnitude and sign of an interaction between mutations, akin to an epistatic coefficient. It was pioneered for use in the study of epistasis in a 2013 study that both provided a primer for the calculation and analyzed several combinatorially complete data sets [44]. It has since been further elaborated on applied to study of higher-order epistasis across a larger sampling of empirical data sets [58, 59].

The Walsh-Hadamard transform implements phenotypic measurements into a vector, then a Hadamard matrix, subsequently scaled by a diagonal matrix. The calculation yields a set of coefficients which measure the degree to which the relationship between genetic information and phenotypes are linear, or second order, third, and so forth. One limitation of the Walsh-Hadamard transform is that its data must be combinatorially complete with no more than two variants at a given locus of information. In this one scenario, the mutations at each of three sites (e.g. the three mutations corresponding to Trimethoprim resistance in *E.coli* dihydrofolate reductase), P21L, A26T, and L28R.

The full data set for the alleles consists of a vector of phenotypic values (resistance to trimethoprim in the case of the DHFR mutants) for all possible combinations of mutations (8 in total), represented by their single amino acid substitutions:

PAL, LAL, PAR, PTL, PTR, LAR, LTL, LTR.

These can be represented in binary notation:

000, 100, 001, 010, 011, 101, 110, 111.

This vector of phenotypes can be multiplied by a (8×8) square matrix, which is the product of a diagonal matrix V and a Hadamard matrix H. These are defined recursively by:

$$V_{n+1} = \begin{pmatrix} \frac{1}{2}Vn & 0 \\ 0 & -V_n \end{pmatrix}, V_0 = 1 \tag{18.1}$$

$$H_{n+1} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}, H_0 = 1 \tag{18.2}$$

n is the number of sites that differ in this enzyme (n = 3 in this setting).

The multiplication gives the following expression:

$$\gamma = V\mathrm{H}\varkappa \tag{18.3}$$

In this scenario, H and V are the matrices described in Eqs. 18.1 and 18.2 and γ is the Walsh coefficient, the measure of the average interaction between parcels of information (mutations in this setting), here measured across environments.

Negative values for an effect suggest that that average effect is negative, positive if it has a beneficial effect on a phenotype (e.g., antibiotic resistance).

Note: while we have provided some details of the calculation above, we encourage those interested in the subtleties of the calculation to refer to several manuscripts, especially Weinreich et al. 2013 [44] and Poelwijk 2016 [58], for a more focused treatment of these methods.

18.5.2 Calculating Higher-Order Interactions

The above formula can be used to calculate the strength of interactions between parcels of information, the mutations corresponding to different amino acid substitutions in an enzyme in our example. But what about higher-order interactions (epistasis beyond pairwise in this case)?

Previous studies have examined how higher-order epistasis manifests in adaptive landscapes that include analogously structured data sets, including other enzymes [44, 55, 59, 60]. Because our focus is on how higher-order epistasis manifests in biological systems, we will offer a means through which one can make these measurements.

For example, in a complete data set comprising eight variants, we can describe the interactions between individual loci and genetic background in binary terms. If we are talking about a combinatorial set of variants with three loci, we can describe the interactions using binary representation.

γ_{000} interaction between the mutations in the wild-type background.
γ_{001} interaction between the "third site" mutation and all other genetic backgrounds.
γ_{010} interaction between the "second site" mutation and all other genetic backgrounds.
γ_{100} interaction between the "first site" mutation and all other genetic backgrounds.
γ_{011} second-order (pairwise) interaction between mutations at the second and third loci.
γ_{101} second-order (pairwise) between mutations at the second and third loci.
γ_{110} second-order (pairwise) between mutations at the first and second loci.
γ_{111} A third-order interaction between mutations at all three loci.

In this set, there is one zeroth order interaction, three first-order interactions, three second-order interactions, and one third-order interaction. The third-order interaction would formally qualify as "higher-order."

In addition, one can take the mean of these epistatic coefficients within order, which can facilitate comparisons between orders. For a given epistatic coefficient we compute an absolute mean epistatic coefficient, E, as in prior studies that have examined higher-order interactions on in silico fitness landscapes [61]:

$$E_i = \frac{|\gamma_i|}{\sum_j |\gamma_j|} \tag{18.4}$$

The absolute value allows us to focus on the magnitude of higher-order interactions. We label them with the term "absolute mean" since we incorporated absolute values and averages in the calculation. This provides mean values for each order, which translates to the overall contribution of, for example, 1st order effects. And we can calculate the higher-order interaction across environments, creating an abstraction called the "mutation effect reaction norm" that highlights how environments influence the effect of mutation interactions [62].

Figure 18.1 is a hypercube representation of the eight mutants, arising from a combination of three different mutations (P21L, A26T, and L28R) in a model enzyme. All eight enzymes were engineered using transgenic methods and their growth and resistance phenotypes measured using experimental methods [55]. Figure 18.2 represents the phenotype data for these eight mutants across the three environmental contexts (A, B, C). From this, we can observe the existence of gene by environment (G x E) interactions, indicated by the fact that different mutants have differing slopes for performance for their growth (trait 1) and resistance (trait 2).

Data for trait 1 and 2 can be measured for all the alleles in this hypergraph, and graphed with respect to the trait values across environment A, B, and C. This is

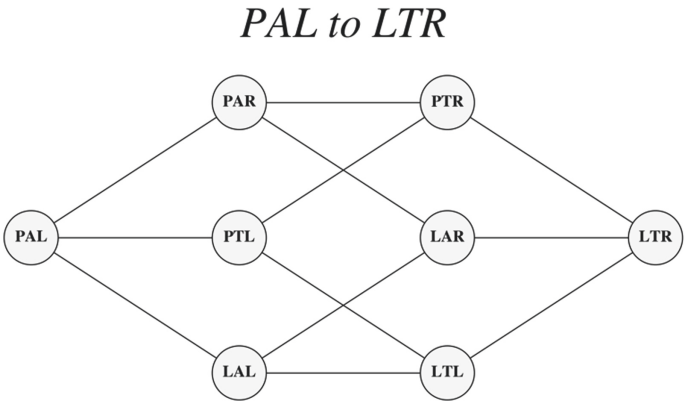


Fig. 18.1 A hypercube representation of the combinatorial set of mutations in the enzyme target of study. Letters correspond to amino acids. In this instance, a bacterial enzyme has two amino variants at each of three sites. Different combinations of these mutations are associated with different values for traits associated with groth rate (trait 1) and antibiotic resistance (trait 2)

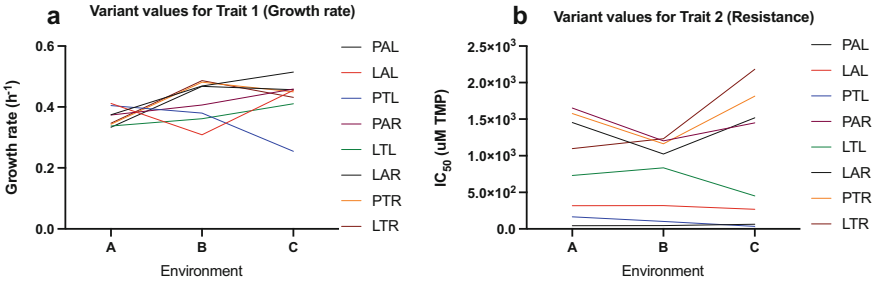


Fig. 18.2 Values for **a** trait 1 (growth) and **b** 2 (resistance) across environments A, B & C

depicted in Fig. 18.2, in an abstraction called a "reaction norm," which is often used to detect gene by environment interactions.

From the data in Fig. 18.2, the epistatic coefficients can be computed as outlined in Eqs. 18.1–18.3. These yield calculations for the average effect of individual interactions between mutations across environments, which can also be depicted in terms of their absolute mean as outlined in Eq. 18.4.

Figure 18.3 depicts the coefficients for different interactions between mutants (Fig. 18.3a and c). For example, the mutation effect corresponding to [*1*] translates

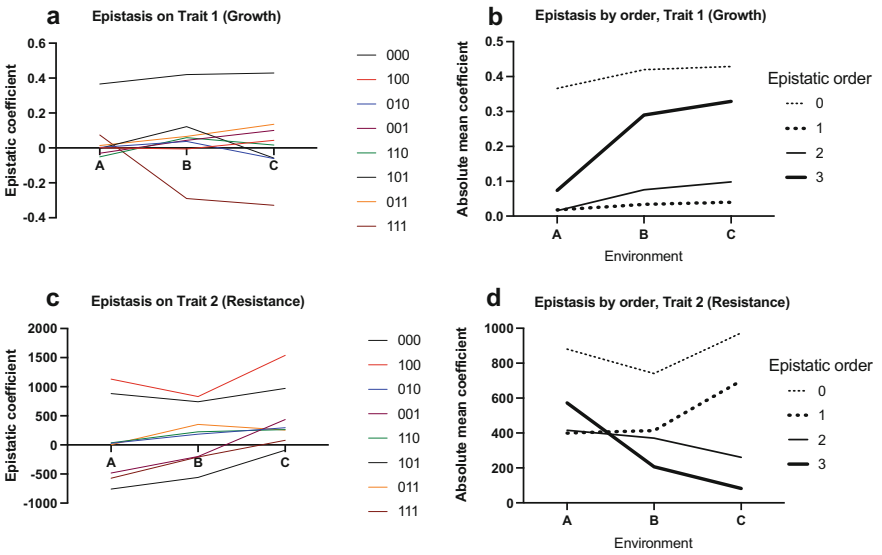


Fig. 18.3 Epistasis across different environments. **a** mutation effect reaction norm for individual mutation interactions, for trait 1 (growth). **b** Absolute mean mutation effects for trait 1, showing effects organized by order. **c** mutation effect reaction norm for individual mutation interactions, for trait 2 (resistance). **b** Absolute mean mutation effects for trait 2, showing effects by order. The notation in the figure only use the subscripts of the effects described in the subscripts of γ described in 19.5.2

to the average effect of adding the second-site mutation (A26T) across available genetic backgrounds. Alternatively, [*11] corresponds to the average phenotypic effect of adding A26T and L28R in combination. Also in Fig. 18.3 is the absolute mean data, where we can observe the overall presence of epistatic effects across environments, organized by order.

What do we learn about how epistasis manifests in an enzyme with a suite of mutations corresponding to different levels of growth and antibiotic resistance? Calculations of epistasis as depicted in Fig. 18.3 communicate the variability of epistatic effects as a function of environment. When there are gene by environment interactions present (as in Fig. 18.2), there are likely epistasis by environment interactions (as in Fig. 18.3).

Even for traits where we expect a clear tradeoff, e.g., growth and antibiotic resistance, the pattern of epistasis differ considerably. As a result, there exists no singular, tractable pattern for how mutations will interact. More specifically, the eminence of higher-order interactions changes across environmental contexts, something that has been observed in other contexts [56, 63, 64]. Given that no locus operates in a genome alone, we might expect the consequences of higher-order interactions to be far greater than the effects measured by single mutations. Not only do mutations interact in surprising ways within loci (as in the results of this data set), but we might also expect genes to interact with other genes in surprising ways (as in gene networks).

The ability of quantitative genetics to predict the trajectory of mean trait values for populations, coupled with our inability to predict the effect of most mutations, poses both challenges and opportunities for the future of genetics. Can we develop theories capable of making predictions relevant for engineering traits via genetic modification, e.g., CRISPR? What additional ethical considerations arise from continuing investment in genetic engineering without a general theory capable of predicting the individual effect of mutations? In the following paragraphs, we advocate for a complex network approach to genetic engineering. One where—far from trying to reduce the effects of mutations—we instead embrace the gestalt when trying to direct the phenotype of an individual through mutational engineering.

18.6 On Higher-Order Interactions and Genetic Modification

Engineering human traits using CRISPR relies on the assumption that scientists can accurately predict phenotype from genotype. Even for diseases caused by mutations in single genes, such as cystic fibrosis [65] or muscular dystrophy [66], CRISPR induced mutations failed to completely restore the “healthy” phenotypes in mouse models or tissue culture.

Many other phenotypes in humans display a property that geneticists term “missing heritability” [67], which means that researchers cannot identify genetic

markers that account for the expected phenotypic similarity between parents and offspring. Despite sequencing thousands of human genomes, we can only predict 6% of the heritability of Type 2 diabetes, 5% for HDL cholesterol, 5% for height, and <3% for early onset myocardial infarction [68]. This "missing heritability" implies that the causal role of genetic interactions, environment, epigenetics, etc. is often as strong as simple changes in one's DNA.

So how can we reconcile the incredible predictive ability of quantitative genetics when applied to animal/plant breeding with our inability to engineer traits using tools like CRISPR? This is where we argue that the modern study of higher-order interactions and genetics currently reside. However, the debate goes back to the early 1900s and the foundations of the field. Specifically, quantitative genetics abstracts away the individual effects of genes and environments (along with their interactions). Instead, it models them as expected effects on the variability of a trait. For example, Huang and Macaky 2016 showed that genetic architecture could not be determined for a principle component analysis of genetic variation (this is despite one's ability to influence trait evolution by selecting on the principal components) [69]. Our modern understanding of the relationship between the individual determinants of traits and their quantitative genetics representation was reviewed by Stinchombe and Hoekstra 2007 [70].

The continued effectiveness of quantitative genetics has led some mathematical geneticists to question the importance of epistasis. Specifically, can we understand and predict evolution without needing to know the underlying causes? Mäki-Tanila and Hill 2014 showed that non-independent interactions between genes increases the additive genetic variation at far higher rates and that it contributes to deviations from additivity [71]. Indeed, they conclude that *"Epistasis may be important in understanding the genetic architecture, for example, of function or human disease, but that does not imply that loci exhibiting it will contribute much genetic variance. Overall we conclude that theoretical predictions and experimental observations of low amounts of epistatic variance in outbred populations are concordant. It is not a likely source of missing heritability, for example, or major influence on predictions of rates of evolution."*

This breakdown in the predictability of how an engineered mutation will affect a phenotype was demonstrated by Guerrero et al. 2019 [56]. This study shows that a mutation's effect depends strongly on genome background, environment, and the interactions between them. As a result, even the sign (positive or negative) of the effect that a mutation will have on a phenotype may be unpredictable. The ability of models that we know are wrong to make accurate predictions is well known outside of mathematical genetics [72] and is indicative of the deep connection between all fields related to complex systems.

18.7 Closing

What are the most important questions, then, about the implications of higher-order interactions for biology? For convenience, we've identified three that capture a broader set of curiosities.

(1) Despite most traits being high-order, many established population genetic models assume pairwise interactions and that high-order interactions can be modeled as the summation of pairwise interactions. How can we accommodate higher-order interactions into this theory?
(2) Fundamentally, we can re-ask where (or what) is the source of missing heritability? Considering the plausibility of higher-order interactions between genetic parcels, how can we "find" this heritability in a manner that doesn't simply relegate the problem to being unsolvable in light of combinatorial explosion?
(3) At a more biophysical and physiological scale, can we use theory from higher-order interactions in other fields to make engineering-level predictions? In the data-driven example in this chapter—a single gene encoding an enzyme, and a small set of mutations—we reveal how higher-order interactions are present and context-dependent. But the methods do allow a mechanistic take on how they manifest and influence a trait of interest. Can we apply such methods to other problems?

The future of higher-order interactions in genetics encompasses these questions and many more. And more broadly, as our understanding of the biological world continue to grow in scope, we can expect the eminence of higher-order interactions to also grow in relevance. While biological information might be highly specialized, with drift and selection responsible for its arrival and dispersal, it is now surrounded by other parcels of information, that all interact in surprising ways, creating a biosphere that is both more corporeal and capricious than scientists and naturalists have appreciated.

References

1. R.M. May, Will a large complex system be stable? Nature **238**, 413–414 (1972)
2. N.D. Martinez, R.J. Williams, J.A. Dunne, M. Pascual, Diversity, complexity, and persistence in large model ecosystems,in *Ecological Networks: Linking Structure to Dynamics in Food Webs* (2006), pp. 163–185
3. U. Dieckmann, R. Law, The dynamical theory of coevolution: a derivation from stochastic ecological processes. J. Math. Biol. **34**, 579–612 (1996)
4. M.M. Mayfield, D.B. Stouffer, Higher-order interactions capture unexplained complexity in diverse communities. Nature Ecol. Evol. **1**, 1–7 (2017)
5. J.M. Levine, J. Bascompte, P.B. Adler, S. Allesina, Beyond pairwise mechanisms of species coexistence in complex communities. Nature **546**, 56–64 (2017)
6. A.D. Letten, D.B. Stouffer, The mechanistic basis for higher-order interactions and non-additivity in competitive communities. Ecol. Lett. **22**, 423–436 (2019)

7. E.L. Preisser, D.I. Bolnick, M.F. Benard, Scared to death? The effects of intimidation and consumption in predator–prey interactions. Ecology **86**, 501–509 (2005)
8. M. Pascual, J.A. Dunne, J.A. Dunne, *Ecological Networks: Linking Structure to Dynamics in Food Webs* (Oxford University Press, 2006)
9. H. Mickalide, S. Kuehn, Higher-order interaction between species inhibits bacterial invasion of a phototroph-predator microbial community. Cell Syst. **9**, 521–533 (2019)
10. R.C. Cobb, R.K. Meentemeyer, D.M. Rizzo, Apparent competition in canopy trees determined by pathogen transmission rather than susceptibility. Ecology **91**, 327–333 (2010)
11. A. Sanchez-Gorostiaga, D. Bajić, M.L. Osborne, J.F. Poyatos, A. Sanchez, High-order interactions distort the functional landscape of microbial consortia. PLoS Biol. **17**, e3000550 (2019)
12. A.L. Gould, V. Zhang, L. Lamberti, E.W. Jones, B. Obadia, N. Korasidis, A. Gavryushkin, J.M. Carlson, N. Beerenwinkel, W.B. Ludington, Microbiome interactions shape host fitness. Proc. Natl. Acad. Sci. **115**, E11951–E11960 (2018)
13. Y. Senay, G. John, S.A. Knutie, C.B. Ogbunugafor, Deconstructing higher-order interactions in the microbiota: a theoretical examination, BioRxiv (2019) 647156
14. E. Tekin, C. White, T.M. Kang, N. Singh, M. Cruz-Loya, R. Damoiseaux, V.M. Savage, P.J. Yeh, Prevalence and patterns of higher-order drug interactions in *Escherichia coli*. NPJ Syst. Biol. Appl. **4**, 1–10 (2018)
15. E. Tekin, P.J. Yeh, V.M. Savage, General form for interaction measures and framework for deriving higher-order emergent effects. Front. Ecol. Evol. **6**, 166 (2018)
16. E. Tekin, V.M. Savage, P.J. Yeh, Measuring higher-order drug interactions: a review of recent approaches. Curr. Opinion Syst. Biol. **4**, 16–23 (2017)
17. K.M. Evans, O. Larouche, S.-J. Watson, S. Farina, M.L. Habegger, M. Friedman, Integration drives rapid phenotypic evolution in flatfishes. Proc. Natl. Acad. Sci. 118 (2021)
18. D.L. Des Marais, R.F. Guerrero, J.R. Lasky, S.V. Scarpino, Topological features of a gene co-expression network predict patterns of natural diversity in environmental response, Proc. Royal Soc. B: Biol. Sci. **284**, 20170914 (2017)
19. W. Bateson, E.R. Saunders, *Reports to the Evolution Committee of the Royal Society: ReportsI-V*, 1902–09, Royal Society (1910)
20. W. Weinberg, Weitere Beitrage zur Theorie der Vererbung. Arch. Rass. Ges. Biol. **7**, 35–49 (1910)
21. W. Bateson, *Mendel's Principles of Heredity* Cambridge University Press, März 1909; 2nd Impr. 3 (1909) 1913
22. M.L. Richmond, Women in the early history of genetics: William Bateson and the Newnham College Mendelians, 1900–1910. Isis **92**, 55–90 (2001)
23. M. Wheldale, The inheritance of flower colour in Antirrhinum majus. Proc. Royal Soc. Londn. Ser. B, Containing Papers of a Biological Character **79**, 288–305 (1907)
24. P.C. Phillips, The language of gene interaction. Genetics **149**, 1167–1171 (1998)
25. W.B. Provine, *The Origins of Theoretical Population Genetics* (University of Chicago Press, 2020)
26. R.A. Fisher, *009: The Correlation Between Relatives on the Supposition of Mendelian Inheritance* (1918)
27. U. Wolf, Identical mutations and phenotypic variation. Hum. Genet. **100**, 305–321 (1997)
28. J.B. Wolf, E.D. Brodie, M.J. Wade, *Epistasis and the Evolutionary Process* (Oxford University Press, 2000)
29. D. Posthuma, A.L. Beem, E.J. De Geus, G.C.M. Van Baal, J.B. Von Hjelmborg, I. Iachine, D.I. Boomsma, Theory and practice in quantitative genetics. Twin Res. Human Genet. **6**, 361–376 (2003)
30. S.B. Carroll, Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. Cell **134**, 25–36 (2008)
31. H.E. Hoekstra, J.A. Coyne, The locus of evolution: evo devo and the genetics of adaptation. Evolution: Int. J. Organic Evol. **61**, 995–1016 (2007)

32. B.C. Daniels, H. Kim, D. Moore, S. Zhou, H.B. Smith, B. Karas, S.A. Kauffman, S.I. Walker, Criticality distinguishes the ensemble of biological regulatory networks. Phys. Rev. Lett. **121**, 138102 (2018)
33. N. Barkai, S. Leibler, Robustness in simple biochemical networks. Nature **387**, 913–917 (1997)
34. R. Cheong, A. Rhee, C.J. Wang, I. Nemenman, A. Levchenko, Information transduction capacity of noisy biochemical signaling networks. Science **334**, 354–358 (2011)
35. R.E. Lenski, J.E. Barrick, C. Ofria, S. Levin, Balancing robustness and evolvability, PLoS Biol. **4**, e428 (2006)
36. A. Wagner, D.A. Fell, The small world inside large metabolic networks. Proc. Royal Soc. Londn. Ser. B: Biol. Sci. **268**, 1803–1810 (2001)
37. S.N. Dorogovtsev, J.F. Mendes, Evolution of networks. Adv. Phys. **51**, 1079–1187 (2002)
38. P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: an explanation of the 1/f noise, Phys. Rev. Lett. **59**, 381–384 (1987). https://doi.org/10.1103/PhysRevLett.59.381
39. R.A. Fisher, *The Genetical Theory of Natural Selection: A Complete Variorum Edition* (OUP Oxford, 1930)
40. R.T. Hietpas, C. Bank, J.D. Jensen, D.N. Bolon, Shifting fitness landscapes in response to altered environments. Evolution **67**, 3512–3522 (2013)
41. J.B.S. Haldane, The cost of natural selection. J. Genet. **55**, 511 (1957)
42. E.A. Boyle, Y.I. Li, J.K. Pritchard, An expanded view of complex traits: from polygenic to omnigenic. Cell **169**, 1177–1186 (2017)
43. T.B. Sackton, D.L. Hartl, Genotypic Context and epistasis in individuals and populations. Cell **166**, 279–287 (2016). https://doi.org/10.1016/j.cell.2016.06.047
44. D.M. Weinreich, Y. Lan, C.S. Wylie, R.B. Heckendorn, Should evolutionary geneticists worry about higher-order epistasis? Curr. Opin. Genet. Dev. **23**, 700–707 (2013). https://doi.org/10.1016/j.gde.2013.10.007
45. Z.R. Sailer, M.J. Harms, *Uninterpretable Interactions: Epistasis as Uncertainty.* BioRxiv 378489 (2018)
46. Z.R. Sailer, M.J. Harms, Detecting high-order epistasis in nonlinear genotype-phenotype maps. Genetics **205**, 1079–1088 (2017)
47. J. Otwinowski, J.B. Plotkin, Inferring fitness landscapes by regression produces biased estimates of epistasis. PNAS **111**, E2301–E2309 (2014). https://doi.org/10.1073/pnas.1400849111
48. J. Otwinowski, D.M. McCandlish, J.B. Plotkin, Inferring the shape of global epistasis. PNAS **115**, E7550–E7558 (2018). https://doi.org/10.1073/pnas.1804015115
49. K. Crona, Rank orders and signed interactions in evolutionary biology. Elife. **9**, e51004 (2020)
50. L. Crawford, P. Zeng, S. Mukherjee, X. Zhou, Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. PLoS Genet. **13**, e1006869 (2017)
51. Y. Jiang, J.C. Reif, Modeling epistasis in genomic selection. Genetics **201**, 759–768 (2015)
52. P. Demetci, W. Cheng, G. Darnell, X. Zhou, S. Ramachandran, L. Crawford, *Multi-scale Inference of Genetic Trait Architecture using Biologically Annotated Neural Networks*, BioRxiv 2020–07 (2021)
53. J.A.G.M. de Visser, J. Krug, Empirical fitness landscapes and the predictability of evolution. Nat. Rev. Genet. **15**, 480–490 (2014). https://doi.org/10.1038/nrg3744
54. F.J. Poelwijk, D.J. Kiviet, D.M. Weinreich, S.J. Tans, Empirical fitness landscapes reveal accessible evolutionary paths. Nature **445**, 383–386 (2007). https://doi.org/10.1038/nature05451
55. J.V. Rodrigues, S. Bershtein, A. Li, E.R. Lozovsky, D.L. Hartl, E.I. Shakhnovich, Biophysical principles predict fitness landscapes of drug resistance. Proc. Natl. Acad. Sci. U.S.A. **113**, E1470-1478 (2016). https://doi.org/10.1073/pnas.1601441113
56. R.F. Guerrero, S.V. Scarpino, J.V. Rodrigues, D.L. Hartl, C.B. Ogbunugafor, Proteostasis environment shapes higher-order epistasis operating on antibiotic resistance. Genetics **212**, 565–575 (2019). https://doi.org/10.1534/genetics.119.302138
57. S. Gottesman, S. Wickner, M.R. Maurizi, Protein quality control: triage by chaperones and proteases. Genes Dev. **11**, 815–823 (1997)

58. F.J. Poelwijk, V. Krishna, R. Ranganathan, The context-dependence of mutations: a linkage of formalisms. PLoS Comput. Biol. **12**, e1004771 (2016)
59. D.M. Weinreich, Y. Lan, J. Jaffe, R.B. Heckendorn, The influence of higher-order epistasis on biological fitness landscape topography. J. Stat. Phys. **172**, 208–225 (2018)
60. E.R. Lozovsky, R.F. Daniels, G.D. Heffernan, D.P. Jacobus, D.L. Hartl, Relevance of higher-order epistasis in drug resistance. Mol. Biol. Evol. **38**, 142–151 (2021). https://doi.org/10.1093/molbev/msaa196
61. V.A. Meszaros, M.D. Miller-Dickson, C.B. Ogbunugafor, Lexical Landscapes as large in silico data for examining advanced properties of fitness landscapes. PLoS ONE **14**, e0220891 (2019). https://doi.org/10.1371/journal.pone.0220891
62. C.B. Ogbunugafor, The mutation effect reaction norm (mu-rn) highlights environmentally dependent mutation effects and epistatic interactions. Evolution (2022). https://doi.org/10.1111/evo.14428
63. S.K. Remold, R.E. Lenski, Pervasive joint influence of epistasis and plasticity on mutational effects in *Escherichia coli*. Nat. Genet. **36**, 423 (2004)
64. K.M. Flynn, T.F. Cooper, F.B. Moore, V.S. Cooper, The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. PLoS Genet **9**, e1003426 (2013)
65. A.L. Firth, T. Menon, G.S. Parker, S.J. Qualls, B.M. Lewis, E. Ke, C.T. Dargitz, R. Wright, A. Khanna, F.H. Gage, I.M. Verma, Functional gene correction for cystic fibrosis in lung epithelial cells generated from patient iPSCs. Cell Rep. **12**, 1385–1390 (2015). https://doi.org/10.1016/j.celrep.2015.07.062
66. C. Long, L. Amoasii, A.A. Mireault, J.R. McAnally, H. Li, E. Sanchez-Ortiz, S. Bhattacharyya, J.M. Shelton, R. Bassel-Duby, E.N. Olson, Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. Science **351**, 400–403 (2016). https://doi.org/10.1126/science.aad5725
67. B. Maher, Personal genomes: the case of the missing heritability. Nature **456**, 18–21 (2008). https://doi.org/10.1038/456018a
68. T.A. Manolio, F.S. Collins, N.J. Cox, D.B. Goldstein, L.A. Hindorff, D.J. Hunter, M.I. McCarthy, E.M. Ramos, L.R. Cardon, A. Chakravarti, J.H. Cho, A.E. Guttmacher, A. Kong, L. Kruglyak, E. Mardis, C.N. Rotimi, M. Slatkin, D. Valle, A.S. Whittemore, M. Boehnke, A.G. Clark, E.E. Eichler, G. Gibson, J.L. Haines, T.F.C. Mackay, S.A. McCarroll, P.M. Visscher, Finding the missing heritability of complex diseases. Nature **461**, 747–753 (2009). https://doi.org/10.1038/nature08494
69. W. Huang, T.F.C. Mackay, The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLoS Genet. **12**, e1006421 (2016). https://doi.org/10.1371/journal.pgen.1006421
70. J.R. Stinchcombe, H.E. Hoekstra, Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity (Edinb). **100**, 158–170 (2008). https://doi.org/10.1038/sj.hdy.6800937
71. A. Mäki-Tanila, W.G. Hill, Influence of gene interaction on complex trait variation with multilocus models. Genetics **198**, 355–367 (2014). https://doi.org/10.1534/genetics.114.165282
72. L. Hébert-Dufresne, S.V. Scarpino, J.-G. Young, Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement. Nat Phys. **16**, 426–431 (2020). https://doi.org/10.1038/s41567-020-0791-2