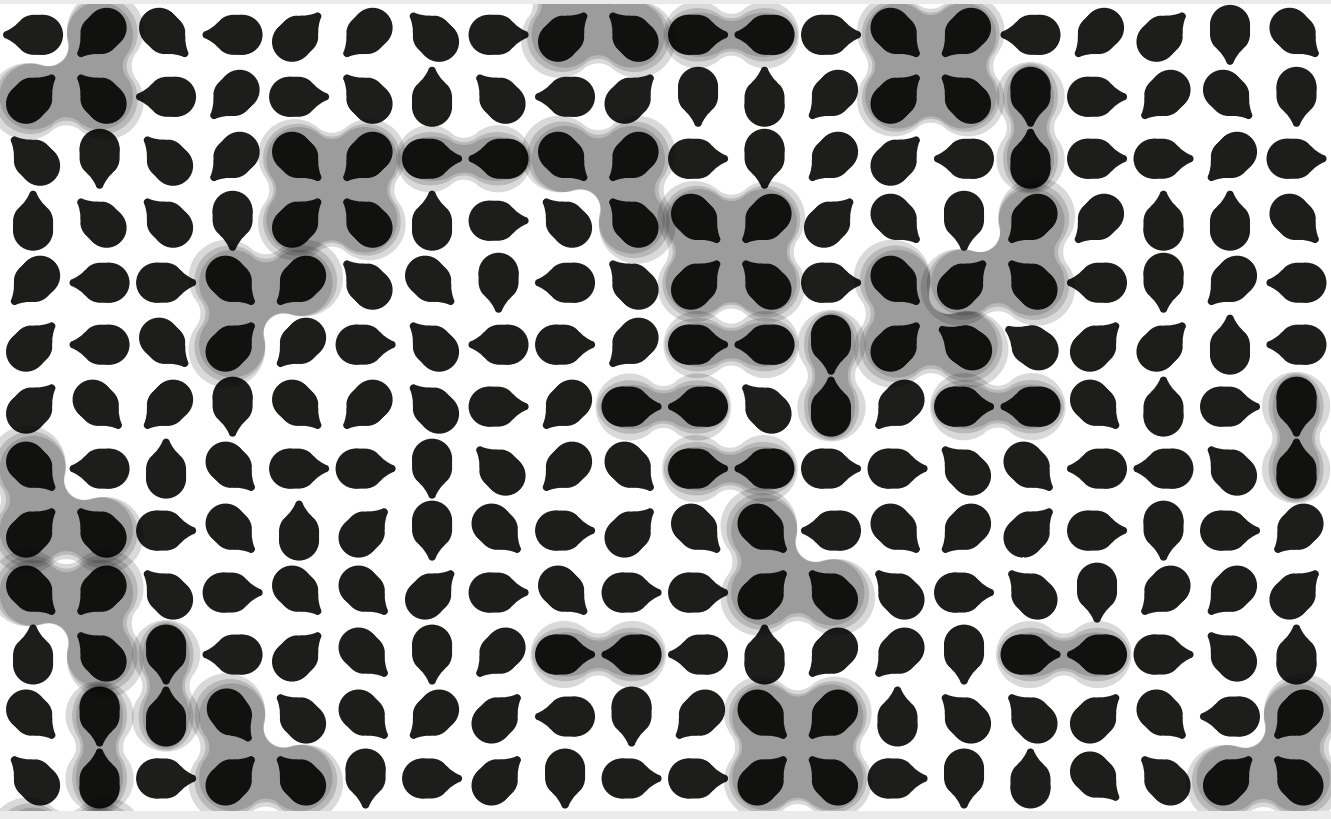
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS

Giulio Burgio

UNIVERSITAT ROVIRA i VIRGILI

Contagion Processes on Higher-Order Networks

Giulio Burgio



DOCTORAL THESIS
2024

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

Giulio Burgio

Contagion Processes on Higher-Order Networks

Ph.D. Thesis

Supervised by:
Dr. Alex Arenas & Dr. Sergio Gómez

Departament d'Enginyeria Informàtica i Matemàtiques

UNIVERSITAT
ROVIRA i VIRGILI

Universitat Rovira i Virgili
Tarragona, Catalonia, Spain

March 2024

Author:
Giulio Burgio

Supervisors:
Dr. Alexandre Arenas Moreno & Dr. Sergio Gómez Jiménez

Universitat Rovira i Virgili

Cover art by: Giulio Burgio

UNIVERSITAT ROVIRA i VIRGILI

We STATE that the present study, entitled *Contagion Processes on Higher-Order Networks*, presented by Giulio Burgio for the award of the degree of Doctor, has been carried out under our supervision at the Departament d'Enginyeria Informàtica i Matemàtiques of this university.

Tarragona, 19th March, 2024

Doctoral Thesis Supervisor,

Dr. Alexandre Arenas Moreno

Doctoral Thesis Supervisor,

Dr. Sergio Gómez Jiménez

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

To all my loved ones

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

ACKNOWLEDGEMENTS

There is no other way I could open this thesis than by expressing my sincere gratitude to the individuals and institutions that have supported me during these years.

First and foremost, I would like to thank my supervisors, Dr. Alex Arenas and Dr. Sergio Gómez, for their care and guidance. Alex, thanks for believing in me since my days at the Alephsys Lab as an Erasmus student—that started it all. I'll always be indebted to you for your unwavering academic and personal support. Sergi, thanks for your patience and kindness, the rigor and the meticulous attention to detail you taught me, and the constant technical help. I learned a lot from both of you, and I am so proud of what we have achieved together. I feel lucky to have had you both as supervisors.

Benja, Giacomo, Lluís, thank you for making me feel part of the group from the very beginning. Each of you has been a unique example to follow. Thanks for sharing your curiosity and scientific enthusiasm with me, for the many heartfelt laughs, and for your friendship. Mattia and Piergiorgio, I have felt closer to home since the 233 began Italian territory. Thanks for our enjoyable discussions, for all the shared memories, and for being my friends.

Thanks to all the other former and current members of the Alephsys Lab, each of you contributed to make my years here memorable. Joan, thanks for your technical support during my first year; I learned Julia from your beautiful scripts. Clara, thanks for all the fun and entertaining discussions and for your visualization tips. Albert, no lunch could be boring with you, and thanks for taking me to climb; that was my first and last time, but it was so fun! Alessio, thanks for helping me through the bureaucracy at the start of the Ph.D., for the great company, and for making me feel closer to our beloved Sicily. Aleix, thanks for the good times we had in Mallorca. David, Lluís ("el nou") and Nil, thank you for our short but enjoyable time together. I'm sure you'll enjoy the lab!

Moving to the other side of the ocean, I had the honor of visiting the Laboratory for Structure and Dynamics at the Vermont Complex Systems Center in Burlington under the supervision of Dr. Laurent Hébert-Dufresne. What a terrific environment—so many unique, kind, and talented people. Laurent,

I am extremely grateful for your continued support and for being such an extraordinary person inside and outside the lab. I learned a lot from you. Juniper, thanks for our exciting discussions. Your and Laurent's hospitality has been unmatchable. Jonathan, thanks for all the hours we spent together racking our brains between physics, sociology, and philosophy; what an exceptionally passionate researcher you are. Thank you (and Vanessa) for hosting me in your wonderful Quebec and for all our memories there (and at the border!). Above all, thanks for being a true friend. Thanks, Guillaume, Sam, and Tim, for the fascinating research we made and the time we shared. To Jean-Gab and Alice for the delicious dinners and game nights at their place, and to all the other former and current members of the VCSC with whom I had the pleasure of spending some good time, thank you.

I owe a special thanks to the European Union and the URV. The Martì i Franqués COFUND Doctoral Programme offered me ideal conditions for developing my thesis project, giving me the opportunity to participate in numerous international conferences and schools and to best disseminate our research. The Programme also allowed me to get to know some fantastic people. Two above all, Camilla and Vittoria, my favorite sociolinguists.

My time in Tarragona would not have been the same without Pao. Thanks for being there during my best and worst moments. You, Yasmina, and Genís welcomed me at your place as if it was mine, never letting me feel alone during the pandemic. Thanks for the great memories we have collected together over these years.

Thank you, Andi, for bearing with me these past few months; they have been hard. I'm deeply grateful for the patience, support, and love you showed me.

I close by giving my deepest thanks to my parents. This journey probably could not have started in the first place without you. Thanks for always respecting my choices and believing in me.

PUBLICATIONS

This thesis includes work by the author that has been published. These publications are the own work of the author of this thesis, and the author has the permission of the publishers to reproduce the contents of these publications for academic purposes. In particular, some results, data, opinions, and figures presented in this thesis have previously appeared as follows:

- Network clique cover approximation to analyze complex contagions through group interactions, G. Burgio, A. Arenas, S. Gómez and J. T. Matamalas, *Communications Physics* **4**(1), 11 (2021).
  https://doi.org/10.1038/s42005-021-00618-z
  (Section 3.1)

- Triadic approximation reveals the role of interaction overlap on the spread of complex contagions on higher-order networks, G. Burgio, S. Gómez and A. Arenas, *Physical Review Letters* **132**(7), 077401 (2024).
  https://doi.org/10.1103/PhysRevLett.132.077401
  (Section 3.2)

- Adaptive hypergraphs and the characteristic scale of higher-order contagions using generalized approximate master equations, G. Burgio, G. St-Onge and L. Hébert-Dufresne, *arXiv preprint arXiv:2307.11268* (2023).
  https://doi.org/10.48550/arXiv.2307.11268
  (Section 3.3)

- Homophily impacts the success of vaccine roll-outs, G. Burgio*, B. Steinegger* and A. Arenas, *Communications Physics* **5**(1), 70 (2022).
  https://doi.org/10.1038/s42005-022-00849-8
  (Section 4.2.1)

- Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics, G. Burgio, B. Steinegger, G. Rapisardi and A. Arenas, *Physical Review Research* **3**(3), 033128 (2021).
  https://doi.org/10.1103/PhysRevResearch.3.033128
  (Section 4.2.2)

- Spreading dynamics in networks under context-dependent behavior, G. Burgio, S. Gómez and A. Arenas, *Physical Review E* **107**(6), 064304 (2023).
  https://doi.org/10.1103/PhysRevE.107.064304
  (Section 4.3)

- Paradoxes in the co-evolution of contagions and institutions, J. St-Onge*, G. Burgio*, S. F. Rosenblatt, T. M. Waring and L. Hébert-Dufresne, *arXiv preprint arXiv:2310.03672* (2023).
  https://doi.org/10.48550/arXiv.2310.03672
  (Section 5)

- Evolution of cooperation in the presence of higher-order interactions: from networks to hypergraphs, G. Burgio, J. T. Matamalas, S. Gómez and A. Arenas, *Entropy* **22**(7), 744 (2020).
  https://doi.org/10.3390/e22070744
  (Section 5)

* These authors contributed equally.

CONTENTS

LIST OF FIGURES

ABSTRACT

We observe contagions at all levels of Nature. The spread of infectious diseases—mediated by the transmission of pathogens—is the paradigmatic example of a *contagion process*. Moreover, behaviors, beliefs, opinions, rumors, or conventions, among others, are contagious entities—mediated by mechanisms such as imitation or peer pressure—populating the human social sphere. Regardless of its nature, a contagion spreads via local interactions among individuals. If and how an extensive spread emerges is thus deeply related to the *complex network* of social interactions.

Differently than in biological contagions, transmission in social contagions often requires exposure to multiple contagious sources (e.g., individuals with a given behavior). Consequently, the contagion dynamics within a *group (higher-order) interaction* involving three or more individuals is generally not decomposable into a collection of pairwise interactions. Given the group organization of social life, accounting for the lack of decomposability becomes crucial to describe and understand social contagions adequately, as well as biological ones, as these typically coevolve with the former.

This thesis, overcoming inherent limitations of traditional dyadic descriptions, develops theoretical approaches to bridge the group organization of local interactions—encoded in *higher-order networks*—to the large-scale behavior of social and biological contagions. Using statistical physics and dynamical systems theory techniques, we characterize various ways group interactions significantly impact contagion dynamics.

We put forward three main contributions. Firstly, by accounting for local dynamical correlations, we demonstrate how both the critical point marking the onset of extensive spreads and the outbreak size are strictly related to the degree of overlap between dyadic and higher-order interactions and to the strength of the latter. Secondly, we develop a general framework that paves the way for the study of adaptive higher-order systems, enabling us to discover novel phenomena excluded in pair-based approaches. Lastly, recognizing the ecological notion of indirect modification as a more general higher-order mechanism, we unveil how the contextual character of prophylactic behavior can remarkably change the evolution of an epidemic.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

1 INTRODUCTION

People organize in groups, and contagions can spread across them. This observation naturally leads to several interrogatives about contagion dynamics—Under which conditions can a contagion extensively spread in a population? If it does, how many and which individuals it reaches? What is the role of groups? These and others are the fundamental questions delimiting the scientific problem we aim to address in this thesis.

To answer them, we will make some bold approximations out of necessity. Indeed, contagions propagate upon either physical proximity and contact or, if mediating information, also virtual interactions; and individuals differ from one another in many relevant aspects, such as their susceptibility to get a contagion and their ability to pass it on to others, as well as in the number, duration, size, and character of the interactions they are involved in. These and several other heterogeneities and factors make contagion processes too difficult to describe by physical models that we can handle or even define. Consequently, we need to resort to simplifying assumptions to make the problem tractable. This necessity, paraphrasing the chemist and philosopher Isabelle Stenger,[1] identifies an epistemological criterion to qualify the problem, hence the system encompassing it and the phenomena it concerns, as *complex*. The vast majority of phenomena we observe in Nature, from systems of strongly interacting particles to ecological and social systems, are complex. As a theoretical study of contagion processes, this thesis is a contribution to that interdisciplinary endeavor nowadays known as complexity science (or just 'complex systems').

1.1 PHYSICS OF COMPLEX SYSTEMS

The study of a complex system requires formulating abstractions sophisticated enough to retain essential properties of the system while neglecting (potentially relevant) others to keep the formulation tractable. Abstraction and quantification are the pillars of physical and mathematical investigation; therefore, it is not surprising that physicists and mathematicians have been pioneers in many aspects of theoretical research in complex systems.

1 Translated from an Edgar Morin's essay [1].

The 1972 'More is different' paper by the Nobel laureate in Physics Philip W. Anderson [2] is, de facto, a manifesto of complex systems science.[2] Starting from the experience of many-body physics (which studies systems made of a large number of strongly interacting particles, e.g., condensed matter), Anderson identifies the fundamental origin of the growth in complexity across the hierarchy of scientific domains, going from elementary particle physics to social sciences, with the corresponding increase of spontaneously broken symmetries. Ensembles of identical atoms constrained to a small enough volume unexpectedly rearrange themselves to form solids that break the continuous translational and rotational symmetries of space embodied in the electrodynamics of the atoms [3]. Specifically, when the solid is crystalline, the continuous symmetries reduce to the discrete symmetries of a geometric lattice. When the solid is instead amorphous (e.g., glasses), these symmetries become only short-ranged, for the geometric disposition changes over a few lattice cells. The symmetry of space can be broken not just geometrically but also by a heterogeneous distribution of information. An obvious example is—in the words of Anderson—the "information-bearing crystalliny" of DNA, for different nucleotides are (aperiodically) disposed along its double helix structure [4]. The symmetry of time is ubiquitously broken, too. Everyday-life examples are the approximate periodic beat of our hearths [5] and the synchrony of applause of an audience [6], while the recently theorized [7, 8] and soon realized [9, 10] time crystals offer a more exotic example. Fast-forwarding to human systems, space and time symmetries are so heavily broken that considering them is pointless—any two persons are ultimately different in every relevant aspect.

In general, each time we move from a more symmetric to a less symmetric system, dealing with it becomes more challenging, in that both the information needed to merely describe it and the dynamic richness it can display increase. This is a pretty obvious consequence of a reduction of symmetry. A less obvious one—which is the actual focus of Anderson's analysis—is its epistemological implication that a fundamental 'theory of everything'—whatever it will be—does not allow us to reconstruct everything we observe in the universe. In fact, it has minimal scope, as not even the spontaneous formation of a crystal can be derived from it, nor can any property of the crystal, like its low-energy excitation spectrum (i.e., sound waves). The manifest hierarchical structure of Nature, whose origin is a matter of eternal debate [11, 12], instead brings us to identify new properties and mechanisms at each level of the hierarchy—"the whole is something else than the sum of its parts" some Gestaltpsychologe would say [13], an expression indeed readapted by Anderson. Such structure—I think—is reflected in the structure of science.[3] As Robert B. Laughlin and David Pines write

2 Although complexity is central to Anderson's discussion, he never uses expressions like 'complexity science' or 'complex system.' This is perhaps one of the reasons why about three decades from its publication passed before the paper got widespread recognition (if quantified via citations) even from the physicists in the complex systems community.

3 In the words of W. C. Wimsatt [12], "Theories come in levels because that's where the entities are. [...] theories of entities at levels provide the biggest bang for a buck. These entities will be theoretically fruitful because of their many causal interactions, and the appropriate choice of entities

in another foundational piece of complexity science [14], "Rather than a Theory of Everything we appear to face a hierarchy of Theories of Things." Within certain space and time scales, a phenomenon is usually understood between the boundaries of two neighboring hierarchy levels, where the lowest one can be effectively regarded as fundamental. The propagation of some content in online platforms, for instance, can be studied regarding the users (humans, bots) as the fundamental units, for the information spreading is determined by the patterns of interactions among them, while factors such as the "virality" of the content effectively act as parameters (possibly changing coupled to the propagation). In other cases, we must resort to higher-level organizing principles such as free-energy minimization or functional development under natural selection to understand the configurations taken by the constituents.

In summary, while the validity of the reductionist view that every manifestation in Nature emerges from a hypothetical fundamental level of reality—described by the sought 'theory of everything'—should be out of the question, the constructionist one that any of those manifestations can be predicted starting from that level is blatantly false. The breakdown of the constructionist paradigm might not even be "just" a practical consequence of our (insurmountable) cognitive and technical limits; there are problems in classical [15–17] and quantum physics [18–20] (e.g., whether or not the energy spectrum of a quantum system has a spectral gap) which have been proven to be logically undecidable, in both an algorithmic (like the Turing's halting problem [21]) and a mathematical (independence from axioms; like the Gödel's incompleteness theorems [22]) sense.

The near decomposability of Nature into hierarchical levels allows for what we call 'physics of complex systems'. If we can consider the constituents and the interactions at a level as fundamental within the space and time scales of interest, then we can—and need to—build effective models at that level; informed mathematical abstractions to make quantitative predictions about a portion of the natural world—physical models, in other words. Dynamical systems, stochastic processes, statistical and nonlinear physics, and agent-based modeling, among others, thus turn out to be essential tools for the theoretical investigation of complex systems at virtually any scale. However, physicists, mathematicians, or computer scientists cannot walk through the intricacies of complex systems by themselves. It is the science of biologists, psychologists, sociologists, economists, or other specialists to make it possible to recognize the natural mechanisms and actors at work and eventually elaborate valuable models. Complexity science is itself a complex system, an interdisciplinary adventure built on the continued cross-pollination among different fields [23]. As physicists in this adventure, we aim to find minimal representations of a system and predict and characterize salient aspects of its behavior from there. Strikingly, thanks to the idealization presupposed by minimal models, it is sometimes possible to draw rigorous parallelisms among phe-

at levels will more often produce naturally segmented systems that are nearly decomposable [...]. Thus language [...] and theories constructed using and refining this language are in this way *responses to* rather than *determiners of* the structure of the world."

nomena of different nature or at even distinct levels, and trace them back to common underlying mechanisms.

1.2 COMPLEX SYSTEMS AS NETWORKED SYSTEMS

Some systems, such as gases at equilibrium, are characterized by a number of microscopic variables (e.g., position and momentum of each gas particle) too large to be traced. Nevertheless, thanks to the statistical independence among those variables, the system's behavior (e.g., temperature, pressure), as related to macroscopic average quantities, is well predictable [24]. This is what the applied mathematician and interdisciplinary research trendsetter Warren Weaver termed *disorganized complexity*, as opposed to the *organized complexity* of the majority of natural systems, where recurrent, nonrandom patterns of interactions create significant correlations among systems' units [25]. Organized complexity lies and evolves somewhere between complete randomness (e.g., a gas) and regularity (e.g., a crystal). Social systems provide a clear example of such a nontrivial organization. People—the system's constituents—interact primarily through their social ties (family members, partners, friends, colleagues, acquaintances, etc.), and all these ties persist over relatively long periods. The combination of the ties of many individuals is what we call a social network [26]. The latter is usually far from trivial. People differ substantially in the number, intensity, and type of social interactions they experience, and social connections are not distributed randomly, resulting in structural features such as cycles and communities or leading some individuals to be more "central" than others. Social networks are, in other words, *complex networks* [27–31]. Figure 1.1 illustrates two social networks from the studies of psychosociologists J. L. Moreno and H. H. Jennings [32, 33], fathers of social network analysis, and perhaps the first (excluding mathematicians studying graph theory) to use networks as a scientific tool.

Natural systems at all scales can be effectively modeled as networked systems, where nodes represent system's constituents and links represent pairwise interdependencies [27–31, 34]. The presence of a (complex) network is, together with dynamical nonlinearity, the main root of complexity. A crucial question of network science [28, 30, 35] is how much we can learn about the behavior of a system solely looking at the underlying network (or networks [36]) and, therefore, whether we can identify common structural principles and properties behind the phenomenology of different systems. Percolation—the problem of finding the size of clusters of connected nodes in a network—provides a minimal example of a (static) complex problem exclusively related to the connectivity structure [37, 38]. In general, complexity stems from the interplay between a network and the dynamics running on top of it [39–42]. Network and dynamics may also co-evolve, giving rise to even more complex behaviors [43–45] (this is perhaps the norm rather than the exception). Recognizing the network organization of many natural systems provided new essential understanding about phenomena across virtually any scientific domain, such as chemistry [46,

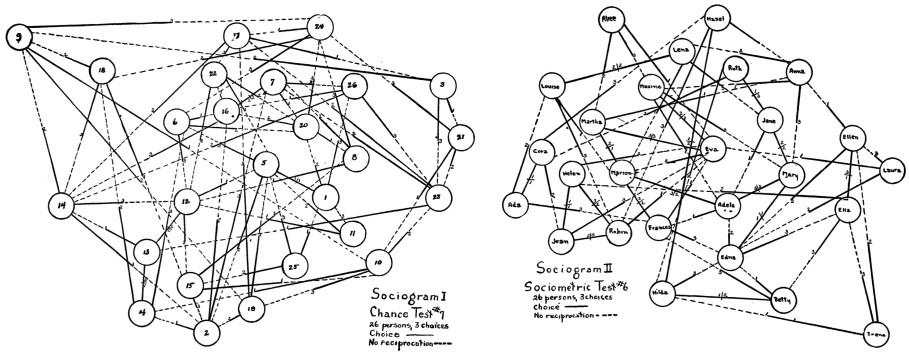


Figure 1.1: Small social networks (sociograms) by J. L. Moreno and H. H. Jennings [33]. The network on the left is generated randomly for statistical comparison with the one on the right, representing preference choices within a group of girls at the New York State Training School for Girls in 1932 [32]. In both networks, each node has three out-going (solid) links representing node's choices.

47], biology [48, 49], medicine [50, 51], ecology [52, 53], linguistics [54], economics [55, 56], finance [57, 58], political sciences [59], and—of course—sociology [60–63].

1.3 CONTAGION PROCESSES AND HIGHER-ORDER NETWORKS

Contagions cover processes of different natures. They can, for instance, be biological—conveying a pathogen—or social—conveying entities such as behaviors or beliefs. In a broad sense, a *contagion process* is the spread of some entity through a population.[4] The entity is transmitted upon interactions between "infectious" and susceptible individuals, potentially leading to a macroscopic, population-level spread. Contagions thus provide a straightforward bridge between physics and the animate world.

Depending on the rates at which individuals catch the contagious entity or "recover" from it, and on other relevant factors, a spread can reach a few individuals or a large proportion of the population. This can be rigorously restated for an infinite population by saying that either an infinitesimal or a finite fraction of the system is reached. In general, the system shifts from an inactive phase of negligible outbreaks to an active phase of extensive outbreaks when contagiousness crosses some critical value. As we will see, the latter specifically marks a *nonequilibrium phase transition* with an absorbing state [64]. The most basic physical models displaying such a transition are found in percolation theory, whose development in the late 50s was motivated also by contagion processes [65, 66]. Models of percolation on lattices as highly simplified descriptions of epidemics (applying to other "isomorphic" processes as well,

4 The words 'contagion', 'spreading', and 'diffusion' are generally regarded as synonyms. However, the former is often used when the medium through which an entity propagates is some kind of population, as is the case for human systems (e.g., people, users).

e.g., forest fires) were first elaborated in the 70s and 80s [67–71], and then extended to more realistic network topologies in the 2000s [72–78]. These and other developments established the study of spreading phenomena as a new research direction in statistical physics as well as in dynamical systems theory, expanding the scope of these fields.

As one may expect, the structural organization of the interactions is crucial for the spreading dynamics. Both its critical (e.g., critical exponents and scaling functions) and noncritical (e.g., outbreak sizes far from criticality) behaviors are intimately related to the properties of the interaction network [64, 79, 80]. It is not too much of a coincidence that the first use that Moreno and Jennings made of their sociograms (as those in Fig. 1.1) is a retrospective analysis of the outbreak of runaways registered in the fall of 1932 at the New York State Training School for Girls [32]. The authors, taking the connections in the social network as robust channels for the flow of social influence, showed that the position of a girl in the network was decisive in determining whether and when she ran away. While the notion of *behavioral contagion* was around from the end of the nineteenth century [81], Moreno and Jennings's study is probably the first to provide evidence for network effects on contagion dynamics, and presumably on any dynamics at all. As the girls in that study, we are typically unaware of who influences our behaviors and to what extent they do so, while we surely ignore our position in such a network of influences. Social networks thus act as invisible hands, shaping some of our ideas and behaviors often beyond our consciousness and control. This ultimately motivates the parallelism, yet with the due differences, between contagions in the social realm and epidemics. We will see how models initially conceived for the latter can be adapted to analyze the spread of behaviors, ideas, rumors, innovations, and other contagious entities populating our societies.

Taking into account networks has been vital to advance our understanding of contagion processes [80, 82–85], revealing the dynamical implications of common structural properties such as the presence of communities or highly connected nodes. Traditionally, networked systems have been described in terms of links representing pairwise interactions (dependencies) among the nodes. However, such a representation cannot account for interactions involving more than two nodes. To see this, imagine, for instance, that we want to describe a conversation among five people during which some dynamics (e.g., opinion formation) may unfold. The event is a group, five-node interaction. A pairwise network representation would reduce it to a collection of two-node interactions, leading to two fundamental problems. First, salient aspects of the dynamics (e.g., peer pressure) may be changed or lost, resulting in inaccurate outcomes. Second, even if the dynamics remained unaffected, the description would become vague in at least two ways: the group interaction, involving the five nodes simultaneously, would be decomposed in ten pairwise interactions that we cannot univocally locate in time (which in most cases also has dynamical consequences); and even if we do not bother about timing, it would be impossible to distinguish groups from collections of pairs, leading static descriptors (e.g., statistics about the number of people per

event) to be ill-defined. To adequately describe non-pairwise interactions and dependencies, we must resort to a more general representation where nodes can be grouped in subsets of unrestricted cardinality, not just in pairs. Such a representation is found in hypergraphs [86], the backbone of the recent branch of complex system science working with so-called *higher-order networks* and concerning systems at all scales [87–90]. In truth, the interest (outside pure mathematics) for hypergraphs, specifically in their equivalent form as bipartite networks, is not very recent [26, 27]. In most cases they have been used to account for highly dense regions (clusters, communities) often observed in real-world networks, unveiling important dynamical effects [76, 78, 91–95]. Nonetheless, the mechanisms defining the dynamics are still pairwise, i.e., decomposable in sets of two-body (two-node) interactions. Instead, when using the expression 'higher-order network', one generally means that the dynamics include 'higher-order', irreducible interactions, each reflected in the network as a generalized connection (called hyperedge) among three or more nodes. Recalling the example of the five-person conversation, one may consider it simply as a collection of dyadic conversations if the dynamics of interest were not affected by the decomposition. In this case, we would refer to the structure including such interaction as a (pairwise) network with clusters/modules. Instead, we would refer to it as a higher-order network if the dynamics were not decomposable.[5]

Theories and experiments in social psychology strongly suggest that dynamics within a group are not reducible to combinations of dyadic interactions [97–107]. The probability of adopting (being "infected" by) a behavior, an opinion, a judgment, or a decision, generally shows some nonlinear dependence on the number/fraction of adopters one is exposed to, thus preventing pairwise decomposition (this requiring linearity) of the contagion dynamics. Godwin and Restle [102], when introducing a dyadic model as one of several possible models of opinion formation, write: "One could argue that all human interaction takes place fundamentally between dyads [...]. Following this line of argument, it seems reasonable to assume that the interaction of an individual with a group of several persons [...] can be decomposed into one-to-one relationships between individuals, and it should be possible to understand the effects of a group as the combined effects of its individual members." Upon comparison of their experimental results with the predictions yielded by this (linear) model, they conclude that their findings doom "any attempt to explain these group data on the basis of any model couched merely in terms of dyadic relationships." They find strong nonlinear effects, although no simple nonlinear relation can be inferred from the data. Different nonlinear dependencies are generally plausible, and each one may or may not apply depending on the case [105, 106].

5 Notice that the sole irreducibility of the interactions does not suffice to qualify the network as 'higher-order'. Even if a node interacted with n other nodes through n distinct dyadic interactions, its dynamical state may depend on some nonlinear (hence irreducible) combination of the states of the other nodes (e.g., see [96]). The dynamics is not pairwise, but the network still is. A higher-order network includes irreducible interactions involving more than two nodes at once.

The irreducibility of group interactions is relevant not only for social contagions but also for biological contagions, at least in two ways. First, people may adopt prophylactic measures (e.g., wearing a face mask) to avoid infection. The nonlinearity of the adoption behavior then reflects on the epidemic spread. Second, models with nonlinear infection rates can work as effective representations for various situations where linearity breaks down [108–113], such as in environments (groups) saturated with infections or when multiple exposures are needed. In the same direction, recent work [114] has shown that a nonlinear infection rate naturally emerges as an effective description when averaging over environments with different infection risks (the differences being ascribable to several factors [115–119], e.g., the types of interactions the environment supposes or allows, or whether it is indoor or outdoor). Notably, such effective models result in simplified descriptions amenable to analytical scrutiny.

Acknowledged the general irreducibility of contagion dynamics in group interactions, this thesis represents our original contribution to advance the current understanding of the phenomenology of contagion processes on higher-order networks. By developing suitable theoretical approaches, we demonstrate how group interactions have a multifaceted impact on contagion dynamics.

1.4 THESIS OUTLINE

The remainder of this dissertation is organized into three main chapters. The first one, Chapter 2, makes—hopefully—the thesis self-contained by laying down the theoretical background about networks and contagion processes needed to follow the contributions presented in the subsequent chapters. Our developments are presented in Chapters 3 and 4. An extensive appendix including calculations, proofs, methods, and additional results assists the three chapters.

Chapter 3 considers contagions with group interactions involving two or more agents simultaneously. We explore higher-order contagions, specifically complex contagion processes based on reinforcement mechanisms, where transmission depends on exposure to multiple, distinguishable sources. All the models introduced in this chapter preserve local dynamical correlations. This crucial aspect, together with adequate modeling choices and methods, enables us to uncover and accurately characterize novel phenomena on static and adaptive higher-order networks.

Chapter 4 expands upon the concept of interaction modification, derived from ecological studies and here recognized as a broader higher-order mechanism applicable to complex systems in various domains. This concept entails the indirect alteration of the dyadic interaction between two agents due to the presence of others. Since the character and magnitude of the alteration depend on the observable traits of the copresent agents, the direct pairwise interaction between the same two agents would generally be different in distinct groups. Interacting in a group, given the co-presence of many agents it implies, thus becomes the necessary ground for such modifications and represents the context within which they occur. The chapter, in particular, delves

into the dynamical implications of context-dependent behavior on the spread of epidemics, revealing some significant effects that descriptions disregarding the group organization of the social sphere overlook.

A short, final chapter, Chapter 5, closes the thesis. There, we recap our results, briefly comment on other developments not discussed in the previous chapters, and identify a few research lines for future work based on the limitations of our and other currently available approaches.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

2 THEORETICAL BACKGROUND

This chapter provides the necessary theoretical background to follow the remainder of this thesis. It offers an introduction to network representations and to the modeling of contagion processes, from the simplest to some state-of-the-art techniques.

2.1 NETWORK REPRESENTATIONS OF COMPLEX SYSTEMS

Loosely speaking, a network is a collection of nodes connected in some way. Representing a system as a network means mapping elements of the system to nodes and relations among such elements to connections, links [28]. A network is thus an interconnected structure where nodes and links have some meaning. In this thesis, given its focus on microscopic physical mechanisms producing macroscopic emergent behaviors, a link always means an *interaction* between system units. The object obtained by stripping a network of the meaning of its elements is what in mathematics is called a *graph*, where nodes are known as vertices and links as edges. Nonetheless, outside pure mathematics, both terminologies are often used interchangeably and are even mixed (e.g., we may say "the edges of the network").

We first briefly introduce graphs and their properties [120], then discuss their higher-order generalization in the form of hypergraphs [86].

2.1.1 *Graphs*

A graph $G = (V, E)$ is a pair formed by a finite set $V = \{1, \dots, N\}$ of *vertices* and a set $E = \{e_1, \dots, e_M\}$ of pairs of vertices called *edges*. As a combinatorial object, a graph is one of the possible ways in which $N = |V|$ elements can be arranged into $M = |E|$ subsets of two elements each. The graph is encoded in its *adjacency matrix* $\boldsymbol{A} \equiv \boldsymbol{A}(G)$, a $N \times N$ matrix whose generic element A_{ij} at row i and column j is $A_{ij} = 1$ if exists $e = \{i, j\} \in E$ (in words, i and j are connected/neighbors) and $A_{ij} = 0$ otherwise. An equivalent representation is given by the *incidence matrix* $\boldsymbol{B} \equiv \boldsymbol{B}(G)$, a $N \times M$ matrix whose generic element is $B_{ij} = 1$ if $i \in e_j$ (we say that e_j is incident on i) and $B_{ij} = 0$ otherwise. Each column of $\boldsymbol{B}$ thus represents an edge of the graph. More generally, the nonzero elements of $\boldsymbol{A}$ and $\boldsymbol{B}$ can be real (or complex) numbers, measuring the

"weight" of an edge or of a vertex within an edge, respectively. In such case, we say that the graph is *weighted*.[1] Furthermore, a relation between two vertices could be not only asymmetric, but unidirectional. This means that for some pair i and j of vertices, there is an edge departing from i and pointing to j, but not the opposite one from j to i, which is encoded as $A_{ij} > 0$ and $A_{ji} = 0$ (in terms of incidence, $B_{jk} > 0$ and $B_{ik} = -B_{jk} < 0$ for some edge e_k). In such case, E is a set of ordered pairs and the graph is said to be *directed*. If every relation is instead symmetric (bidirectional), then $A_{ij} = A_{ji}$ for any pair i and j of vertices (i.e., $\boldsymbol{A}$ is symmetric; equivalently, $B_{ik} = B_{jk}$ for any edge e_k connecting i and j) and the graph is called *undirected*. In such case, as we implicitly assumed at the beginning, E is just a set of subsets of two elements each. Hereafter, unless stated differently, a graph is understood to be undirected and unweighted, in which case the adjacency and incidence matrices are binary matrices.[2]

Let us list some basic concepts associated to graphs that we will often find throughout the thesis. Generalizations to weighted or directed graphs are straightforward and reported whenever needed.

A subset of vertices where each vertex is connected to any other vertex in the subset is called a *clique*. If the subset consists of n vertices, we say it is a *n-clique*. A n-clique is thus a *complete* graph of n vertices (a 2-clique is just an edge). If a clique is not included in any larger clique, we call it *maximal*. A *n-cycle* is a closed chain of n edges joining n vertices (in particular, a 3-cycle is a 3-clique). A *tree* is a graph with no cycles.

Given a vertex i, the *degree* k_i is the number of edges incident on i, i.e., its number of neighbors. In terms of the adjacency and incidence matrices, we have $k_i = \sum_{j=1}^{N} A_{ji} = \sum_{j=1}^{M} B_{ij}$. Counting the fraction of nodes with degree k in the graph, we get its degree distribution p_k. A graph where all the vertices have the same degree k is called a *k-regular* graph.

Another basic quantity is the *clustering coefficient*, a measure of how frequently any two neighbors of a node are themselves neighbors, so that the three form a closed triplet (i.e., a 3-cycle). The local clustering coefficient C_i of vertex i is defined as $C_i = \sum_{j<k} A_{ij}A_{jk}A_{ki} / \binom{k_i}{2}$, i.e., as the ratio between the number of closed triplets including i and the maximum value $\binom{k_i}{2}$ that such number can take (number of unordered pairs out of k_i elements). Related global quantities are the mean of the distribution of the local clustering coefficient, and the fraction of triplets in the graph which are closed (known as transitivity).

Lastly, a graph is said *connected* if starting from any vertex it is possible to reach any other vertex in the graph through a sequence of edges (a *path*) connecting them. Each connected subgraph of a graph is called a *connected component*. If the graph is disconnected, then there are two or more connected components. A connected component whose number of vertices scales at least linearly with N in the large N limit is

1 Notice that this term is usually used to denote weights on edges only.
2 The edge set could contain repeated edges (e.g., $e_1 = e_2 = \{i, j\}$), defining a *multigraph*; or loops, which are edges connecting a vertex to itself. We do not consider these cases.

called a *giant connected component*, as it will include a finite fraction of the vertices in the graph.

Having insisted on the pairwise nature of graphs, we hope it is natural at this point thinking of more general objects by relaxing the pairwise condition.

2.1.2 *Hypergraphs*

A *hypergraph* $H = (V, E)$ is a pair consisting of a finite set $V = \{1, \dots, N\}$ of vertices and a set $E = \{e_1, \dots, e_M\}$ of subsets of vertices. These subsets can have any cardinality and are called *hyperedges*. The hypergraph H is thus a subset of the power-set 2^V. A hyperedge of cardinality n is also called a *n-edge*. The maximum and minimum cardinality of the edges in H are called the *rank* and *co-rank* of H, respectively, $\text{rank}(H) = m_{\max}$ and $\text{co-rank}(H) = m_{\min}$. If $m_{\max}(H) = m_{\min}(H) = m$, the hypergraph is said to be *m-uniform*. Any graph is thus a 2-uniform hypergraph. A hyperedge not included in any other hyperedge is said to be maximal. Figure 2.1 displays a small hypergraph.

A hypergraph can be encoded in a series of adjacency tensors $\{\boldsymbol{A^{(m)}}\}_{m=m_{\min},\dots,m_{\max}}$, each one accounting for the connectivity at the level of m-edges (for a graph, $\boldsymbol{A} \equiv \boldsymbol{A^{(2)}}$ is the adjacency matrix). Then, the element $A^{(m)}_{i_1,\dots,i_m} = 1$ if exists $e = \{i_1, \dots, i_m\} \in E$ and $A^{(m)}_{i_1,\dots,i_m} = 0$ otherwise. A much more concise codification of the hypergraph can be given in terms of the incidence matrix $\boldsymbol{B}$, which naturally represents connections of any order through its columns. As for graphs, it holds $B_{ij} = 1$ if $i \in e_j$ and $B_{ij} = 0$ otherwise. This is due to the equivalence between hypergraphs and bipartite graphs (see Fig. 2.1). A *bipartite* graph is a graph where the vertex set is partitioned into two subsets representing two types of vertices, and so that edges only connect vertices of different type. The equivalence then immediately follows by mapping the two subsets of vertices of the bipartite graph to the vertex set and the edge set of the hypergraph. The inclusion of a vertex into a hyperedge corresponds, in the bipartite graph, to that same vertex being connected to the vertex representing the hyperedge. The incidence matrix of the hypergraph is thus nothing but the adjacency matrix of the corresponding bipartite graph.

Since a vertex can be included in hyperedges of different cardinalities, the most basic definition of the degree of a node i is the *m-degree* $k_i^{(m)}$, being this the number of m-edges incident on i. This can be expressed as

$$k_i^{(m)} = \frac{1}{(m-1)!} \sum_{j_1,\dots,j_{m-1}=1}^{N} A^{(m)}_{i,j_1,\dots,j_{m-1}} = \sum_{j=1}^{M} B_{ij} \delta_{|e_j|,m} \,,$$

being $\delta_{a,b}$ the Kronecker delta. The *generalized degree* (or *hyperdegree*) k_i of vertex i, defined as the number of hyperedges of any cardinality incident on it, is then given by $k_i = \sum_{m=m_{\min}}^{m_{\max}} k_i^{(m)} = \sum_{j=1}^{M} B_{ij}$.

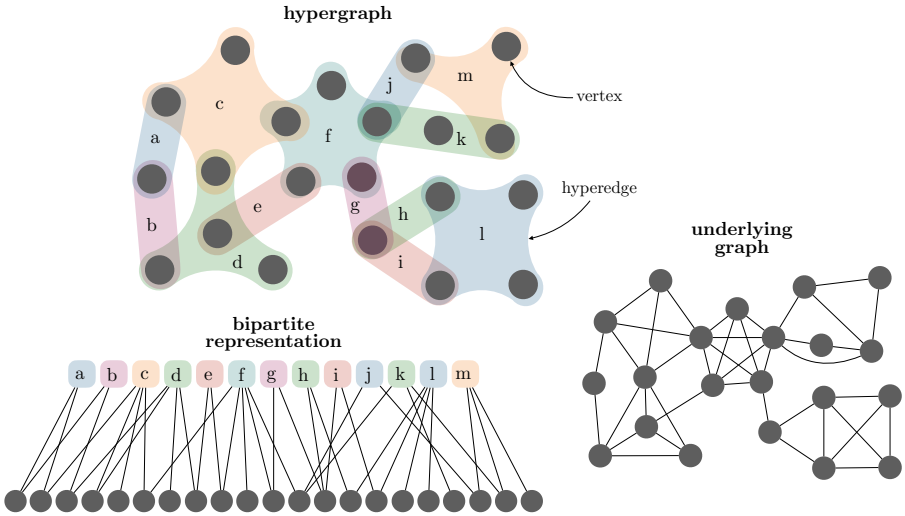


Figure 2.1: A hypergraph, its equivalent representation as a bipartite graph, and its underlying graph. Notice that the hypergraph in this example is linear.

A hypergraph H is called *simple* if no hyperedge is subset of any other hyperedge. This means that H has no repeated hyperedges and that can be obtained from another hypergraph by removing all the hyperedges which are subsets of at least another hyperedge. If H is simple and any two hyperedges share no more than one vertex, then H is said to be *linear*.[3] We will make massive use of the linear property in the next chapters. Another class of hypergraphs is that of (*abstract*) *simplicial complexes*, to which Sec. 2.1.2.1 is dedicated.

To a hypergraph H we can also associate a graph $G(V, E_2)$, called *2-section* or *underlying graph* of H, whose edge set E_2 contains all the edges subsets of at least a hyperedge in H (see Fig. 2.1). This means that a n-edge in H induces a n-clique in G. The (giant) connected components of H are clearly those of G, thus H is connected if and only if G is connected. The degree $k_i(G)$ of vertex i in G is bounded by its degree sequence $\{k_i^{(m)}\}_m$ in H, but generally not determined by it. An edge in G can indeed be induced by multiple hyperedges in H, implying that $k_i(G)$ depends on the local organization of the hyperedges incident on i, not just on their amount. Accordingly, an exception are linear hypergraphs, for which $k_i(G) = \sum_{m=m_{\min}}^{m_{\max}} (m-1) k_i^{(m)}$.

2.1.2.1 Simplicial complexes

Unlike generic hypergraphs, simplicial complexes received substantial attention from a broader community beyond pure mathematics. The main reason is that, besides being a combinatorial object, a simplicial complex is also a geometric object. As such,

3 Notice that a graph is trivially linear. In this sense, we may say that linear hypergraphs (of rank at least 3) are, among all hypergraphs, the "closest" to graphs.

it is the union of a set of simplices of various dimensions "glued together", where a n-dimensional simplex (or n-simplex) is defined as the convex hull of $n+1$ points, i.e., a generalization of a triangle. A 0-simplex is just a point, a vertex; a 1-simplex is a line segment together with the two vertices it connects; a 2-simplex is a triangle with its three edges (segments) and three vertices; a 3-simplex is a tetrahedron with its four faces (triangles), six edges (segments) and four vertices; and so on for higher dimensions. A simplicial complex results from a so-called triangulation of a topological space, a homeomorphism (a bijective, continuous function whose inverse is also continuous) mapping the latter into a piecewise linear space—the simplicial complex. Thanks to their dual nature, simplicial complexes thus provide a combinatorial approach to understanding the shape and structure of topological spaces, allowing for the translation of topological problems into algebraic ones, the essential idea behind algebraic topology [121]. For this reason, they play a pivotal role in topological data analysis, which aims at finding topological features of datasets by analyzing the simplicial complexes generated by triangulating the data points [122]. As a transversal framework, topological data analysis finds application in the most disparate fields, such as biology [123], neuroscience [124], fluid dynamics [125], machine learning [126], and the study of spreading processes [127], to name a few.

This said, many authors in the network science community present simplicial complexes as objects somewhat distinct from hypergraphs, even when regarding and using them solely as combinatorial entities. Strictly speaking, the term *simplicial complex* refers to the geometric realization of what is called an *abstract simplicial complex*, which is the actual combinatorial object, hence a specific kind of hypergraph. Nevertheless, the two terms are often used interchangeably outside mathematics. For conciseness, even though we only consider the combinatorial content, we will refer to abstract simplicial complexes simply as 'simplicial complexes' or 'complexes'.

A simplicial complex $\mathcal{K}$ on the vertex set V is a hypergraph (a subset of the power-set 2^V) endowed with the additional *hereditary property* (or *downward closure*): given $f \in \mathcal{K}$ and $f' \subseteq f$, then $f' \in \mathcal{K}$. The elements of $\mathcal{K}$ are called *faces*, and a n-dimensional face (or *n-face*) is a subset of V made of $n+1$ vertices. A n-face is thus a $(n+1)$-edge. In words, the hereditary property states that given a face (hyperedge) is included in the complex, then all its sub-faces (sub-hyperedges) are included too.[4] The power-set 2^f of a face f is what we called a *simplex*. For instance, the face $f = \{i,j,k\}$ generates the simplex $s = \{\{i,j,k\},\{i,j\},\{i,k\},\{j,k\},\{i\},\{j\},\{k\}\}$. If f is a n-face, 2^f is a *n-simplex*. If d is the maximum dimension of the faces in $\mathcal{K}$, then $\mathcal{K}$ is d-dimensional and is called *simplicial d-complex* (a hypergraph of rank $d+1$). A maximal simplex, not face of any other simplex, is called a *facet*. If all the facets in $\mathcal{K}$ have dimension d, $\mathcal{K}$ is a *homogeneous* d-complex.

A complex $\mathcal{K}$ is said to be q-connected if, given any two simplices $s, s' \subset \mathcal{K}$, there exists a sequence of simplices connecting s and s' such that any two adjacent simplices

4 This implies that the empty set is also part of a simplicial complex. However, as we do not assign any physical meaning to the empty set, we can just forget about it.

of the sequence have (at least) a q-face in common [128]. Notice that 0-connectedness means connectedness in the usual sense of graphs. Moreover, we conveniently say that $\mathcal{K}$ is *strictly q-connected* if $\mathcal{K}$ is q-connected but no pair of simplices is $(q+1)$-connected. Notice that the hypergraph whose hyperedges are the facets of a strictly 0-connected simplicial complex is linear.

2.1.3 *Some network models*

Let us now go through some key models of random graphs while discussing basic structural features discovered in networks derived from real-world systems.

Perhaps the simplest of these models is the Erdős-Rényi model. In its $G(N,p)$ formulation [129], it starts with N unconnected vertices and places an edge with probability p for each pair of vertices. Since the presence or absence of an edge is the result of a Bernoulli trial, the probability that a vertex has degree k is simply the probability of k successes in $N-1$ trials; hence

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-1-k} \,. \tag{2.1}$$

Being the degree a sum of independent and identically distributed random variables, average and raw second moment of the degree distribution read $\langle k \rangle = p(N-1)$ and $\langle k^2 \rangle = \langle k \rangle (1-p) + \langle k \rangle^2$. P. Erdős and A. Rényi proved that $G(N,p)$ undergoes a phase transition in the limit $N \to \infty$ at $\langle k \rangle = 1$, where a giant connected component emerges [130]. Taking the sparse network limit in Eq. (2.1), that is, $N \to \infty$ while keeping $\langle k \rangle$ fixed (hence $p = \langle k \rangle / (N-1) \to 0$), one finds that $p_k \simeq \langle k \rangle^k e^{-\langle k \rangle} / k!$, which is a Poisson distribution. The standard deviation equals $\sqrt{\langle k \rangle} < \langle k \rangle$, implying that the bulk of the distribution lies in the interval $\langle k \rangle \pm \sqrt{\langle k \rangle}$. The average $\langle k \rangle$ defines a typical scale for the degree of the nodes in a Erdős-Rényi network. The connections are, in other words, distributed in a relatively homogeneous way.

Networks generated from real-world data spanning systems of very different nature, from biological to socio-technical systems, show however degree distributions that are usually far from being binomial or poissonian. Instead, they are often well approximated by heavy-tailed distributions such as power-law and log-normal distributions [131], whose tails are not exponentially bounded. A generative model able to produce power-law degree distributions was proposed by A.-L. Barabási and R. Albert in 1999 [132].

In the Barabási-Albert model, starting from a small clique, a new vertex is added at each time step and randomly connected to m already present vertices. Each of the m vertices is sampled with a probability proportional to its current degree, a mechanism

known as *preferential attachment* (or *cumulative advantage* [133]). This generates a "rich-get-richer" growth dynamics that leads the degree distribution to take the form [134]

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)}\,, \tag{2.2}$$

which, for large k, converges to a power-law $k^{-\gamma}$ with exponent $\gamma = 3$.[5] Due to the scale invariance of power laws, which are thus "free" from any characteristic scale, a network with a power-law degree distribution is also referred to as a "scale-free network". Loosely speaking, such a network possesses a small minority of highly connected nodes (called "hubs" in jargon), as opposed to a large majority of lowly-connected ones. Approximating p_k to a pure power-law $k^{-\gamma}$, one can show that its n-th raw moment $\langle k^n \rangle$ diverges as a polynomial for $N \to \infty$ if $n > \gamma - 1$ (and logarithmically if $n = \gamma - 1$), remaining instead finite for $n < \gamma - 1$. In particular, the standard deviation diverges for $\gamma \leqslant 3$, implying that the more connected nodes (hubs) can have a number of connections of several orders of magnitude larger than the average node—there is no typical scale for the degree.[6] In scale-free networks the connections are therefore distributed heterogeneously among the nodes; more so the smaller is γ.

Based on the simplicity of the growth mechanism of their model and the evidence of power-law degree distributions in different empirical networks (World Wide Web, US power-grid network, actor collaboration network [132], and scientific citation network [136]), Barabási and Albert argued that power-law degree distributions are a universal feature of real networks [132]. Their work produced great excitement within the network science community, leading to the discovery of other scale-free networks, such as the Internet [137], metabolic networks [138], online social networks [139], and several others. Nonetheless, criticisms directed to their claim stating power-laws' universality were not long in arriving. Much of the debate stems from the difficulty of fitting power-laws to distributions derived from natural systems, which are often of too small size. In some cases, the distribution could be a power-law but be masked by the finite size of the sample (i.e., it would appear as a power-law if the system was bigger) [140], while in others, a different distribution could be a better fit [131]. The debate continues, yet there is no doubt that many real networks have heavy-tailed degree distributions [141].

Another structural feature observed in natural systems is the presence of local clustering. In social networks, in particular, if two individuals share a mutual friend, there is a higher probability that the two will form a social connection [142], a mechanism known as *triadic closure* [143]. The first random model able to generate networks with high local clustering while still respecting the so-called *small-world* property observed in social networks (and not only) is the one proposed in 1998 by D. J. Watts and S. H. Strogatz [144]. The small-world property, widely popularized as the 'six degrees of

5 Modifications of the attachment mechanism have then been shown to produce power-laws with different exponents γ [30].

6 Some authors [135] use the term "scale-free network" only in this case.

separation' [145], states that the average distance between two nodes in the network scales with the network size N not faster than $\log(N)$. Small-worldness characterizes networks in several scientific domains [146]. Many random networks, including those produced by the Erdős-Rényi and the Barabási-Albert models, boast this property but lack clustering [28]. The Watts-Strogatz model produces both. However, yielding networks interpolating between regular lattices and Erdős-Rényi networks, the Watts-Strogatz model cannot generate heavy-tailed distributions.

The Dorogovtsev-Mendes-Samukhin (or Dorogovtsev-Mendes) model [147] gives a straightforward rule able to generate random networks with either heavy-tailed distributions (implying small-worldness) and high local clustering. Started the network as a 3-clique; the rule makes the network grow by adding a vertex at each time step and connecting it to both vertices at the end of a randomly chosen edge. The addition of the vertex thus implies the addition of a 3-clique. Consequently, the network has a high local clustering coefficient and no tree-like regions. On the other hand, since the probability of finding a vertex at the end of a randomly picked edge is linearly proportional to the degree of the vertex, the rule naturally leads to preferential attachment. In the limit of infinite time steps, the degree distribution indeed converges to Eq. (2.2) evaluated at $m = 2$ (the number of vertices at the end of an edge).[7]

To close this section, let us introduce the *configuration model*, which allows the creation of networks with an arbitrary degree distribution [148–150]. It is done by directly specifying the degree of each vertex $i \in V$, all stored in the degree sequence $\{k_i\}_{i\in V}$. Assigning k_i half-edges ("stubs") to the generic vertex i, pairs of half-edges are drawn uniformly at random and combined to form full edges until no unmatched stubs are left.[8] Each degree sequence thus specifies an entire ensemble of random networks. Being the matching random, networks built from this basic version of the configuration model are also locally tree-like, therefore with vanishing clustering. Nonetheless, the model can be easily generalized to produce structures with 3-cliques; or, more generally, with some local motif of interest [151–154]. It all depends on the meaning we give to the stubs. We may specify a sequence whose i-th element represents the number of those motifs a vertex participates in and then suitably match such "generalized stubs". For instance, following this approach, Miller [78] builds networks with different levels of clustering by defining two degree sequences, a standard one for edges and another one for 3-cliques. We will use a similar technique to generate random hypergraphs with desired generalized degree sequences.

7 Notice that, for any finite value of the time step t, this behavior is observed for $k \ll \sqrt{t}$. The distribution has heavy tails anyway.

8 Multi-edges and loops may form in the process. The probability that this happens is, however, negligible for a large number of nodes as long as the second moment of the degree distribution does not diverge. In such case a structural cut-off (e.g., $k_{\max} \sim \sqrt{(N)}$) can be added.

2.2 CONTAGION PROCESSES

The word 'contagion' has its roots in the Latin verb 'contingere', from the assimilation of 'com', meaning 'with, together', and 'tangere', meaning 'to touch', thus carrying the idea of contact. The earliest usage of 'contagio' dates back to ancient Rome, as testified by its presence in Lucretius' epic poem "De Rerum Natura" (estimated 1st century BCE), where it is used to describe the transmission of diseases from one person to another through contact or proximity. The word probably did not leave its disease-related meaning until the end of the 19th century, with the seminal works by sociologists Gabriel Tarde and Gustave Le Bon. In Tarde's atomistic view of society, as he explains in his 'The Laws of Imitation' [155], social life involves the constant imitation of ideas, behaviors, and innovations among individuals—he writes, "Socially, everything is just inventions and imitations"—, giving rise to a complex web of interpersonal influences. Through his work, which unfortunately went unnoticed by many later scholars, he laid down concepts and mechanisms foundational for modern theories in sociology and evolutionary economics [156]. Although it had effectively been introduced by Tarde and Le Bon (the latter introduced the term 'behavioral contagion' [81]), the term *social contagion* was first used in 1939 by Herbert Blumer [157], one of the fathers of the study of spontaneous collective behaviors. From the second half of the 20th century, the term 'contagion' became indeed of common use in sociology to describe the spread of behaviors, opinions, rumors, innovations, norms [143, 158–168]. Between the end of the century and the early 2000s, it started to be used to indicate the spread of financial crises or shocks, either due to spillovers between interdependent market economies or to the behavior of financial agents like investors [169], and word-of-mouth marketing [170–172]. Eventually, with the rise of the digital age and social media, the term has further expanded to describe the spread of information, trends, or viral content online [173–176].

In a broad sense, a *contagion process* (more concisely, a contagion) is the spread of some entity through a population. The entity can be a pathogen, a behavior, an opinion, or anything that can be transmitted between the units of a system. Accordingly, the 'contact' conveying the contagion needs not to be physical as in infectious spreading; rather, it is understood as a generic interaction upon which an entity is transmitted.

Contagion processes can be classified as *simple* or *complex*. Different, non-equivalent criteria can be chosen to define the two. One criterion is based on how multiple exposures combine to determine the probability of getting infected. A contagion is simple if each exposure acts independently from any other exposure [112]. If $p\Delta t$ is the probability for a unit to get infected upon a single exposure during the time interval Δt, k exposures to infected units combine to give the total probability $f_k(p)$ of getting infected as $f_k(p) = 1 - (1 - p\Delta t)^k$. At the lowest order in Δt (continuous-time limit), this becomes $f_k(p) = kp\Delta t$. One thus may refer to simple contagions also as 'linear' contagions, for the infection probability has—in continuous time—a linear dependence on

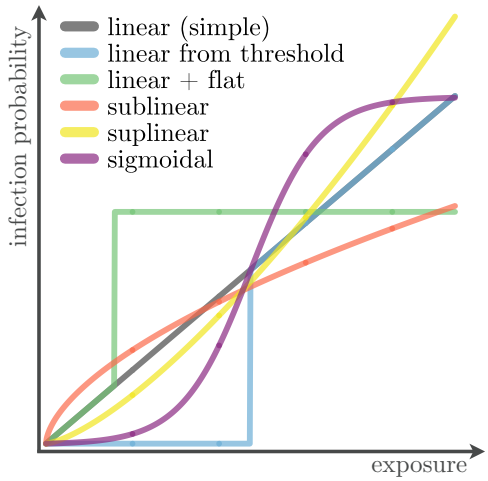


Figure 2.2: Examples of functional dependences of the infection probability on the amount of exposures. The latter can either refer to the absolute number of contagious exposures or to the fraction of the exposures which are contagious among those experienced in a given time span. Time is assumed continuous. Simple and complex contagions are in this figure discerned according to the first of the two definitions given in this section.

the number of exposures. Complex contagions are—trivially—those in which multiple exposures do not act independently, but their effect is instead interdependent [112]. We may call them also 'nonlinear', for—even in continuous time—the probability of getting infected after k exposures is no longer a linear function of k. The paradigmatic example of a complex contagion is one in which the infection probability is characterized by a threshold $\bar{k}$, such that at least $\bar{k}$ exposures are needed for transmission to be possible [165, 177] or to occur with much higher probability [178]. In general, a complex contagion is based on some positive or negative synergy mechanism among different exposures. Possible functional forms for the infection probability are illustrated in Fig. 2.2.

Another criterion to distinguish simple from complex contagions emphasizes the multiplicity of the contagious sources to which a unit is exposed [177] rather than the multiplicity of the exposures. According to this criterion, a contagion is simple if transmission is possible after one or more exposures (which may or may not combine synergistically) to the *same* contagious unit. A complex contagion is one where transmission requires exposure to two or more *different* contagious units. This classification thus shifts the boundary between simple and complex contagions given before, as some synergistic (nonlinear) contagions are now regarded as simple. On the other hand, contagions that are complex according to the second definition are also complex for the first one. The second is a stricter criterion. Models with indistinguishable units clearly cannot discriminate contagions based on the second definition.

Biological contagions are 'simple' based on the second criterion, for the distinguishability of the infection sources has no meaning at the biological level. What matters is the dose received, not the source of the dose. However, if a minimum infective dose is required for an infection, some of these contagions may be complex according to the first criterion [113]. Nonetheless, it is reasonable to assert that biological contagions can be broadly categorized as simple contagions based on both definitions [177].

Conversely, mounting empirical evidence supports the hypothesis that social contagions are often 'complex' according to the second criterion, hence to the first [164, 165, 173, 177, 179, 180]. Made some exceptions, such as the diffusion of job information [143] or of highly contagious rumors credible on first hearing [177], socio-cultural traits as many behaviors, opinions, or innovations, necessitate some form of social reinforcement for them to be successfully passed on [177]. The reason is that adopting them entails covering a cost of some sort (e.g., deriving from joining a protest, wearing a face mask, adopting an unpopular opinion, or investing in some new technology) that might not be rewarded eventually. Individuals are thus reluctant to bear that cost when no others have done it yet or without having enough reasons to believe that enough others will follow. Therefore, influence from two or more sources becomes crucial to transmitting adoption. Consequently, a minimum "critical mass" of adopters is often needed for many social entities to spread widely [181].

2.2.1 *The law of mass action*

Under the historical path we briefly outlined, the first mathematical analysis of contagions is considered the one put forward by Daniel Bernoulli in the second half of the 18th century [182], when he tried to estimate the benefits of variolation, the precursor of the smallpox vaccine, assuming a constant annual infection risk and case fatality risk across all age groups. As a culmination of a series of studies, developed during the 19th century until the early 20th century, aimed at understanding the contagion dynamics behind the observed epidemic curves [183–185], the works of R. Ross [186–188] and A. G. McKendrick [189–192] laid down the foundations of the modern theory of mathematical epidemiology. Their crucial contribution was transferring from chemistry to epidemiology the so-called *law of mass action*, which states that the rate of a chemical reaction is proportional to the product of the concentrations of the reactants [193, 194]. In epidemiological terms, infected and susceptible people are the "reactants" that, upon contact, produce additional infected people. According to the law of mass action, people are assumed to interact based on uniformly random contacts, analogously to how molecules are assumed to collide under Boltzmann's molecular chaos hypothesis in statistical physics. Consequently, the rate at which new infections occur is directly proportional to the fractions of infected and susceptible people in the population. We discuss the implications of the mass-action hypothesis in the sections

below, where we present basic models of simple (Sec. 2.2.2) and complex (Sec. 2.2.3) contagions.

2.2.2 *Simple contagions*

2.2.2.1 SIR model

The susceptible-infected-recovered (SIR) model is perhaps the most popular and used contagion model. It is a special case of the theory introduced in 1927 by W. O. Kermack and A. G. McKendrick [190]. In this model, a susceptible individual becomes infected at a rate β upon contact with an infected individual, as can be formalized in the stoichiometric equation $S + I \xrightarrow{\beta} 2I$. Therefore, $\beta\Delta t$ is the probability that the infection occurs during the small time interval Δt. Clearly, this probability should be proportional to the number of contacts k an individual has during the interval. Redefining β so that it is now a per-contact infection rate, the infection probability becomes $\beta k\Delta t$ (equivalently, we may regard β as a per-node infection probability and k as a contact rate). Moreover, infected individuals recover and obtain permanent immunity (or die; in both cases they are removed from the dynamics) at rate μ. Let $S(t)$, $I(t)$ and $R(t)$ be the fraction of susceptible, infected and recovered individuals in the population at time t, respectively, the dynamics is described by the following set of ordinary differential equations (ODEs),

$$\dot{S}(t) = -\beta k S(t) I(t) \,, \tag{2.3a}$$

$$\dot{I}(t) = \beta k S(t) I(t) - \mu I(t) \,, \tag{2.3b}$$

$$\dot{R}(t) = \mu I(t) \,, \tag{2.3c}$$

where $\dot{x}(t) \equiv dx(t)/dt$. Given the constraint $S(t) + I(t) + R(t) = 1$, only two of these equations are independent. Notice that an equivalent description in terms of numbers—not fractions—is obtained by multiplying both sides of Eqs. (2.3) by the population size N. Since we are treating the state variables as continuous, it is implicit in Eqs. (2.3) the infinite-size limit $N \to \infty$.

As it is often the case for nonlinear systems, although its apparent simplicity, we lack an explicit solution of Eqs. (2.3). We can however express two fractions (say S and I) in terms of the remaining one (R). Substituting $I(t)$ from Eq. (2.3c) into Eq. (2.3a), the latter becomes $\dot{S}(t) = -(\beta k/\mu) S(t) \dot{R}(t)$. This is a separable equation which, recast in the form $dS(t)/S(t) = -(\beta k/\mu) dR(t)$ and integrated from time 0 to t, gives

$$S(t) = S(0) \exp\left(-\frac{\beta k}{\mu} (R(t) - R(0)) \right) \,. \tag{2.4}$$

From this, $I(t) = 1 - S(t) - R(t)$ is just a function of $R(t)$, with parameters $S(0)$ and $R(0)$. Let us suppose that only few infected individuals are present in the population at $t = 0$. This means taking the limit $I(0) \to 0$, hence, $S(0) \to 1 - R(0)$. Besides,

imposing $\dot{R}(t) = 0$ implies that I must be zero at equilibrium. In the large time limit, $t \to \infty$, once the system approached stationarity, we thus have $S_\infty = 1 - R_\infty$, where we defined $S_\infty \equiv S(t \to \infty)$ and $R_\infty \equiv R(t \to \infty)$. The latter is known as the *final attack rate* of the epidemic, as it is the fraction of individuals that got the infection—recovering eventually. Therefore, at equilibrium, from Eq. (2.4) we get

$$R_\infty = 1 - (1 - R(0)) \exp\left[-\frac{\beta k}{\mu}\left(R_\infty - R(0)\right)\right] . \tag{2.5}$$

We see that $R_\infty = R(0)$ ($S_\infty = 1 - R(0)$) is a trivial solution of Eq. (2.5). In this case the infinitesimal initial seed of infected individuals is not able to produce an extensive outbreak. On the other hand, we see that taking $R_\infty = 1$, the r.h.s. of Eq. (2.5) is smaller than the l.h.s. for any finite value of $\beta k/\mu$. Since both sides are continuous functions of R_∞, a nontrivial solution $R_\infty > R(0)$ exists if and only if the r.h.s. grows faster than the l.h.s. at $R_\infty = R(0)$. Differentiating Eq. (2.5) at $R_\infty = R(0)$, one finds that this occurs when

$$\beta k(1 - R(0))/\mu > 1 . \tag{2.6}$$

When fulfilled, this condition implies that a finite fraction of individuals gets infected during the outbreak. To see whether this equilibrium is stable or not, let us perform a linear stability analysis of Eqs. (2.3). Linearizing around a disease-free state $(S, I, R) = (1 - R(0), 0, R(0))$, Eq. (2.3b) becomes $\dot{I}(t) = (\beta k(1 - R(0)) - \mu)I(t)$, which has solution $I(t) = I(0)\exp[(\beta k(1 - R(0)) - \mu)t]$. This increases exponentially when the condition in Eq. (2.6) is satisfied, proving that the disease-free state is unstable. If instead $\beta k(1 - R(0))/\mu < 1$, the system goes back exponentially to the disease-free state, which is thus stable in this case. In summary, $\beta k(1 - R(0))/\mu = 1$ is the threshold condition separating the disease-free, *inactive phase*, from the epidemic, *active phase*. The threshold value $\beta_{\text{cr}} = \mu/[k(1 - R(0))]$ is called *epidemic* (*invasion*) *threshold*, for an epidemic develops (invades the population) if and only if $\beta > \beta_{\text{cr}}$. Moreover, if the latter holds, imposing $\dot{I}(t) = 0$ yields the state of the system at $t = \hat{t}$, the time at which the infected fraction $I(t)$ is maximal. From Eq. (2.3b), we find $\hat{S} \equiv S(\hat{t}) = \beta_{\text{cr}}/\beta$, which plugged in Eq. (2.4) yields $\hat{R} \equiv R(\hat{t}) = R(0) + (\beta_{\text{cr}}/\beta)\ln(S(0)\beta/\beta_{\text{cr}})$, and in turn, $\hat{I} \equiv I(\hat{t}) = 1 - \hat{S} - \hat{R}$. The phenomenology of the system is summarized in Fig. 2.3.

Let us make some general considerations starting from this analysis. First, from the point of view of the theory of dynamical systems (see Appendix A.1 for the basics of linear stability analysis), a disease-free state is stable if none of the eigenvalues of the Jacobian matrix associated to Eqs. (2.3), evaluated around $(S, I, R) = (1 - R(0), 0, R(0))$, has a positive real part. The matrix has two zero eigenvalues (recall there is only one free state variable) and another one given by $\beta k(1 - R(0)) - \mu$, which is in fact positive if and only if Eq. (2.6) is satisfied. The two equilibria $R_\infty = R(0)$ and $R_\infty > R(0)$ exist for any value of the parameters and they exchange their stability at $\beta k(1 - R(0))/\mu = 1$. The threshold condition thus marks a transcritical bifurcation.

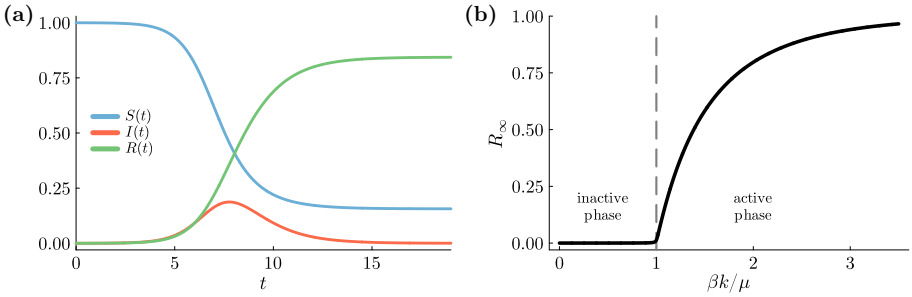


Figure 2.3: Phenomenology of the SIR model. (a) Time evolution of the fraction of susceptible, $S(t)$, infected, $I(t)$, and recovered, $R(t)$, individuals in the population for $\beta k = 2.2$ and $\mu = 1$ ($R_0 = 2.2$). (b) Final attack rate, R_∞, as a function of $\beta k/\mu$. The system is initialized setting $I(0) = 10^{-4}$ and $S(0) = 1 - I(0)$ ($R(0) = 0$).

Second, the deterministic dynamics above, describing the evolution of the average state of the system (specified by $S(t)$, $I(t)$, and $R(t)$), holds only in the infinite-size limit, $N \to \infty$. For any finite N, instead, there are non-negligible statistical fluctuations around that average (see Sec. 2.2.4 for a detailed discussion). As a consequence, there is a nonzero probability that the number of infected individuals deviates from its infinite-size behavior and fluctuates down to zero at any moment. Once the system falls into an inactive state $(S, I, R) = (1 - R, 0, R)$, it cannot leave it, for no state transition can happen without infectious sources, no matter the size of the system. That is, an inactive state is an *absorbing state*. In a finite system there is always a finite probability flux from a non-absorbing (active) state to some absorbing one (inactive), but there is no flux in the opposite direction. In the language of statistical physics, the equilibrium condition of *detailed balance* is thus broken, and the transcritical bifurcation corresponds to a *nonequilibrium phase transition* [64]. All of this holds true for all the models considered in the remainder of this thesis, as they all admit at least one absorbing state.

Third, the quantity $\beta k/\mu$ has a precise meaning. The ratio $1/\mu$ is the average infection period, i.e., the average time interval between the time an individual gets infected and the time it recovers. Throughout this period, the infected individual transmits the infection to each of its susceptible contacts with probability β/μ. In a population where all other individuals are susceptible ($R(0) = 0$), the infected individual therefore infects, on average, $\beta k/\mu$ individuals. This quantity, called *basic reproduction number* and usually denoted by R_0, is thus the average number of secondary cases caused by a primary case [195]. When $R_0 = \beta k/\mu > 1$, the primary case generates more than one other case before recovering. Each of these secondary cases will do the same (the pool of susceptible individuals is approximately still the same), giving rise to an initial exponential growth of cases. More generally, if not the whole population is initially susceptible, for instance because a fraction $R(0)$ of it is vaccinated

or isolated, then $S(0) = 1 - R(0) < 1$ ($R(0) > 0$), and the quantity $R_0 S(0)$ is called just *reproduction number*. The general condition $R_0 S(0) = 1$ is thus an alternative and clearly interpretable way of stating the threshold condition for the spectrum of the Jacobian matrix. We refer the reader to Appendix A.1.1 for a detailed and more rigorous discussion about the (basic) reproduction number, its definition as the dominant eigenvalue of the so-called *next generation matrix* [196], and the relation between the latter and the Jacobian matrix.

Lastly, we may ask where is the contact network in the model. There isn't actually one, but rather an infinite number of them. According to the hypothesis of a random contact process, each individual does not interact persistently with the same set of individuals. Contacts rather occur uniformly at random between any two individuals in the population. In network terms, contacts can be thought of as links being each time reshuffled. In fact, we only specified the number (or rate) of contacts k, what in a network is the degree of a node. And k is the same for every node, so the network is k-regular. At every instant, however, there is a new k-regular network, which can be thought of as a new realization of the configuration model. Implicitly, we are in fact making an assumption of time-scale separation, where the time scale at which the network of contacts evolve is much smaller than the time scale at which the contagion occurs. In network science, this assumption is known as the *annealed network* approach, in analogy with annealed disorder in solid state physics [197, 198]. We will see this approximation in action repeatedly throughout this thesis.

2.2.2.2 SIS model

The SIR model is adequate when infection confers permanent immunity to the recovered individual. This is the case for the smallpox or the plague, but not, for instance, for syphilis or influenza [82]. In these cases, recovery provides immunity only for a limited period, after which the individual is again susceptible. Social and cultural traits are generally not definitive. Behaviors, opinions, trends, can be adopted, abandoned and readopted. In a social context, "reinfection" is thus usually possible.

The most simple example with reinfection is the susceptible-infected-susceptible (SIS) model, where immunity is lost immediately. The equations for the SIS can be obtained from those for the SIR by simply identifying the R state with the S state. The flux entering $R(t)$ (Eq. (2.3c)) is thus redirected into $S(t)$ (Eq. (2.3a)). Using $S(t) = 1 - I(t)$, the SIS is governed by the following equation,

$$\dot{I}(t) = [\beta k(1 - I(t)) - \mu] I(t) \ . \tag{2.7}$$

Contrarily to the SIR model, we can find an explicit formula for either the time evolution and the equilibria of the SIS model. Imposing $\dot{I}(t) = 0$ we find two equilibria. The trivial, absorbing $I = 0$ and the nontrivial $I^\star = 1 - 1/R_0$, where, as in the SIR model, $R_0 = \beta k/\mu$ is the basic reproduction number. The nonzero solution is known as the *endemic* equilibrium and denotes the situation in which the infection remains

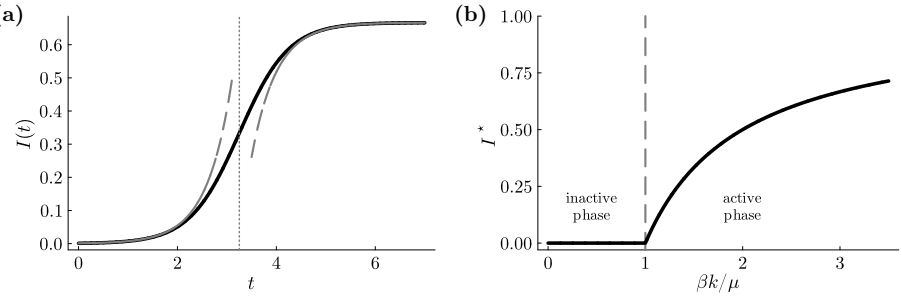


Figure 2.4: Phenomenology of the SIS model. (a) Time evolution of the fraction $I(t)$ of infected individual in the population (notice $S(t) = 1 - I(t)$) for $\beta k = 3$ and $\mu = 1$ ($R_0 = 3$). The dashed curves indicate the two functions which approximate $I(t)$ during the initial exponential growth and the final exponential saturation. The vertical dotted line marks the inflection point $\tilde{t} = (\beta k I^\star)^{-1} \log(I^\star / I(0) - 1)$. Set $I(0) = 10^{-4}$, the system converges to the endemic state $I(t \to \infty) = I^\star = 1 - 1/R_0$. (b) Endemic equilibrium as a function of $\beta k/\mu$.

constantly present in the population. The system converges to the active, endemic state when $R_0 > 1$, and to the inactive state when $R_0 < 1$, undergoing a transcritical bifurcation at $R_0 = 1$. Notice that the absorbing state is always reached after a sufficiently long time when the system has a finite size. The transcritical bifurcation marks a nonequilibrium phase transition and the active solution is a metastable state, i.e., a state where the system persists for a time interval diverging (exponentially) with system's size. For a large enough system, we can therefore regard the active solution effectively as stable (albeit small fluctuations around it will observed).

To solve Eq. (2.7), we make the change of variable $I(t) = u(t)^{-1}$, assuming that $I(t)$ never vanishes, hence $R_0 > 1$. Substituting $I(t) = u(t)^{-1}$ into Eq. (2.7) and multiplying both sides by $u(t)^2$, we get the first-order equation $\dot{u}(t) = -(\beta k - \mu)u(t) + \beta k$. Solving first the homogeneous equation and then plugging the solution in the complete equation, one easily gets the solution $u(t)$, hence $I(t) = u(t)^{-1}$. This reads

$$I(t) = I(0) \frac{I^\star}{I(0) + (I^\star - I(0)) \exp(-\beta k I^\star t)} \,. \tag{2.8}$$

Taking $I(0) \to 0$, we see that $I(t)$ approximates as follows,

$$I(t) \approx \begin{cases} I(0) e^{\beta k I^\star t} & \text{for } t \ll \tilde{t} \,, \\ I^\star \left[1 - \left(\dfrac{I^\star}{I(0)} - 1 \right) e^{-\beta k I^\star t} \right] & \text{for } t \gg \tilde{t} \,, \end{cases} \tag{2.9}$$

where $\tilde{t} = (\beta k I^\star)^{-1} \log(I^\star / I(0) - 1)$. That is, $I(t)$ interpolates over time between two exponential curves, an initial growth and a final saturation. Differentiating twice Eq. (2.8) and requiring $\ddot{I}(t) = 0$, one finds that $t = \tilde{t}$ is exactly the inflection point of the curve $I(t)$, with $I(\tilde{t}) = I^\star/2$. Notice also that the first expression in Eq. (2.9) is

the solution of the linear limit of Eq. (2.7), $\dot{I}(t) \simeq \mu(R_0 - 1)I(t)$. When $R_0 < 1$, $I(t)$ goes to zero exponentially. Finally, if $R_0 = 1$ (i.e., $\beta k = \mu$), we find from Eq. (2.7) the solution $I(t) = I(0)[1 + I(0)\beta k t]^{-1}$, vanishing as $1/t$. The phenomenology of the model is reported in Fig. 2.4.

2.2.3 *Complex contagions*

2.2.3.1 Dodds-Watts generalized contagion model

Dodds and Watts [112] defined a generalized model of contagion encompassing and interpolating between models of simple and complex contagions, according to the first (but not the second) definition given in Sec. 2.2. To this end, the authors introduce a memory of past exposures to the contagious entity.

An individual can be susceptible, infected or recovered, as in the SIR model. At each discrete time step t, each individual i interacts with another individual j chosen uniformly at random from the population (mass-action). If i is susceptible and j is infected, i receives, with probability β, a positive dose $s_i(t)$ randomly drawn from a dose size distribution $f(s)$. Each individual i cumulates the doses received over the past T time steps, resulting in a cumulative dose $S_i(t) = \sum_{t'=t-T+1}^{t} s_i(t')$. Each individual is endowed with a dose threshold s_i^* drawn from $g(s^*)$ at $t = 0$ and fixed for all future times. If i is susceptible and $S_i(t) \geqslant s_i^*$, i becomes infected. If infected, i recovers with probability μ once $S_i(t) < s_i^*$. Once recovered, it becomes susceptible again with probability γ.

The probability $P_{\text{inf}}(K)$ that a generic susceptible individual becomes infected after $K \leqslant T$ encounters with infected individuals over the last T steps, reads

$$P_{\text{inf}}(K) = \sum_{k=1}^{K} \binom{K}{k} \beta^k (1-\beta)^{K-k} P_k \,, \tag{2.10}$$

where

$$P_k = \int_0^\infty ds^* g(s^*) p_k(s^*) \tag{2.11}$$

is the probability that a generic individual gets infected after receiving k positive doses, being $p_k(s^*)$ the probability that the sum of k doses drawn from $f(d)$ is equal to or greater than s^*.

The standard SIR model is recovered taking $f(s) = \delta(s - \bar{s})$ and $g(s^*) = \delta(s^* - \bar{s}^*)$ with $\bar{s} = \bar{s}^*$, so that $P_k = 1$ for any k and $P_{\text{inf}}(K) = 1 - (1-\beta)^K$, which is one minus the probability of not receiving the infection from any of the K contacts. Taking $\bar{s} < \bar{s}^*$ we get a sharp threshold model ($P_k = 1$ if $k \geqslant \bar{s}^*/\bar{s}$ and 0 otherwise), whereas generic, smooth threshold models are obtained choosing distributed dose sizes and thresholds (P_k a generic increasing function of k). The resulting dynamics is quite complex even for the simplest choices. Taking however $\mu = \gamma = 1$, we get rid of the R compartment, obtaining a SIS-like dynamics. Then, an individual is currently

infected with probability P_k if received a number k of positive doses over the past T steps, being susceptible otherwise. Since the joint probability that the individual chosen uniformly at random for contact is infected and that such individual transmits a positive dose is βI, we get the following implicit equation

$$I = \sum_{k=1}^{T} \binom{T}{k} (\beta I)^k (1-\beta I)^{T-k} P_k \equiv F(I) \,, \tag{2.12}$$

valid at stationarity. The solutions of Eq. (2.12) are the equilibria of the system. The inactive state $I^\star = 0$ is always a solution. Let us suppose then $I \neq 0$. Using $(1-x)^n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} x^i$, collecting the terms of the same order in I, and dividing both sides by I, Eq. (2.12) becomes

$$0 = T\beta P_1 - 1 + \sum_{k=2}^{T} \binom{T}{k} h_k \beta^k I^{k-1} \,, \tag{2.13}$$

where

$$h_k = \sum_{l=1}^{k} (-1)^{k-l} \binom{k}{l} P_l \,. \tag{2.14}$$

Let n be the smallest value of $k \geqslant 2$ such that $h_k \neq 0$. Then, considering I small, solving Eq. (2.13) at the lowest order in I (which is proportional to $h_n I^{n-1}$), yields

$$I^\star \simeq \left[\frac{1 - T\beta P_1}{\binom{T}{n} h_n \beta^n} \right]^{\frac{1}{n-1}} \,. \tag{2.15}$$

First of all, $I^\star$ goes to zero for $\beta \to \beta_{\rm cr} = (TP_1)^{-1}$. For this to be the invasion threshold, $I^\star$ must be positive and stable for $\beta \to \beta_{\rm cr}^+$. Positivity requires $h_n < 0$. For instance, supposed $n = 2$, this means $P_1 > P_2/2$. Below $\beta_{\rm cr}$, the solution is instead negative, thus unphysical, and the only equilibrium is the inactive state $I^\star = 0$. The latter is stable if and only if $dF/dI < 1$ at $I^\star = 0$. Consistently, we find $dF/dI = \beta/\beta_{\rm cr}$, so that the inactive state is stable for $\beta < \beta_{\rm cr}$ and unstable for $\beta > \beta_{\rm cr}$. In the latter case the positive solution provided by Eq. (2.15) is the stable equilibrium. Therefore, the point $\beta = \beta_{\rm cr}$ marks a transcritical bifurcation where the zero and the nonzero solutions exchange stability. If instead $h_n > 0$ (e.g., $P_1 < P_2/2$ for $n = 2$), Eq. (2.15) provides a positive solution for $\beta < \beta_{\rm cr}$ and a negative one for $\beta > \beta_{\rm cr}$. Given that the inactive state is stable for $\beta < \beta_{\rm cr}$ independently from h_n, and that the positive solution existing for $\beta < \beta_{\rm cr}$ is the nearest nonzero equilibrium to the inactive state, it follows that this solution is unstable. Taking the limit $\beta \to \beta_{\rm cr}$, at first order in $\beta - \beta_{\rm cr}$, we can rewrite the positive solutions (stable or not) given by Eq. (2.15) as

$$I^\star \simeq \left[\binom{T}{n} \beta_{\rm cr}^{n+1} \right]^{-\frac{1}{n-1}} \left| \frac{\beta - \beta_{\rm cr}}{h_n} \right|^{\frac{1}{n-1}} \sim |\beta - \beta_{\rm cr}|^{\xi} \,, \tag{2.16}$$

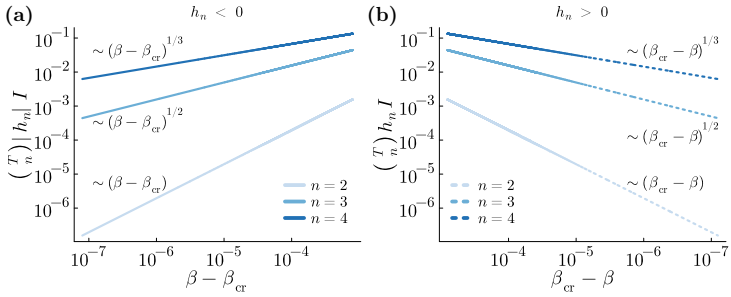


Figure 2.5: Critical scaling, as given by Eq. (2.16), of the positive solutions of Eq. (2.12) around the invasion threshold $\beta = \beta_{cr}$, for different values of n. (a) $h_n < 0$ (including class I models, $h_2 < 0$), for which the positive solutions emerging from the continuous phase transition at $\beta = \beta_{cr}$ are stable (solid lines). (b) $h_n > 0$ (including class II models, $h_2 > 0$), for which the positive solutions represent the unstable lower branch of equilibria emanating at $\beta = \beta_p < \beta_{cr}$ and crossing zero at $\beta = \beta_{cr}$, where a discontinuous transition exists.

where we introduced the critical exponent $\xi = 1/(n-1)$. The latter implies that the nature of the transcritical bifurcation depends on n, which is in turn determined by how the different probabilities P_k compare with each other. Figure 2.5 reports the solutions given by Eq. (2.16), illustrating the scaling for different n as dictated by ξ. The bifurcation has the same nature of the one found in standard SIR and SIS models, for which $\xi = 1$, when $n = 2$, i.e., for $P_1 > P_2/2$. Accordingly, the condition $P_1 > P_2/2$ identifies those models that in Dodds *et al.* [112] are called class I, or 'epidemic threshold' models. Complex contagions in this regime qualitatively behave as simple contagions.

As already observed, when instead $P_1 < P_2/2$ (hence $n = 2$), the solution identified by Eq. (2.16) is positive in a region below β_{cr}, where it is unstable. Being Eq. (2.13) of order $T - 1$ in I, other $T - 2$ solutions exist, some of which may be physical (i.e., real and positive). Assuming only one of them is acceptable, this third solution lies above the latter and is thus (locally) stable. This implies the existence of another threshold $\beta_p < \beta_{cr}$, called *persistence threshold*, marking a saddle-node bifurcation from which the two nonzero solutions emanate.[9] The stable equilibria form an upper branch of solutions whose value increases with β, whereas the lower branch identified by the unstable equilibria decreases, crossing zero at $\beta = \beta_{cr}$. The interval $\beta \in [\beta_p, \beta_{cr}]$ corresponds to a bistability region where, for each β, there are two locally stable equilibria separated by an unstable one. To reach the upper solution, then the initial fraction $I(0)$ of infected individuals must lie above the lower solution at the considered β. The lower solution represents the "critical mass" of infected individuals needed to reach the active state. Going through $\beta = \beta_{cr}$, then $I^\star$ jumps from zero to the upper solution

9 If more than two nonzero physical solutions exist, additional saddle-node bifurcations appear. Nonetheless, this is usually excluded by plausible values of the probabilities P_k.

(now globally stable), describing a discontinuous phase transition. Since the critical mass vanishes approaching $\beta = \beta_{\text{cr}}$, Dodds *et al.* [112] call these class II models, identified by the condition $P_1 < P_2/2$, also 'vanishing critical mass' models. Boasting this critical-mass effect, complex contagions described by class II models have a qualitatively different behavior than any simple contagion.

At the intersection of the two classes, corresponding to the case $P_1 = P_2/2$ ($h_2 = 0$), the transcritical and saddle-node bifurcations (the persistence and invasion thresholds) may or not coincide, leading the phase transition to be continuous or discontinuous respectively. These models are regarded as class I in Dodds *et al.* [112]. Equation (2.16), however, reveals that such classification is partially imprecise. Supposing that the transition at the transcritical bifurcation is continuous, then the nature of the bifurcation is different than in simple contagion models, for $\xi = 1/(n-1) < 1$ (recall $n > 2$), as acknowledged in Ref. [112] itself. The imprecision lies in assuming that the condition $P_1 = P_2/2$ guarantees the continuity of the transition. As indeed observed when analyzing Eq. (2.16), a necessary and sufficient condition for the transition to be continuous (discontinuous) is that $h_n < 0$ ($h_n > 0$). Therefore, claiming that the transition for a model with $P_1 = P_2/2$ is continuous necessitates verifying that $h_3 < 0$. If $h_3 > 0$, the transition is discontinuous; if $h_3 = 0$, we must look at the sign of h_4. More generally, we must check the sign of the first nonzero h_k, which is h_n by definition. Nonetheless, being $h_2 = 0$ ($P_1 = P_2/2$) a special case, having also $h_m = 0$ for $m > 2$ requires a very fine tuning of the probabilities P_k. Indeed, one can prove by induction that for the generic h_k to be zero, $P_1 = P_l/l$ must hold for any $l \in \{2, \dots, k\}$.[10] Inserting $P_l = lP_1$ in the expression for h_k, this can be recast in the form $h_k \propto \sum_{l=1}^{k} (-1)^l \binom{k}{l} l$, which is identically zero.[11]

A class III of models, called 'pure critical mass' models in Ref. [112], also exists, as identified by the condition $P_1 < 1/T$, implying $\beta_{\text{cr}} > 1$. Being the invasion threshold inaccessible, the inactive state is always (locally) stable. Therefore, if a saddle-node bifurcation exists, the lower branch of unstable equilibria remains nonzero for any β. This means that a strictly positive critical mass of infected individuals is always required for the system to reach an active state (the upper branch). The condition for the saddle-node bifurcation to appear, and with it the stable branch of positive solutions, depends on the involved combination of the probabilities P_k as given in the r.h.s. of Eq. (2.15).

The three classes of models can be clearly discerned by plotting the curve $I \mapsto F(I)$ against the identity line $I \mapsto I$. We illustrate this in Fig. 2.6. The intersection points are the equilibria. An equilibrium $I^\star$ is then stable (unstable) if $F(I)$ crosses the identity

10 A pathological case exists; the one in which $P_1 = P_l/l$ for any l from 2 up to T. From Eqs. (2.12) and (2.13) follows that the inactive state is then the only equilibrium, unless $\beta = \beta_{\text{cr}}$ and $P_T = 1$ (implying $\beta_{\text{cr}} = 1$), in which case any $I \in [0, 1]$ is a neutrally stable equilibrium.

11 *Proof.* Using $\binom{k}{l} = \frac{k}{l}\binom{k-1}{l-1}$, it follows $h_k \propto \sum_{l=0}^{k-1} c_{k-1,l}$, with $c_{k-1,l} = (-1)^l \binom{k-1}{l}$. Using the Pascal's rule $\binom{k-1}{l} = \binom{k-2}{l} + \binom{k-2}{l-1}$, we have $c_{k-1,l} = c_{k-2,l} - c_{k-2,l-1}$. It follows $h_k \propto c_{k-1,0} + c_{k-1,k-1} + \sum_{l=1}^{k-2} c_{k-2,l} - \sum_{l=1}^{k-2} c_{k-2,l-1} = c_{k-1,0} + c_{k-1,k-1} - c_{k-2,0} + c_{k-2,k-2}$, obtained by adding and subtracting both $c_{k-2,0}$ and $c_{k-2,k-2}$. From $c_{k-1,0} = c_{k-2,0} = 1$ and $c_{k-1,k-1} = -c_{k-2,k-2}$, then $h_k = 0$.

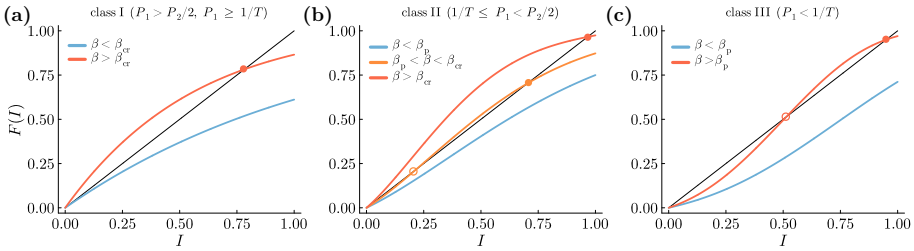


Figure 2.6: Stability plots for the Dodds-Watts model. $F(I)$ against I for (a) class I ($P_1 > P_2/2$, $P_1 \geqslant 1/T$), (b) class II ($1/T \leqslant P_1 < P_2/2$), and (c) class III ($P_1 < 1/T$) models. The identity line $I \mapsto I$ is represented by a solid black line, whereas colored solid curves denote $I \mapsto F(I)$ for various values of β below or above the invasion and persistence thresholds of each class. Filled and empty circles indicate respectively stable and unstable nonzero equilibria, identified by $I = F(I)$.

line from above (below) at the respective intersection, as this corresponds to $dF/dI < 1$ ($dF/dI > 1$) at $I^\star$.

2.2.3.2 Simplicial contagion model

Introduced by Iacopini *et al.* [199], the simplicial contagion model recognizes group, many-body interactions as an alternative mechanism responsible for complex contagions. Differently from previous models, where pairwise influences are separate interactions and are thus cumulated over time, this model accounts for the synergistic effect given by the *simultaneous* presence of two or more contagious units in a group. Contagions described by this model are thus complex according to both definitions given in Sec. 2.2. The model is defined for interaction structures represented by simplicial complexes, where groups are mapped into simplices, hence its name. Nonetheless, the authors analyze the qualitative behavior of the system through a simple ODE that, relying on the mass-action assumption, generalizes the ODE describing a SIS process, Eq. (2.7).

In this model, in addition to two-body interactions (links, 1-faces) in which a susceptible individual gets the infection from an infected one at rate $\beta^{(1)}$, a susceptible individual gets also infected at rate $\beta^{(n)}$ in a $(n+1)$-body interaction (n-face) with n simultaneously infected individuals, i.e., while taking part to a group where at least n members are infected. Supposed m is the size of the largest groups in the system ($m-1$ would be the dimension of the simplicial complex), the evolution of the fraction of infected individuals in the population, $I(t)$, is then governed by

$$\dot{I}(t) = -\mu I(t) + (1 - I(t)) \sum_{n=1}^{m-1} \beta^{(n)} k^{(n)} I(t)^n \,, \tag{2.17}$$

being μ the recovery rate and $k^{(n)}$ the number of $(n+1)$-body interactions involving an individual in any time interval Δt. Equation (2.17) is an ODE of order m. Imposing stationarity we thus get an algebraic equation of order m in I. Besides the trivial solution $I^\star = 0$, $m-1$ nonzero (but not necessarily physical) solutions exist. The pairwise SIS model is recovered setting $m = 2$.

Let us take $m = 3$ to get the minimal non-pairwise model. Linearizing around $I = 0$, Eq. (2.17) becomes $\dot{I} \simeq (-\mu + \beta^{(1)}k^{(1)})I$, implying that the inactive state $I^\star = 0$ changes stability at $\beta^{(1)} = \mu/k^{(1)} \equiv \beta_{\rm cr}$, being unstable above (notice that this holds for any m). The latter is the same invasion threshold found in the pairwise SIS model. In other words, in this mean-field approximation, the stability of the inactive state is predicted to be independent from any higher-order infection rate.

The two nonzero equilibria are immediately found as the solution of a quadratic equation, and read

$$I^\star_\pm = \frac{\beta^{(2)}k^{(2)} - \beta^{(1)}k^{(1)} \pm \sqrt{(\beta^{(2)}k^{(2)} - \beta^{(1)}k^{(1)})^2 - 4\beta^{(2)}k^{(2)}(\mu - \beta^{(1)}k^{(1)})}}{2\beta^{(2)}k^{(2)}} \,. \tag{2.18}$$

These are real for $\beta^{(1)} \geqslant \left(2\sqrt{\mu\beta^{(2)}k^{(2)}} - \beta^{(2)}k^{(2)}\right)/k^{(1)} \equiv \beta_{\rm p}$. The threshold $\beta_{\rm p}$ takes its maximum at $\beta^{(2)} = \mu/k^{(2)}$, at which it equals $\beta_{\rm cr}$, being strictly smaller otherwise. At $\beta^{(1)} = \beta_{\rm p}$, $I^\star_\pm = 1 - \sqrt{\mu/\beta^{(2)}k^{(2)}}$, which is nonnegative if and only if $\beta^{(2)} \geqslant \mu/k^{(2)}$. Besides, $I^\star_+$ increases and $I^\star_-$ decreases with $\beta^{(1)}$. Being $I^\star_\pm$ positive at $\beta^{(1)} = \beta_{\rm p}$ when $\beta^{(2)} > \mu/k^{(2)}$, $I^\star_+$ remains positive for any $\beta^{(1)} \geqslant \beta_{\rm p}$, whereas $I^\star_-$ decreases with $\beta^{(1)}$ and crosses zero (from above) at $\beta^{(1)} = \beta_{\rm cr}$, being negative thereafter. When $\beta^{(2)} \leqslant \mu/k^{(2)}$, instead, since $I^\star_\pm$ is nonpositive at $\beta^{(1)} = \beta_{\rm p}$, $I^\star_-$ is negative for any higher value of $\beta^{(1)}$, whereas $I^\star_+$ increases with $\beta^{(1)}$ and crosses zero (from below) at $\beta^{(1)} = \beta_{\rm cr}$, remaining positive from there on.

Linearizing around $I^\star_\pm$ by taking $I = I^\star_\pm + \delta I$, one gets $\dot{\delta I} \simeq -I^\star_\pm(I^\star_\pm - I^\star_\mp)\delta I$, whose sign is equal to $\mp{\rm sgn}(\delta I)$, given $I^\star_\pm \geqslant 0$ and $I^\star_+ \geqslant I^\star_-$. Consequently, when physical, $I^\star_+$ is stable and $I^\star_-$ is unstable.

The results are summarized in Fig. 2.7. When $\beta^{(1)} < \beta_{\rm p}$, there is only one equilibrium, the inactive state $I^\star = 0$. If $\beta^{(2)} > \mu/k^{(2)}$ and $\beta^{(1)} \in [\beta_{\rm p}, \beta_{\rm cr})$, the inactive state is still locally stable, while two solutions $I^\star_+, I^\star_- > 0$ emerge, being respectively stable and unstable. Then, when $\beta^{(1)} > \beta_{\rm cr}$, the inactive state is unstable, $I^\star_-$ negative, hence nonphysical, and the only stable physical equilibrium is $I^\star_+$.

This is the same qualitative phenomenology predicted by the Dodds-Watts model discussed in the previous section. Epidemic-threshold (class I) models are found for $\beta^{(2)} \leqslant \mu/k^{(2)}$, as the system undergoes a unique, transcritical bifurcation at $\beta^{(1)} = \beta_{\rm cr}$, where $I^\star = 0$ and $I^\star_+$ exchange stability. Vanishing-critical-mass (class II) models are instead found for $\beta^{(2)} > \mu/k^{(2)}$, as a saddle-node bifurcation emerges at $\beta^{(1)} = \beta_{\rm p}$

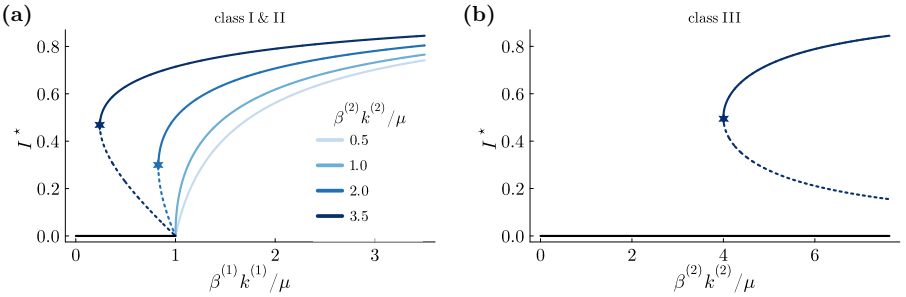


Figure 2.7: Phase diagram of the simplicial contagion model. (a) Class I ($\beta^{(2)} \leqslant \mu/k^{(2)}$) and class II ($\beta^{(2)} > \mu/k^{(2)}$) models according to the classification given in Sec. 2.2.3.1. The two classes merge at the tricritical point $(\beta^{(1)}, \beta^{(2)}) = (\mu/k^{(1)}, \mu/k^{(2)})$. (b) Class III model, obtained setting $\beta^{(1)} = 0$. Blue solid (dotted) curves denote the stable (unstable) branch $I^\star_+$ ($I^\star_-$), whereas $I^\star = 0$ is reported by a black solid line whenever stable. The saddle-node bifurcations at β_{p} are marked by star-shaped points.

and $I^\star_\pm > 0$, implying bistability in the interval $\beta^{(1)} \in [\beta_{\mathrm{p}}, \beta_{\mathrm{cr}})$.[12] The critical mass $I^\star_-$, needed by the system to converge to the upper solution $I^\star_+$, indeed vanishes at $\beta^{(1)} = \beta_{\mathrm{cr}}$, above which $I^\star_+ > 0$ is the only stable equilibrium. Class I and class II coalesce at the tricritical point $(\beta^{(1)}, \beta^{(2)}) = (\mu/k^{(1)}, \mu/k^{(2)})$, where the transcritical and the saddle-node bifurcations merge.[13] Lastly, pure-critical-mass (class III) models correspond to $\beta^{(1)} = 0$, for which $I^\star = 0$ is (locally) stable (linearizing, $\dot{I} \simeq -\mu I < 0$) and an unbounded bistability region exists, $\beta^{(2)} \in [4\mu/k^{(2)}, \infty)$, for the lower solution $I^\star_-$ decreases with $\beta^{(2)}$ but never touches zero. Notice that the simplicial contagion model has been recently extended to 4-body interactions, finding regions of tristability [200]. Again, such regions are found also in the Dodds-Watts model by increasing the period T over which doses are cumulated [112].

Given this qualitative equivalence, one might be tempted to conclude that knowing whether the multiple exposures need to be towards different sources is irrelevant to the spread of a complex contagion. Accounting for group interactions, where the distinguishability of the sources matters, seems then a marginal detail. In this regard, it is crucial to notice that Eq. (2.17) is agnostic about whether the n exposures making up the force of infection ($\beta^{(n)} I(t)^n$) come from n different sources or if some of them repeatedly come from the same source. Therefore, even though the model is defined to describe contagions that are complex according to the second (hence first) definition given in Sec. 2.2, the above analysis, by relying on the mass-action assumption, holds identical also for contagions that are complex by the first definition only. For the same

12 Interestingly, $\beta^{(2)} \geqslant 4\mu/k^{(2)}$ implies $\beta_{\mathrm{p}} \leqslant 0$. In such case, once reached the upper solution, the system persists in an active state even when $\beta^{(1)} = 0$, exclusively supported by three-body infections.

13 The critical exponent ξ of the bifurcation becomes $1/2$ at merging, as in the Dodds-Watts model when $h_3 < 0$.

reason, the analysis is blind to how pairwise and group interactions are organized in the interaction structure, thus making the same predictions for all those hypergraphs sharing the same set of average degrees. Such indeterminacies are resolved by making individuals explicitly distinguishable, as when identifying them with nodes on a quenched network (see Sec. 3.1), or implicitly so by accounting for the dynamical correlations that a network creates (see Sec. 3.2), correlations that the mass-action hypothesis thoroughly neglects. By doing so, we will reveal dynamical effects exclusively ascribable to the higher-order organization of the interactions. This means that, even if we chose the parameters in the Dodds-Watts model and the simplicial contagion model in order for them to give the same outcome (e.g., same equilibrium prevalence), we could still detect the presence of group interactions, for the outcome would be altered by changing the arrangement (but not the number) of the interactions in the structure.

2.2.4 *From Markov chains to compartmental models*

The models we analyzed in Secs. 2.2.2.1-2.2.3.2 are deterministic models, described by ODEs. Nonetheless, being a model of contagion an effective theory, we talked about transition (infection and recovery) probabilities. Let us therefore briefly clarify the connection between the probabilistic and the deterministic description.

Let us consider a system made of N units. Denoting with $x_i(t)$ the state of unit i at time t, $x_i(t)$ is a random variable taking $\Sigma \in \mathbb{N}$ possible values. Therefore, the system's state vector $x(t) = (x_1(t), \dots, x_N(t)) \in \mathcal{X}$ can be found in one of $|\mathcal{X}| = \Sigma^N$ possible configurations, being $\mathcal{X}$ the system's state space. Let $P(x,t) \equiv P(x(t) = x|x(0) = x_0)$ be the probability that the system is in state x at time t given it started in state x_0 at time $t = 0$. Supposing time to be continuous, if for any x and t, $P(x, t + dt)$ only depends on the state distribution $\{P(x,t)\}_{x\in\mathcal{X}}$ at time t (Markov property), then the trajectory $x(t)$ defines a *continuous-time Markov chain*, also known as *jump process* [201]. The temporal evolution of $P(x,t)$ is thus governed by the following *master equation* (Kolmogorov forward equation) [202],

$$\dot{P}(x,t) = \sum_{x'\in\mathcal{X}} \left[W(x|x')P(x',t) - W(x'|x)P(x,t)\right] , \tag{2.19}$$

where $W(x|x')$ is the transition rate from state x' to state x. The two terms in the sum represent the probability fluxes towards and from state x, respectively. Having assumed that the transition rates do not depend explicitly on time (they do only implicitly, via the system's states), Eq. (2.19) defines a *homogeneous* Markov chain. Once chosen x_0 and specified a form for the transition rates, the integration of the master equation provides the exact temporal evolution of the entire probability distribution $\{P(x,t)\}_{x\in\mathcal{X}}$. There is a number of $\Sigma^N - 1 \approx \Sigma^N$ equations of the form Eq. (2.19) that we need to integrate. This number, as an exponential of N, becomes ridiculously

high even for the simplest model with $\Sigma = 2$ and a population of few dozens of units, preventing the computational integration of the master equation for even small systems.

In many cases, however, we are interested in the state of aggregates rather than of single units. Thus, we can reduce the problem's dimensionality by grouping the units in m *compartments* (i.e., sub-populations). The compartments are treated as equivalence classes, so any two units in the same class should have any of their properties identically distributed at any time t. Further individual properties which may influence the units' behavior are consequently neglected. The resulting model is called a *compartmental model*. In a social system, for example, the classes are typically defined by individual features such as age, income, or employment status, in addition to the dynamic state of individuals. Notice that in Secs. 2.2.2.1 and 2.2.2.2, we implicitly assumed $m = 1$, meaning that all the individuals have identical features—hence, there is no point in retaining them—and can only differ for their dynamic state—which is what discern the compartments. Since the individual features are typically fixed or change very few times during the period we observe a system, we can conveniently think of the generic compartment k as partitioned in Σ sub-compartments, each corresponding to a specific dynamic state (e.g., the two sub-compartments of the susceptible and infected individuals within the compartment of individuals of high income). If the size of each compartment is extensive (i.e., $\propto N$), then one can prove that the total number of equations to integrate no longer scales exponentially, but as a power of N, precisely as $N^{\Sigma-1}/[(\Sigma-1)!]^m$. This derives from the fact that the system's state is now entirely specified by the state configuration of each compartment, which is given by counting how many units there are in each of the Σ sub-compartments. Denoted with $n_i^{(k)}$ the number of units of compartment k in state i, the state of the system is then $n = (n_1^{(1)}, \dots, n_\Sigma^{(1)}, \dots, n_1^{(m)}, \dots, n_\Sigma^{(m)})$. In other words, the state of each compartment (hence of the system overall) is described by a set of birth-death processes [84]. For instance, for infection dynamics, an infection event corresponds to the "birth" of an infected unit in a given compartment, leading the counts of susceptible and infected units in that compartment to decrease and increase by one, respectively. The stochastic evolution of the system is governed by a master equation analogous to Eq. (2.19), substituted x with n.

The last step to get to a deterministic description in the form of ODEs consists in taking the infinite-size limit, $N \to \infty$, as guaranteed by the following crucial results. Oppenheim *et al.* [203] proved for some paradigmatic examples of birth-death processes, including SIR-like (hence SIS-like) dynamics as special cases, that $P(n,t)$ becomes, at order $1/N$, a delta function centered at $n(t) = \langle n(t) \rangle \equiv (\langle n_1^{(1)}(t) \rangle, \dots, \langle n_\Sigma^{(1)}(t) \rangle, \dots, \langle n_1^{(m)}(t) \rangle, \dots, \langle n_\Sigma^{(m)}(t) \rangle)$, $\forall t$. Therefore, for $N \to \infty$, the process is exactly described by a system of ODEs whose variables are the average numbers (or fractions) $\{\langle n_i^{(k)}(t) \rangle\}_{i=1,\dots,\Sigma}^{k=1,\dots,m}$. For the models in Secs. 2.2.2.1 and 2.2.2.2, for which $m = 1$, those averages are $S(t)$, $I(t)$ and $R(t)$. If Ref. [203] considered polynomial transition rates (e.g., $\propto \beta S(t) I(t)$), Kurtz [201, 204, 205] extended these results

to a much broader class of homogeneous Markov chains, by only requiring the transition rates $W(n|n')$ to be bounded functions in any finite interval of $\mathbb{R}^N$.

In summary, partitioned our system in m extensive compartments and taken the $N \to \infty$ limit, the homogeneous Markov chain describing the system's dynamics can be replaced by a deterministic model for state averages. The resulting model is thus the so-called *thermodynamic limit* of the stochastic description of the finite-size system. Since the state of each part of the system and the coupling between parts is described in terms of average quantities, we call the model a *mean-field* theory, in analogy to the theories developed in statistical physics [206]. Each unit in the system indeed "feels" the influence of other units through the average states of those units—through mean fields. Independently from whether the resulting description is exact or not, we hereafter qualify a quantity as a mean-field one whenever it is only defined in terms of an average over some set of state variables (or over the probability distribution of a state variable), thus neglecting the fluctuations around it.

Everything we said for continuous-time descriptions can be restated in discrete time. In this case, a system is said to satisfy the Markov property if $P(x, t+1)$ depends on the state distribution $\{P(x,t)\}_{x\in\mathcal{X}}$ at time step t, but not on that at previous time steps ($t-1$, $t-2$, etc.). The trajectory $x(t)$ is then called a *discrete-time Markov chain* or *jump chain* [201].

2.3 CONTAGION PROCESSES ON NETWORKS

The thermodynamic limit discussed in the previous section is valid based on the assumption that the current state of the system fully specifies the transition rates of the Markov chain. The simplest form for the transition rates is provided by the law of mass-action, which, assuming that the interactions occur uniformly at random in analogy to collisions in a gas of uncorrelated particles, implies that the rates are just proportional to the product of the densities of units in the considered compartments. However, how units in different compartments interact among them generally correlates with the properties defining the membership of the units to those compartments. Therefore, the partition allows us to bring some additional information about the interaction patterns among the units to obtain a more refined description. For instance, if some data indicate that people interact more often with people of similar age, then it would be helpful to divide the population into age classes and account for such assortativity to get a more realistic description of the contact patterns. Nonetheless, even if the mixing among compartments is not uniform, as long as it is random (hence entirely determined by the densities of units in the compartments), the connection between the stochastic description in terms of a Markov chain and the deterministic one in terms of ODEs is still valid. As we will see later on, the rigid presence of a network makes that connection crumble.

For now, let us explore the case of a generic heterogeneous mixing, which has been referred to as the annealed network assumption at the end of Sec. 2.2.2.1. We

can imagine a network of contacts being present at a given instant. However, the connections are constantly reshuffled at a much higher rate than the frequency at which contagion-related events occur. Consequently, the interactions are randomly redrawn each time, possibly following some generic, nonuniform pattern.

2.3.1 *Annealed networks*

Let us suppose to partition a population in m compartments (e.g., age classes) and, for the sake of simplicity, assume that the units can be found in one of two states, as in a SIS-like process. Let S_i and I_i be the fraction of susceptible and infected units in compartment $i \in \{1, \dots, m\}$, respectively, so that $S_i + I_i = 1$. The state of the system is thus specified by the vector $I = (I_1, \dots, I_m)$. Denoting with $C_{ij} = [\boldsymbol{C}]_{ij}$ the number of contacts a unit in compartment i has with units in compartment j in the time interval Δt, the probability for a susceptible unit in i to get the infection by a unit in j is proportional to $\beta C_{ij} \Delta t$. Assuming as before no correlation between the states of the units, the evolution of the system is governed in the thermodynamic limit by a set of m ODEs of the form,

$$\dot{I}_i(t) = -\mu I_i(t) + (1 - I_i(t)) \beta \underbrace{\sum_{j=1}^{m} C_{ij} I_j(t)}_{f_i(t)} = -\mu I_i(t) + (1 - I_i(t)) \beta f_i(t) \,, \tag{2.20}$$

where $\beta f_i(t)$ is the so-called *force of infection* for units in compartment i. Alone, $f_i(t)$ is the total number of infectious contacts experienced by a unit in i at time t. Notice that Eq. (2.7) is recovered from Eq. (2.20) by taking $m = 1$.

Imposing stationarity, i.e., $\dot{I} = 0$, we find the relation

$$I_i = \frac{\frac{\beta}{\mu} f_i}{1 + \frac{\beta}{\mu} f_i} \,, \tag{2.21}$$

which, as one can expect, says that the probability for a unit in compartment i to be infected increases with f_i, hence with the respective force of infection. Consequently, at equilibrium, infections are generally distributed in a nonuniform way within the population. In proximity of the epidemic threshold, where $I_i, f_i \in \mathcal{O}(1/N)$, Eq. (2.21) becomes linear,

$$I_i \simeq \frac{\beta}{\mu} f_i \,, \tag{2.22}$$

implying that the inactive state $I = 0$ looses stability if and only if the vector $f = (f_1, \dots, f_m) = 0$ does. Indeed, once linearized Eq. (2.20) around $I = 0$, one finds that both I and f satisfy the same equation, of the form

$$\dot{x} = [\beta \boldsymbol{C} - \mu \mathbb{1}] \, x \,. \tag{2.23}$$

Accordingly, the inactive state changes stability when

$$R_0 = \frac{\beta}{\mu}\Lambda_{\max}(\boldsymbol{C}) = 1 \,, \tag{2.24}$$

being $\Lambda_{\max}(\boldsymbol{C})$ the dominant eigenvalue of the contact matrix $\boldsymbol{C}$. The epidemic threshold thus reads $\beta_{\rm cr} = \mu/\Lambda_{\max}(\boldsymbol{C})$. The value of $\Lambda_{\max}(\boldsymbol{C})$ depends on how contacts are distributed within the population.

To get some insight, let us rewrite the entries of the contact matrix as $C_{ij} = k_i\theta_{ij}$, where k_i is the number of contacts for a unit in compartment i, and $\theta_{ij} = C_{ij}/k_i$ is the probability that a contact of a unit in i is with a unit in j. Through the entries of the row stochastic matrix $\boldsymbol{\theta}$, we can model any contact pattern among the compartments. For instance, by defining $\theta_{ij} = \delta_{ij}\kappa_i/k_i + (1-\delta_{ij})(1-\kappa_i/k_i)/(m-1)$, we can model patterns going from full assortativity ($\kappa_i = k_i$) to full disassortativity ($\kappa_i = 0$). In the full assortative limit, $\kappa_i \to k_i$, $\forall i$, one can easily show that $\Lambda_{\max}(\boldsymbol{C}) = k_{\max}$ at the lowest order in $k_i - \kappa_i$, being $k_{\max} \equiv \max\{k_1, \dots, k_m\}$. The epidemic threshold is therefore determined by the compartment(s) with the highest number of contacts. In the full disassortative limit, $\kappa_i \to 0$, $\forall i$, the computation of $\Lambda_{\max}(\boldsymbol{C})$ is far from trivial when $m > 2$, for the force of infection coming from compartment i goes back to i via several different paths across compartments. For $m = 2$, one can easily show that $\Lambda_{\max}(\boldsymbol{C}) = (k_1k_2)^{1/2}$, implying that the epidemic threshold depends on the geometric mean of the number of contacts of the two compartments. For $m > 2$, labeled the compartments in descending order with respect to the number of contacts, so that $k_1 = k_{\max}$ and $k_2 \leqslant k_1$, if $k_i \ll k_1$ for any $i \in \{3, \dots, m\}$, then one can show that $\Lambda_{\max}(\boldsymbol{C}) \simeq c(k_1k_2)^{1/2}$, where $c < 1$ ($c = 1$ only in the trivial case $k_i = k$, $\forall i$). The epidemic threshold thus decreases when going from the disassortative to the assortative limit. Besides, differentiating $f_i = k_i \sum_{j=1}^m \theta_{ij} I_j$ with respect to k_i, a straightforward calculation shows that f_i grows with k_i for any κ. From Eq. (2.21) then follows that, at equilibrium, I_i increases with f_i, hence with k_i. Infections are therefore mostly localized around those units which interact the most. Accordingly, the epidemic threshold decreases when highly interacting units interact preferentially among them.

2.3.1.1 Proportionate mixing

To isolate the effect that the heterogeneity in the number of contacts can have on the contagion dynamics, let us model θ_{ij} through the hypothesis of *proportionate mixing* [83], according to which the contacts occur uniformly at random (neutral mixing). The probability of interacting with a unit in compartment j is thus proportional to the total number of contacts emanating from compartment j. Denoted with p_j the fraction of units in compartment j, we thus have $\theta_{ij} = p_jk_j/\langle k\rangle$, being $\langle k\rangle = \sum_{i=1}^m p_ik_i$ the average number of contacts of a unit in the system. The quantity f_i takes so the form $f_i = k_i\langle k(I)\rangle/\langle k\rangle$, defined the mean-field quantity $\langle k(I)\rangle \equiv \sum_{j=1}^m p_jk_jI_j$, the average number of contacts of an infected unit. Consistently with what shown above, Eq. (2.21)

implies that I_i increases with k_i. Compartments are distinguished only by the number of interactions their members take part to. In network terms, compartments coincide here with degree classes.

Since $I_i > 0$ if and only if $\langle k(I)\rangle > 0$, $\forall i$, we can determine the epidemic threshold by writing down a dynamic equation for $\langle k(I)\rangle$ in proximity of $\langle k(I)\rangle = 0$ (i.e., of $I = 0$). At the lowest order in $\langle k(I)\rangle$, we obtain

$$\langle k\dot{(I)}\rangle(t) = \sum_{j=1}^{m} p_j k_j \dot{I}_j(t) = \left[-\mu + \beta\frac{\langle k^2\rangle}{\langle k\rangle}\right]\langle k(I)\rangle(t)\ . \tag{2.25}$$

Equation (2.25) provides

$$R_0 = \frac{\beta}{\mu}\frac{\langle k^2\rangle}{\langle k\rangle}\ , \tag{2.26}$$

from which, imposing the threshold condition $R_0 = 1$, we find $\beta_{\rm cr} = \mu\langle k\rangle/\langle k^2\rangle$. Fixed $\langle k\rangle$, the more dispersed is the distribution of the numbers $\{k_i\}$, the larger is $\langle k^2\rangle$, and therefore the smaller is the epidemic threshold. This result was derived in the 80s by mathematical epidemiologists who understood that to explain the increase of sexually transmitted diseases registered in the United States during the 60s and 70s, it was crucial to account for the observed heterogeneity in the number of sexual partners [207, 208]. For instance, Anderson *et al.* [208] found Eq. (2.25) by modeling the spread of the human immunodeficiency virus (HIV), responsible for the acquired immunodeficiency syndrome (AIDS). Nonetheless, the result remained confined to epidemiology for many years. In 2001, the physicists R. Pastor-Satorras and A. Vespignani [209], apparently unaware of these works in epidemiology, derived Eq. (2.25) in the context of a SIS-like contact process in networks, pushing its dynamical implications to the limit. Building on the then recent discovery of scale-free networks in real systems, they assumed a power-law distribution for the number of contacts (degree), $\sim k^{-\gamma}$, showing that, for $\gamma \leqslant 3$ (recall from Sec. 2.1.3 that $\langle k\rangle/\langle k^2\rangle \to 0$ in this case), the epidemic threshold vanishes in the thermodynamic limit. In other words, they proved that no matter how small is the infection probability β, such a heavy-tailed distribution makes the inactive state unstable—an outbreak is always supercritical. Thanks to the transversal language and formalism of network theory, the impact of this result extended across several disciplines, in the full spirit of complex systems research.

Clearly, the epidemic threshold is truly zero only for an infinite population, being instead always nonzero for finite, real systems. Moreover, as discussed in Sec. 2.1.3, it is still debated whether some real-world networks are actually scale-free or not. Nonetheless, many of them boast a heavy-tailed distribution for which $\langle k^2\rangle$ is large compared to $\langle k\rangle$, implying a much smaller value for the epidemic threshold than the one predicted by a simple homogeneous approximation (i.e., $\langle k^2\rangle \approx \langle k\rangle^2$). The vanishing of the epidemic threshold in the thermodynamic limit thus represents the extreme instance of a key qualitative result, the decrease of the epidemic threshold due to a larger heterogeneity in the activity (degree) of the system's units (nodes).

2.3.2 *Quenched networks*

As we have seen, the annealed network formalism (i.e., heterogeneous random mixing) can reveal primary effects of the heterogeneities present in the system. The heterogeneity can refer not only to the properties of the single units, as for the number of interactions (degree) they have, but also to the organization of such interactions within the system, as the frequency at which units interact among them generally correlates with their individual properties (assortativity of the mixing). This formalism is a refinement of the law of mass action. It assumes that if a unit in compartment i interacts at a given time with some unit in compartment j, the latter is chosen uniformly at random within compartment j. If the size of the compartments diverges, the probability that a unit samples twice another unit vanishes, as if the interaction network was entirely redrawn at each time. As a result, the states of different units do not correlate. The statistical fluctuations around the average state of each compartment become negligible, implying that the thermodynamic limit of the process is a deterministic system of ODEs for the average states.

Differently from particles in an ideal gas, units of most natural systems rarely interact at random. Taking human societies as an illustrative example, people interact repeatedly through their social ties. For instance, if it is generally true that we interact more often with people of age similar to ours [210–214], we do not randomly sample the subset of people of our similar age in a population to determine who we interact with—as the annealed formalism assumes. Our friends, partners, and colleagues are likely to be included in that subset, but they do not change constantly; instead, they persist over time scales that are usually much larger than those at which many processes of interest occur, such as disease or information spreading. Similarly, if a web page in the World Wide Web points to a page with specific content (e.g., a fashion blog containing a hyperlink to the page of some fashion house), the pointed page will not be constantly substituted with a new page of similar content, it will rather endure in time, usually. Such temporal persistency motivates describing the interactions occurring in these and many other systems as rigid connections. In other words, we can better model many natural systems as static networks where the properties of the units and their respective interactions are "quenched", "frozen". Regarding the set of those properties as being sampled from some—very complicated and inaccessible—probability distribution, the corresponding network is indeed a quenched disordered system [215], the disorder being encoded in the network itself and, possibly, in the weight of its links. We thus call a static network also *quenched network*.

As quenched disordered systems in general, processes in quenched networks are much more challenging to analyze than in the annealed counterparts. In particular, the infinite-size limit of a Markov chain is not only not guaranteed to be a deterministic system of ODEs for the state averages, but we do not expect it to be so. The connection between the two is guaranteed only for (homogeneous) Markov chains whose transition rates are entirely determined by the current state composition of

each compartment, i.e., by how many units of a compartment are in each of the possible states. In the presence of a static network, however, it is not just the number of units in each state that determines the transition rates but how units in the various states are distributed over the network. The information about such distribution is not contained in the state of the compartments. Because of the neglected information, the transition rates, therefore, become random variables, for they could assume different values even though the state of the compartments remained the same, making the process non-markovian. We can readily illustrate this with the simplest of the contagion models, a binary SI-like dynamics where units are only distinguished by their state. Defined the system's state as the number $N_I(t)$ of units which are in state I at time t, the rate at which $N_I(t)$ changes is proportional to the number $N_{SI}(t)$ of S-I edges in the network. The knowledge of $N_I(t)$, however, does not provide sufficient information, as S and I units can be arranged to yield different values for $N_{SI}(t)$.[14] Still, generally, $N_{SI}(t)$ is not determined by just its value and the value of $N_I(t)$ at the previous instant. As Fig. 2.8 shows with a simple example, $N_{SI}(t)$ depends on the previous precise location of the I units in the network, i.e., on the exact state of the system. A process that is a Markov chain in terms of the exact state of the (networked) system thus generally translates into some non-markovian dynamics when described solely in terms of state averages (or other aggregated variables, e.g., N_{SI}). We have no good reason to expect the resulting description to approach, in the infinite-size limit, a deterministic system of ODEs. In essence, the structural constraints representing the persistent interactions in a static network create local dynamical correlations, which the annealed approach washes out. Figure 2.9 provides an example of a case where the latter fails spectacularly.

To further clarify how the correspondence between the stochastic description (Markov chain) and the deterministic one (ODEs for state averages) breaks down for static networks, we can imagine starting from a perfectly homogeneous population, where we neglect the units' heterogeneities, and then inform the description by accounting for more and more details about those properties of the units which are relevant for the process of interest. In the beginning, since the N units are all equivalent, they have no preference to interact with other specific units, and the interactions occur uniformly at random. The mass-action assumption is satisfied, and a system of ODEs is the exact thermodynamic limit. Accounting for some heterogeneities, we then divide the population into a few compartments (e.g., in a social system, solely based on age), say m, so that each compartment is still an extensive subsystem (of average size N/m) and units within a compartment are equivalent. Thus, two units generally interact more or less frequently than expected under uniform mixing depending on the compartments they belong to (i.e., their distinguishing properties). Being the compartments extensive, the refinement of the mass-action assumption given

14 Knowledge of $N_I(t)$ would suffice if the network was complete, since $N_{SI}(t) = N_I(t)(N - N_I(t))$ in this case. We can say that a complete network is equivalent to a homogeneous well-mixing where each node potentially interacts with any other node.

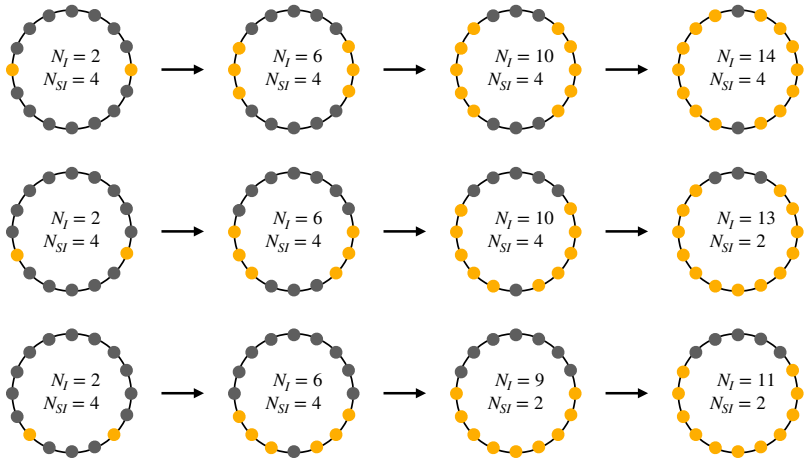


Figure 2.8: Example illustrating the impossibility to exactly describe the evolution of a networked system by following global quantities. A determinstic SI-like process is considered, where infected nodes (orange) pass the infection to susceptible nodes (grey) with probability one from one time step to the next (from left to right). Although the three initial configurations of the system are identical in terms of the global quantities N_I, the number of infected nodes, and N_{SI}, the number of S-I edges, they can transition to configurations with different N_I and N_{SI}. Knowledge of the latter is thus insufficient to determine the evolution of the system, and even a deterministic dynamics becomes stochastic when described in terms of such aggregated quantities.

by the annealed formalism provides the proper thermodynamic limit. By refining the description, we get to the point where the population is partitioned in a number m of compartments that is no longer independent from N; the compartments become sub-extensive ($\sim N^{\alpha}$, $0 < \alpha < 1$). Consequently, compared to the previous cases, the statistical fluctuations vanish at a much slower rate with N, and we need to consider much higher system sizes for the deterministic description to be satisfactory.[15] Eventually, when $m = \nu N \in \mathcal{O}(N)$, $\nu \leqslant 1$, the system is divided into compartments of fixed average size (e.g., households), $1/\nu$. Being the compartments finite, a unit has a non-vanishing probability of sampling (interacting) more than once (with) a given unit in any finite time interval. Statistical fluctuations do not vanish with N, as the states of the units become correlated to some extent, and we do not expect a deterministic description in terms of compartments' state averages to hold in the infinite-size limit. By increasing ν, the compartments become smaller, interactions more frequently repeated, and the states of the units more correlated. When eventually $\nu = 1$, each compartment consists of just one unit—everyone is unique in this world—, implying

15 Strictly speaking, when sending N to infinity, the fluctuations in a sub-extensive compartment, while still vanishing, become infinitely larger than those in an extensive compartment, as the latter are infinitesimals of higher order.

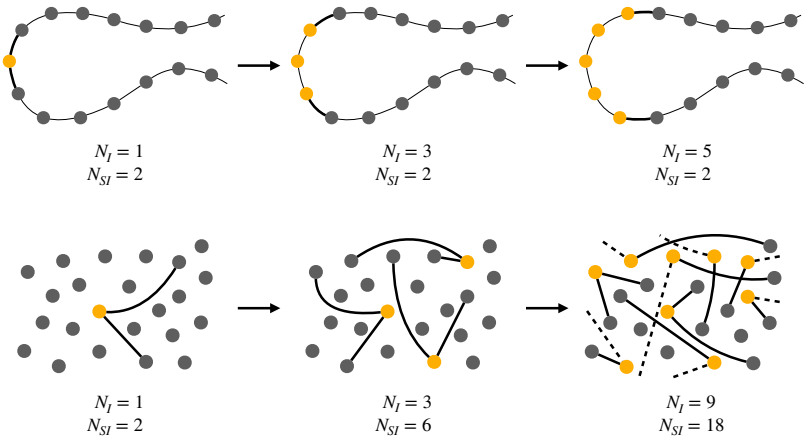


Figure 2.9: Example of how an annealed description can completely fail to predict the dynamics on a quenched network. We consider the same dynamics as in Fig. 2.8. The first row depicts the temporal evolution (of a portion) of a (infinite) chain-like network, where each node interacts with its two first neighbors along the chain. The second row represents the temporal evolution of the corresponding annealed network, which assumes that two nodes each node interacts with during a time step are sampled uniformly at random. For better readiness, only *S-I* edges are reported in this case (as dashed lines if connecting to nodes outside the shown subset). The annealed approach wrongly predicts N_I to grow exponentially ($N_I(t+1) = N_I(t) + N_{SI}(t) = 3N_I(t)$, as long as $N_I \ll N$), whereas the chain-like structure leads N_I to grow linearly ($N_I(t+1) = N_I(t) + N_{SI}(t) = N_I(t) + 2$).

that a unit always interacts with the same set of other units, maximizing correlations. This is a static network.[16]

Real-world networked systems are generally somewhat dynamic, however. For instance, people usually have relationships representing their (meta-) stable backbone of social interactions, but they also participate in random or spurious interactions that hardly or never repeat. However, also these interactions can be meaningful for specific processes, among which epidemics (e.g., getting an infection from a stranger in a public place) and information spreading (e.g., seeing a post published on an online platform by a user one does not follow). Real networks are usually neither

16 Clearly, whether a network is static or not is independent of whether the relevant information we have about the units is sufficiently refined to discern them from each other. If, on the one hand, such an amount of information is often unavailable, on the other, we can usually reconstruct the network without disposing of that information. For instance, we may retrieve the network of "who-follows-whom" or "who-is-friend-of-whom" in a social platform, which is the infrastructure on which information propagates in that system, without knowing anything about the individual properties of the single users. It is from the modeling perspective that such information is important. If we lack information about individual properties relevant to the dynamics (e.g., knowing that users have different susceptibilities to specific kinds of information), then we are limited in how accurately we can model the process of interest; nonetheless, we may still be able to find the underlying network.

static nor fully annealed, but strong temporal correlations typically exist between network configurations at subsequent times. Such correlations often make static networks good representations, albeit accounting for temporal variations would provide a more accurate—but even more complex—description. We do not consider temporal network models in this thesis (although we sometimes use temporal network data), so we refer to the review [45] and book [216] (and references therein) by Holme and Saramäki for a thorough discussion about time-varying networks.

In summary, an exact description of a Markov chain unfolding on a static network requires, in principle, the full knowledge of the state of the system, namely, the state of the single units together with their location in the network. At least formally, we could write down a generic master equation of the form of Eq. (2.19) for the system's state. However, the system's state space $\mathcal{X}$ has a prohibitive dimension. Indeed, for any of the possible Σ^N units' state assignments, there are generally multiple non-isomorphic ways to make such assignments.[17] Solely finding such non-isomorphic assignments is already computationally hard (it can be related to the problem of defective coloring [217]), let alone integrating the master equations (if an assignment has $s(N)$ average non-isomorphic realizations in a network of N nodes, then $\mathcal{X}$ has dimension $s(N)\Sigma^N$).

When dealing with processes on static (or non-randomly varying) networks, we must accept that an exact description is generally out of reach. Nonetheless, we can formulate approximate models of various sophistication that often provide a satisfactory description. The fundamental aspect of the approximation such models introduce is how good they are in preserving—or better, not neglecting—local dynamical correlations. Some key models at the base of the developments presented in this thesis, are discussed below.

2.3.2.1 Node-based approximation: Microscopic Markov chain approach

We call an approximation 'node-based' if it regards the states of the nodes in the network as uncorrelated random variables, thus neglecting any dynamical correlation among them. The exact evolution of the system is approximated through a set of N microscopic equations, each governing the expected state of a single node. That is, fluctuations around the average state of a node are neglected. The evolution of the system is described by $(\Sigma - 1)N$ equations (one state probability is fixed by normalization at the node level), whose integration is thus scalable to large systems. If the quenched interaction structure is retained, the resulting model is then usually referred to as 'individual-based' or 'quenched' mean-field approach [135], discussed below.

17 To make sure what we mean here by "non-isomorphic" is clear, take the smallest non-complete network, a chain *a*-*b*-*c* of three nodes, where *b* is the central one. Suppose one node is in state 1 and the other two are in state 2. Given the asymmetry between the central and peripheral nodes, the non-isomorphic state configurations are two, 2-1-2, and one between the isomorphic pair, 1-2-2 and 2-2-1. If also *a* and *c* were connected (i.e., in the effective absence of the network), then each label assignment would be isomorphic to any of the other two.

Node-based models have been proposed considering time either continuous [218, 219] or discrete [220–222]. As one of the models developed in this thesis builds on the discrete-time model presented in Gomez *et al.* [222], called microscopic Markov chain approach (MMCA), we choose to introduce the latter. Nonetheless, an equivalent phenomenology holds in continuous time, which can be recovered by properly sending the time step duration to zero.

We illustrate the MMCA in its original formulation used to model a spreading process following a SIS dynamics.[18] Denoted with $P_i^I(t)$ the probability for node i to be infected at time t (hence $P_i^S(t) = 1 - P_i^I(t)$ is the probability it is susceptible), $P_i^I(t)$ obeys the following difference equation

$$P_i^I(t+1) = (1-\mu)P_i^I(t) + P_i^S(t)(1-q_i(t)) \,, \tag{2.27}$$

being $q_i(t)$ the probability that node i is not infected by any of its neighbors through the time step from t to $t+1$, reading

$$q_i(t) = \prod_{j=1}^{N} \left(1 - A_{ij}\beta P_j^I(t)\right) \,, \tag{2.28}$$

where $\boldsymbol{A}$ is the adjacency matrix. The N equations of the form of Eq. (2.27) are iterated until the system reaches a stationary state, at which $P_i^I(t+1) = P_i^I(t) \equiv P_i^I$, $\forall i$. The equilibria can be found by imposing stationarity in Eq. (2.27) and solving the resulting nonlinear equation,

$$P_i^I = \frac{1-q_i}{\mu + 1 - q_i} \,, \tag{2.29}$$

analogous to Eq. (2.21). It is immediate to see that the disease-free state ($P_i^I = 0$, $\forall i$) is an equilibrium. The latter is stable if and only if β is lower than the epidemic threshold, $\beta_{\rm cr}$; otherwise, there exists one endemic equilibrium for each $\beta > \beta_{\rm cr}$. To find $\beta_{\rm cr}$, in Ref. [222], the authors perturb the disease-free state by taking $P_i^I \in \mathcal{O}(\epsilon)$, $\forall i$, given $\epsilon \to 0$, and plug this into Eq. (2.29). Neglecting $\mathcal{O}(\epsilon^2)$ terms, the probability for node i to get infected becomes $1 - q_i \simeq \beta \sum_{j=1}^{N} A_{ij} P_j^I$, which is proportional to the average number of infected nodes in its neighborhood,[19] a local mean-field quantity. In terms of the state vector $P^I = (P_1^I, \dots, P_N^I)$, the linearized Eq. (2.29) can then be recast in the form of an eigenvalue equation,

$$\boldsymbol{A}P^I = \frac{\mu}{\beta} P^I \,. \tag{2.30}$$

18 In Ref. [222], the authors consider that a node can recover *and* be reinfected within the same time step. We do not consider this possibility here, nor in the following models. In fact, it does not affect the epidemic threshold. Moreover, the associated joint probability for the two events is negligible in continuous time.

19 Notice that this is the form taken by the probability of getting infected in continuous time, where Eq. (2.27) becomes $\dot{P}_i^I(t) = -\mu P_i^I(t) + P_i^S(t)\beta \sum_{j=1}^{N} A_{ij} P_j^I(t)$.

Equation (2.30) says that μ/β is an eigenvalue of the adjacency matrix. Since we are looking for the smallest value of β at which the disease-free state looses stability, the epidemic threshold is provided by the dominant eigenvalue $\Lambda_{\max}(\boldsymbol{A})$, as

$$\beta_{\mathrm{cr}} = \frac{\mu}{\Lambda_{\max}(\boldsymbol{A})} \,, \tag{2.31}$$

having assumed $\Lambda_{\max}(\boldsymbol{A})$ is real. The connection with Eq. (2.24) is easily established. Reducing each of the m compartments in the annealed formalism to one of the N nodes in the network (hence $m = N$) , the contact matrix $\boldsymbol{C}$ becomes the adjacency matrix $\boldsymbol{A}$. The latter is the microscopically detailed version of the former.

It might not be immediately clear, however, why the procedure above provides the epidemic threshold. To this end, notice that, given it solves Eq. (2.30)—which is the linearized form of Eq. (2.29)—, P^I is an equilibrium state of the system. The latter is nonzero but each of its entries is of the order of ϵ. Given that $\epsilon \to 0$, P^I is thus arbitrarily close to zero, namely, to the disease-free state. For $\beta < \beta_{\mathrm{cr}}$, Eq. (2.30) cannot be satisfied, meaning that the disease-free state is stable. Right at $\beta = \beta_{\mathrm{cr}}$, instead, the system does not go back to the disease-free state, nor diverges from it, but stays infinitesimally close to it. In other words, $\beta = \beta_{\mathrm{cr}}$ makes the disease-free state become neutrally stable, thus marking the epidemic onset.

We can arrive to Eq. (2.31) in a more straightforward and rigorous way through standard dynamical system theory. Linearizing Eq. (2.27) and expressing it in terms of the state vector P^I, we have

$$P^I(t+1) = \boldsymbol{J} P^I(t) \,, \tag{2.32}$$

having defined the $N \times N$ Jacobian matrix $\boldsymbol{J} = \beta \boldsymbol{A} + (1-\mu)\mathbb{1}$. The epidemic threshold is then the value of β at which the dominant eigenvalue of $\boldsymbol{J}$ has modulus equal to 1, i.e., $|\Lambda_{\max}(\boldsymbol{J})| = 1$. This translates into $\beta_{\mathrm{cr}}|\Lambda_{\max}(\boldsymbol{A})| + 1 - \mu = 1$, from which $\beta_{\mathrm{cr}} = \mu/|\Lambda_{\max}(\boldsymbol{A})|$ is found. Equation (2.31) is then recovered if $\Lambda_{\max}(\boldsymbol{A})$ is real. Sufficient conditions for $\Lambda_{\max}(\boldsymbol{A})$ to be real will be discussed in a moment. For now, let us assume it is real. Then, $\Lambda_{\max}(\boldsymbol{J})$ is real too. At the epidemic threshold, $\Lambda_{\max}(\boldsymbol{J}) = 1$ implies that the dominant eigenvector is a fixed point of the linearized dynamics. Over time, a starting state $P^I(0)$ sufficiently close to the disease-free equilibrium (in principle, there could be more than one stable equilibrium), will align to the dominant eigenvector. This observation allows us to clarify how the procedure used in Ref. [222] is able to identify the epidemic threshold. What is implicitly assumed between Eqs. (2.29) and (2.31) is that the transient time needed for an initial state to converge to the dominant eigenvector has already passed. Indeed, Eq. (2.30) says that P^I is an eigenvector of $\boldsymbol{A}$ of eigenvalue μ/β. This implies that P^I is also an eigenvector of $\boldsymbol{J}$, of eigenvalue exactly 1, meaning that P^I is a stationary state, consistently with what assumed to get to Eq. (2.30). This equivalence between requiring $\Lambda_{\max}(\boldsymbol{J}) = 1$ and imposing stationarity in order to find the epidemic threshold, will turn out to

be key to compute the threshold in more refined models accounting for dynamical correlations.

PERRON-FROBENIUS THEOREM. The correspondence established above relies on $\Lambda_{\max}(\boldsymbol{J})$, hence $\Lambda_{\max}(\boldsymbol{A})$, being real. If the dominant eigenvalue of $\boldsymbol{J}$ has modulus 1 but is not real (or, more generally, none of its dominant eigenvalues is real), then the corresponding eigenvector is not a fixed point of the dynamics, as it has a time-dependent phase associated to it, nor can be identified as a probability vector. In such case, we cannot guarantee that imposing stationarity yields always a valid epidemic threshold. Therefore, it would be useful to have a sufficient condition establishing when a dominant eigenvalue of $\boldsymbol{J}$ is real. For instance, this is the case for undirected structures, for which $\boldsymbol{A}$ (or any other matrix related to it) is symmetric and has therefore a real spectrum. However, there is no need to invoke such a strong condition. Rather, we only need its dominant eigenvalue to be real, not the entire spectrum. In this respect, a fundamental result is the Perron-Frobenius theorem [223]. The theorem guarantees the existence of a positive real dominant eigenvalue for any real square matrix which is non-negative and irreducible. Specifically, a matrix $\boldsymbol{M}$ is called irreducible if it does not have any proper linear subspace which is invariant, that is, such that any vector in the subspace is still in the subspace after $\boldsymbol{M}$ acted on it. This is not really the most handy definition. Luckily, there is an equivalent defining property of irreducibility which directly refer to the backbone of the interaction structure. Indeed, it can be shown that $\boldsymbol{M}$ is irreducible if and only if the graph whose edge set contains a directed edge from node i to node j if $M_{ij} \neq 0$, is strongly connected [223]. This is a much weaker constraint than symmetry. Accordingly, weighted networks—thus directed networks, as a special case—which are (weakly) connected,[20] all satisfy the Perron-Frobenius theorem.

Nonetheless, it should be noted that the theorem only gives a sufficient condition. The dominant eigenvalue may still be real even if its associated matrix does not satisfy the hypothesis of the theorem. Therefore, even in such case, one may still give a try to the procedure described above and then check whether it provided the right threshold.

2.3.2.2 Pair-based approximation: Epidemic link equations

In the previous section, we assumed the states of the nodes to be independent random variables. However, the states of two adjacent nodes are mutually dependent, for they can directly infect each other. A node can influence the state of potentially any other node reachable from it, not just the state of its neighbors. If, at some point, a node is infected, it means that it has been reached by an infection chain generated somewhere in the network of contacts. Therefore, the states of two nodes in the same connected network component always correlate to some extent. Nonetheless, due to

20 If not, then the theorem applies to each of the connected components.

the stochasticity of the propagation dynamics, the further two nodes are from each other, the more the nodes and thus the randomness between them. We thus expect the mutual influence between two nodes to decrease with their distance. Not rarely, even node-based (uncorrelated-states) models can perform surprisingly well, suggesting that dynamical correlations might not be very persistent. This is usually the case for random networks. Since the diameter of such networks is relatively small, the mutual relative influence between two nodes is strongly reduced by the numerous (short) paths through which many other nodes can influence them. However, even in random networks, node-based models lose much accuracy around critical points, where strong correlations develop [84].

A remarkable improvement is represented by pair-based models [224–229], also known as 'pair quenched' mean-field theories. Tracking the evolution of the joint state of the two nodes in any edge, these models account for the dynamical correlations between any two neighbors in the network. The latter can be found in Σ^2 states. For instance, if nodes can be in state S and I, an edge can be in one of the four states, SS, SI, IS, and II. To illustrate the pair approximation, we refer to the Epidemic link equations (ELE) model [229], which is a direct extension of the MMCA.

In writing Eq. (2.27), we made the implicit assumption that, given node i is in state S, the probability that a neighbor j is in state I is independent from the state of i. If we denote with $P_{j|i}^{I|S}$ the conditional probability that j is in state I given i is in state S, this means assuming that $P_{j|i}^{I|S} = P_j^I$, which, from $P_{j|i}^{I|S} = P_{ij}^{SI}/P_i^S$, is equivalent to $P_{ij}^{SI} = P_i^S P_j^I$. Any pair approximation model rejects this assumption. For the ELE model, specifically, we still have Eq. (2.27), but now the probability for node i to not get infected, $q_i(t)$, reads

$$q_i(t) = \prod_{j=1}^{N} \left(1 - A_{ij}\beta \frac{P_{ij}^{SI}(t)}{P_i^S(t)}\right) . \tag{2.33}$$

If we assumed $P_{ij}^{SI} = P_i^S P_j^I$, we would recover Eq. (2.28). Using Eq. (2.33), the system of dynamical equations is not closed. One needs to describe how the edge state probabilities evolve. Taking into account all the possible changes of state of the two nodes i and j over the discrete time step, P_{ij}^{SI} and P_{ij}^{II} evolve according to

$$\begin{aligned} P_{ij}^{SI}(t+1) = {} & q_{i(j)}(t)(1-q_{j(i)}(t))P_{ij}^{SS}(t) + (1-\mu)(1-\beta)q_{i(j)}(t)P_{ij}^{SI}(t) \\ & + \mu[1-(1-\beta)q_{j(i)}(t)]P_{ij}^{IS}(t) + \mu(1-\mu)P_{ij}^{II}(t) , \end{aligned} \tag{2.34}$$

$$\begin{aligned} P_{ij}^{II}(t+1) = {} & (1-q_{i(j)}(t))(1-q_{j(i)}(t))P_{ij}^{SS}(t) \\ & + (1-\mu)[1-(1-\beta)q_{i(j)}(t)]P_{ij}^{SI}(t) \\ & + (1-\mu)[1-(1-\beta)q_{j(i)}(t)]P_{ij}^{IS}(t) + (1-\mu)^2 P_{ij}^{II}(t) , \end{aligned} \tag{2.35}$$

having denoted with $q_{i(j)}(t)$ the probability that i is not infected by any of its neighbors, excluding j, that is

$$q_{i(j)}(t) = \prod_{\substack{l=1 \\ l\neq j}}^{N} \left(1 - A_{il}\beta \frac{P_{il}^{SI}}{P_i^S(t)}\right) \; ; \tag{2.36}$$

which is nothing but Eq. (2.33) with the contribution from j removed from the product. Please, notice that $P_{ij}^{\sigma_i\sigma_j} \equiv P_{ji}^{\sigma_j\sigma_i}$ (e.g., $P_{ij}^{IS} \equiv P_{ji}^{SI}$). The equation for P_{ij}^{IS} is simply found by exchanging i and j in Eq. (2.34), whereas P_{ij}^{SS} is already fixed by normalization as $P_{ij}^{SS} = 1 - P_{ij}^{SI} - P_{ij}^{IS} - P_{ij}^{II}$. In truth, we only need one equation per edge, due to the two constraints $P_i^I = P_{ij}^{IS} + P_{ij}^{II}$ and $P_j^I = P_{ij}^{SI} + P_{ij}^{II}$, expressing the node probabilities as marginals over the edge probabilities. For instance, we may express P_{ij}^{IS} and P_{ij}^{SI} as $P_{ij}^{IS} = P_i^I - P_{ij}^{II}$ and $P_{ij}^{SI} = P_j^I - P_{ij}^{II}$, respectively, and follow just the evolution of P_{ij}^{II}. Consequently, if L is the number of edges, we only need to solve $N + L$ equations.

Equations (2.27), (2.34) and (2.35), assisted by Eqs. (2.33) and (2.36), define a closed system of equations, that we can thus integrate. In order to close it, in fact, we tacitly made—again—an assumption of independence. In writing down Eq. (2.36), we neglected the information we have on the state of node j. In principle, the probability multiplying β should be $P_{l|ij}^{I|S\sigma_j} = P_{ijl}^{S\sigma_j I}/P_{ij}^{S\sigma_j}$. Equation (2.36) thus follows if $P_{ijl}^{S\sigma_j I} = P_{ij}^{S\sigma_j} P_{il}^{SI}/P_i^S$. This means assuming that the states of nodes j and l are, given the state of the middle node i, conditionally independent. More generally, for any triplet *j-i-l*,

$$P_{ijl}^{\sigma_i\sigma_j\sigma_l} = \frac{P_{ij}^{\sigma_i\sigma_j} P_{il}^{\sigma_i\sigma_l}}{P_i^{\sigma_i}} \; . \tag{2.37}$$

Equation (2.37) is generally referred to as the *pair approximation* closure [227], for the state of triplets, and therefore of any larger subgraph, is approximated in terms of the states of the pairs composing them. By preserving dynamical correlations within pairs, pair-based models provide more accurate predictions than node-based models, especially in sparser random networks with no or few small cycles [84]. In particular, the pair-based approach is exact for the SIR model on tree networks [228], for any two branches departing from any node are statistically independent once conditioned on the state of the node.

Let us now derive an expression for the epidemic threshold by linearizing the dynamic equations around the disease-free state. This is done by taking $P_i^I, P_{ij}^{SI}, P_{ij}^{II} \in \mathcal{O}(\epsilon)$, with $\epsilon \to 0$, that is, considering all the probabilities of states with at least one infected node to be infinitesimal. Crucially, in order to correctly predict the epidemic threshold, we must be consistent with the pair approximation. This means considering P_{ij}^{II} to be of the same order ϵ as P_{ij}^{SI}, and not of order ϵ^2, as it would be implied

by the node-based factorization $P_{ij}^{II} = P_i^I P_j^I$. For what said before, we only need to linearize the equations for P_i^I and P_{ij}^{II}. These are

$$P_i^I(t+1) = (1-\mu)P_i^I(t) + \beta \sum_{j=1}^{N} A_{ij}(P_j^I(t) - P_{ij}^{II}(t)) \,, \tag{2.38}$$

$$P_{ij}^{II}(t+1) = \beta(1-\mu)(P_i^I(t) + P_j^I(t) - 2P_{ij}^{II}(t)) + (1-\mu)^2 P_{ij}^{II}(t) \,, \tag{2.39}$$

and define a linear system of $N+L$ equations. The dominant eigenvalue of the associated $(N+L)\times(N+L)$ Jacobian matrix $\boldsymbol{J}$ provides the threshold condition. We can, however, find the epidemic threshold from a reduced $N \times N$ Jacobian matrix. At the threshold, it holds $\Lambda_{\max}(\boldsymbol{J}) = 1$, hence the dominant eigenvector is a fixed point of $\boldsymbol{J}$. In the large time limit, $t \to \infty$, the state of the system thus aligns to the dominant eigenvector and approaches stationarity. Imposing $P_{ij}^{II}(t+1) = P_{ij}^{II}(t)$ in Eq. (2.39), solving for P_{ij}^{II}, and substituting in Eq. (2.38), the dynamical equation for the state vector $P^I = (P_1^I, \dots, P_N^I)$, becomes

$$\boldsymbol{J'} P^I(t+1) = \boldsymbol{J'} P^I(t) \,, \tag{2.40}$$

where $\boldsymbol{J'}$ denotes the Jacobian matrix of the reduced system,

$$\boldsymbol{J'} = \beta X \boldsymbol{A} - \beta(1-X)\boldsymbol{D} + (1-\mu)\mathbb{1} \,, \tag{2.41}$$

being $X = 1 - \beta(1-\mu)/(\mu(2-\mu) + 2\beta(1-\mu))$ and $\boldsymbol{D} = \mathrm{diag}(k_1, \dots, k_N)$ a diagonal matrix with the degrees of the nodes as entries. The epidemic threshold is found as the smallest value of β satisfying the equation $\Lambda_{\max}(\boldsymbol{J'}) = 1$. To notice that a reduced Jacobian matrix related to $\boldsymbol{J'}$ is found in continuous time [227], which can be indeed recovered taking the continuous-time limit of the ELE model.

We can make the condition $\Lambda_{\max}(\boldsymbol{J'}) = 1$ more explicit by defining the matrix

$$\boldsymbol{B} = X\boldsymbol{A} - (1-X)\boldsymbol{D} \,, \tag{2.42}$$

so that $\boldsymbol{J'} = \beta\boldsymbol{B} + (1-\mu)\mathbb{1}$. Since $\boldsymbol{J'}$ and $\boldsymbol{B}$ just differ for a matrix proportional to the identity, they share the same set of eigenvectors. In particular, from $\Lambda_{\max}(\boldsymbol{J'}) = 1$, it follows $\Lambda_{\max}(\boldsymbol{B}) = \mu/\beta$, hence

$$\beta_{\mathrm{cr}} = \frac{\mu}{\Lambda_{\max}(\boldsymbol{B})} \,. \tag{2.43}$$

We see that matrix $\boldsymbol{B}$ plays in the ELE the role played by $\boldsymbol{A}$ in the MMCA. The two matrices are indeed strictly related and both convey the effect of the quenched structure. Nonetheless, the entries of $\boldsymbol{B}$ are functions of β (and μ), and it wouldn't have been so obvious to prove Eq. (2.43) if we hadn't known it derived from $\Lambda_{\max}(\boldsymbol{J'}) = 1$.

In fact, following Ref. [229], by imposing stationarity in Eqs. (2.38) and (2.39), and solving for P^I, we would have arrived directly to the following eigenvalue equation,

$$\boldsymbol{B}P^I = \frac{\mu}{\beta}P^I \ . \tag{2.44}$$

Equation (2.44) says that μ/β must be an eigenvalue of $\boldsymbol{B}$. However, the fact that $\boldsymbol{B}$ is itself a (continuous) function of β, rises the question of whether the smallest β solving Eq. (2.44) is given by $\Lambda_{\max}(\boldsymbol{B})$ or by some other eigenvalue. However, we observe that $\mu/\beta \to \infty$ for $\beta \to 0$, while $\boldsymbol{B} \to \boldsymbol{A}$, hence $\Lambda_{\max}(\boldsymbol{B}) \in \mathbb{R}^+$. Since P^I is positive (each entry is of the order of ϵ), $\Lambda_{\max}(\boldsymbol{B})$ must be positive for any β and μ (this can be also proved numerically, as done in Ref. [229]). Consequently, increasing β from zero, the first eigenvalue of $\boldsymbol{B}$ that the (decreasing) function μ/β crosses, is indeed $\Lambda_{\max}(\boldsymbol{B})$; and Eq. (2.43) is recovered.

In truth, once got to Eq. (2.44), we can derive an alternative, yet equivalent, condition for $\beta_{\rm cr}$ in a more robust and elegant way. This is an anticipation and a special case of the procedure presented in Sec. B.1.1. We first rewrite Eq. (2.44) as

$$\boldsymbol{M}P^I = \widetilde{\boldsymbol{D}}P^I \ , \tag{2.45}$$

where we defined $\boldsymbol{M} = \beta X \boldsymbol{A}$ and $\widetilde{\boldsymbol{D}} = \mu \mathbb{1} + \beta(1-X)\boldsymbol{D}$. Since $X \in [0,1]$, matrix $\boldsymbol{M}$, as $\boldsymbol{A}$, is non-negative. The same can be said for $\widetilde{\boldsymbol{D}}$. Then, assuming the network is connected, $\boldsymbol{A}$ is irreducible, so is $\boldsymbol{M}$, and $\boldsymbol{D}$ is invertible, implying $\widetilde{\boldsymbol{D}}$ is invertible too. We can thus apply $\widetilde{\boldsymbol{D}}^{-1}$ to both sides of Eq. (2.45) and get a new eigenvalue equation,

$$\widetilde{\boldsymbol{M}}P^I = P^I \ , \tag{2.46}$$

where $\widetilde{\boldsymbol{M}} = \widetilde{\boldsymbol{D}}^{-1}\boldsymbol{M}$. The matrix $\widetilde{\boldsymbol{M}}$, of elements $\widetilde{M}_{ij} = M_{ij}(\widetilde{D}_{ii})^{-1}$, is also non-negative and irreducible. The Perron-Frobenius theorem thus guarantees that $\widetilde{\boldsymbol{M}}$ has a real and positive dominant eigenvalue and establishes that the associated dominant eigenvector is the only one with all entries being positive [223]. Since P^I is positive, the only possibility for P^I to solve Eq. (2.46) is to be the dominant eigenvector of $\widetilde{\boldsymbol{M}}$. This finally implies the condition $\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = 1$, which establishes an implicit equation for $\beta_{\rm cr}$. Since $\widetilde{\boldsymbol{M}}$ depends nonlinearly on β, there could be more than one value of β solving the implicit equation. The threshold $\beta_{\rm cr}$ is the smallest of the solutions.

The ELE model has been generalized by Matamalas *et al.* [230] to describe the simplicial contagion model [199], discussed in Sec. 2.2.3.2, in quenched simplicial 2-complexes. A detailed analysis of the resulting 'simplicial ELE' model, is provided in Appendix A.3.

2.3.2.3 A hierarchy of closures

Closing the equations at the level of nodes in Sec. 2.3.2.1 required assuming statistical independence between the states of neighboring nodes. Doing the same at the

level of pairs in Sec. 2.3.2.2 required assuming statistical independence between the states of the neighbors of a node. We expressed two-node (edge) state probabilities in terms of node state probabilities in the former case and three-node (triplet) state probabilities in terms of two-node state probabilities in the latter. We could write dynamic equations for three-node states to refine the description further. The equations would contain four-node states. We can then choose to model the latter in terms of the states of smaller subgraphs (of three and two nodes) or write dynamic equations for them, too, which will, in turn, include five-node states. That is, we can write exact equations for n-node state probabilities that, in order to be closed and thus usable, need knowledge about $(n+1)$-node state probabilities. Unless specific cases allowed by the interplay between the structure of the network and the considered dynamics (e.g., pair approximation for SIR model in tree networks [228]), expressing the $(n+1)$-node state probabilities in terms of n-node state ones represents an approximation. Consequently, to obtain an exact *and* closed model, we generally need to provide dynamic equations for the state probabilities of subgraphs of any size, eventually falling back into the exact description in terms of N-node state probabilities for the entire system.

Such iterative relation between the dynamic evolution of the state probability distribution at one order and the knowledge (hence the dynamic evolution) of the state probability distribution at the next higher order is related to the Bogoliubov–Born–Green–Kirkwood–Yvon (BBGKY) hierarchy in statistical physics [24], where the evolution of the n-particle probability density function needs knowledge of the $(n+1)$-particle probability density function. Boltzmann's assumption of molecular chaos, which factorizes the two-particle density in the product of two one-particle densities by assuming statistical independence, is the lowest-order truncation of the BBGKY hierarchy. The corresponding truncation for dynamical processes on networks is the one made in a node-based model. The next-order truncation gives the first-order correction to the Boltzmann transport equation and a pair-based model, respectively. Going up in the hierarchy, we obtain increasingly accurate models, but they are also harder to integrate and, generally, analyze. Climbing the hierarchy is a trade-off game.

CUT-VERTICES CLOSURE. Analogously to how a pair-based approximation gives an exact description of a SIR-like process in tree networks, a triplet-based approximation describes a SIR-like process exactly in networks with no cycles of more than three nodes. These are in truth special cases of a more general result proven in Kiss *et al.* [231] and based on cut-vertices, i.e., vertices whose removal disconnects the network. The result simply follows by observing that, in a SIR-like model, (i) chosen a closure that covers the network by means of subgraphs joined by cut-vertices, the only state probabilities appearing in the dynamic equations are those where the cut-vertex joining any two subgraphs is in the susceptible state; and (ii) that any two subgraphs joining at a cut-vertex have independent states if the cut-vertex is susceptible, for no infection—no influence—went through the latter, which would otherwise be infec-

ted or recovered. Then, given two generic vertex sets, $\{v_1, \ldots, v_n\}$ and $\{v_n, \ldots, v_m\}$, inducing two subgraphs joining at the cut-vertex v_n, the closure

$$P^{\sigma_{v_1} \ldots S \ldots \sigma_{v_m}}_{v_1,\ldots,v_n,\ldots,v_m} = \frac{P^{\sigma_{v_1} \ldots S}_{v_1,\ldots,v_n} P^{S \ldots \sigma_{v_m}}_{v_n,\ldots,v_m}}{P^{S}_{v_n}} \tag{2.47}$$

yields an exact model—an exact truncation of the hierarchy of closures.

For a SIS-like model, condition (ii) above breaks down, for the susceptibility of a cut-vertex connecting two otherwise disconnected subgraphs implies independence between the states of the latter only before the first time the cut-vertex recovers. Indeed, the latter may receive the infection from one side and transmit it to the other so that the states of the nodes at the two sides are correlated even after the cut-vertex returns to the susceptible state. In other words, the susceptibility of the cut-vertex no longer guarantees that an infection has not previously gone through it. There is currently no recipe for designing highly accurate closures for SIS-like and similar dynamics with recursive states in generic networks. Nonetheless, quenched network models using the cut-vertices closure in Eq. (2.47) give usually satisfactory descriptions in random networks [84].

The most important class of quenched models of contagion processes that we do not discuss here, for which we refer to the reviews by Kiss *et al.* [84] and Wang *et al.* [85], is the one based on percolation theory [37]. Such models are mainly conceived for non-recurrent dynamics like a SIR. By mapping the probability that an edge transmits an infection to the occupation probability of that edge in (directed) percolation, one can calculate the size distribution of the outbreaks (out-components) and the probability that an extensive outbreak occurs (a giant strongly connected component exists), in addition to the size of the latter and the invasion threshold [75, 77, 232–238]. These approaches are highly accurate in configuration networks and, because of every node being a cut-vertex, exact in tree networks. Extensions to recurrent dynamics like a SIS have also been proposed [239, 240].

2.3.3 *Annealed approaches to quenched networks*

Models retaining the quenched structure can achieve high levels of accuracy in predicting the state of the whole system and single subsystems and units. They also allow us to explore the dynamical implications of any specific topology. However, achieving this level of detail—and thus accuracy—comes at the cost of reduced mathematical tractability. Moreover, data to reconstruct the fine network structure of natural systems are usually unavailable [241], forcing us to resort to some coarse-grained representation (e.g., based only on a degree distribution) where connections are aleatory, not rigid. Consequently, either to gain more analytical insights or out of necessity, one often needs to approximate the quenched structure through some annealed representation.

The crudest of such approximations is the homogeneous mean-field theory, where each unit is considered equivalent to any other system unit, completely neglecting the network. Of this kind, when regarded as approximations of processes on quenched structures, are the models discussed in Sec. 2.2. A more refined description is given by heterogeneous mean-field theories, like those presented in Sec. 2.3.1, by accounting for potential differences in the probability that units in different classes interact. Accordingly, it is assumed that such interaction is exclusively determined by which classes the units belong to, i.e., by their static properties (e.g., number of contacts or vaccination status), but not by their dynamical states. Both homogeneous and heterogeneous mean-field approximations thus neglect the dynamical correlations created by the quenched structure, leading often to poor predictions [84, 242].

Dynamical correlations are partially tracked by pair mean-field approximations [243–248], describing the evolution of average two-node state probabilities (e.g., proportion of links in state *SI*). These can also be adapted to account for the presence of clustering (3-cycles), leading to 'clustered' or 'triangular' pair approximations [249–251], although the equations for pairs become remarkably complicated. A better (and more accurate) description is, in this case, provided by a mean-field triplet (or 'triple', 'triadic') approximations [250, 252, 253] which, tracking average three-node state probabilities, naturally discern 3-cycles from open triplets. Contrary to what one might expect, we will show in Sec. 3.2 that such an approximation can be more tractable than a pair-based one, allowing us to obtain new insights about the dynamical implications of a clustered structure.

Such mean-field approximations account for average dynamical correlations within a single type of motifs (pairs, triads, etc.) incident on a node. An interesting alternative is the one offered by approximate master equations (AME) models, which, as introduced by Marceau *et al.* [254], approximate the state of the network in terms of the states of neighborhoods. In this way, AME retain average correlations within neighborhoods, hence between links incident on a node. Such AME [242, 254] are sometimes called 'node-centered' AME to differentiate them from the 'group-centered' AME by Hébert-Dufresne *et al.* [93]. The latter approximate the network state through the states of cliques[21] of arbitrary size, thus generalizing (heterogeneous) pair mean-field approximations. Such group-centered description can be effectively understood as an AME model by regarding cliques as nodes of a bipartite network so that the neighborhood of a clique consists of the nodes it includes. Below, we delve into both AME models, as these are the building blocks of the general model we propose in Sec. 3.3.

21 The authors consider the more general case in which random removals of internal links dilute cliques.

2.3.3.1 Node-centered approximate master equations

The idea behind node-centered AME is to keep track of not just the state of a node, but of the state of its neighborhood too. In this way, this approach is able to partially retain local dynamical correlations.

We assume hereafter that nodes can be either susceptible or infected. Let $S_{k,l}(t)$ be the joint probability that a node is susceptible, has degree k, and $l \leqslant k$ of its neighbors are infected, at time t. Let then $\beta(k,l)$ be the rate at which such a node becomes infected. The quantity l is sometimes called 'effective' or 'active' degree. Analogously, let $I_{k,l}(t)$ be the joint probability that a node is infected, has degree k, and $l \leqslant k$ of its neighbors are infected, at time t; and $\mu(k,l)$ be the rate at which it becomes susceptible. Suitably choosing the functional form for the transition rates $\beta(k,l)$ and $\mu(k,l)$, and forgetting about the meaning of 'susceptible' and 'infected', this formalism can be used to accurately describe general binary-state dynamics [242]. Nevertheless, in most contagion models, (i) recovery is a fully endogenous process of a node, hence $\mu(k,l) = \mu$, $\forall k,l$; and (ii) infection depends exclusively on the number of infectious contacts, that is $\beta(k,l) = \beta(l)$, $\forall k$.[22] Dropping the dependence on k from the infection rate does not however simplify the formalism, therefore we keep it general.

With this definitions, the dynamics is a jump process described by the following system of equations,

$$\begin{aligned}
\dot{S}_{k,l}(t) = &- \beta(k,l) S_{k,l}(t) + \mu I_{k,l}(t) \\
&- \theta_S(t)(k-l) S_{k,l}(t) + \theta_S(t)(k-l+1) S_{k,l-1}(t) \\
&- \mu l S_{k,l}(t) + \mu(l+1) S_{k,l+1}(t) \,, \qquad (2.48) \\
\dot{I}_{k,l}(t) = &- \mu I_{k,l}(t) + \beta(k,l) S_{k,l}(t) \\
&- \theta_I(t)(k-l) I_{k,l}(t) + \theta_I(t)(k-l+1) I_{k,l-1}(t) \\
&- \mu l I_{k,l}(t) + \mu(l+1) I_{k,l+1}(t) \,, \qquad (2.49)
\end{aligned}$$

with the constraint $\sum_{k,l}[S_{k,l}(t) + I_{k,l}(t)] = \sum_k p_k = 1$, being p_k the degree distribution. The first two terms in each equation represent processes involving a node directly. The remainder represent instead processes involving the neighbors of a node. In particular, the quantities $\theta_S(t)$ and $\theta_I(t)$ represent the probabilities that a susceptible neighbor of, respectively, a susceptible and infected node, gets infected. Since we lack any additional information about the neighbors of the susceptible neighbor, besides one of

22 This last point makes clear why heterogeneous mean-field theories, which distinguish nodes based solely on their degree, are often poor approximations for contagion process in quenched networks. Even though, as seen in Sec. 2.3.1, the probability for a node to get infected can correlate with its number of contacts (k), what really matters is how many of those contacts are infectious (l) and can thus have an effect on the node's dynamics.

them being respectively susceptible and infected, those probabilities are computed through mean-field arguments; and read

$$\theta_S(t) = \frac{\sum_{k,l} \beta(k,l)(k-l)S_{k,l}(t)}{\sum_{k,l}(k-l)S_{k,l}(t)} , \tag{2.50}$$

$$\theta_I(t) = \frac{\sum_{k,l} \beta(k,l) l S_{k,l}(t)}{\sum_{k,l} l S_{k,l}(t)} . \tag{2.51}$$

These are averages of the infection rate $\beta(k,l)$ conditioned on the fact that the susceptible neighbor is reached from, respectively, a susceptible and infected node. The probability that a neighbor of a susceptible (infected) node is a susceptible node of degree k and effective degree l is proportional to the fraction $S_{k,l}(t)$ of such nodes and to the number $k-l$ (l) of susceptible (infected) neighbors such a node has. The possibility of distinguishing $\theta_S(t)$ and $\theta_I(t)$ is precisely the refinement made by these AME when compared to a heterogeneous mean-field theory where only the degree k is considered.

The node-centered AME, Eqs. (2.48) and (2.49), describe the state transitions of a node in class (k,l) exactly, as a master equation would do. The latter would describe exactly also the dynamics of the neighbors of such node, which would in turn require knowing the state of the neighbors of those neighbors. Lacking this information, Eqs. (2.48) and (2.49) approximate the neighbors' dynamics through the mean fields $\theta_S(t)$ and $\theta_I(t)$. In a nutshell, taking fully into account the local perspective of a node, these AME approximate the dynamics over the entire network by decomposing it in terms of classes (k,l) of neighborhoods.

Node-centered AME are highly accurate in random networks [242, 254, 255] and—in accordance to its original purpose [254]—enable us to properly model adaptive systems where a node rearranges its connections in response to the state of its neighborhood.

2.3.3.2 Group-centered approximate master equations

Neglecting information about neighbors of neighbors, node-centered AME are not suited to describe processes on higher-order networks where cliques/hyperedges of three or more nodes exist. Shifting the focus from single nodes to groups (cliques/hyperedges) of nodes, group-centered AME are precisely designed for this task. Instead of decomposing the dynamics of the system in classes of neighborhoods, we decompose it in classes of groups, discerning these by their size and by the number of infected members. Group-centered AME are thus a generalization of the (heterogeneous) mean-field pair approximation to account for dynamical correlations within cliques/hyperedges of arbitrary size. Not accounting however for correlations within neighborhoods of nodes, they are less accurate in pairwise networks than node-centered AME.

We track the evolution of the probability $C_{n,i}(t)$ that a group has size n and possesses $i \leqslant n$ infected members at time t. Supposing that the group sizes are drawn from a distribution g_n, we have $\sum_{n,i} C_{n,i}(t) = \sum_n g_n = 1$. Besides, the node dynamics is described tracking the probabilities $S_m(t)$ and $I_m(t)$ that at time t a node belongs to m groups (we say has 'membership' m) and is respectively susceptible and infected. That is, extending pairwise heterogeneous mean-field theories, nodes are distinguished by their generalized degree m. This is drawn from p_m, hence $\sum_m [S_m(t) + I_m(t)] = \sum_m p_m = 1$. Denoted with $\beta(n,i)$ the infection rate in a group of size n with i infected members, the resulting jump process is described by the following system of equations

$$\dot{S}_m(t) = -m\theta(t)S_m(t) + \mu(p_m - S_m(t)) \,, \tag{2.52}$$

$$\begin{aligned}\dot{C}_{n,i}(t) = &-(n-i)\beta(n,i)C_{n,i}(t) - \mu i C_{n,i}(t) \\ &-(n-i+1)\beta(n,i-1)C_{n,i-1}(t) - \mu(i+1)C_{n,i+1}(t) \\ &-(n-i)\phi(t)\theta(t)C_{n,i}(t) + (n-i+1)\phi(t)\theta(t)C_{n,i-1}(t) \,,\end{aligned} \tag{2.53}$$

where

$$\theta(t) = \frac{\sum_{n,i}\beta(n,i)(n-i)C_{n,i}(t)}{\sum_{n,i}\beta(n,i)C_{n,i}(t)} \,, \tag{2.54}$$

$$\phi(t) = \frac{\sum_m (m-1)mS_m(t)}{\sum_m mS_m(t)} \,. \tag{2.55}$$

The mean field $\theta(t)$ is the average infection rate over the groups a susceptible node is member of. The probability that the latter is member of a group of size n and i infected members is proportional to the number $n-i$ of susceptible members in that group and to the fraction $C_{n,i}(t)$ of such groups, hence to $(n-i)C_{n,i}(t)$. The other mean field, $\phi(t)$, then provides the expected excess membership of a susceptible node, i.e., the average number of groups to which such a node belongs to in addition to the considered group. The probability that a clique includes a susceptible node of excess membership $m-1$ is simply proportional to membership m of the node and to the fraction $S_m(t)$ of such nodes, hence to $mS_m(t)$.

The terms in the first two lines in Eq. (2.53) represent state transitions due to interactions occurring within the considered group. As in a master equation, these transitions are described exactly. The approximation comes then from the terms in the third line, where the state transitions of the susceptible nodes in a group due to their membership to other groups are modeled via the mean-field quantity $\phi(t)\theta(t)$. As expected by construction, group-centered AME show high accuracy in random structures built from a bipartite configuration model [93, 256–258].[23]

23 This works assigning m stubs to each node of membership m and n stubs to each group of size n, and then randomly matching node-stubs with group-stubs.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

3

HIGHER-ORDER CONTAGION PROCESSES

One of the most remarkable achievements of network science has been the characterization of the dependence of contagion processes on the structural properties of the interaction network through which they spread [80, 82–85, 209, 259–261]. In particular, many real-world networks of interest, especially those found in human social systems, boast a clustered and cycles-rich structure. Triadic closures are indeed renowned to be a distinguishing feature of such systems [144, 262], together with the presence of larger dense communities [263, 264]. Households and workplaces are common contact-based examples of that, while online social communities and groups in messaging apps are information-based ones. Such dense substructures are (almost) cliques in a network representation and their presence can substantially change the spreading behavior. In the case of simple contagion processes, for which a single active neighbor suffices for transmission, the redundancy of connections within the cliques of clustered structures is responsible for the increase, relatively to their randomized counterpart, of the value of the invasion threshold, and for the reduction of the final outbreak size [78, 93].

Apart from a collection of dyadic interactions, cliques can be also regarded as the pairwise projection of richer substructures representing higher-order, group interactions, involving more than two agents (nodes) at once. In fact, from few years on, a growing branch of literature dedicated to the study of various dynamical processes involving group interactions [87, 88, 265–270], has been showing that such interactions can heavily affect the dynamics. Neglecting them can lead to inaccurate predictions even at the qualitative level.

In this chapter, we are interested in complex contagion processes according to the second definition given in Sec. 2.2. Specifically, in complex contagions mediated by group interactions, where the outcome of a potentially spreading interaction depends on how many different contagious agents take part to it simultaneously, and not just on the sheer number of them. We refer to such processes generically as *higher-order contagions*.

The presence of complex contagion mechanisms has been assessed in various contexts [96], but largely in online social networks and forums [175, 271] thanks to their

unique data traceability. Nonetheless, whether or not such mechanisms also substantiate in the form of group interactions, in other words, if and how much the simultaneity of the exposure to multiple sources matters, is still an open question. A major reason for this is that higher-order relational data have been only very recently started to be recorded and analyzed as such, being instead traditionally decomposed in pairwise relations. Nonetheless, as mentioned in Sec. 1.3, many socio-psychological studies on small groups support the irreducibility of group dynamics, hence the existence of group-based complex contagions. The research presented in this chapter contributes to the ongoing intense theoretical exploration of higher-order mechanisms [88, 89, 272].

Several studies have lately appeared which unveiled the dynamic effects of group interactions on complex contagion [113, 199, 200, 273–281]. Crucially, all these studies assume statistical independence among nodes' states. The assumption greatly eases the analysis, allowing for basic characterizations of the contagion dynamics by assuming well-mixing or using particularly symmetric quenched structures. The most striking finding of these models is that group interactions can induce critical mass effects and bistability (hence hysteresis)[1], as seen in Sec. 2.2.3.2. As observed there, higher-order interactions provide an alternative path to the phenomenology already known from complex contagion models on pairwise networks, where exposures are cumulated over several dyadic interactions (see Sec. 2.2.3.1). [2] This naturally prompts the question of whether additional dynamic effects can arise exclusively from higher-order interactions and their structural organization within the system. The answer is *yes*. As we will see in this chapter, the discovered effects can only be (reliably) predicted by a theory that preserves a sufficient amount of local dynamical correlations.

We start in Sec. 3.1 by generalizing the MMCA and ELE to build a quenched model able to account for dynamical correlations within cliques/hyperedges of potentially any order. In the attempt to do so, we will realize that the applicability of discrete-time models is restricted to a specific class of networks. This will lead us to introduce a constrained notion of graph cover, for which a heuristic algorithm is proposed. Focusing eventually on simplicial 2-complexes, our model will provide a second look at the behavior of higher-order contagions while achieving a notably higher predictive accuracy when compared to other models in the literature. The analysis is further extended in Sec. 3.2, where a mean-field model at the level of 3-node states will enable us to demonstrate that and how the dynamics of higher-order contagions is strictly related to the way interactions are organized in the structure, hence to the chosen higher-order representation. We close the chapter by developing in Sec. 3.3 a generalized AMEs framework able to consider dynamical correlations within and across groups.

1 Additionally, St-Onge *et al.* [113] proved that a temporal window of super-exponential growth emerge when combining the notion of a minimal infective dose (as in Sec. 2.2.3.1) with a large heterogeneity in the duration of group interactions.

2 Ferraz de Arruda *et al.* [281] also find multistability (beyond bistability), which the Dodds-Watts model already admits, and intermittency, which originates from community structure and not directly from higher-order interactions.

Crucially, this will pave the way for studying adaptive higher-order networks, specifically, representations of systems wherein agents can strategically reorganize their group memberships in response to local dynamics—a modeling possibility simply beyond the scope of earlier approaches. We will uncover intriguing effects stemming from the group structure, hence excluded in pairwise networks.

3.1 HIGHER-ORDER CONTAGIONS REVISITED: THE IMPORTANCE OF DYNAMICAL CORRELATIONS

In this section, we investigate the behavior of higher-order contagions spreading on quenched simplicial complexes. We approximate the dynamics as a simplicial SIS process [199]. Accordingly, considering a discrete-time formulation, we assume that infected individuals recover with probability μ during a time step ($I \xrightarrow{\mu} S$), while a susceptible individual becomes infected with probability $\beta^{(1)}$ during a 2-body interaction (a 1-face) with an infected individual ($S + I \xrightarrow{\beta^{(1)}} 2I$) and with probability $\beta^{(2)}$ during a 3-body interaction (a 2-face) with two infected individuals ($S + 2I \xrightarrow{\beta^{(2)}} 3I$). More generally, a susceptible node in a group of $n+1$ nodes (an n-face) gets the infection with probability $\beta^{(n)}$ when all the other n nodes are infected. Due to the hereditary property of simplicial complexes, if n nodes are infected, and thus can pass the infection as a group of n nodes, then also any subset of that group can: for any $k \leqslant n$, there are $\binom{n}{k}$ subsets passing the infection with probability $\beta^{(k)}$ as a group of k nodes. In other words, to each state configuration of the n nodes (i.e., which of them are infected) corresponds a set of concurrent channels of infections, and these channels are correlated.

In this setting, previous models using annealed, node-based (uncorrelated states) approximations found that the invasion threshold, $\beta_{\mathrm{cr}}^{(1)}$, at which the inactive (fully susceptible) state becomes unstable, is independent from the second-order coupling $\beta^{(2)}$ (and from any higher one, $\{\beta^{(n)}\}_{n>2}$), no matter the properties of the interaction structure [199, 275]. This finding is however in clear disagreement with numerical simulations [199, 277, 282].

For simplicial 2-complexes, Matamalas *et al.* [230] proposed a quenched pair-based model, here referred to as 'simplicial ELE' model, accounting for 2-nodes and, partially, 3-nodes dynamical correlations. The model proves to be notably accurate in networks with either homogeneous and heterogeneous degrees. However, as shown in Appendix A.3, it suffers of subtle inconsistencies that, while making it usable for any 2-dimensional SC, also impede the computation of the critical point.

Here, looking for a discrete-time model holding for simplicial complexes of any dimension, we reveal that fixing those inconsistencies puts a topological condition on the interaction structure the model can describe; namely, that two group interactions can only share one node. In order to satisfy this constraint, we introduce the notion of edge-disjoint edge clique cover (EECC) of a simplicial complex and a heuristic

to find it. By means of our Microscopic Epidemic Clique Equations (MECLE), we provide a 2-fold extension of the existent discrete-time (complex) contagion models, accounting for higher-order correlations and group interactions (if any) at the same time. We prove those dynamical correlations to be essential to describe how the critical point depends on the higher-order couplings. Lastly, different approaches to treat group interactions sharing multiple nodes are also discussed, while hinting at easy adaptations of ours.

3.1.1 *An overcounting problem in discrete-time models*

At the end of Sec. A.3, we wrote down in Eq. (A.14) the probability for a node i to be infected when in a simplex with two other nodes, j and l. We noted that, independently from the moment closure one chooses, Eq. (A.14) gives the exact expression for a discrete-time model, where many nodes can change state during the same time step, having this a finite duration. Sticking to the same example of an isolated triangle $\{i,j,l\}$, yet without specifying the contagion dynamics, the infection probability $\mathbb{P}_{ijl}^{(i)}$ for node i takes the general form

$$\mathbb{P}_{ijl}^{(i)} = (1-q_1)\left(P_{jl|i}^{IS|S} + P_{jl|i}^{SI|S}\right) + (1-q_2)P_{jl|i}^{II|S} \,, \tag{3.1}$$

where q_1 and q_2 are the probabilities that i does not get infected when, respectively, only one or both of the two other nodes are infected. Equation (A.14) is recovered by taking $q_1 = 1-\beta$ and $q_2 = (1-\beta)^2\left(1-\beta^{\triangle}\right)$. Now, if we suppose that also exists a triangle $\{i,j,k\}$, sharing the two nodes i and j (i.e., the edge $\{i,j\}$) with the previous one, and try again to write down an expression for the overall probability for node i to get infected, we immediately recognize that the shared edge implies the contribution from node j to be double-counted. Indeed, the overall infection probability for i would be given by

$$1-\left(1-\mathbb{P}_{ijl}^{(i)}\right)\left(1-\mathbb{P}_{ijk}^{(i)}\right) = \mathbb{P}_{ijl}^{(i)} + \mathbb{P}_{ijk}^{(i)} - \mathbb{P}_{ijl}^{(i)}\mathbb{P}_{ijk}^{(i)} \,. \tag{3.2}$$

Since j belongs to both triangles, the pairwise infection coming from j would appear twice, proportional to $P_{jl|i}^{IS|S}$ and $P_{jk|i}^{IS|S}$ [3]. Please, notice that there is nothing special about edges in being over-counted. In general, if s simplices have a set of $m \geqslant 2$ nodes in common, i.e, they share a $(m-1)$-face, all the interactions included in the face are counted s times, instead of once. If $m > 2$, then the over-counted interactions are also non-pairwise.

3 This issue—and this only—can be solved by just accepting repeated interactions as part of the model through a reinterpretation of the interaction structure. In the current example, it would mean that the two triangles $\{i,j,l\}$ and $\{i,j,k\}$ represent two separate groups, one in which i and j gather with l, and one in which they gather with k. The two groups form in different moments within the same time step, during which the i-j interaction occurs two times. If then we want i and j to interact also as an isolated pair, we must add a maximal edge $\{i,j\}$, which would lead i and j to interact three times in a time step.

Shared edges (or larger subsets) bring however another problem. Analogously to what observed in Sec. A.3, the product $\mathbb{P}^{(i)}_{ijl}\mathbb{P}^{(i)}_{ijk}$ in Eq. (3.2) yields terms of the form $P^{I\sigma_l|S}_{jl|i}P^{S\sigma_k|S}_{jk|i}$ and $P^{I\sigma_l|S}_{jl|i}P^{I\sigma_k|S}_{jk|i}$ (et similia). The former are simply inconsistent; the latter wrongly assume that the state of a node (j) in a simplex is independent from the state of the same node in another simplex, as if the state of the node was redrawn each time. A correct formulation must therefore avoid such terms (which also prevent a proper marginalization of the equations, as seen for instance in Appendix A.3). To this end, we should disentangle the contribution of each node from the rest. However, in discrete time, this is impossible already for an extremely simple dynamics as the simplicial SIS we are considering. As a consequence, a microscopic discrete-time model whose equations are closed at the level of simplices, cannot be correct if the latter share more than one node. Indeed, all the aforementioned problems stem from the existence of a shared edge, and would disappear by simply replacing one of the two shared nodes (say j) with a node (say h) different from all the others.

We are thus left with two possibilities: (i) closing the equations at the level of maximal simplices under the hypothesis that any two of them share at most one node (as in Eq. (3.3)); or (ii) we close them at the level of generic maximal subgraphs (e.g., the 4-node subgraph of the example above), up to the point at which any two of such subgraphs share no more than one node. It is not too hard to convince ourselves that option (ii) is not a viable choice. Indeed, some of the subgraphs would consist of four or more nodes (if not, then we would fall back to option (i)) and the number of non-isomorphic configurations rapidly grows with the size of the subgraphs. Each non-isomorphic configuration yields a set of dynamical equations valid only for that specific subgraph. Additionally, the number of possible states for a subgraph increases exponentially with its size. While the computational cost of keeping track of the state evolution of such subgraphs becomes rapidly prohibitive, just writing down the equations for each non-isomorphic configuration, at each size, seems a practically unfeasible task, even considering small subgraphs only[4]. We therefore choose option (i) and leverage the perfect symmetry of simplices to get equations valid for simplices of any order. This choice implies dealing with simplicial complexes that are strictly 0-connected[5]. The fact that, in general, there could be simplices sharing more nodes, opens the problem of finding the strictly 0-connected simplicial complex which best approximates the original one. By "best" we mean the one that minimizes the number of disrupted shared faces. We address this problem in the next section. Please, observe that all the analysis above holds independently from the moment closure we could choose, as we indeed made no reference to any specific one.

If we wanted to keep simplices with more than one node in common, the only always valid alternative would be opting for a continuous-time formulation. Taking

4 Eventually, option (ii) is not even applicable to 1-connected simplicial complexes. In such case the equations would no longer be microscopic, as they would actually coincide with the exact macroscopic equations describing the state of the entire system.

5 Or, more generally, whose connected components are strictly 0-connected.

the example above, consisting of the two simplices $\{i,j,l\}$ and $\{i,j,k\}$, the rate at which node i gets infected is correctly written as

$$r\left(P^{I|S}_{j|i}+P^{I|S}_{l|i}+P^{I|S}_{k|i}\right)+r_{\triangle}\left(P^{II|S}_{jl|i}+P^{II|S}_{jk|i}\right)\,,$$

where r and $r_{\triangle}$ are the rates at which a node gets infected via a 2- and 3-node interaction, respectively. The reason why in continuous time there are neither over-counted interactions nor unwanted terms, is because the probability for two or more microscopic processes to occur within the same infinitesimal time interval is negligible. As a consequence, only one among the several infection channels occurs with finite probability within a small time interval. In turn, this allows us to disentangle the various channels. We show this in Appendix C.1.4, where a continuous-time model with 2- and 3-node interactions is provided.

In the following, sticking to a discrete-time formulation, we assume that the interaction structure is represented by a SC. We first address the problem of ensuring the strict 0-connectedness of the structure, while trying to preserve as many higher-order interactions as possible. Besides, since the dynamical equations for the state of an n-clique are just a special case of those for the state of a $(n+1)$-simplex, we take advantage of this to keep track of the dynamical correlations within, not just simplices, but cliques too.

3.1.1.1 Edge-disjoint edge clique cover

Let us indicate with $\mathcal{K}$ the simplicial complex and with $\mathcal{K}^{(1)}$ its underlying graph. To accounts for whether an n-clique c in $\mathcal{K}^{(1)}$ is the projection of a $(n-1)$-simplex—thus including higher-order interactions—or it is just an n-clique also in $\mathcal{K}$—including only pairwise interactions—, we introduce a binary variable g, the group classifier, defined to be 1 or 0, respectively; by which c is regarded as a (g,n)-clique. Since a 1-simplex is equivalent to a 2-clique, we choose to assign $g=0$ to any 2-clique, so that $n\geqslant 3$ when $g=1$. From now on, unless explicitly specified, we refer to the g-classified cliques in $\mathcal{K}^{(1)}$ simply as 'cliques'. These are the building blocks of our description.

To avoid the over-counting, in fact, we aim at covering all the edges of $\mathcal{K}^{(1)}$ by means of a set of cliques such that any two of them share at most one node, giving what we call here an *edge-disjoint edge clique cover*. Cliques sharing more than one node are consequently decomposed in lower-order cliques. This comes with no essential repercussions when the decomposed cliques have all $g=0$. Otherwise, some group interactions must be ignored. Covering the structure by cliques only, we are able to give a unique equation holding for any order n; an impossible task via generic, less symmetric subgraphs. Furthermore, since we want to capture as much correlations and group interactions as possible, we want the cover to consist of the least possible number of cliques. In this way, larger cliques—thus correlations and interactions of higher order—are preferably included in the cover. Finding such minimal set of edge-disjoint cliques is closely related to the edge clique cover problem, known to be NP-

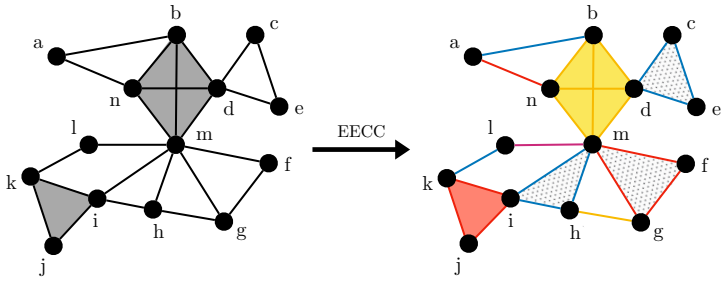


Figure 3.1: Edge-disjoint edge clique cover (EECC) of a small simplicial complex (SC). The SC, shown on the left, consists of fourteen nodes, identified via letters, connected by one 3-simplex, one 2-simplex, and fourteen 1-simplices. Grey areas indicate r-simplices with $r \geqslant 2$, including $r+1$ nodes each. The EECC of the simplicial complex is shown on the right, where colored and dotted areas are used to visualize, respectively, the $(1,r)$-cliques and $(0,r)$-cliques in it, with colors carrying no specific meanings. The simplicial complex is decomposed in: one $(1,4)$-clique, $\{b,d,m,n\}$; one $(1,3)$-clique, $\{i,j,k\}$; three $(0,3)$-cliques, $\{c,d,e\}$, $\{f,g,m\}$, $\{h,i,m\}$; and five $(0,2)$-cliques, $\{a,b\}$, $\{a,n\}$, $\{g,h\}$, $\{k,l\}$, $\{l,m\}$. The underlying subgraph induced by the subset $\{a,b,d,m,n\}$ originally consists of a $(1,4)$-clique and a $(0,3)$-clique. To get the EECC, the former must be included, whereas the $(0,3)$-clique is decomposed into two (non-maximal) $(0,2)$-cliques. Similarly, the underlying subgraph induced by the subset $\{f,g,h,i,m\}$ is made of three overlapping $(0,3)$-cliques, and the EECC is in this case obtained by first including $\{f,g,m\}$ and $\{h,i,m\}$ (and then the remaining edge $\{g,h\}$), instead of $\{g,h,m\}$.

complete [283]. Heuristics are thus necessary to estimate the solution in large graphs. For convenience, hereafter we reserve the acronym EECC (edge-disjoint edge clique cover) to sets which are solutions of the problem. If all the maximal cliques in $\mathcal{K}^{(1)}$ are edge-disjoint, then $\mathcal{K}^{(1)}$ admits a unique EECC, simply given by the set of the maximal cliques in it. Otherwise, $\mathcal{K}^{(1)}$ generally admits multiple EECCs. See Fig. 3.1 for illustration.

To estimate an EECC, we propose a greedy heuristic which proved to be the best among several others we conceived for this task. Given $\mathcal{K}^{(1)}$ is just a graph, let us consider a generic graph G. The heuristic proceeds as follows:

1. Find the set C of all the maximal cliques in G;
2. Include in the EECC and remove from C all the elements in C that do not share edges with other elements in C;
3. While C is not empty,
    a) For every maximal clique $c \in C$, compute the score r_c, defined as the fraction of edges that c shares with the other elements in C;

b) Consider the elements of C with the lowest score; include in the EECC and remove from C (one, randomly chosen, of) the element(s) of highest order among them.

The rationale behind step 3 is preferring those cliques which share less edges relatively to their size, i.e., to the potential number of edges they could share. Indeed, a clique c_1 has a lower score than a clique c_2 when (i) c_1 and c_2 are of the same size but c_1 shares a lower number of its edges; (ii) c_1 and c_2 share the same number of edges, but c_1 is larger. Therefore, generally, preferring cliques with lower scores brings relatively larger cliques into the cover, while sacrificing adjacent cliques whose inclusion would lead to the disruption of more cliques. The cover shown in Fig. 3.1 is the unique EECC in that case and is the one found by our greedy algorithm. Noteworthy, when dealing with modular structures—as those representing many real social systems—in which communities of nodes are loosely connected among them, the search for an (edge-disjoint) edge clique cover can be speeded up. Indeed, each time two regions of the structure are joined by bridging cliques (which are evidently maximal), the problem of finding the optimal cover of the whole structure reduces to that of finding it in each of the two, smaller regions.

As we observed, whenever the maximal cliques in $\mathcal{K}^{(1)}$ are not all edge-disjoint, some cliques are forced to be decomposed in sub-cliques during step 3. Since decomposing a $(1,\cdot)$-clique in (strictly 0-connected) sub-cliques means also neglecting some group interactions, whenever a $(0,\cdot)$-clique and a $(1,\cdot)$-clique are not edge-disjoint, we prefer to include the latter in the cover. Thus, being $\mathcal{K}$ a simplicial complex and $\mathcal{K}^{(1)}$ its g-classified underlying graph, a EECC of $\mathcal{K}$ is estimated as follows:

1. Consider the subgraph $G_1 \subseteq \mathcal{K}^{(1)}$ induced by the nodes in the maximal $(1,\cdot)$-cliques in $\mathcal{K}$. Estimate a EECC of G_1; let us call $\mathcal{S}(\mathcal{K})$ the resulting EECC;
2. Consider the subgraph $G_0 = \mathcal{K}^{(1)} \backslash G_1$. Estimate a EECC of G_0; let us call it $\mathcal{C}(\mathcal{K})$;
3. $\mathcal{D}(\mathcal{K}) \equiv (\mathcal{S}(\mathcal{K}), \mathcal{C}(\mathcal{K}))$ is the estimated EECC of $\mathcal{K}$.

A basic question is about the dependence of the prediction made by the model when different EECCs are estimated for a given structure. Multiple EECCs can indeed exist for the same structure and, even when the EECC is unique, a heuristic could reach one of many sub-optimal solutions (local minima). Since the estimated EECC is the structure on top of which a dynamical process is eventually considered, an indispensable quality of the heuristic is estimating EECCs of little variability. Otherwise, the predictions made by a model could vary substantially depending on the estimated EECC. In Appendix C.1.1, we test the contagion model presented in the next section in structures with a high proportion of non-edge-disjoint cliques, for which very different EECCs could be in principle estimated. Despite this, the predictions made by the model turn out to be remarkably consistent across different estimated covers, corroborating the robustness of our method. This suggests that, given a SC, it is safe to make use of the first EECC computed for it—at least for the contagion

dynamics herein considered. In fact, this can be expected given that the heuristic is potentially not deterministic only in step 3.b, when forced to choose among different cliques of largest size among those of lowest score. At every iteration, such cliques are usually very few and, as long as the structure is not too small, it is sensible to expect that they impact future iterations in a similar fashion. For this reasons, we expect the heuristic to estimate similar EECCs for a given structure, and therefore the evolutions predicted also for other dynamical processes to be robust across different realizations.

Calling m_1 the maximum order of the cliques we want to include in $\mathcal{S}(\mathcal{K})$ (i.e., considering simplices of dimension up to $m_1 - 1$ in $\mathcal{K}$), m_1 could be smaller than ω_1, the maximum order of the cliques in G_1. In such case, when looking for an EECC of G_1, those maximal cliques in G_1 of order greater than m_1 are decomposed in edge-disjoint sub-cliques (corresponding to sub-simplices in $\mathcal{K}$) of variable order $m' \in \{1, \dots, m_1\}$. Clearly, the higher is m_1, the higher is the order of the group interactions (and of the correlations within them) included in the description. Overall, as long as the proportion of obviated group interactions is small enough, the deviations from the complete dynamics are expected to be comparatively negligible; or, alternatively, the error made by including 1-connected simplices should be marginal.

In the building of an EECC, different values $m_0 \geqslant 2$ (with $m_0 \leqslant \omega_0$, being ω_0 the maximum order of the cliques in G_0) for the maximum order to be considered for $(0, \cdot)$-cliques can also be chosen. The higher is m_0, the higher is the order of the captured dynamical correlations within the cliques in $\mathcal{K}$. In summary, the couple (m_0, m_1) identifies the considered implementation of the model.

3.1.2 *Linear maximal-cliques closure*

Preliminarily to the definition of the model, let us first introduce the moment closure approximation we will use in the next section and throughout the rest of thesis. This is based on the condition that any two maximal cliques in the underlying graph of the hypergraph (not necessarily a SC) are edge-disjoint. This means that the hypergraph is linear at the level of maximal cliques, hence the name *linear maximal-cliques closure*.

As a generalization of the quenched pair approximation, we then assume the states of any two adjacent maximal cliques to be conditionally independent given the state of the node they share. Let $P_{i_1 \dots i_n}^{\sigma_{i_1} \dots \sigma_{i_n}}$ be the state probability for the clique $\{i_1, \dots, i_n\}$, i.e., the probability that node i_1 is in state σ_{i_1}, node i_2 is in state σ_{i_2}, and so on. Then, given two maximal cliques, $\{i, j_1, \dots, j_n\}$ and $\{i, k_1, \dots, k_m\}$, having node i in common, their joint state probability factorizes in the product of the state probabilities of the single cliques, one being conditioned on node i, i.e.,

$$P_{ij_1 \dots j_n k_1 \dots k_m}^{\sigma_i \sigma_{j_1} \dots \sigma_{j_n} \sigma_{k_1} \dots \sigma_{k_m}} = \frac{P_{ij_1 \dots j_n}^{\sigma_i \sigma_{j_1} \dots \sigma_{j_n}} P_{ik_1 \dots k_m}^{\sigma_i \sigma_{k_1} \dots \sigma_{k_m}}}{P_i^{\sigma_i}} = P_{j_1 \dots j_n | i}^{\sigma_{j_1} \dots \sigma_{j_n} | \sigma_i} P_{ik_1 \dots k_m}^{\sigma_i \sigma_{k_1} \dots \sigma_{k_m}} , \tag{3.3}$$

where the underline identifies i as the shared node and $P^{\sigma_{j_1}\dots\sigma_{j_n}|\sigma_i}_{j_1\dots\underline{j_n}|i}$ denotes the state probability for the subset $\{j_1,\dots,j_n\}$ conditioned on the state of node i. Notice that Eq. (3.3) is the special case of the cut-vertices closure given in Eq. (2.47) when the partition in subgraphs induced by the cut-vertices consists of maximal cliques only. Nonetheless, we find convenient give the closure in Eq. (3.3) its own name considered the extensive use we make of it throughout this thesis (independently from whether the shared node is or not a cut-vertex).

According to Eq. (3.3), the maximal cliques become the building blocks of the dynamics, accounting for the correlations among the states of the nodes they include. The dynamical correlations between adjacent cliques are conveyed by the state of the node they share. We expect that the larger the adjacent cliques are, the smaller the influence the common node has on all the others, and thus, the more accurate the moment closure is. The larger the adjacent cliques are, the smaller the influence the common node has on all the others, and thus, the more accurate the moment closure is.. The improvement is expected to be particularly prominent for random structures built from a configuration model. Indeed, since the coarse-grained network constructed by associating maximal cliques to 'super-nodes' is locally tree-like in this case, the original nodes often act as mighty shortcuts between different parts of the system, so that conditioning on their state makes adjacent cliques nearly independent. Nonetheless, while increasing the number of nodes encompassed by a closure allows us to retain more correlations and thus provide better predictions, there is currently no way of estimating how high the order of the closure must be to attain a preset level of accuracy.

3.1.3 *The MECLE model*

Let us build here the dynamical equations for a generic (g,n)-clique, $\{i_1,\dots,i_n\}$, in $\mathcal{D}(\mathcal{K})$, with $n \leqslant m \equiv \max\{m_0,m_1\}$. Let $P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g}$ be the associated state probability. In particular, for the SIS process, $\{\sigma_{i_1},\dots,\sigma_{i_n}\} \in \{S,I\}^n$. Clearly, summing over all the possible state configurations, it must hold

$$\sum_{\{\sigma_{i_1},\dots,\sigma_{i_n}\}} P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g} = 1\,. \tag{3.4}$$

For a shorter notation, let us denote with $\{\sigma_{i_k}\}_{k=1,\dots,n}$ the state of the clique at time t and with $\{\sigma'_{i_k}\}_{k=1,\dots,n}$ its state at time $t+1$. Assuming the system is markovian, the state of the system at time t contains all the information needed to predict the state at time $t+1$. The *microscopic epidemic clique equations* (MECLE), governing the evolution of the state of the (g,n)-clique $\{i_1,\dots,i_n\}$, read

$$P^{\sigma'_{i_1}\dots\sigma'_{i_n}}_{i_1\dots i_n,g}(t+1) = \sum_{\{\sigma_{i_1},\dots,\sigma_{i_n}\}} P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g}(t)\ \Phi_g\left(\{\sigma'_{i_k}\};\{\beta^{(s)}\},\mu\middle|\{\sigma_{i_k}\}\right)\,, \tag{3.5}$$

where

$$\Phi_g\left(\{\sigma'_{i_k}\};\{\beta^{(s)}\},\mu\middle|\{\sigma_{i_k}\}\right)=\prod_{k=1}^{n}\phi_{i_k,g}\left(\sigma'_{i_k};\{\beta^{(s)}\},\mu\middle|\{\sigma_{i_k}\}\right) \tag{3.6}$$

is the probability of transitioning to the arrival state $\{\sigma'_{i_k}\}_{k=1,\dots,n}$ conditioned on starting in state $\{\sigma_{i_k}\}_{k=1,\dots,n}$. It is understood that Φ_g is computed at time t. This is expressed as a product over the single-node transition probabilities $\{\phi_{i_k,g}\}_{k=1,\dots,n}$. The single-node transitions are thus independent events within the time step. This derives from the implicit assumption that all the events within a given time step are simultaneous, and therefore not causally related. Introducing the indicator function $\mathbf{1}_p$, returning 1 if proposition p is true and 0 otherwise, $\phi_{i_k,g}$ is written as

$$\begin{aligned}\phi_{i_k,g}\left(\sigma'_{i_k};\{\beta^{(s)}\},\mu\middle|\{\sigma_{i_k}\}\right)&=\mathbf{1}_{(\sigma_{i_k}=I,\ \sigma'_{i_k}=I)}\,(1-\mu)+\mathbf{1}_{(\sigma_{i_k}=I,\ \sigma'_{i_k}=S)}\ \mu\\&\quad+\mathbf{1}_{(\sigma_{i_k}=S,\ \sigma'_{i_k}=I)}(1-Q_{i_k})+\mathbf{1}_{(\sigma_{i_k}=S,\ \sigma'_{i_k}=S)}\ Q_{i_k}\end{aligned} \tag{3.7}$$

where $Q_{i_k}\equiv Q_{i_k}\left(\sigma'_{i_k};\{\beta^{(s)}\}\middle|\{\sigma_{i_k}\}\right)$ is the probability that node i_k does not get the infection. Denoting with $N_I=\left|\{i_{k=1,\dots,n}|\sigma_{i_k}=I\}\right|$ the number of infected nodes in the starting state, Q_{i_k} is given by

$$Q_{i_k}=w^{(n-1)}_{N_I,g}\,q^{(n-1)}_{i_k(\neg i_k),g}\prod_{(g',r)\neq(g,n)}q^{(r)}_{i_k,g'}\ , \tag{3.8}$$

where the product runs over the couples (g',r) such that $1\leqslant r\leqslant m_0-1$ for $g'=0$, and $2\leqslant r\leqslant m_1-1$ for $g'=1$, while excluding the couple (g,n). Here, $w^{(n-1)}_{l,g}\equiv w^{(n-1)}_{l,g}\left(\{\beta^{(s)}\}\right)$ is the probability that a susceptible node (e.g., i_k) does not get infected within a (g,n)-clique (resp., $\{i_1,\dots,i_n\}$) where l out of the $n-1$ other nodes are infected. For the simplicial SIS model, we have

$$w^{(n-1)}_{l,g}=\mathbf{1}_{g=0}\left(1-\beta^{(1)}\right)^{l}+\mathbf{1}_{g=1}\prod_{s=1}^{l}\left(1-\beta^{(s)}\right)^{\binom{l}{s}}\ . \tag{3.9}$$

Notice that the (no-)infection kernel, $w^{(n-1)}_{l,g}$, can be adapted to model other contagion dynamics. Accordingly, the MECLE are by no means restricted to the SIS dynamics considered here. Lastly, $q^{(r)}_{i_k,g}\equiv q^{(r)}_{i_k,g}\left(\{\beta^{(s)}\}\right)$ is the probability that node i_k does not get infected via any of the $(g,r+1)$-cliques incident on it, that is

$$q^{(r)}_{i_k,g}=\prod_{\{j_1,\dots,j_r\}\in\Gamma^{(r)}_{i_k,g}}\left[1-\sum_{l=1}^{r}\frac{1-w^{(r)}_{l,g}}{l!\,(r-l)!}\sum_{k_1\neq\dots\neq k_r=1}^{r}P^{I\dots IS\dots S|S}_{j_{k_1}\dots j_{k_l}j_{k_{l+1}}\dots j_{k_r}|i_k,g}\right]\ , \tag{3.10}$$

being $\Gamma^{(r)}_{i_k,g}$ the set of r-tuples of indexes corresponding to subsets of r nodes forming a $(g,r+1)$-clique with i_k (if $\Gamma^{(r)}_{i_k,g}=\varnothing$, then $q^{(r)}_{i_k,g}$ is put to 1). The quantity $q^{(n-1)}_{i_k(\neg i_k),g}$ coincides with $q^{(n-1)}_{i_k,g}$ except for excluding the considered (g,n)-clique $\{i_1,\dots,i_n\}$ from

the product, as its contribution is already accounted for by $w^{(n-1)}_{N_I,g}$. To get Eq. (3.10), note we have used the linear maximal-cliques closure, Eq. (3.3).

At this point, the probability P^I_i for the single node i to be infected, can be computed either marginalizing it from Eq. (3.5), by summing over the state probabilities of any (g,n)-clique including i where $\sigma_i = I$ (e.g., for $(g,n) = (0,2)$, $P^I_i = P^{IS}_{ij,0} + P^{II}_{ij,0}$), or directly taking $n = 1$ in Eq. (3.5) (notice $w^{(0)}_{N_I,g} = 1$ and $q^{(0)}_{i,g} = 1$). Either way, one gets,

$$P^I_i(t+1) = P^I_i(t)(1-\mu) + P^S_i(t)\left(1 - \prod_{(g,r)} q^{(r)}_{i,g}\right) . \tag{3.11}$$

By averaging $P^I_i(t)$ over all the nodes, we find the expected prevalence, $\rho(t)$, i.e., the expected proportion of infected nodes at time t,

$$\rho(t) = \frac{1}{N}\sum_{i=1}^{N} P^I_i(t) . \tag{3.12}$$

This is the order parameter of the system.

At any time step t, Eq. (3.11) yields one constraint for each of the $n \geqslant 2$ nodes in a $(\cdot,n)$-clique, leaving $2^n - n - 1$ independent state probabilities to be determined, one being fixed by the normalization, Eq. (3.4). Therefore, if $\mathcal{K}$ has N nodes and $\mathcal{D}(\mathcal{K})$ consists of $C^{(n)}$ $(\cdot,n)$-cliques, $n \leqslant m$, the MECLE reduce to a system of (at most) $n_{\text{eq}} = N + \sum_{n=2}^{m}(2^n - n - 1)C^{(n)}$ independent equations.

With the MECLE in our hands, we can frame the existent models in the (m_0, m_1) notation. As $m_1 = 2$ implies $\mathcal{S}(\mathcal{K}) = \varnothing$, $\{(m_0,2)\}_{2\leqslant m_0 \leqslant \omega_0}$ is the class of models accounting for correlations within cliques of order up to m_0 on pairwise networks. In particular, $(2,2)$ gives the ELE [229]. The original, node-based MMCA [222], is then recovered from $(2,2)$ by neglecting pairwise correlations assuming $P^{SI}_{ij} = P^S_i P^I_j$ (i.e., $P^{I|S}_{j|i} = P^I_j$ in Eq. (3.10)). Considering many-body interactions ($m_1 \geqslant 3$), the simplicial ELE and MMCA [230] fall outside the MECLE framework due their internal inconsistency (see Sec. A.3). The MECLE (and their continuous-time counterpart given in Appendix C.1.4) thus define the first model able to consistently account for both higher-order dynamical correlations and higher-order interactions in quenched structures. Given the complexity of the MECLE, we show their $(3,3)$ implementation explicitly in Appendix C.1.2. This is the specific model we use in the next section to study the contagion process in the presence of 2- and 3-body interactions, organized in simplicial 2-complexes. The $(3,3)$ model accounts for 3-node correlations within either $(0,3)$- and $(1,3)$-cliques.

Lastly, linearizing the MECLE as shown in Appendix B.1.1, we derive an implicit equation, Eq. (B.12), relating the invasion threshold, $\beta^{(1)}_{\text{cr}}$, to the dominant eigenvalue of a $N \times N$ matrix encoding both the structural and the dynamical details of the system, which allows us to compute $\beta^{(1)}_{\text{cr}}$ analytically. Equation (B.12) holds for structures boasting cliques and interactions of any order, and it is not restricted to the simpli-

cial contagion considered here—as neither are the MECLE—, given the (no-)infection kernel is completely defined by Eq. (3.9), which can be modified to represent other contagion dynamics.

3.1.3.1 Approaches for groups sharing multiple nodes

Allowing group interactions to share two or more nodes comes with the drawback of impeding the computation of the invasion threshold through linearization (see Appendix B.1.1). Forgoing the latter, let us briefly discuss other possible approaches in discrete time.

One option are the simplicial ELE [230], which admit generic simplicial 2-complexes. Despite their internal inconsistency, the simplicial ELE turn out to be accurate when the strictly 0-connectedness is heavily broken (and the MECLE are unreliable), at least for less sparse simplicial complexes and small enough values of $\beta^{(2)}$ [230], for which the inconsistency has weaker consequences.

An alternative approach, is to opt for other hypergraphs by dropping out the hereditary property. In particular, referring to the class of simple hypergraphs [86], i.e., those in which a hyperedge—what in a simplicial complex is a face—cannot be subset of any other hyperedge, one can resort to a partially consistent approach. As anticipated in Sec. 3.1.1, in such structures, to each configuration of the group corresponds a unique channel of infection, so the necessity of constraining higher-order interactions to share no more than one node can be relaxed. The structure of the MECLE remains unchanged. One only needs to specify a suitable form for $w_{l,1}^{(r)}$, being $g = 1$ for hyperedges including more than two nodes. For example, requiring for transmission all the other nodes in a hyperedge to be infected, we would write $w_{l,1}^{(r)} = \mathbf{1}_{l=r}\left(1-\beta^{(r)}\right)$. Furthermore, for linear hypergraphs [86], in which two hyperedges can only share one node, the invasion threshold is again calculable as done in Appendix B.1.1.

3.1.4 *Results for simplicial 2-complexes*

We apply here the MECLE to the case of 2-dimensional SCs. These provide a representation for a population where individuals interact via either 2- or 3-body interactions. The simplicial complexes are constructed from some graph (the future underlying graph) by converting into $(1,3)$-cliques (i.e., 2-faces) a fraction $p_\triangle$ of the 3-cliques allowed by the EECC of the graph, while considering as $(0,3)$-cliques the remaining fraction, $1-p_\triangle$. Specifically, if all the 3-cliques are converted ($p_\triangle = 1$), the resulting simplicial complex is the clique complex of the original graph. A clique complex of a graph is the simplicial complex built by promoting each clique (e.g., an n-clique) in the graph to a simplex over the same nodes (resp., a $(n-1)$-simplex) in the SC. A 2-dimensional clique complex is specifically a simplicial complex where every 3-clique in its underlying graph corresponds to a 2-simplex in the complex. To better appreciate the improvement brought by the MECLE with respect to previous models,

we mainly show results for clique complexes, although notable improvements can be generally found for wide ranges of $p_\triangle$ depending on the used structure. Conveniently, if $p_\triangle$ is not specified then the structure is understood to be a clique complex. By the way, indicating with $\bar{k}^{(g,r)}$ the average number of $(g, n+1)$-cliques incident on a node, a simplicial complex is a clique complex if $\bar{k}^{(0,r)} = 0$ for $r > 1$ ($\bar{k}^{(0,2)} = 0$ in 2-dimensional SCs).

We identify a generated simplicial complex adding 'SC' to the name of the graph (model) used for it. Among those we use, a 'periodic triangular SC' is a simplicial complex built from a periodic triangular lattice, while a 'Dorogovtsev-Mendes SC' is a simplicial complex built upon a graph given by the Dorogovtsev-Mendes generative model [147]. In the latter, both the $(\cdot, 1)$- and $(\cdot, 2)$-degree distributions nearly follow power-laws with exponent around 3. We also refer to 'random SC's, obtained as follows. First we generate a simplicial 2-complex through the random simplicial complex model [199]. This yields homogeneous, sparse simplicial complexes with very few 1-connected triangles. Nonetheless, to exclude them and ensure the simplicial complex to be strictly 0-connected, we consider its underlying graph, estimate its EECC, and finally convert 3-cliques into 2-faces. The resulting simplicial complexes have $(\cdot, 1)$- and $(\cdot, 2)$-degree Poissonian distributions. All the used simplicial complexes consists of around $N = 10^4$ nodes.

We assess the performance of the MECLE and of the other discrete-time markovian models in the literature, i.e., the simplicial ELE and MMCA [230], by comparing their predictions for the prevalence, ρ, with those provided by Monte Carlo (MC) simulations. Each simulation is carried out according to the quasistationary state (QS) method [284, 285], which allows for the estimation of the asymptotic behavior of finite-sized systems in the presence of an absorbing state ($\rho = 0$). See Appendix A.2 for details about the method. The system is given a transient time of 10^5 time steps, and then ρ is computed as an average over 2×10^4 additional time steps.

Figure 3.2 shows that the improvement brought by the MECLE with respect to the simplicial ELE is substantial in both predicting the epidemic prevalence and, even more, locating the critical point. For the latter we report in the caption the values of the relative absolute error $\varepsilon_{\beta^{(1)}} = |\beta^{(1)}_{\text{model}}/\beta^{(1)}_{\text{MC}} - 1|$ produced by the two models, where $\beta^{(1)}_{\text{model}}$ is the threshold they predict and $\beta^{(1)}_{\text{MC}}$ the one resulting from MC simulations. The improvement is notable across all the considered SCs, despite their strong structural dissimilarities. We report in Fig. 3.2 also the predictions made by the MMCA and MF approximations of both the MECLE and the simplicial ELE. While all these approximations are, as expected, much less accurate than the more refined models (sometimes even wrongly predicting the nature of the phase transition, like seen in panel (c)), they are not equivalent among them. Indeed, the MMCA and MF approximations of the MECLE are equally or more accurate than their counterparts derived from the simplicial ELE.

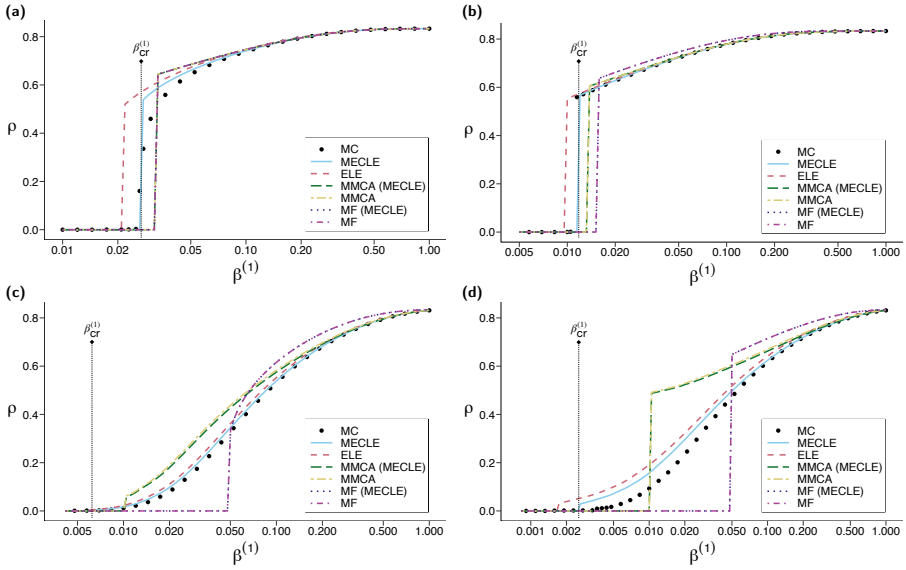


Figure 3.2: Epidemic prevalence ρ as function of the pairwise infection probability $\beta^{(1)}$ on different 2-dimensional SCs. Notice the x-axis is in logarithmic scale to better appreciate the results for small values of $\beta^{(1)}$. Results obtained from Monte Carlo (MC) simulations are depicted by dots, while lines represent the analytically computed prevalence using the indicated models. MMCA and MMCA(MECLE) refer to the Microscopic Markov Chain approximation of, respectively, the simplicial ELE model and the MECLE model, as obtained by considering the state probabilities of the nodes as uncorrelated; while MF and MF(MECLE) refer to their node-based, homogeneous mean-field approximations (see Appendix B.1.1.1). Note that MF and MF(MECLE) are indistinguishable at the used scale. The value of the invasion threshold predicted by the MECLE, as computed in Appendix B.1.1, is marked with a vertical dotted line. The recovery probability is fixed to $\mu = 0.2$. (a) Periodic triangular simplicial complex with $\bar{k}^{(0,1)} = 0.00$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 3.00$, and 3-body infection probability $\beta^{(2)} = 0.25$; the relative absolute error in locating the invasion threshold is $\varepsilon_{\beta^{(1)}} \approx 0.08$ for MECLE and $\varepsilon_{\beta^{(1)}} \approx 0.12$ for simplicial ELE. (b) Random simplicial complex with $\bar{k}^{(0,1)} = 4.10$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 3.95$, and $\beta^{(2)} = 0.15$; $\varepsilon_{\beta^{(1)}} \approx 0.06$ for MECLE and $\varepsilon_{\beta^{(1)}} \approx 0.09$ for simplicial ELE. (c) Dorogovtsev-Mendes simplicial complex with $\bar{k}^{(0,1)} = 1.10$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 1.45$, and $\beta^{(2)} = 0.25$; $\varepsilon_{\beta^{(1)}} \approx 0.07$ for MECLE and $\varepsilon_{\beta^{(1)}} \approx 0.16$ for simplicial ELE. (d) Same as (c) but with $\beta^{(2)} = 0.50$; $\varepsilon_{\beta^{(1)}} \approx 0.23$ for MECLE and $\varepsilon_{\beta^{(1)}} \approx 0.50$ for simplicial ELE.

We then leverage the prominent jump of the discontinuous transition[6] found for random simplicial complexes (Fig. 3.2(b)), to illustrate the existence of a hysteresis

6 This jump is very small for Dorogovtsev-Mendes simplicial complexes due to the degree heterogeneity. Indeed, as proved by Juhn *et al.* [274], the discontinuity is expected to disappear for enough heterogeneous scale-free networks.

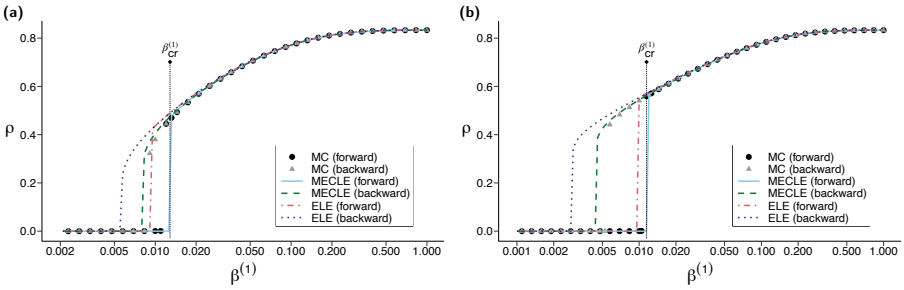


Figure 3.3: Hysteresis cycle of the epidemic prevalence ρ with respect to the pairwise infection probability $\beta^{(1)}$ on different 2-dimensional SCs. The 'forward' and 'backward' curves are obtained through small equilibrium transformations taking as initial value ρ_0 the (quasi-)equilibrium value of ρ got at the next smaller and next greater value of $\beta^{(1)}$, respectively. The hysteresis cycle reveals the bistable region, located between the leftmost, persistence threshold, and the rightmost, invasion threshold. The computed value of the latter is marked for the MECLE with a vertical dotted line. The recovery probability is fixed to $\mu = 0.2$. (a) Random simplicial complex with $\bar{k}^{(0,1)} = 4.10$, $\bar{k}^{(0,2)} = 1.58$ and $\bar{k}^{(1,2)} = 2.37$ ($p_\triangle = 0.6$), and 3-body infection probability $\beta^{(2)} = 0.25$; the relative errors in locating the critical threshold are $\varepsilon^{\text{for.}}_{\beta^{(1)}} \approx 0.14$ and $\varepsilon^{\text{back.}}_{\beta^{(1)}} \approx 0.10$ for MECLE, and $\varepsilon^{\text{for.}}_{\beta^{(1)}} \approx 0.18$ and $\varepsilon^{\text{back.}}_{\beta^{(1)}} \approx 0.38$ for simplicial ELE. (b) Random simplicial complex with $\bar{k}^{(0,1)} = 4.10$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 3.95$ ($p_\triangle = 1.0$), and $\beta^{(2)} = 0.15$; $\varepsilon^{\text{for.}}_{\beta^{(1)}} \approx 0.06$ and $\varepsilon^{\text{back.}}_{\beta^{(1)}} \approx 0.15$ for MECLE, and $\varepsilon^{\text{for.}}_{\beta^{(1)}} \approx 0.09$ and $\varepsilon^{\text{back.}}_{\beta^{(1)}} \approx 0.47$ for simplicial ELE.

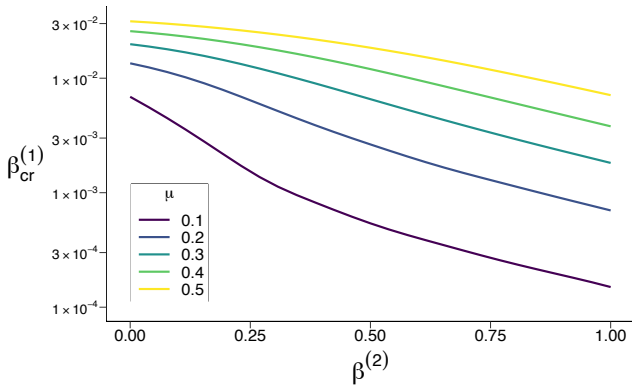


Figure 3.4: Dependence of the invasion threshold on the 3-body infection probability $\beta^{(2)}$. Computed as in Appendix B.1.1, $\beta^{(1)}_{\text{cr}}$ is shown against $\beta^{(2)}$ for a Dorogovtsev-Mendes simplicial complex with $\bar{k}^{(0,1)} = 1.10$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 1.45$. Note that node-based models, disregarding dynamical correlations, would wrongly predict $\beta^{(1)}_{\text{cr}}$ to be independent from $\beta^{(2)}$ (see Appendix B.1.1.1).

cycle, enclosing a bistable region, in Fig. 3.3. Again, the simplicial ELE are outperformed by the MECLE. This is even more evident for the 'backward' curves and can be expected from the analysis in Sec. A.3. There we showed that the overestimation of the probability of getting infected made by the simplicial ELE grows, not only with $\beta^{(2)}$, but also with the probability that two nodes in a triangle are infected—activating the 3-body infection channel. Such probability takes relatively high values at the upper backward branch, making the overestimation worse, and in turn resulting in the persistence threshold being shifted towards smaller values.

More interestingly, our analysis clearly shows that node-based models fail to pinpoint the invasion threshold at even a qualitative level. Indeed, by neglecting dynamical correlations, they predict (see Appendix B.1.1.1) $\beta_{\rm cr}^{(1)}$ to be independent from the values of the higher-order infection probabilities, $\{\beta^{(s)}\}_{s>1}$. In particular, in the case of simplicial 2-complexes, increasing enough $\beta^{(2)}$ results in the appearance of a bistable region also in those models [199, 275], but the $\beta_{\rm cr}^{(1)}$ they predict is unaffected. On the contrary, due to dynamical correlations, increasing $\beta^{(2)}$ does lead $\beta_{\rm cr}^{(1)}$ to decrease, as arises from MC simulations and correctly predicted by the MECLE. We show this dependence in Fig. 3.4 for Dorogovtsev-Mendes SCs, while it can already be grasped by comparing panels (c) and (d) in Fig. 3.2. While an explicit equation for $\beta_{\rm cr}^{(1)}$ is generally inaccessible for complex interaction structures, in Appendix C.1.3 we get closer to it for particularly symmetric structures. We analyze regular SCs, of which the periodic triangular simplicial complex is a special instance, and simplicial complexes built upon Friendship graphs (a generalization of star graphs), respectively representative of the two extremes of structural homogeneity and heterogeneity. We prove that $\beta_{\rm cr}^{(1)}$ decays with $\beta^{(2)}$, and that the dependence considerably increases with the structural heterogeneity, in accordance with what reported in Figs. 3.2 and 3.4. The decrease exists independently from the size N of the system, excluding it from being a finite-size effect.

To notice that the dependence of the invasion threshold on $\beta^{(2)}$ had shown up already in previously reported numerical simulations in simplicial complexes built from real data [199]. However, it went seemingly overlooked, eventually leading to claim [286] that only pairwise interactions govern the value of $\beta_{\rm cr}^{(1)}$. Even though we have considered a discrete-time model, instead of the continuous-time models upon which that claim is based on, we can prove that the existence of the qualitative shift predicted by our analysis has nothing to do with the character of time. The continuous-time limit of the MECLE is easily derived as in Appendix C.1.4, where we show that $\beta^{(2)}$ indeed lowers the threshold, and to a similar extent than in the discrete-time model (compare Figs. 3.4 and C.5). As in the discrete-time formulation, if dynamical correlations are preserved, the terms responsible for the higher-order infection channels do not disappear from the linearized equations around the inactive state, eventually contributing to the value of the invasion threshold.

To close the analysis, we consider clique complexes with some fraction $p_{\rm s}$ of edges shared by two or more 2-faces. We test in this way how the MECLE behave when

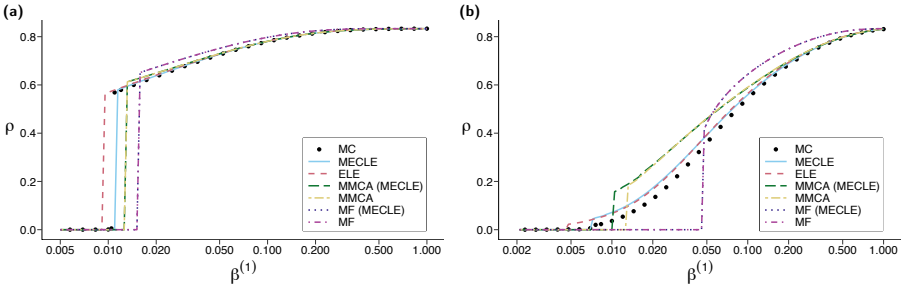


Figure 3.5: Epidemic prevalence ρ as function of the link infection probability $\beta^{(1)}$ on 2-dimensional simplicial complexes with low percentage p_s of shared non-maximal edges. p_s is computed as the fraction of non-maximal edges which are included in more than one 2-face. Besides, with $\bar{s}$ we indicate the average number of 2-faces in which the edges corresponding to the fraction p_s are included ($\bar{s} \geqslant 2$). The recovery probability is fixed to $\mu = 0.2$. (a) Random simplicial complex with $\bar{k}^{(0,1)} = 2.25$, $\bar{k}^{(0,2)} = 0.00$ and $\bar{k}^{(1,2)} = 5.20$; $p_s = 0.03$, $\bar{s} = 2.01$, and 3-body infection probability $\beta^{(2)} = 0.15$. (b) Dorogovtsev-Mendes simplicial complex with $\bar{k}^{(0,1)} = 1.10$, and $\bar{k}^{(0,2)} = 0.00$ $\bar{k}^{(1,2)} = 1.60$, $p_s = 0.05$, $\bar{s} = 2.06$, and $\beta^{(2)} = 0.25$.

strict 0-connectedness is slightly broken. As shown in Fig. 3.5, the MECLE perform comparably to or still better than the simplicial ELE when p_s is low enough.

3.1.5 *Summary*

The Microscopic Epidemic Clique Equations we developed constitute a clique-based quenched model for either simple and complex contagion processes able to preserve dynamical correlations within cliques/hyperedges of virtually any order. By doing so, the model is able to reproduce numerical results with unprecedented accuracy, and predicts how either first-order and higher-order infection probabilities contribute to destabilize the inactive state.

The observed dependence of the invasion threshold, $\beta_{cr}^{(1)}$, on $\beta^{(2)}$ and the ability of the MECLE to predict it, reveals the necessity of accounting for higher-order dynamical correlations. In the considered dynamics, an interaction (infection channel) of order s, with coupling $\beta^{(s)}$, requires s nodes to be active (infected). Therefore, in order to preserve its contribution near the threshold, the state probability with s active nodes must not be neglected. Near the invasion threshold, higher-order interactions rely exclusively on comparatively higher-order correlations, so that their effect when varying the higher-order couplings, $\{\beta^{(s)}\}_{s>1}$, is observed only if such correlations are preserved. Accordingly, such dependence eludes node-based (MF and MMCA) approximations of the system, while pair-based ones (like the simplicial ELE), by partially accounting for second-order correlations within triangles, do predict this dependence, but with low accuracy. Importantly, since the contribution of higher-order

couplings to the invasion threshold stems from dynamical correlations, we expect it to diminish when the structure is not quenched but varies in time, eventually disappearing for a fully annealed structure. This has been in fact confirmed by a later study [287].

3.2 THE ROLE OF INTERACTION OVERLAP ON THE SPREAD OF HIGHER-ORDER CONTAGIONS

The microscopic details and the order of the correlations the MECLE account for, allowed us to obtain remarkably accurate predictions in several simplicial 2-complexes. Despite the complexity of the model, it has been possible to find an implicit equation to compute the invasion threshold, $\beta_{\rm cr}^{(1)}$, analytically (see Appendix B.1.1). Unfortunately, an explicit formula remains inaccessible even for very simple structures. The functional relation connecting $\beta_{\rm cr}^{(1)}$ to $\beta^{(2)}$ (and higher-order couplings) and to the generalized degree distributions is therefore still unknown. Moreover, so far we studied only simplicial complexes, and we don't know what to expect for generic hypergraphs. Simplicial complexes are indeed strongly constrained hypergraphs due to the requirement that all the subsets of any hyperedge (face) need to be included in the hypergraph too. Equally constrained, but in the opposite direction, are simple hypergraphs, where none of the subsets of any hyperedge can be included in the hypergraph. Simplicial complexes and simple hypergraphs thus represent the two extremes of a spectrum of "intermediate" hypergraphs where hyperedges generically include some of their subsets. Recent research suggests that empirical hypergraphs are found along the entire spectrum [288–290]. From a physical point of view, going through this spectrum means considering structures with different degrees of overlap between interactions of different orders. Such overlap is maximal in simplicial complexes and zero in simple hypergraphs: the existence of an n-body interaction, in the former, implies the presence of all the interactions of lower order within the same set of nodes; in the latter, conversely, it implies their absence. This naturally leads to the question about if and how the dynamics is affected by the degree of overlap.

In this section, we try to address all these open questions for rank-3 hypergraphs. We do it by first deriving a mean-field triadic approximation that, going beyond standard node-based and pair-based approximations, allows us to easily explore the aforementioned spectrum of structures. This improved mean-field model reveals how the stability of the inactive state depends on both $\beta^{(1)}$ and $\beta^{(2)}$, and relates this dependence to the degree of overlap between 2- and 3-body interactions. We derive a closed-form formula for the critical surface and demonstrate the overlap to have a double-edged effect: it lowers the invasion threshold, but makes the outbreaks generally smaller.

3.2.1 *Mean-field triadic approximation*

We consider random hypergraphs that are sparse to the extent that the probability for two maximal cliques in its underlying graph to share more than one node vanishes in the thermodynamic limit[7]. As we want to model generic rank-3 hypergraphs with various degrees of overlap between 2- and 3-edges, we need to account for the fact that a 3-clique in the underlying graph, when projected back to the hypergraph, comes in one of three flavors: a 3-cycle (i.e., a length-3 cycle), conveying three 2-body interactions, a 3-edge, conveying a single 3-body interaction, or a 2-simplex (or triangle), conveying all of them. To this end, we introduce the following binary tensors: $A^{(1)}$, such that $A^{(1)}_{ij} = 1$ if the maximal 2-edge (link) $\{i,j\}$ exists; $A^{(1,0)}$ and $A^{(0,1)}$, such that $A^{(1,0)}_{ijl} = 1$ ($A^{(1,0)}_{ijl} = 0$) and $A^{(0,1)}_{ijl} = 0$ ($A^{(0,1)}_{ijl} = 1$) if $\{i,j,l\}$ is a 3-cycle (3-edge); and such that $A^{(1,0)}_{ijl} A^{(0,1)}_{ijl} = 1$ if $\{i,j,l\}$ is a triangle. Triangles are therefore accounted for by the tensor $A^{(1,1)} = A^{(1,0)} \odot A^{(0,1)}$, indicated with '$\odot$' the element-wise product. With these definitions, the limit case of a simplicial complex is recovered imposing that, for any 3-clique $\{i,j,l\}$, $A^{(0,1)}_{ijl} = 1 \Rightarrow A^{(1,0)}_{ijl} = 1$ (hence $A^{(1,1)}_{ijl} = 1$), so that the existence of a 3-edge implies the existence of the 2-edges it includes. The opposite limit of a simple hypergraph is instead recovered requiring $A^{(0,1)}_{ijl} = 1 \Rightarrow A^{(1,0)}_{ijl} = 0$ (hence $A^{(1,0)}_{ijl} = 0$), so that the 2-edges are all maximal. By the way, notice that a simple rank-3 hypergraph is also a linear hypergraph, for any two hyperedges will share at most one node. We prefer using 'linear' hereafter, as it explicitly reflects the assumed sparsity condition and makes a direct connection to the linear maximal-cliques closure (Eq. (3.3)) we use. Figure 3.6(a) displays some illustrative examples of hypergraphs built according to the presented formalism.

Considering up to 3-body interactions, we account for the evolution of the state probability $P_i^{\sigma_i}$ for node i to be in state σ_i, $P_{ij}^{\sigma_i\sigma_j}$ for the maximal link $\{i,j\}$ to be in state $\sigma_i\sigma_j$, $P_{ijl}^{\sigma_i\sigma_j\sigma_l}$ for the (maximal) 3-clique $\{i,j,l\}$ to be in state $\sigma_i\sigma_j\sigma_l$. The state probability of any other local structure is approximated in terms of the maximal 2- and 3-cliques composing it through the linear maximal-cliques closure, Eq. (3.3). This means approximating the state of the following motifs: two adjacent maximal 2-cliques, a maximal

7 The number of maximal edges of any cardinality must scale slower than quadratically with the system size.

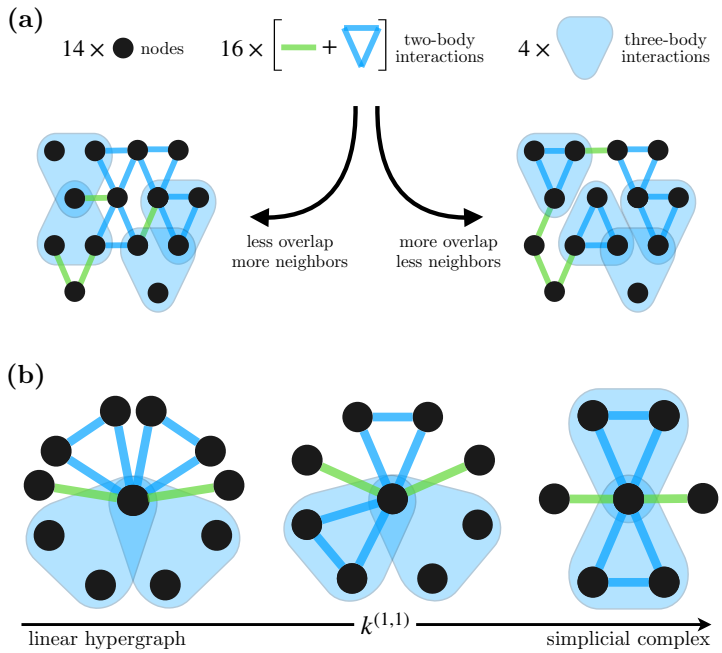


Figure 3.6: (a) Two hypergraphs built from the same number of nodes, 2-body interactions ($\kappa^{(1)} = 16$) and 3-body interactions ($\kappa^{(2)} = 4$), but arranged to have different numbers of 3-cycles, 3-edges, and triangles. From left to right, the degree of overlap grows from $1/4$ to $3/4$, while the number of neighbors per node drops from $25/14$ to $19/14$. (b) Example of how the neighborhood of a focal node changes with $k^{(1,1)}$ while keeping $\kappa^{(1)} = 4$ and $\kappa^{(2)} = 2$ fixed. Varying $k^{(1,1)}$ from 0 to $\kappa^{(2)}$, the degree of overlap goes from 0 (linear hypergraph) to 1 (simplicial complex), correspondingly leading to a smaller and more redundant neighborhood.

2-clique adjacent to a 3-clique, and two adjacent 3-cliques. Their state is thus decomposed as

$$P_{ij\underline{jl}}^{\sigma_i\sigma_j\sigma_l} = \frac{P_{ij}^{\sigma_i\sigma_j} P_{jl}^{\sigma_j\sigma_l}}{P_j^{\sigma_j}} , \tag{3.13a}$$

$$P_{ij\underline{jlh}}^{\sigma_i\sigma_j\sigma_l\sigma_h} = \frac{P_{ij}^{\sigma_i\sigma_j} P_{jlh}^{\sigma_j\sigma_l\sigma_h}}{P_j^{\sigma_j}} , \tag{3.13b}$$

$$P_{ij\underline{jl}hk}^{\sigma_i\sigma_j\sigma_l\sigma_h\sigma_k} = \frac{P_{ijl}^{\sigma_i\sigma_j\sigma_l} P_{lhk}^{\sigma_l\sigma_h\sigma_k}}{P_l^{\sigma_l}} . \tag{3.13c}$$

In the following, we refer to Eqs. (3.13) as the (mean-field) *triadic approximation*.

Having shown in Sec. 3.1 that a continuous-time formulation yields the same phenomenology, in the following we refer to such a formulation to further simplify the

model. We first write down the exact microscopic equations for the SIS process with infection rates $\beta^{(1)}$ and $\beta^{(2)}$ for 2- and 3-body interactions, respectively. The equations are then closed using the triadic approximation. Rescaled time by the recovery rate, μ, without loss of generality, the process is eventually described by the following system,

$$\begin{aligned}
\dot{P}_i^I =& -P_i^I + \beta^{(1)} \sum_j A_{ij}^{(1)} P_{ij}^{SI} \\
&+ \frac{1}{2} \sum_{j,l} \left[A_{ijl}^{(1,0)} \beta^{(1)} (P_{ijl}^{SSI} + P_{ijl}^{SIS} + 2P_{ijl}^{SII}) + A_{ijl}^{(0,1)} \beta^{(2)} P_{ijl}^{SII} \right] && \text{(3.14a)} \\
\dot{P}_{ij}^{SI} =& -(1+\beta^{(1)}) P_{ij}^{SI} + P_{ij}^{II} - \beta^{(1)} \sum_{l \neq j} A_{il}^{(1)} P_{j\underline{il}}^{ISI} + \beta^{(1)} \sum_{l \neq i} A_{jl}^{(1)} P_{i\underline{jl}}^{SSI} \\
&- \frac{1}{2} \sum_{l,h} \left[A_{ilh}^{(1,0)} \beta^{(1)} (P_{j\underline{ilh}}^{ISIS} + P_{j\underline{ilh}}^{ISSI} + 2P_{j\underline{ilh}}^{ISII}) + A_{ilh}^{(0,1)} \beta^{(2)} P_{j\underline{ilh}}^{ISII} \right] \\
&+ \frac{1}{2} \sum_{l,h} [i \leftrightarrow j] && \text{(3.14b)} \\
\dot{P}_{ijl}^{SSI} =& -(1 + 2A_{ijl}^{(1,0)} \beta^{(1)}) P_{ijl}^{SSI} + P_{ijl}^{ISI} + P_{ijl}^{SII} \\
&- \beta^{(1)} \sum_{h \neq j,l} A_{ih}^{(1)} P_{jl\underline{ih}}^{SISI} - \beta^{(1)} \sum_{h \neq i,l} A_{jh}^{(1)} P_{il\underline{jh}}^{SISI} + \beta^{(1)} \sum_{h \neq i,j} A_{lh}^{(1)} P_{ij\underline{lh}}^{SSSI} \\
&- \frac{1}{2} \sum_{h,k \neq j,l} \left[A_{ihk}^{(1,0)} \beta^{(1)} (P_{jl\underline{ihk}}^{SISIS} + P_{jl\underline{ihk}}^{SISSI} + 2P_{jl\underline{ihk}}^{SISII}) + A_{ihk}^{(0,1)} \beta^{(2)} P_{jl\underline{ihk}}^{SISII} \right] \\
&- \frac{1}{2} \sum_{h,k \neq i,l} [i \leftrightarrow j] + \frac{1}{2} \sum_{h,k \neq i,j} [i \leftrightarrow l] && \text{(3.14c)} \\
\dot{P}_{ijl}^{SII} =& -(2 + 2A_{ijl}^{(1,0)} \beta^{(1)} + A_{ijl}^{(0,1)} \beta^{(2)}) P_{ijl}^{SII} + A_{ijl}^{(1,0)} \beta^{(1)} (P_{ijl}^{SSI} + P_{ijl}^{SIS}) + P_{ijl}^{III} \\
&- \beta^{(1)} \sum_{h \neq j,l} A_{ih}^{(1)} P_{jl\underline{ih}}^{IISI} + \beta^{(1)} \sum_{h \neq i,l} A_{jh}^{(1)} P_{il\underline{jh}}^{SISI} + \beta^{(1)} \sum_{h \neq i,j} A_{lh}^{(1)} P_{ij\underline{lh}}^{SISI} \\
&- \frac{1}{2} \sum_{h,k \neq j,l} \left[A_{ihk}^{(1,0)} \beta^{(1)} (P_{jl\underline{ihk}}^{IISIS} + P_{jl\underline{ihk}}^{IISSI} + 2P_{jl\underline{ihk}}^{IISII}) + A_{ihk}^{(0,1)} \beta^{(2)} P_{jl\underline{ihk}}^{IISII} \right] \\
&+ \frac{1}{2} \sum_{h,k \neq i,l} [i \leftrightarrow j] + \frac{1}{2} \sum_{h,k \neq i,j} [i \leftrightarrow l] && \text{(3.14d)}
\end{aligned}$$

where $[i \leftrightarrow j]$ ($[i \leftrightarrow l]$) denotes that obtained by swapping i and j (l) within the square brackets in the previous explicit term and taking both i and j (l) in state S. The other state probabilities are found as $P_i^S = 1 - P_i^I$, $P_{ij}^{SS} = 1 - P_i^I - P_{ij}^{SI}$, $P_{ij}^{II} = P_i^I - P_{ij}^{IS}$, $P_{ijl}^{SSS} = 1 - P_i^I - P_{ijl}^{SII} - P_{ijl}^{SSI} - P_{ijl}^{SIS}$, and $P_{ijl}^{III} = P_i^I - P_{ijl}^{ISI} - P_{ijl}^{IIS} - P_{ijl}^{ISS}$. As said, Eqs. (3.14) are closed through Eqs. (3.13). The system consists then of $N + L + 2(T^{(1,0)} + T^{(0,1)} + T^{(1,1)})$ equations, being N, L, $T^{(1,0)}$, $T^{(0,1)}$, and $T^{(1,1)}$, the number of nodes, maximal links, 3-cycles, 3-edges, and triangles, respectively.

To make this model analytically treatable, we perform a mean-field approximation by regarding all the nodes and cliques as equivalent to their average counterparts. Accordingly, every node is assumed to be part of the same number of maximal links,

$k^{(1)}$, 3-cliques, $k^{(1,0)}$, 3-edges, $k^{(0,1)}$, and triangles, $k^{(1,1)}$; and thus participates in $\kappa^{(1)} = k^{(1)} + 2(k^{(1,0)} + k^{(1,1)})$ 2-body interactions and $\kappa^{(2)} = k^{(0,1)} + k^{(1,1)}$ 3-body interactions. The state probabilities P_i^σ, $P_{ij}^{\sigma\sigma'}$ and $P_{ijl}^{\sigma\sigma'\sigma''}$, with $\sigma, \sigma', \sigma'' \in \{S, I\}$, are taken equal to their respective averages, $P^\sigma = \sum_i P_i^\sigma / N$, $P^{\sigma\sigma'} = \sum_{i,j} A_{ij}^{(1)} P_{ij}^{\sigma\sigma'} / Nk^{(1)}$, and $P_x^{\sigma\sigma'\sigma''} = \sum_{i,j,l} A_{ijl}^x P_{ijl}^{\sigma\sigma'\sigma''} / 2Nk^x$, where the index $x \in \{(1,0),(0,1),(1,1)\}$ indicates the type of the considered 3-clique (notice that, being the structure undirected, permuting the superscripts σ, σ', and σ'', has no effect).

Using the indicator function, $\mathbf{1}_p$, the reduced system reads

$$\begin{aligned}
\dot{P}^I =& -P^I + \beta^{(1)} k^{(1)} P^{SI} + 2\beta^{(1)} \left[k^{(1,0)} (P_{(1,0)}^{SSI} + P_{(1,0)}^{SII}) + k^{(1,1)} (P_{(1,1)}^{SSI} + P_{(1,1)}^{SII}) \right] \\
&+ \beta^{(2)} \left[k^{(0,1)} P_{(0,1)}^{SII} + k^{(1,1)} P_{(1,1)}^{SII} \right] \qquad (3.15a) \\
\dot{P}^{SI} =& -(1+\beta^{(1)}) P^{SI} + P^{II} - \beta^{(1)} (k^{(1)} - 1) P^{SI} \frac{P^{SI} - P^{SS}}{P^S} \\
&- 2\beta^{(1)} \left[k^{(1,0)} (P_{(1,0)}^{SSI} + P_{(1,0)}^{SII}) + k^{(1,1)} (P_{(1,1)}^{SSI} + P_{(1,1)}^{SII}) \right] \frac{P^{SI} - P^{SS}}{P^S} \\
&- \beta^{(2)} \left[k^{(0,1)} P_{(0,1)}^{SII} + k^{(1,1)} P_{(1,1)}^{SII} \right] \frac{P^{SI} - P^{SS}}{P^S} \qquad (3.15b) \\
\dot{P}_x^{SSI} =& -(1 + 2\beta^{(1)} \mathbf{1}_{x \neq (0,1)}) P_x^{SSI} + 2P_x^{SII} - \beta^{(1)} k^{(1)} P^{SI} \frac{2P_x^{SSI} - P_x^{SSS}}{P^S} \\
&- 2\beta^{(1)} (k^{(1,0)} - \mathbf{1}_{x=(1,0)}) (P_{(1,0)}^{SSI} + P_{(1,0)}^{SII}) \frac{2P_x^{SSI} - P_x^{SSS}}{P^S} \\
&- 2\beta^{(1)} (k^{(1,1)} - \mathbf{1}_{x=(1,1)}) (P_{(1,1)}^{SSI} + P_{(1,1)}^{SII}) \frac{2P_x^{SSI} - P_x^{SSS}}{P^S} \\
&- \beta^{(2)} \left[(k^{(0,1)} - \mathbf{1}_{x=(0,1)}) P_{(0,1)}^{SII} + (k^{(1,1)} - \mathbf{1}_{x=(1,1)}) P_{(1,1)}^{SII} \right] \frac{2P_x^{SSI} - P_x^{SSS}}{P^S} \qquad (3.15c) \\
\dot{P}_x^{SII} =& -(2 + 2\beta^{(1)} \mathbf{1}_{x \neq (0,1)} + \beta^{(2)} \mathbf{1}_{x \neq (1,0)}) P_x^{SII} + 2\beta^{(1)} \mathbf{1}_{x \neq (0,1)} P_x^{SSI} \\
&+ P_x^{III} - \beta^{(1)} k^{(1)} P^{SI} \frac{P_x^{SII} - 2P_x^{SSI}}{P^S} \\
&- 2\beta^{(1)} (k^{(1,0)} - \mathbf{1}_{x=(1,0)}) (P_{(1,0)}^{SSI} + P_{(1,0)}^{SII}) \frac{P_x^{SII} - 2P_x^{SSI}}{P^S} \\
&- 2\beta^{(1)} (k^{(1,1)} - \mathbf{1}_{x=(1,1)}) (P_{(1,1)}^{SSI} + P_{(1,1)}^{SII}) \frac{P_x^{SII} - 2P_x^{SSI}}{P^S} \\
&- \beta^{(2)} \left[(k^{(0,1)} - \mathbf{1}_{x=(0,1)}) P_{(0,1)}^{SII} + (k^{(1,1)} - \mathbf{1}_{x=(1,1)}) P_{(1,1)}^{SII} \right] \frac{P_x^{SII} - 2P_x^{SSI}}{P^S} \qquad (3.15d)
\end{aligned}$$

where $P^S = 1 - P^I$, $P^{SS} = 1 - P^I - P^{SI}$, $P^{II} = P^I - P^{SI}$, $P_x^{SSS} = 1 - P^I - P_x^{SII} - 2P_x^{SSI}$, $P_x^{III} = P^I - P_x^{SSI} - 2P_x^{SII}$.

To correctly locate the invasion threshold, we linearize Eqs. (3.15) around the inactive state by regarding of the same infinitesimal order ϵ the probabilities of infected states, i.e., $P^I, P^{SI}, P^{II}, P_x^{SSI}, P_x^{SII}, P_x^{III} \in \mathcal{O}(\epsilon)$. At this order, Eqs. (3.15) become

$$\begin{aligned}
\dot{P}^I =& -P^I + \beta^{(1)}k^{(1)}P^{SI} + 2\beta^{(1)}\left[k^{(1,0)}(P^{SSI}_{(1,0)} + P^{SII}_{(1,0)}) + k^{(1,1)}(P^{SSI}_{(1,1)} + P^{SII}_{(1,1)})\right] \\
&+ \beta^{(2)}\left[k^{(0,1)}P^{SII}_{(0,1)} + k^{(1,1)}P^{SII}_{(1,1)}\right] && \text{(3.16a)} \\
\dot{P}^{SI} =& -(1+\beta^{(1)})P^{SI} + P^{II} + \beta^{(1)}(k^{(1)}-1)P^{SI} \\
&+ 2\beta^{(1)}\left[k^{(1,0)}(P^{SSI}_{(1,0)} + P^{SII}_{(1,0)}) + k^{(1,1)}(P^{SSI}_{(1,1)} + P^{SII}_{(1,1)})\right] \\
&+ \beta^{(2)}\left[k^{(0,1)}P^{SII}_{(0,1)} + k^{(1,1)}P^{SII}_{(1,1)}\right] && \text{(3.16b)} \\
\dot{P}^{SSI}_x =& -(1+2\beta^{(1)}\mathbf{1}_{x\neq(0,1)})P^{SSI}_x + 2P^{SII}_x + \beta^{(1)}k^{(1)}P^{SI} \\
&+ 2\beta^{(1)}(k^{(1,0)} - \mathbf{1}_{x=(1,0)})(P^{SSI}_{(1,0)} + P^{SII}_{(1,0)}) \\
&+ 2\beta^{(1)}(k^{(1,1)} - \mathbf{1}_{x=(1,1)})(P^{SSI}_{(1,1)} + P^{SII}_{(1,1)}) \\
&+ \beta^{(2)}\left[(k^{(0,1)} - \mathbf{1}_{x=(0,1)})P^{SII}_{(0,1)} + (k^{(1,1)} - \mathbf{1}_{x=(1,1)})P^{SII}_{(1,1)}\right] && \text{(3.16c)} \\
\dot{P}^{SII}_x =& -(2+2\beta^{(1)}\mathbf{1}_{x\neq(0,1)} + \beta^{(2)}\mathbf{1}_{x\neq(1,0)})P^{SII}_x + 2\beta^{(1)}\mathbf{1}_{x\neq(0,1)}P^{SSI}_x + P^{III}_x && \text{(3.16d)}
\end{aligned}$$

Imposing stationarity in Eqs. (3.16), we get a homogeneous system of eight equations for the eight variables, $P^I, P^{SI}, P^{SSI}_x, P^{SII}_x$, $x \in \{(1,0),(0,1),(1,1)\}$. We find the relations

$$P^{SI} = \frac{P^I}{1+\beta^{(1)}}\,, \tag{3.17}$$

$$P^{SSI}_x = \frac{P^I}{1+\mathbf{1}_{x\neq(0,1)}\beta^{(1)}(2+2\beta^{(1)}+\mathbf{1}_{x\neq(1,0)}\beta^{(2)})}\,, \tag{3.18}$$

$$P^{SII}_x = \mathbf{1}_{x\neq(0,1)}\beta^{(1)}P^{SSI}_x\,, \tag{3.19}$$

which plugged back into the r.h.s. of Eq. (3.16a), provide us with the following critical condition,

$$k^{(1)}\frac{\beta^{(1)}}{1+\beta^{(1)}} + k^{(1,0)}\frac{2\beta^{(1)}(1+\beta^{(1)})}{1+2\beta^{(1)}(1+\beta^{(1)})} + k^{(1,1)}\frac{\beta^{(1)}(2+2\beta^{(1)}+\beta^{(2)})}{1+\beta^{(1)}(2+2\beta^{(1)}+\beta^{(2)})} = 1\,. \tag{3.20}$$

Equation (3.20) defines the critical surface in the parameter space at which the inactive state loses stability. We can express the invasion threshold as one of the two critical values, $\beta^{(1)}_{\rm cr}$ and $\beta^{(2)}_{\rm cr}$, which solve Eq. (3.20). We can read each of the terms in the l.h.s. as the probability that some infection process (proportional to combinations of rates $\beta^{(1)}$ and $\beta^{(2)}$) occurs before recovery (rate 1), within a specific type of group.

3.2.2 *Results*

We analyze at this point how either the critical surface, hence the invasion threshold, and the prevalence depend on the degree of overlap between 2- and 3-body interactions. To isolate the effect of the overlap, we fix the numbers $\kappa^{(1)}$ and $\kappa^{(2)}$ of 2- and 3-body interactions per node, respectively, and the number of isolated 2-body interactions per node, $k^{(1)}$, and we increase $k^{(1,1)}$ from 0 to $\kappa^{(2)} \leqslant (\kappa^{(1)} - k^{(1)})/2$ (the upper bound guarantees that $k^{(1,0)}$ is non-negative). This implies correspondingly decreasing $k^{(1,0)} = (\kappa^{(1)} - k^{(1)})/2 - k^{(1,1)}$ and $k^{(0,1)} = \kappa^{(2)} - k^{(1,1)}$. We get zero overlap—a linear hypergraph—taking $k^{(1,1)} = 0$ and maximal overlap—a simplicial complex—using $k^{(1,1)} = \kappa^{(2)}$. In other words, given the numbers of 2- and 3-body interactions in the system, we imagine to arrange them to produce different degrees of overlap, exploring in this way the spectrum of rank-3 hypergraphs going from a linear hypergraph to a simplicial complex (see Fig. 3.6). Our goal is then to understand how different arrangements, i.e., degrees of overlaps, affect the dynamics.

Let us begin by analizing the critical surface, Eq. (3.20), illustrated in Fig. 3.7(a). First of all, it is immediate to see that Eq. (3.20) does not admit solutions when either $\beta^{(1)} = 0$ or, equivalently, there are only 3-body interactions ($\kappa^{(2)} = k^{(0,1)}$, $\kappa^{(1)} = 0$).[8] Indeed, the latter enter Eqs. (3.16) only through $P^{SII}_{(0,1)}$, which, however, we find to be zero from Eq. (3.19). Consequently, 3-edges do not figure in Eq. (3.20), meaning that the invasion threshold does not depend on $\beta^{(2)}$ in linear hypergraphs. The only 3-body interactions appearing in Eq. (3.20) are those within triangles. This reveals that (i) 2-body interactions are needed to destabilize the inactive state and that (ii) a 3-body interaction cannot affect the stability of the inactive state unless "activated" by the presence of 2-body interactions within the same subset of nodes. That activation occurs in triangles.

Suppose $k^{(1,1)} > 0$ then. Since each term in the l.h.s. of Eq. (3.20) is a strictly increasing function of the infection rates, the larger is $\beta^{(2)}$ ($\beta^{(1)}$), the smaller is $\beta^{(1)} = \beta^{(1)}_{\rm cr}$ ($\beta^{(2)} = \beta^{(2)}_{\rm cr}$) that solves Eq. (3.20). Extending the findings presented in Sec. 3.1, this proves that the pairwise invasion threshold $\beta^{(1)}_{\rm cr}$ decreases with $\beta^{(2)}$ in any non-linear hypergraph, not just in simplicial complexes: this effect is therefore the rule, not the exception. Moreover, when $k^{(1,1)} > 1$, it is easy to see that $\beta^{(1)}_{\rm cr}$ can be made arbitrarily small by increasing $\beta^{(2)}$. Indeed, fixed $\beta^{(1)}$, increasing $\beta^{(2)}$ makes the third term in the l.h.s. of Eq. (3.20) converge to $k^{(1,1)}$. Since the other terms are non-negative, then, if $k^{(1,1)} > 1$, $\beta^{(1)}$ must decrease sufficiently fast to (over-)compensate the effect of $\beta^{(2)}$ and thus solve the equation. In the large $\beta^{(2)}$ limit, $\beta^{(1)}_{\rm cr}$ hence approaches zero. If instead $k^{(1,1)} < 1$, the third term is smaller than 1 for any $\beta^{(2)}$, therefore the sum of the first two terms must be positive, implying that $\beta^{(1)}$ cannot be too small to be solution. Accordingly, we observe in Fig. 3.7(a) that, on the critical surface, $\beta^{(1)}$ must be sufficiently larger than 0 when $k^{(1,1)} < 1$, whereas $\beta^{(1)}$ approaches zero if correspondingly $\beta^{(2)} \to \infty$ when $k^{(1,1)} > 1$ (which can be also seen by solving

8 Evidence for this has been recently found only numerically [277, 291].

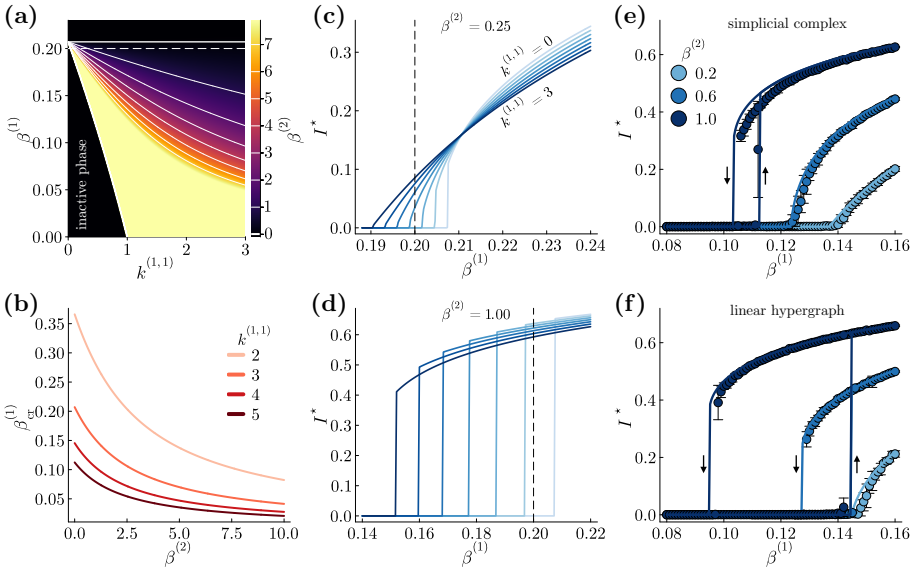


Figure 3.7: Phenomenology predicted by the mean-field triadic approximation. (a) Critical surface $(k^{(1,1)}, \beta^{(1)}, \beta^{(2)})_{\text{cr}}$ defined by Eq. (3.20) for hypergraphs with $\kappa^{(1)} = 6$ 2-body ($k^{(1)} = 0$) and $\kappa^{(2)} = 3$ 3-body interactions per node. To keep $k^{(1)}$, $\kappa^{(1)}$ and $\kappa^{(2)}$ fixed, $k^{(1,0)}$ and $k^{(0,1)}$ are respectively decreased when increasing $k^{(1,1)}$ (see main text). The white solid curves correspond to the values of $\beta^{(2)}$ indicated in the colorbar (limited at almost 8 for better readability), the line $\beta^{(2)} = 0$ being thicker. The latter represents the threshold $\beta^{(1)}_{(1,0)} \approx 0.207$. The dashed line denotes $\beta^{(1)}_{(1)} = 0.2$, associated to the locally tree-like network with same $\kappa^{(1)}$. (b) Critical threshold $\beta^{(1)}_{\text{cr}} \equiv \beta^{(1)}_{(1,1)}$, Eq. (3.21), for homogeneous simplicial 2-complexes. (c)–(d) Equilibrium prevalence, $I^\star$, for $\beta^{(2)} \in \{0.25, 1.00\}$, and $k^{(1,1)}$ from 0 (linear hypergraph; lightest shade) to 3 (simplicial complex; darkest shade) in steps of 0.5. The dashed line indicates $\beta^{(1)}_{(1)} = 0.2$. (e)–(f) Comparison of the model (solid lines) with numerical simulations performed on random regular hypergraphs with $N = 5000$, $\kappa^{(1)} = 8$, $k^{(1)} = 2$, $\kappa^{(2)} = 3$, and $k^{(1,1)} = 3$ (simplicial complex) and $k^{(1,1)} = 0$ (linear hypergraph). Points and error bars denote averages and standard errors computed over 20 random initializations. The arrows help to distinguish the forward and backward curves in hysteresis cycles.

Eq. (3.20) with respect to $\beta^{(2)}$ and taking the limit $\beta^{(1)} \to 0$). Therefore, depending on $k^{(1,1)}$ being higher or lower than 1, which determines whether the triangles can or not percolate the structure, 3-body interactions can or not bring the invasion threshold $\beta^{(1)}_{\text{cr}}$ arbitrarily close to zero. In this sense, $k^{(1,1)} = 1$ is a critical condition for the magnitude of the effects that 3-body interactions can have.

Taking for a moment $\beta^{(2)} = 0$ in Eq. (3.20), we find that the invasion threshold for a simple contagion lies in the interval $[\beta^{(1)}_{(1)}, \beta^{(1)}_{(1,0)}]$, being $\beta^{(1)}_{(1)} = 1/(\kappa^{(1)} - 1)$ the invasion threshold for a locally tree-like network ($\kappa^{(2)} = 0$, $\kappa^{(1)} = k^{(1)}$) and $\beta^{(1)}_{(1,0)} = \left[\sqrt{1 + 4/(\kappa^{(1)} - 2)} - 1\right]/2$ the threshold for a (clustered) 3-cycle-based network ($\kappa^{(2)} = 0$, $\kappa^{(1)} = 2k^{(1,0)}$). It is easy to demonstrate that, in agreement with previous studies proving clustering to raise the critical point of simple [78, 93, 249] and slightly nonlinear [258] contagions, $\beta^{(1)}_{(1,0)} > \beta^{(1)}_{(1)}$ for any $\kappa^{(1)}$.

Considering instead a triangle-based network ($\kappa^{(2)} = k^{(1,1)}$, $\kappa^{(1)} = 2k^{(1,1)}$), i.e., a homogeneous simplicial 2-complex, we find that the invasion threshold, $\beta^{(1)}_{(1,1)}$, reads

$$\beta^{(1)}_{(1,1)} = \frac{\beta^{(2)} + 2}{4}\left[\sqrt{1 + \frac{16}{(\kappa^{(1)} - 2)(\beta^{(2)} + 2)^2}} - 1\right] \tag{3.21}$$

Figure 3.7(b) shows $\beta^{(1)}_{(1,1)}$ against $\beta^{(2)}$ for different values of $\kappa^{(1)}$. As anticipated, $\beta^{(1)}_{(1,1)}$ decreases, vanishing as $1/\beta^{(2)}$ for large $\beta^{(2)}$. In particular, $\beta^{(1)}_{(1,1)}$ becomes smaller than $\beta^{(1)}_{(1)}$, the threshold for the locally tree-like network with same $\kappa^{(1)}$, if and only if $\beta^{(2)} > 2/(\kappa^{(1)} - 2)$. Please, notice that Eq. (3.21) reflects the fact that extensive contagions are possible—a phase transition exists—only for $\kappa^{(1)} > 2$ (i.e., $k^{(1,1)} > 1$), which is a necessary condition for a giant connected component to exist. [9]

Summing up, we proved that the invasion threshold is lowered by $\beta^{(2)}$ as long as there is some overlap between 2- and 3-body interactions. Since 3-edges yield a negligible contribution around the inactive state, exchanging them for triangles helps the spread to thrive. Consequently, the higher is the degree of overlap, the larger is the decrease induced by $\beta^{(2)}$. As shown in detail in Fig. 3.7(a), either $\beta^{(1)}_{\text{cr}}$ and $\beta^{(2)}_{\text{cr}}$ decrease with $k^{(1,1)}$, taking the lowest values in a simplicial complex and the highest in a linear hypergraph (for which, being $k^{(1,1)} = 0$, $\beta^{(1)}_{\text{cr}} = \beta^{(1)}_{(1,0)}$ and is thus unaffected by $\beta^{(2)}$).

We now move to analyze the effect that the overlap has on the equilibrium prevalence. A larger overlap (i.e., $k^{(1,1)}$) implies a smaller and more redundant neighborhood (see Fig. 3.6). Except for small enough pairwise infection rates, more specifically, near or below the threshold $\beta^{(1)}_{(1,0)}$ for clustered graphs (or for linear hypergraphs), Figs. 3.7(c) and (d) show that such redundancy is detrimental for the spread. Increasing the overlap indeed implies that some potentially infectious channels are "wasted" in relation to maximizing the spread, as they lead to nodes which are already infected. This is analogous to the effect of clustering in graphs, with the difference that increasing clustering is detrimental for spreading even at early stages, for it increases

9 A recent work [292] found explicit formulas for either $\beta^{(1)}_{(1,1)}$ and $\beta^{(1)}_{(1,0)}$. These are however less accurate than ours and break down under extreme sparseness, for they provide finite invasion thresholds even for $k^{(1,1)} \leqslant 1$ and $k^{(1,0)} \leqslant 1$, respectively. Their formulas converge to ours in the dense regimes, $k^{(1,1)} \gg 1$ and $k^{(1,0)} \gg 1$.

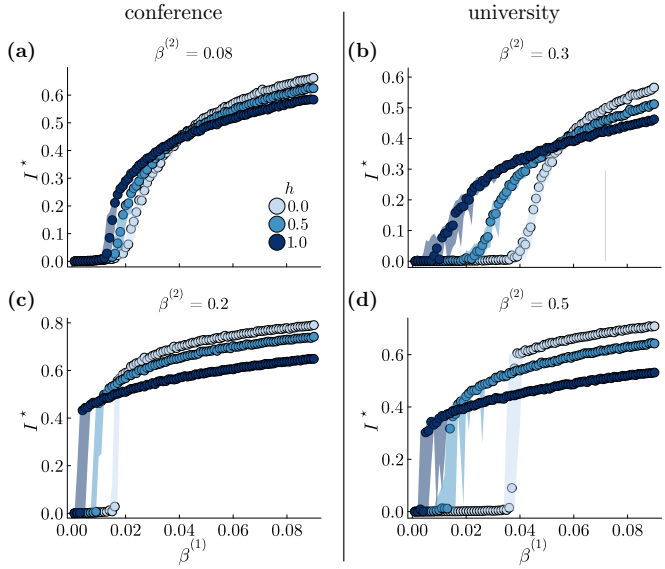


Figure 3.8: Monte Carlo simulations performed on the hypergraphs constructed from the conference's [293] ((a) and (c)) and the university campus' [294] ((b) and (d)) datasets. Points denote medians computed over 20 random initializations; ribbons cover from the 5th to the 95th percentile. The added 3-body interactions form triangles with probability $h = 0$ (linear hypergraph), $h = 0.5$, and $h = 1$ (simplicial complex).

the invasion threshold. The largest spreads are thus generally found for linear hypergraphs, these ensuring the least-redundant, widest neighborhoods. Finally, as can be observed in Fig. 3.7(c), solely varying the overlap can change the nature of the phase transition. More precisely, the value of $\beta^{(2)}$ at which the transition turns discontinuous becomes lower by increasing the overlap.

We test the model on random regular hypergraphs. These are generated through a configuration model in which we first assign the same numbers $k^{(1)}$, $k^{(0,1)}$, $k^{(1,0)}$, and $k^{(1,1)}$, of generalized stubs (degrees) to every node, and then match the stubs at random. In Figs. 3.7(e) and (f), in particular, we show the results for a simplicial complex and a linear hypergraph. Their comparison confirms all of our qualitative predictions, with even a good quantitative agreement.

We further test those predictions on hypergraphs constructed from real-world datasets, containing record of face-to-face interactions during a conference [293], and proximity data within a university campus [294]. We refer the reader to Appendix C.2.1 for the procedure we used to convert each dataset into a binary network. The hypergraphs are then constructed by adding 3-body interactions to either 3-cycles with probability h, converting them in triangles, or to randomly selected triplets of unconnected nodes otherwise, forming 3-edges. As shown in Appendix C.2.1, these hypergraphs are het-

erogeneous and assortative in degree, and are also dense, leading many 3-cliques to share two nodes. Even though the basic assumptions of homogeneity and sparseness that we made are heavily violated in more realistic structures, the numerical results reported in Fig. 3.8 show that the qualitative phenomenology uncovered by the mean-field theory remains valid. This confirms its *structural* origin: the overlap between 3- and 2-body interactions. Accordingly, we conjecture the same qualitative picture to hold for other contagion models, as already verified for a simplicial SIR process (see Appendix C.2.2).

3.2.3 *Summary*

Through a more refined mean-field model, we revealed a fundamental relation between the behavior of complex contagion processes and the way interactions are arranged in the higher-order structure. Extending beyond traditional node- and pair-based approximations, our analysis establishes how 3-body interactions contribute to destabilizing the inactive state, proving their contribution is contingent on overlapping with 2-body interactions. We demonstrated that simplicial complexes and linear hypergraphs—having maximal and no overlap, respectively—exert diametrically opposed dynamical effects. The former lower the critical point, while often resulting in smaller spreads; the latter heighten the critical point, yet typically leading to larger spreads. Complementing recent findings in synchronization [295], our investigation underscores the necessity of identifying the most suitable representation for specific higher-order processes.

A clear limitation of our approach is its scalability, for writing down equations for a general clique-based mean-field theory is impracticable.[10] Nonetheless, it is possible to interpret the form of every term in the expression for the critical surface, Eq. (3.20). Each one corresponds to a certain type of clique and coincides with the probability that an infected unit produces, before recovering, secondary infections through any of the possible infection channels within the clique. We therefore conjecture that, without needing to build the dynamic equations, we can find the correct expression for the critical surface for hypergraphs of arbitrary rank by listing all the infections channels an infected node can generate in an otherwise susceptible clique.

10 Group-centered AME models, as the one we will develop in Sec. 3.3, circumvent such limitation by introducing a subtle approximation. They assume that a node, instead of randomly sampling a given number of interactions of each size and type (e.g., $k^{(1)}$, $k^{(1,0)}$, $k^{(0,1)}$ and $k^{(1,1)}$), redistributes its membership over all groups according to the group size (and type) distribution. Consequently, (heterogeneous) clique-based mean-field theories are more accurate than group-centered AME. The difference becomes evident for very sparse structure, being negligible otherwise.

3.3 ADAPTIVE HYPERGRAPHS AND THE CHARACTERISTIC SCALE OF HIGHER-ORDER CONTAGIONS

The results in Secs. 3.1 and 3.2 prove the importance of accounting for dynamical correlations to provide reliable predictions for contagion processes, even at the qualitative level. This holds for quenched structures but also under the partial annealing assumed by the mean-field model in Sec. 3.2. Besides, degree heterogeneity has important dynamical implications by itself, even when the network is fully annealed. Indeed, any variance in degree implies that sampling a random connection is different from sampling a random node (see Sec. 2.3.1.1), which is the statistical bias behind the friendship paradox where "your friends have more friends than you do" [30, 296]. The same is true in higher-order networks when sampling groups (hyperedges) at random [93]—"your friends participate in more groups than you do". Therefore, because of their degree, not all nodes in the network follow the same dynamics, and neither do so all nodes with the same degree for different neighborhoods are possible.

An accurate description of dynamical processes on complex, group-structured networks thus needs to account for local heterogeneities (variance in degree and group size) and dynamical correlations. To capture these effects, we resort to approximate master equations (AME) [93, 242, 254, 256, 292, 297–300], discussed in Sec. 2.3.3. We recall that, compared to standard heterogeneous mean-field models [135, 275], the AME frameworks stand out by tracking the entire distribution of potential dynamical states within neighborhoods/groups, not just average states. This allows AME models to capture local heterogeneities, even for nodes of the same degree or groups of the same size. Recent results have shown that these state distributions can be very heterogeneous and even bimodal, explaining blind spots of mean-field approaches that fail to capture important dynamical regimes like mesoscopic localization, establishing that around the critical region the contagion localizes in the larger groups [301–303].

Node-centered AME (Sec. 2.3.3.1) retain dynamical correlations within the neighborhood of a node, hence partially between adjacent links, but cannot account for group structure. The latter is considered by group-centered AME (Sec. 2.3.3.2), which preserve correlations within groups, but not between groups. Capturing both the higher-order structure and the correlations around a node is out of the scope of current models. Apart from describing the dynamics accurately, this is critical for systems in which adaptive behavior is allowed, for an agent modifies its connections in response to the state of its neighborhood, hence to the composition of the groups therein.

To fill this gap, here we extend the AME framework by suitably integrating previous approaches. Crucially, a naïve merging of the node- and group-centered AME models would result in an impractical description. On the one hand, this would imply classifying groups based on either their size and a state vector of $\Sigma - 1$ elements, being Σ the number of available states for a node. On the other, nodes would be distinguished by both their state and a degree vector where each entry corresponds to a

particular state configuration of the set of the other members, leading to a rapid combinatorial explosion.[11] We are thus forced to coarse-grain in some way the state space of the neighborhood of a node. We look for an approximation capable of drastically reducing the dimension of the description while still capturing the relevant dynamical correlations in the system.

We will focus on binary state dynamics (e.g., SIS) and show that the description can be effectively reduced. To do so, we will keep track of the groups' level of activity (infectiousness). This will unlock a series of theoretical questions: When—as modelers—should we consider a group *active* to predict the dynamics best? How correlated are two groups of a node if one is large and consistently active and the other small and sporadically so? If adaptive behavior is allowed, under which conditions does it act as a global control against contagions in group-structured systems? Lastly, is the best definition of an active group the same for modelers trying to describe the contagion accurately and for agents hoping to avoid it?

3.3.1 *The GAME model*

Let us consider a general binary-state dynamics on infinite-size random networks with groups (i.e., random hypergraphs where groups are encoded as hyperedges). Nodes can either be susceptible (S) or infected (I) and have a membership (or hyperdegree) m drawn from the distribution g_m. Groups are of various size n drawn from the distribution p_n. We partition the groups and the nodes according to their local properties. Specifically, we track $C_{n,i}(t) \in [0, p_n]$, the fraction of groups of size n with $i \in \{0, \dots, n\}$ infected nodes at time t. We also track $S_{m,l}$ and $I_{m,l} \in [0, g_m]$, the fraction of susceptible and infected nodes with membership m and $l \in \{0, \dots, m\}$ incident *active* groups. For a given node, a group to which it belongs is active when it contains at least $\bar{i}$ infected nodes other than itself. Accordingly, we call l ($m - l$) the active (inactive) membership of a node. These three types of compartments are also to be interpreted as joint probabilities, i.e., $C_{n,i} \equiv P(n, i)$, $S_{m,l} \equiv P(S, m, l)$, and $I_{m,l} \equiv P(I, m, l)$, with normalization $\sum_{n,i} C_{n,i} = \sum_n p_n = 1$ and $\sum_{m,l}(S_{m,l} + I_{m,l}) = \sum_m g_m = 1$. Unless specified, sums run over all possible values.

With the aim of building a mean-field theory, we introduce the *effective* infection (recovery) rate $\tilde{\beta}_{n,i}$ ($\tilde{\alpha}_{n,i}$) for a node within a group of size n with i infected members, and the effective infection (recovery) rate $\tilde{\beta}_{m,l}$ ($\tilde{\alpha}_{m,l}$) for a node of membership m, l of which are active. From these definitions, we introduce the following set of *generalized approximate master equations* (GAME), schematized in Fig. 3.9,

11 Take the simplest case in which $\Sigma = 2$ and the structure is a m-regular n-uniform hypergraph, so that all nodes have hyperdegree m and all groups have size n. The nodes' degree vector would have n entries (e.g., number of active nodes among the other $n - 1$ in the group), implying that there are as many node classes (hence dynamic equations) as partitions of m into n distinguishable parts (in the sense that, e.g., (n_1, n_2) and (n_2, n_1) count as different partitions). With $n = 3$, we would already get 9, 15, and 18 node classes for respectively $m = 3, 4, 5$. For the same sequence of m, $n = 4$ would yield 20, 41, and 56 classes.

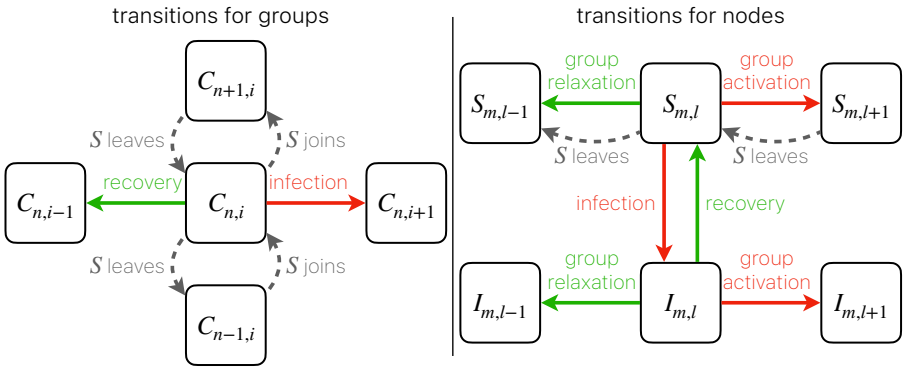


Figure 3.9: States and transitions for the generalized approximate master equations. Groups are identified by their size n and number of infectious members i. Nodes are identified by their group membership m and the number ℓ of those that they consider active (i.e., having at least $\bar{\imath}$ infected nodes among the other members). GAME transitions, Eqs. (3.22), correspond to binary-state ($\{S, I\}$) contagion dynamics and are indicated by red and green arrows. Adaptive (A-GAME) transitions, Eqs. (3.45), are indicated by dashed arrows.

$$\begin{aligned}
\dot{C}_{n,i} &= \tilde{\alpha}_{n,i+1}(i+1)C_{n,i+1} - \tilde{\alpha}_{n,i}iC_{n,i} \\
&\quad + \tilde{\beta}_{n,i-1}(n-i+1)C_{n,i-1} - \tilde{\beta}_{n,i}(n-i)C_{n,i} && \text{(3.22a)} \\
\dot{S}_{m,l} &= \tilde{\alpha}_{m,l}I_{m,l} - \tilde{\beta}_{m,l}S_{m,l} \\
&\quad + \theta_S\left[(m-l+1)S_{m,l-1} - (m-l)S_{m,l}\right] && \text{(3.22b)} \\
&\quad + \phi_S\left[(l+1)S_{m,l+1} - lS_{m,l}\right] , \\
\dot{I}_{m,l} &= -\tilde{\alpha}_{m,l}I_{m,l} + \tilde{\beta}_{m,l}S_{m,l} \\
&\quad + \theta_I\left[(m-l+1)I_{m,l-1} - (m-l)I_{m,l}\right] && \text{(3.22c)} \\
&\quad + \phi_I\left[(l+1)I_{m,l+1} - lI_{m,l}\right] .
\end{aligned}$$

The four mean fields are calculated as

$$\theta_S = \frac{\sum_n (n-\bar{\imath}+1)(n-\bar{\imath})C_{n,\bar{\imath}-1}\tilde{\beta}_{n,\bar{\imath}-1}}{\sum_{n,i\leqslant\bar{\imath}-1}(n-i)C_{n,i}} , \tag{3.23a}$$

$$\phi_S = \frac{\sum_n (n-\bar{\imath})\bar{\imath}C_{n,\bar{\imath}}\tilde{\alpha}_{n,\bar{\imath}}}{\sum_{n,i\geqslant\bar{\imath}}(n-i)C_{n,i}} , \tag{3.23b}$$

$$\theta_I = \frac{\sum_n \bar{\imath}(n-\bar{\imath})C_{n,\bar{\imath}}\tilde{\beta}_{n,\bar{\imath}}}{\sum_{n,i\leqslant\bar{\imath}} iC_{n,i}} , \tag{3.23c}$$

$$\phi_I = \frac{\sum_n (\bar{\imath}+1)\bar{\imath}C_{n,\bar{\imath}+1}\tilde{\alpha}_{n,\bar{\imath}+1}}{\sum_{n,i\geqslant\bar{\imath}+1} iC_{n,i}} . \tag{3.23d}$$

These quantities are the average rates at which inactive groups become active (θ_X) or vice versa (ϕ_X) given that we know the state ($X \in \{S, I\}$) of one of their members. We therefore sum over all groups eligible for the transition (e.g., with one too few or too many infectious nodes) and count the number of nodes therein whose state matches that of the node of interest. This gives us a biased distribution over states, renormalized with the sum in the denominator, and over which we average the local rate of transition. For instance, consider θ_S in Eq. (3.23a), which is the average rate at which a random inactive group becomes active from the perspective of one of its susceptible members. Since the probability that a group includes a susceptible node is proportional to the number of such nodes in the group (which is $n - i$ if the group has size n), the probability that an inactive group containing the susceptible node has size n and $i = \bar{i} - 1$ (i.e., is at the edge of becoming active) is $(n - \bar{i} + 1)C_{n,\bar{i}-1} / \sum_{n,i\leqslant\bar{i}-1}(n - i)C_{n,i}$. Such a group has $n - \bar{i}$ susceptible nodes and each of them gets infected with rate $\bar{\beta}_{n,\bar{i}-1}$, thus the local infection rate reads $(n - \bar{i})\bar{\beta}_{n,\bar{i}-1}$. Summing over group sizes, we get Eq. (3.23a). Equations (3.23)(b)-(d) are found analogously.

To close the GAME, we need to estimate the previously introduced effective transition rates, $\bar{\alpha}_{n,i}$, $\bar{\beta}_{n,i}$, $\tilde{\alpha}_{m,l}$ and $\tilde{\beta}_{m,l}$. We calculate these rates with mean-field arguments, yet the form of this calculation depends on the nature of the dynamics. Specifically, we use two approaches based on whether the dynamics operate at the node or group level. In a *node-based* dynamics, the transition rates for a node are functions of the states of all of its neighbors, independently from how those neighbors are scattered in different groups. For instance, this means that the rate at which a node gets infected only depends on the number of infectious neighbors. In a *group-based* dynamics, instead, each group brings an independent contribution so that an overall transition rate is given by the transition rates summed over all the groups a node belongs to. This translates into a node getting infected independently through each group based on the number of infectious members therein. In other words, node-based and group-based dynamics model mechanisms which are potentially nonlinear on the state of entire neighborhoods and single groups, respectively. The two approaches become equivalent for dynamics like simple contagions, which are linear in the states of nodes (hence, groups).

3.3.1.1 Node-based dynamics

We consider general continuous-time Markov processes where susceptible nodes transition to the infected state at rate $\beta(k, \ell)$, being k the (pairwise) degree of the node and $\ell \in \{0, \dots, k\}$ its infected degree. Similarly, infected nodes transition to the susceptible state at rate $\alpha(k, \ell)$. Degree k and infected degree ℓ are total quantities of a node summed over all groups it belongs to. The dynamics therefore ignores how these quantities are distributed over groups. Within this general node-based process, we can calculate the mean-field transition rates as follows.

Let us first consider the effective infection rate $\bar{\beta}_{n,i}$. If we pick a susceptible node in a (focal) group of size n, of which i are infected, the degree of this node can be decomposed as $k = n - 1 + r$, where r is the excess degree, due to memberships to other groups. Similarly, the infected degree can be decomposed as $\ell = i + s$, where s is the excess infected degree. If r and s were specified, then the infection rate of this susceptible node would simply be $\beta(n - 1 + r, i + s)$. In this version of the GAME, we approximate $\bar{\beta}_{n,i}$ by averaging over a joint distribution

$$\bar{\beta}_{n,i} = \sum_{r,s} \beta(n - 1 + r, i + s) P(r, s|n, i, S) \ , \tag{3.24}$$

where the distribution $P(r, s|n, i, S)$ is to be determined. It is the probability that a susceptible node in a group of size n with i infected nodes has excess degree r and infected excess degree s. For this task, we leverage the properties of probability generating functions (PGFs), discussed in Appendix A.4. Specifically, we need an expression for the following (bivariate) PGF,

$$E_S^i(x, y) = \sum_{r,s} P(r, s|n, i, S) x^r y^s \ . \tag{3.25}$$

Known $E_S^i(x, y)$, we can extract the distribution $P(r, s|n, i, S)$ from it and so compute the effective rate $\bar{\beta}_{n,i}$. To highlight that the latter is obtained averaging over the distribution associated with $E_S^i(x, y)$, we write

$$\bar{\beta}_{n,i} \equiv \langle \beta(n - 1 + r, i + s) \rangle_{E_S^i} \ . \tag{3.26}$$

To find $P(r, s|n, i, S)$, let us first derive an alternative expression for $E_S^i(x, y)$ in terms of the state variables, $\{C_{n,i}\}$, $\{S_{m,l}\}$ and $\{I_{m,l}\}$.

If we take a random *external* group to which the susceptible node belongs, the contribution to its degree and infected degree is associated with two PGFs, depending on whether or not the group is active for the node. If it is not (i.e., the group contains less than $\bar{i}$ infected nodes), then the appropriate PGF to use is

$$K_S^{i<\bar{i}}(x, y) = \frac{\sum_{n,i<\bar{i}} (n - i) C_{n,i} x^{n-1} y^i}{\sum_{n,i<\bar{i}} (n - i) C_{n,i}} \ ; \tag{3.27}$$

otherwise, it is

$$K_S^{i\geqslant\bar{i}}(x, y) = \frac{\sum_{n,i\geqslant\bar{i}} (n - i) C_{n,i} x^{n-1} y^i}{\sum_{n,i\geqslant\bar{i}} (n - i) C_{n,i}} \ . \tag{3.28}$$

Through the auxiliary variables x and y, these PGFs count the number of other members $(n - 1)$ and infected members (i) in a group, respectively, weighting these numbers with respect to the probability (proportional to $(n - i)C_{n,i}$) that a susceptible node is part of such a group. Notice how the PGFs above encode all the moments of the probability distribution of the size and state of the groups around a (susceptible) node,

conditioned on the state of the latter. For instance, the average number of infected nodes a susceptible node finds in an inactive group is given by $\partial K_S^{i<\bar{i}}(x,y)/\partial y|_{x,y=1}$.

Equations (3.27) and (3.28) provide information about a single, random external group. But how many groups does the node belong to? Its excess membership and excess active membership are still unspecified. To compute them, we leverage again the information about whether the focal group is inactive ($i < \bar{i}$) or active ($i \geqslant \bar{i}$). Depending on this, a PGF of different form is used, i.e.,

$$G_S^i(x,y) = \begin{cases} \dfrac{\sum_{m,l}(m-l)S_{m,l}x^{m-l-1}y^l}{\sum_{m,l}(m-l)S_{m,l}} & \text{if } i < \bar{i}\,, \\ \dfrac{\sum_{m,l} lS_{m,l}x^{m-l}y^{l-1}}{\sum_{m,l} lS_{m,l}} & \text{if } i \geqslant \bar{i}\,, \end{cases} \tag{3.29}$$

where the variables x and y here count the number of inactive and active external groups, respectively. If the focal group is inactive, then the probability for the susceptible node to be part of it is proportional to $(m-l)S_{m,l}$; if active, such probability is instead proportional to $lS_{m,l}$. The moments of the probability distribution of the excess (active) membership of a susceptible node, conditioned on this belonging to an inactive or active group, are all encoded in $G_S^i(x,y)$. The conditioned, average excess active and inactive membership of the node, for example, are simply found as $\partial G_S^i(x,y)/\partial y|_{x,y=1}$ and $\partial G_S^i(x,y)/\partial x|_{x,y=1}$, respectively. Their sum then provides the conditioned, average excess membership.

Finally, assuming that the contributions to r and s from different groups are statistically independent[12] (besides being identically distributed, all drawn from $K_S^{i<\bar{i}}$ and $K_S^{i\geqslant\bar{i}}$), using the properties of PGFs, it follows that $E_S^i(x,y)$ can be expressed through the composition

$$E_S^i(x,y) = G_S^i\left(K_S^{i<\bar{i}}(x,y), K_S^{i\geqslant\bar{i}}(x,y)\right)\,. \tag{3.30}$$

The distributions associated to $K_S^{i<\bar{i}}$, $K_S^{i\geqslant\bar{i}}$ and G_S^i are fully determined by the knowledge of the set of states $\{C_{n,i}\}$ and $\{S_{m,l}\}$. Initialized the system, we know those PGFs for each x and y, hence E_S^i using Eq. (3.30). We can then invert Eq. (3.25) to get $P(r,s|n,i,S)$ using a discrete Fourier transform (refer to Appendix A.4.1 to see how) and find $\bar{\beta}_{n,i}$. Known the latter, the system can be integrated through the infinitesimal interval dt to get the new set of states, which can be thus fed back into Eq. (3.30) to repeat the procedure. In this way, the system can be integrated over any time span.

12 Notice here the implicit assumption that the different groups have in common only the focal node. Any two groups sharing at least another node would be directly correlated, breaking the assumption. As in Sec. 3.2, the structure is thus considered to be sufficiently sparse so that any two maximal hyperedges have a vanishing probability of sharing more than one node.

Analogous steps yield the effective recovery rate $\bar{\alpha}_{n,i}$ for an infected node, leading to

$$K_I^{i\leqslant\bar{i}}(x,y) = \frac{\sum_{n,i\leqslant\bar{i}} iC_{n,i}x^{n-1}y^{i-1}}{\sum_{n,i\leqslant\bar{i}} iC_{n,i}} , \tag{3.31}$$

$$K_I^{i>\bar{i}}(x,y) = \frac{\sum_{n,i>\bar{i}} iC_{n,i}x^{n-1}y^{i-1}}{\sum_{n,i>\bar{i}} iC_{n,i}} , \tag{3.32}$$

$$G_I^i(x,y) = \begin{cases} \dfrac{\sum_{m,l}(m-l)I_{m,l}x^{m-l-1}y^l}{\sum_{m,l}(m-l)I_{m,l}} & \text{if } i \leqslant \bar{i} , \\ \dfrac{\sum_{m,l} lI_{m,l}x^{m-l}y^{l-1}}{\sum_{m,l} lI_{m,l}} & \text{if } i > \bar{i} , \end{cases} \tag{3.33}$$

$$E_I^i(x,y) = G_I^i\left(K_I^{i\leqslant\bar{i}}(x,y), K_I^{i>\bar{i}}(x,y)\right) , \tag{3.34}$$

and eventually to

$$\bar{\alpha}_{n,i} \equiv \langle \alpha(n-1+r, i-1+s)\rangle_{E_I^i} . \tag{3.35}$$

Notice that, from the perspective of an infected node, a group is active if there are more than $\bar{i}$ infected nodes among the other members.

The construction of the PGFs to compute $\tilde{\alpha}_{m,l}$ and $\tilde{\beta}_{m,l}$ is straightforward, for the memberships are already given. For an infected node with membership m and l of them active, we need the distribution $P(k,\ell|m,l,I)$ for its degree k and infected degree ℓ. The associated PGF reads

$$E_I^{m,l}(x,y) = \sum_{k,\ell} P(k,\ell|m,l,I)x^k y^\ell = \left[K_I^{i\leqslant\bar{i}}(x,y)\right]^{m-l}\left[K_I^{i>\bar{i}}(x,y)\right]^l , \tag{3.36}$$

where, for the second equality, we assumed that the contributions to k and ℓ from different groups are independent. The effective recovery rate is then

$$\tilde{\alpha}_{m,l} \equiv \langle \alpha(k,\ell)\rangle_{E_I^{m,l}} . \tag{3.37}$$

Analogously, for a susceptible node, we have

$$E_S^{m,l}(x,y) = \sum_{k,\ell} P(k,\ell|m,l,S)x^k y^\ell = \left[K_S^{i<\bar{i}}(x,y)\right]^{m-l}\left[K_S^{i\geqslant\bar{i}}(x,y)\right]^l , \tag{3.38}$$

and the effective infection rate reads

$$\tilde{\beta}_{m,l} \equiv \langle \beta(k,\ell)\rangle_{E_S^{m,l}} . \tag{3.39}$$

3.3.1.2 Group-based dynamics

We now consider the case where groups are the main actors responsible for transitions. Specifically, a susceptible node in a group of size n with i infectious members receives

an infection rate of $\lambda(n,i)$ from this group. The overall, effective infection rate $\bar{\beta}_{n,i}$ thus becomes $\lambda(n,i)+\bar{\lambda}_i$, where $\bar{\lambda}_i$ is the average infection rate from all the external groups, reading

$$\bar{\lambda}_i = \begin{cases} \dfrac{\sum_{m,l}(m-l)S_{m,l}\left[(m-l-1)\bar{\lambda}_{i<\bar{i}}+l\bar{\lambda}_{i\geqslant\bar{i}}\right]}{\sum_{m,l}(m-l)S_{m,l}} & \text{if } i<\bar{i} \\ \dfrac{\sum_{m,l} lS_{m,l}\left[(m-l)\bar{\lambda}_{i<\bar{i}}+(l-1)\bar{\lambda}_{i\geqslant\bar{i}}\right]}{\sum_{m,l} lS_{m,l}} & \text{if } i\geqslant\bar{i} \end{cases}$$
$$= \left[\bar{\lambda}_{i<\bar{i}}\frac{\partial}{\partial x}+\bar{\lambda}_{i\geqslant\bar{i}}\frac{\partial}{\partial y}\right]_{x,y=1} G_S^i(x,y)\ . \tag{3.40}$$

The quantities $\bar{\lambda}_{i<\bar{i}}$ and $\bar{\lambda}_{i\geqslant\bar{i}}$ are the average infection rates from a random external inactive or active group (for a susceptible node), respectively, and read

$$\bar{\lambda}_{i<\bar{i}} = \frac{\sum_{n,i<\bar{i}}(n-i)C_{n,i}\lambda(n,i)}{\sum_{n,i<\bar{i}}(n-i)C_{n,i}} = \langle\lambda(n,i)\rangle_{K_S^{i<\bar{i}}}\ , \tag{3.41}$$

$$\bar{\lambda}_{i\geqslant\bar{i}} = \frac{\sum_{n,i\geqslant\bar{i}}(n-i)C_{n,i}\lambda(n,i)}{\sum_{n,i\geqslant\bar{i}}(n-i)C_{n,i}} = \langle\lambda(n,i)\rangle_{K_S^{i\geqslant\bar{i}}}\ . \tag{3.42}$$

From Eq. (3.40) we see that $\bar{\lambda}_i$ results from the sum of two terms, each one given by the product between the average infection rate in an external group of a given activity state and the average number of such groups a susceptible node belongs to. Notice that, although the rates $\bar{\lambda}_i$, $\bar{\lambda}_{i<\bar{i}}$ and $\bar{\lambda}_{i\geqslant\bar{i}}$ can be expressed in terms of PGFs, we don't need the latter to compute the rates, for these are just sums over state variables. The reason for this simplification is that the overall effect on a node is just a linear combination of the effects produced by each group the node belongs to, whereas the effects coming from each neighbor could be combined nonlinearly in the node-based dynamics.

Next, the effective rate $\tilde{\beta}_{m,l}$ simply reads

$$\tilde{\beta}_{m,l} = (m-l)\bar{\lambda}_{i<\bar{i}}+l\bar{\lambda}_{i\geqslant\bar{i}}\ , \tag{3.43}$$

as we have full knowledge of the memberships m and l.

The recovery rates $\bar{\alpha}_{n,i}$ and $\tilde{\alpha}_{m,l}$ are computed analogously by estimating the average recovery rates $\bar{\mu}_{i\leqslant\bar{i}}$ and $\bar{\mu}_{i>\bar{i}}$ of infected nodes in inactive and active groups, respectively, given the within-group recovery rate $\mu(n,i)$.

3.3.1.3 Results

In the following, we test our model using diverse contagion functions able to produce either standard simple contagions and threshold (complex) contagions where a minimum number of infectious contacts (either across the neighborhood or within a group) is required for transmission. In particular, continuing with our exploration of

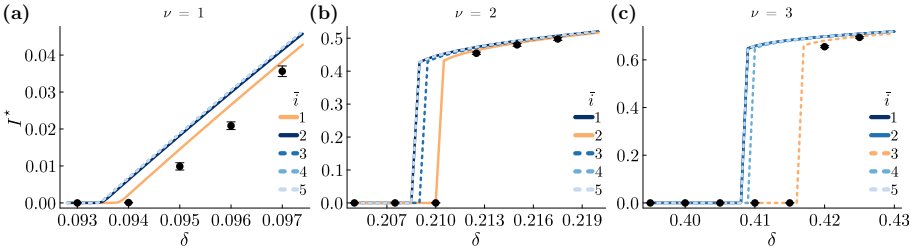


Figure 3.10: Equilibrium prevalence, $I^\star$, obtained on random 3-regular 5-uniform hypergraphs under group-based dynamics (Eqs. (3.40)-(3.43)) considering the threshold-linear infection rate in Eq. (3.44), for (a) $\nu = 1$ (initial prevalence, $I(0) = 0.05$), (b) $\nu = 2$ ($I(0) = 0.8$) and (c) $\nu = 3$ ($I(0) = 0.8$). Solid and dashed lines represent the results obtained integrating Eqs. (3.22) for $\bar{i} \in \{1, \dots, 5\}$, while points and error bars (when visible) denote averages and standard errors over 20 random realizations resulting from Monte Carlo simulations performed on hypergraphs with $N = 5 \times 10^4$ nodes. The most accurate model is always the one having $\bar{i} = \nu$ (orange).

higher-order contagions, we consider the group-based dynamics. For the latter we use the following function for the within-group infection rate

$$\lambda(n,i) = \begin{cases} \delta i & \text{if } i \geqslant \nu \\ 0 & \text{otherwise} \end{cases}, \tag{3.44}$$

where δ is the per-node infection rate. The contagion is simple (linear) for $\nu = 1$ and complex (linear-above-threshold) for $\nu > 1$. In particular, $\nu > 1$ identifies models where the inactive state is always (locally) stable (i.e, these are class III models according to the classification by Dodds *et al.* [112]; see Sec. 2.2.3.1). In such cases the critical point is thus a persistence threshold. We focus on a SIS-like dynamics from now on, so that recovery is a spontaneous node transition, for which we use a simple constant rate $\tilde{\alpha}_{n,i} = \tilde{\alpha}_{m,l} = 1$ (i.e., without loss of generality, we rescale time by the constant recovery rate).

Figure 3.10 illustrates the high accuracy of the GAME against simulations on static hypergraphs. We refer in particular to 3-regular 5-uniform hypergraphs, where all groups include 5 nodes and each node participates to 3 groups. Using all the values of $\bar{i}$ from 1 to 5, the results clearly show which characteristic activity level best captures the dynamics. We find that against both simple and complex contagions, the optimal description is the one whose characteristic activity level $\bar{i}$ matches the minimum number of infectious members required for transmission, ν. One could expect that, especially far from the critical point, a higher $\bar{i}$ might be useful to distinguish large groups that are actively transmitting from small groups where nodes were infected through some other groups. Yet, the most straightforward answer is the best: The optimal model is the one that tags as *active* the groups that *can* transmit.

A higher $\bar{i}$ might become preferable under heterogeneous group size, which allow the contagion to localize around large groups [301–303]. Further investigation is required to clarify this point.

3.3.2 *A-GAME: Adaptive structures*

Thanks to the formalism developed in Sec. 3.3.1, able to account for dynamical correlations within and across groups by tracking their activity level, the GAME provide a more accurate description than less refined mean-field models. If, on the one hand, such improvement marks an incremental step, on the other, the very same formalism opens unprecedented possibilities for the modeling of adaptive systems. In fact, the GAME's true potential is released when tracing the activity of the groups is not just a descriptive advantage but is instead the premise for specific physical mechanisms. One of the most natural of such mechanisms is perhaps strategic rewiring, where nodes (agents) rewire away from (leave) some groups to join others as an adaptive response to a process—not necessarily a contagion—unfolding through the system. At least a partial knowledge of the (activity) state of the groups is required to describe interesting adaptive responses as those observed in social, cultural, and economic systems [43]; otherwise, one could only model naïve agents leaving and joining groups entirely at random. The adaptive GAME (A-GAME) presented below account for that knowledge, allowing for the modeling of adaptive group-structured systems, namely, *adaptive hypergraphs*.

Focusing on contagion processes, we consider here the case where susceptible nodes have the possibility to rewire away from active groups. That is, we confer the activity threshold $\bar{i}$ the physical meaning of being the minimum number of infected nodes in a group that a susceptible node no longer tolerates, making it willing to leave the group. In this context, we thus refer to $\bar{i}$ as the tolerance threshold. This higher-order adaptation mechanism is a 2-fold generalization of the adaptive pairwise model by Gross *et al.* [304] (further extended still in pairwise settings by other researchers [254, 305–308]). There, a node rewires away from an infected neighbor (i.e., from an active link) to attach to a susceptible node (resp., to an inactive link), always. Therefore, apart from neglecting the group organization, this model makes the implicit assumption that nodes always possess the information (and the resources) required to perform a favorable rewiring.[13] Here, we relax that assumption by defining the probability η that rewiring is targeted towards inactive groups instead of random groups in any state. The probability η, hereafter referred to as the rewiring accuracy, effectively measures the information nodes have about groups before joining them.

13 Closer to our work, Tunc *et al.* [308] considers the model in Ref. [304] in a network with communities. There, the agents are assumed to rewire their active links while ensuring the preservation of the community structure. This requires not only information about the location of susceptible nodes in the network but also about the connectivity among communities. That agents have such global knowledge looks quite unrealistic.

Within the A-GAME framework, the model by Gross *et al.* [304] thus identifies as the exceptional case in which all groups are pairs *and* nodes have perfect information. Furthermore, with group rewiring, while the average group size and average group membership are conserved, the active membership of nodes and the size distribution of groups are allowed to change due to the adaptive mechanism. As a consequence, not only does the distribution of the pairwise degree generally change, but even its average does, a possibility that the pairwise case excludes [304]. Accordingly, our model can, in principle, interpolate between a complete network with a giant infinite group and a sparse regular network.

The rewiring rate, denoted with γ, is an effective parameter whose value can include several factors deciding how much reactive agents are, such as the information about the groups and the costs of various nature entailed by changing connections. Regardless of the exact adaptive mechanism, the annealed calculation of the mean-field quantities remains the same as in Eqs. (3.23), for the memory-less system does not capture correlations due to recent rewiring effects. Rewiring brings the following transitions to our system,

$$\dot{C}^{\mathrm{a}}_{n,i} = \gamma \mathbf{1}_{i\geqslant\bar{\imath}} \left[(n+1-i)C_{n+1,i} - (n-i)C_{n,i}\right] + \gamma \left(\frac{\eta \mathbf{1}_{i<\bar{\imath}}}{C_{i<\bar{\imath}}} + 1 - \eta\right) \left[\Omega_{S|i\geqslant\bar{\imath}} C_{n-1,i} - \Omega_{S|i\geqslant\bar{\imath}} C_{n,i}\right] , \tag{3.45a}$$

$$\dot{S}^{\mathrm{a}}_{m,l} = \gamma \left[\eta + (1-\eta)C_{i<\bar{\imath}}\right] \left[(l+1)S_{m,l+1} - lS_{m,l}\right] . \tag{3.45b}$$

These transitions, tagged with the superscript 'a' for *adaptive*, are added to Eqs. (3.22a) and (3.22b), and graphically represented by the dashed lines in Fig. 3.9. The terms in Eq. (3.45a) account for susceptible nodes leaving and joining groups, respectively, while Eq. (3.45b) only needs to account for susceptible nodes rewiring away from their active groups. To calculate these rates, we have defined $C_{i<\bar{\imath}} = \sum_{n,i<\bar{\imath}} C_{n,i}$ and $\Omega_{S|i\geqslant\bar{\imath}} = \sum_{n,i\geqslant\bar{\imath}} (n-i)C_{n,i}$, and used the indicator function $\mathbf{1}_x$ to condition on the activity of the groups. Note that we formulate a more general adaptive model where both susceptible and infectious nodes can rewire [309] in Appendix B.2.1.

3.3.2.1 Results

We experiment with adaptive hypergraphs on both simple ($\nu = 1$) and complex ($\nu = 2$) contagions. We consider 3-regular hypergraphs initialized with group sizes following a truncated Poisson distribution with mode $n = 4$ and support $\{2, \dots, 8\}$. Since we allow nodes to form groups of size at most $n = 8$, the values of the tolerance threshold $\bar{\imath}$ we need to explore goes from 1 to 8, as there is no rewiring for $\bar{\imath} \geqslant 8$ (no group can be tagged as active). Already for a simple contagion, taking $\bar{\imath} = \nu = 1$, Fig. 3.11 highlights a few notable findings.

As in pairwise networks [304], there exists a bistable region for fast enough rewiring. The bistability originates from the intricate interplay between infections and

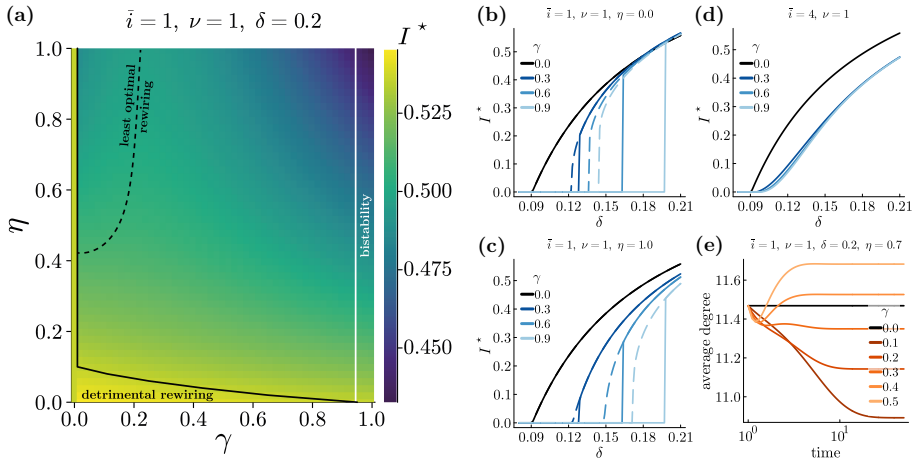


Figure 3.11: Results for adaptive hypergraphs (Eqs. (3.45)) under a simple contagion. The initial size distribution is a truncated Poisson with mode $n = 4$ and support $\{2, \dots, 8\}$ ($n = 8$ being the maximum allowed group size), yielding an average size $\langle n \rangle = 4.196$. All nodes have (constant) membership $m = 3$. (a) Phase diagram showing the equilibrium prevalence $I^\star$ against rewiring rate γ and rewiring accuracy η, when strategy $\bar{i} = \nu = 1$ is used; $I(0) = 0.45$ for the upper branch and $I(0) = 10^{-5}$ for the lower branch (for which $I^\star = 0$ on the right of the vertical white line). Apart from a bistable region when rewiring is fast enough, we find (i) a low-accuracy region where rewiring is detrimental (on the left of the black line) and (ii) a slow-rewiring region where, for high enough accuracy, prevalence presents a local maximum, defining a curve of least optimal rewiring rate (black dashed line); see Fig. C.8(a) to better appreciate these regions. ((b) and (c)) Lower (solid) and upper (dashed) branches of the prevalence $I^\star$ versus the per-node infection rate δ using different rewiring rates γ for, respectively, random ($\eta = 0.0$) and perfectly targeted ($\eta = 1.0$) rewiring. In either case, a faster rewiring (larger γ) increases the invasion threshold (which, notice, is independent from the accuracy η) while enlarging the bistable region enclosed by the hysteresis cycle. Additionally, we find that increasing accuracy heightens the persistence threshold, leading the bistable region to shrink. (d) Interestingly, increasing the tolerance threshold above 1, as to $\bar{i} = 4 \approx \langle n \rangle$, suppresses bistability by pushing the invasion threshold down to the value it takes with no rewiring ($\gamma = 0$), hence making the transition continuous. No dependence on η is observed in this case for the prevalence. (e) Time evolution of the average (pairwise) degree around the local maximum prevalence in (a); specifically, taking $\gamma \leqslant 0.5$ and $\eta = 0.7$. A very slow but nonzero rewiring (here, $\gamma = 0.1$), minimizes the number of contacts, whereas a sufficiently fast rewiring (approaching $\gamma = 0.5$) allows to better avoid contagion (while even increasing connectivity). An intermediate rewiring is instead relatively too fast to minimize contacts and too slow to escape contagion, leading to the local maximum in prevalence found between $\gamma = 0.2$ and 0.3.

rewiring. By increasing the rewiring rate γ, susceptible nodes avoid infectious ones more reactively, making the system less vulnerable to new outbreaks—the invasion threshold heightens. On the flip side, a faster rewiring also implies that susceptible nodes tend to cluster together more and increase their connectivity, in this way helping to sustain the contagion once prevalence became high enough—the persistence threshold departs more from the invasion threshold. Consequently, as in Gross *et al.* [304], the bistability region is widened by γ. While Ref. [304] only considered perfect rewiring ($\eta = 1$), we see that the bistable region, not only exists also for $\eta < 1$, but is actually widened by reducing accuracy. On one hand, the invasion threshold is independent from the latter, for leaving an infected group leads with certainty[14] to a fully susceptible one when the system is near the inactive state. On the other hand, decreasing η reduces the persistence threshold, for susceptible nodes are less able to isolate from infected ones (see Fig. C.9 and the discussion therein for details). Decreasing either η or γ reduces the persistence threshold of complex contagions ($\nu \geqslant 2$) too (see Fig. C.8).

As shown in Fig. 3.11(d), the bistability is suppressed by considering higher tolerances, $\bar{\imath} > 1$. In fact, since susceptible nodes do not rewire away in the presence of one infected node, the probability of rewiring vanishes near eradication. Consequently, the invasion threshold (hence the persistence one) reduces to the one found with no rewiring and the phase transition becomes continuous[15] (although with a different critical scaling, whose precise evaluation will need further investigation).

More importantly, the higher-order organization leads to two previously unseen phenomena. First, a low-accuracy region of *detrimental rewiring*, where prevalence is higher than with no rewiring. Second, a slow-rewiring region where, for high enough accuracy, prevalence is lower than with no rewiring, but it first increases with γ before eventually decreasing, defining a *least optimal rewiring rate* for each accuracy level η. As we will see, both phenomena require the existence of groups of different size, and are therefore excluded in pairwise networks.

To better understand the rich phenomenology of higher-order adaptation, let us vary the tolerance threshold $\bar{\imath}$ and see how this affects prevalence. The results are presented in Fig. 3.12 for either simple ($\nu = 1$) and complex ($\nu = 2$) contagions. We identify two optimal strategies for agents to elude contagions:

1. *Avoid contagious groups*. This is optimal when targeting is both fast (high γ) and accurate (high η). To do so, nodes have to mimic modelers by setting $\bar{\imath} = \nu$. Being enough reactive and precise, they can manage to escape infection without necessarily lower the connectivity of the structure;

14 Recall the assumption of infinitely many finite-sized groups.

15 We expect a similar phenomenon in pairwise networks when a susceptible node rewires away from those edges connecting to infected nodes only if such edges are at least two. It would then connect to different nodes. We find this generalization for pairwise networks however less natural, for two reasons at least. Not only having good information about the state of more edges (neighbors) might already be improbable, but then establishing multiple new connections at once might be inaccessible (e.g., because too costly).

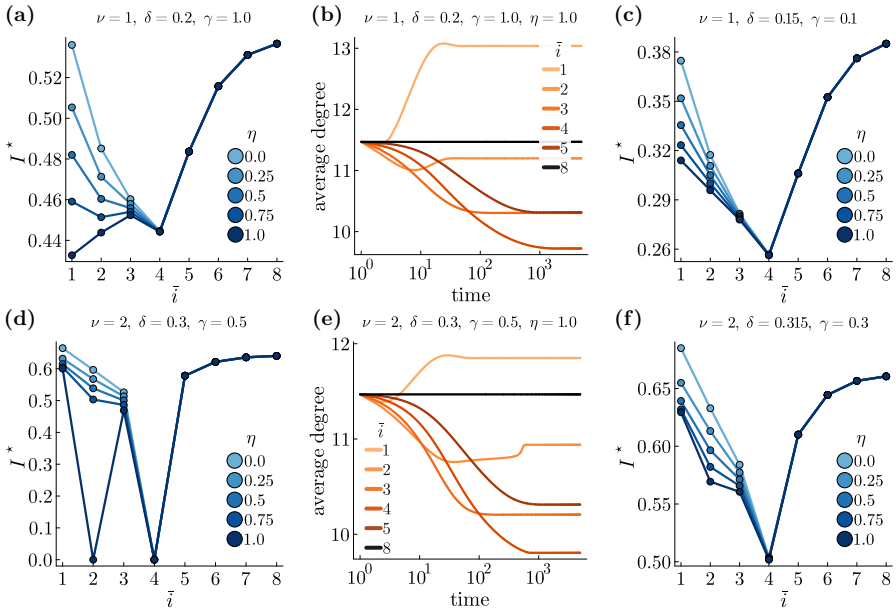


Figure 3.12: Dynamical effects of the tolerance threshold $\bar{\imath}$. The hypergraphs are initialized as in Fig. 3.11. For either simple ($\nu = 1$) and complex ($\nu = 2$) contagion, varying $\bar{\imath}$ we find two main strategies for the agents to reduce (and possibly eradicate) contagions. At high γ ((a) and (d)), if η is high too, it is optimal for agents to target the dynamics by setting $\bar{\imath} = \nu$, which is also how modelers can optimize their model as seen in Fig. 3.10. At low η, or low γ ((c) and (f)), it is optimal to target instead the structure and minimize degree by setting $\bar{\imath} = 4 \approx \langle n \rangle$. ((b) and (e)) Time evolution of the average (pairwise) degree corresponding to the results under perfect targeting ($\eta = 1.0$) in (a) and (d), respectively. Strategies $\bar{\imath} = \nu$ and $\bar{\imath} = 4 \approx \langle n \rangle$ are the two best, but whereas the second targets the structure by reducing the number of contacts, the first targets the dynamics, while not necessarily decreasing connectivity.

2. *Avoid large groups*. This is optimal when rewiring is slow (low γ) or targeting is poor (low η). By setting $\bar{\imath} \approx \langle n \rangle$, the average group size, nodes rewire away from groups of size larger than average ($\langle n \rangle + 1$, at least), thereby minimizing their average degree and the probability of getting infected.

Importantly, to say whether the rewiring is 'fast' or 'slow' and 'accurate' or 'poor', we have to compare the system with the state of its static counterpart. There is therefore no a priori boundary defining which strategy is optimal. In fact, in intermediate regimes, both strategies can work just as well but— surprisingly—an intermediary strategy ($\nu < \bar{\imath} < \langle n \rangle$) may not (see Figs. 3.12(a) and (d)). We hypothesize that this is because the mechanisms underlying these strategies are actually in opposition.

In Figs. 3.12(b) and (e), we show the average degree of the adaptive hypergraphs under different rewiring strategies. Notice that the average degree is given

by $\langle m \rangle \langle n(n-1) \rangle / \langle n \rangle$, and thus increases with group size heterogeneity, being average group size $\langle n \rangle$ and membership $\langle m \rangle$ conserved quantities. We see that the $\bar{i} = \nu$ strategy works despite slightly decreasing ($\nu = 2$) or even increasing ($\nu = 1$) the connectivity of the system eventually[16]. As expected, the $\bar{i} \approx \langle n \rangle$ strategy works by having susceptible agents avoid groups larger than average and therefore create a more uniform sparse network, actually minimizing the average degree. In accordance to our hypothesis, the two adaptive strategies work in different ways, targeting either dynamics or structure.

LEAST-OPTIMAL REWIRING RATE. A similar logic explains the region of least-optimal rewiring rates observed in Fig. 3.11(a) for sufficiently high accuracy. Larger groups, having more members that can be infected and being the infection rate proportional to the number of infected members, reach the tolerance threshold, $i = \bar{i}$, faster on average. Slow rewiring (low γ) then allows the number i of infected members to significantly correlate with group size n before the typical rewiring time ($1/\gamma$). As a consequence, when rewiring is accurate enough, susceptible agents preferentially migrate to smaller groups, decreasing the average degree. Conversely, fast rewiring (high γ) makes targeting the dynamics the optimal strategy (without necessarily decreasing connectivity). In between, we find a least optimal rewiring rate that is too slow to avoid contagion—target dynamics—but also too fast to homogenize group size and minimize degree—target structure—, for rewiring comes too early for i and n to be sufficiently correlated.

DETRIMENTAL REWIRING. The correlation between size and infected members of a group is also at the base of the detrimental rewiring found in the region of low accuracy, where rewiring is nearly performed uniformly at random. In this region, targeting dynamics is not just sub-optimal, it is the worst strategy. As seen in Fig. 3.11(a), such region is wider for slow rewiring and shrinks, until disappearing, by making it faster. When rewiring is slow, that correlation is strong, thus escaping from a group whenever just one member is infected ($\bar{i} = \nu = 1$) implies, most probably, rewiring away from a smaller group to reach a larger and more infectious one. Increasing the rewiring rate, that correlation becomes smaller, and so does the probability of connecting to a group in a worse situation. Further increasing the rewiring speed, the strategy of leaving a group whenever it turns infectious eventually pays off, for most probably leads to a susceptible group. Notice, as Fig. 3.12(d) shows, that a strategy $\bar{i} < \nu$ is also detrimental for low accuracy, but in a trivial way: nodes are ex-

16 As observed in Figs. 3.12(b) and (e), the average degree always decreases at the beginning, even under strategies like $\bar{i} = 1$ that eventually increase it. Indeed, since the probability at $t = 0$ that a group includes $\bar{i}$ or more infected members is proportional to its size, susceptible nodes starts to rewire mostly from the larger groups to join smaller groups, temporarily making the structure more uniform. Accordingly, such transient disappears when the initial prevalence is close to zero.

ceedingly intolerant, risking of leaving (still) innocuous groups to join infectious ones.

Lastly, it is evident from Fig. 3.12 how only the strategies using $\bar{i} < \langle n \rangle$ are affected by the accuracy η, especially the one that best targets dynamics, $\bar{i} = \nu$. From $\bar{i} \approx \langle n \rangle$ on, instead, the results are practically independent from η. Somewhat counterintuitive, this independence is simply explained by first noticing that, given a $\bar{i}$ that approximately equals or exceeds $\langle n \rangle$, a group can be left by a susceptible node only if its size is strictly larger than average. Then, if accuracy is high, the susceptible node targets less infectious groups, which on average are also smaller, for they possess less infected nodes than the starting group either statistically, due to the correlation between n and i, or by necessity, having size $n \leqslant \bar{i}$. On the other hand, if accuracy is low and rewiring is thus almost uniform, the node is anyway more likely to end up in smaller groups[17], which are in fact less infectious. Either strategically (high accuracy) or by chance (low accuracy), the outcome of rewiring is effectively the same. Due to the correlation between n and i, that equivalence no longer holds for strategies with $\bar{i} < \langle n \rangle$: accuracy matters, as increasing it lowers the probability of joining a larger and more infectious group when leaving a smaller one.

3.3.3 *Summary*

We studied contagions on static and adaptive hypergraphs by developing a general model to capture dynamical correlations within and across groups. To do so, we introduced the notion of characteristic scale $\bar{i}$ of a contagion to tag groups as active or inactive based on the number of infectious nodes i they contain. Our GAME thus comprises previous AME frameworks as special cases, recovered when one collapses groups to pairwise edges or considers all groups as equivalently active. Hence, whether the aim is to describe binary-state dynamics on pairwise or higher-order networks, whether static or adaptive, it is in the GAME.

We then asked three questions. What is the characteristic scale $\bar{i}$ such that our mathematics best fit simulations on static structures? What is the tolerance threshold $\bar{i}$ allowing agents to best avoid the contagion? How similar are those two answers?

We found that modelers should use a characteristic scale $\bar{i} = \nu$ set by the number of infectious neighbors ν necessary for infection. However, agents have more options if they want to avoid the contagion. When rewiring is fast and accurate, agents can act as modelers and set $\bar{i} = \nu$ to minimize the contagion events without necessarily

17 Notice that this is the case for non-skewed group size distributions like the truncated Poisson we used here. In fact, what ultimately matters is how the size of the group the node leaves compares with the median. If the median is smaller than the mean (positive skew), a uniform sampling of the groups leads, on average, to a group smaller than the one left only if the latter had size larger than the median (not necessarily than the mean). On the contrary, when the median is larger than the mean (negative skew), leaving a group of size larger than the mean leads, on average, to a larger group if the former group had size smaller than the median. Therefore, the independency from η holds, more generally, for groups with size above the median.

minimizing also their connectivity—this can even increase. When rewiring is slow or inaccurate, agents should instead aim to minimize their degree by rewiring away from large groups based on $\bar{i} \approx \langle n \rangle$, the average group size. Setting the right tolerance level can have a massive impact on the dynamics, as it can push the system below the invasion/persistence threshold.

Altogether, we introduced adaptive hypergraphs, which are not as constrained as most adaptive network models. Their average degree is not fixed over time, enabling them to organize in diverse ways. Nevertheless, the conserved quantities of average hyperdegree and hyperedge size allowed us to formulate an accurate and general model based on approximate master equations. We believe this is a promising start to the study of self-organized group-structured systems.

4

CONTEXT-DEPENDENT CONTAGION PROCESSES

In the previous chapters, we studied the effect of group interactions that, by directly involving more than two units of the system, cannot be reduced to combinations of pairwise interactions [88, 90]. In truth, this irreducibility can still hold when all the interactions in the system directly involve only two units. Indeed, even when an interaction is pairwise (e.g., an infectious transmission), if the behavior of the involved units is modified by the co-presence of other units (e.g., other people) and the behavior affects the interaction, then the latter cannot be considered in isolation. The group as a whole (the two directly interacting units and the remaining, co-present ones) defines a *context* which alters the direct interactions taking place within it. These higher-order interactions are known as *interaction modifications* (or trait-mediated indirect interactions) in ecology [310–313], whereby the throphical relationship between two species is altered by the presence of a third one. For instance, if the presence of species A induces some defensive behavior in species B (even though B may not be a prey of A), then the rate at which B predates species C is lower than it would if A was absent. By altering the behavior of B, A modifies the B-C interaction, specifically weakening it and thus helping C. Crucially, this is different from the indirect effect that a predator of B, say D, has on the predation rate of B on C. The effect would be in this case the result of a chain of direct trophical interactions, in which D decreases B's population and thus lowers the B-C predation rate, actually facilitating C's survival [310].

Going beyond ecology, we can realize that examples of interaction modifications can be also found in other kinds of systems. In opinion formation, the direct influence that an individual can exert on another can be modified by the opinions [314] or the individual characteristics [315] of the others in the group. In sociolinguistics, when an individual moves to a new region, the learning rate of the local language depends on how often the locals use that language in the presence of that individual [316]. They could bear the cost of switching to another language (e.g., English) for a more inclusive conversation, but eventually slowing down the diffusion of the local one; or could instead stick to it, enforcing its diffusion. In neuronal systems, glias manipulate neurotransmission and synaptic connections between neurons [317, 318]. In biochemical systems, enzymes increase the reaction rate among certain species, while

inhibitors blocking the activity of enzymes, in turn decrease that rate [319]. Lastly, a percolation process (yet to be mapped to a physical one) defined by interaction modifications where the activity of nodes regulates the activation of links, has been shown to generate time-varying networks where the size of the giant connected component undergoes a period doubling and a route to chaos [320].

In the following sections, we will focus on spreading dynamics, specifically on epidemic spreading. Recent studies have emphasized the importance of considering the group organization of contacts in understanding and controlling a contagion process [113, 257, 273, 321]. In this regard, the adoption of nonpharmaceutical prophylactic behaviors, such as face-mask wearing and physical distancing, proves to be a viable and effective strategy for epidemic control [322–331]. Existing studies on (adaptive) network models for the spread of contagious diseases have considered agents whose behavior depends either on individual characteristics, on the observed behavior among peers, and on external sources of information, within both well-mixed and structured populations (see references [62-89] in the review by Benson et al. [332], and the review by Wang et al. [333]).

However, these studies do not distinguish between potentially infectious contacts that occur in isolation versus those that occur in the co-presence of other individuals. This distinction becomes essential when accounting for behavioral adoption. Indeed, even though each transmission concerns only two people, the likelihood of transmission is indirectly affected by the way contacts are organized within groups, for the adoption of the prophylactic behavior (e.g., wearing a face-mask) depends on the level of adoption an individual observes in the entire group [99, 334, 335]. Adoption is thus mutable, contextual: an individual may exhibit one behavior or another depending on the current context. We can make a complete parallel with the species' example above. For instance, if individual A joins individuals B and C, and A wears a face-mask, this may induce B and/or C to wear it too. If, say, B was infectious, the behavioral change (wearing the mask) induced by observing A's behavior will have decreased the probability of a transmission (from B) occurring in the group. Again, this is different from the indirect effect of an interaction chain in which a B's contact, say individual D, by wearing a mask reduces their own probability of getting an infection and thus of infecting B, in turn lowering the probability for B to infect any of their contacts (e.g., A and C).

To account for context-dependency, we first need a way to describe a context. As shown by the examples above, the context is defined by either the intrinsic properties and/or the dynamical state of the involved—co-present—units. Therefore, modeling contexts and their effects requires a theory able to account for the interaction patterns among two or more units of different type and/or state. We provide the most basic instance of such a theory in the Sec. 4.1 to then use it to study contagion processes in Secs. 4.2 and 4.3. To first show the pure effect of different interaction patterns, in Sec. 4.2, we consider the degenerate case of context-independent behavior, where the spreading between agents is affected by fixed individual traits, namely whether the

agent adopt or not a "permanent" prophylactic measure (e.g., get vaccinated). There we show how varying the mixing between adopters and non-adopters gives rise to three dynamical regimes, with interesting practical consequences. Then, in Sec. 4.3, we finally explore the dynamical implications of a context-dependent behavior, revealing a rich phenomenology which radically depends on either epidemiological and sociological factors. The introduction of context-dependency stimulates the discourse on the interplay between behavior and epidemics along a new dimension of complexity, hopefully providing novel mechanisms to interpret, and perhaps better control, real epidemics. We close the chapter by giving in Sec. 4.4 a hint of a new framework that would allow to deal with systems characterized by generic interaction modifications. We point out a flexible representation for such systems in the form of a multilayer, multipartite network.

4.1 HETEROGENEOUS MIXING IN HIGHER-ORDER NETWORKS

Understood a context as a set of co-present units with their properties (states, intrinsic characteristics, etc.), it can be described as a configuration of an edge of a hypergraph. The entire hypergraph, through its edge set, encodes the set of configurations—one for each edge—simultaneously present in the system, and thus contains a huge amount of information in the form of which unit belongs to which edge and with which values of their properties. All this information is usually unavailable and, anyway, often unnecessary to observe the dynamical effects of context. Rather, retaining the essential information makes such effects not only observable but also mathematically tractable, allowing us to obtain general understanding.

For the current purposes of our work, it is fine assuming that the properties of a unit can be represented via a categorical variable, the 'type' of the unit. We thus ask the following. Given a hypergraph and a partition of its nodes into different types, we want the minimal description able to account for the probability distributions for each type to belong to edges of various size and type-composition (i.e., to various contexts). To this end, we can neglect the exact edge set of the hypergraph and refer instead to its annealed version preserving those probability distributions. This is done by making the probability for a given edge to occur be a function of the type of the nodes it involves. Put differently, we operate a dimensional reduction based on the assumption that the type of the nodes is the sole property needed to infer the interaction structure. In doing so, we find and parametrize an ensemble of hypergraphs with the same edge size and type-composition distributions, of which the original hypergraph is a particular realization. By taking into account the (potentially) heterogeneous mixing among nodes of different type, such description allows for the study of the effect of variable local contexts on the dynamical processes of interest.

We first present the general theory valid for any hypergraph, and then provide explicit formulas for rank-3 hypergraphs. We consider the case in which nodes are partitioned into two types, which is what we need for the contagion models in Secs. 4.2

and 4.3. The generalization of the model to account for any number of types is conceptually straightforward, albeit the notation can become quite cumbersome.

4.1.1 *General case*

Consider a set (a population) of N nodes (agents) where each of them is assigned one of two types. Given the meaning the type of an agent will take in the following sections, without loss of generality, we label the two types as 'A' and 'N'. Let ρ be the proportion of A-type nodes, hence $1-\rho$ the proportion of N-type's. We assume that nodes are grouped at random (e.g., at each time step of a given process), with probabilities depending solely on the types of the nodes involved in the groups. Accordingly, the value of the higher-order adjacency tensor $A^{(n)}_{i_1,\dots,i_n}$, encoding whether the n-edge incident on the nodes $\{i_1,\dots,i_n\}$ exists or not, depends only on the set $\{X_{i_1},\dots,X_{i_n}\}$ of the nodes' types, and can be thus replaced with its expected value over the ensemble of hypergraphs parametrized by N, ρ, and a set of mixing parameters (defined below) regulating the average level of type-(dis)assortativity in a group. To this end, we need an expression for the number $g^{(n)}_{n_\mathrm{A},n_\mathrm{N}}$ of $(n_\mathrm{A}+n_\mathrm{N})$-edges composed of n_A A-type nodes and n_N N-type ones, being subsets of n-edges (i.e., $n_\mathrm{A}+n_\mathrm{N}\leqslant n\leqslant n_{\max}$). From this we can then derive an expression for the probability that a node of a certain type takes part in a group of any given type-composition.

Denoted with $e^{(n)}$ the total number of n-edges (e.g., formed at a given time) in the hypergraph, $g^{(n)}_{n_\mathrm{A},n_\mathrm{N}}$ is related to it through

$$p^{(n)}_{n_\mathrm{A},n_\mathrm{N}}=\frac{g^{(n)}_{n_\mathrm{A},n_\mathrm{N}}}{\binom{n}{n_\mathrm{A}+n_\mathrm{N}}e^{(n)}}\,, \tag{4.1}$$

being $p^{(n)}_{n_\mathrm{A},n_\mathrm{N}}$ the probability that sampling $n_\mathrm{A}+n_\mathrm{N}$ nodes from a randomly chosen n-edge (there are $\binom{n}{n_\mathrm{A}+n_\mathrm{N}}e^{(n)}$ ways of doing it), the sample consists of n_A A-type's and n_N N-type's. Also, let $p^{(n)}_{l_\mathrm{A},l_\mathrm{N}|m_\mathrm{A},m_\mathrm{N}}$ be the conditional probability that, given a set of m_A A-type and m_N N-type nodes in an n-edge, sampling a set of other $l_\mathrm{A}+l_\mathrm{N}$ nodes in it ($m_\mathrm{A}+l_\mathrm{A}+m_\mathrm{N}+l_\mathrm{N}\leqslant n$), the sample consists of l_A A-type's and l_N N-type's. Since there are $\binom{m_\mathrm{X}+l_\mathrm{X}}{m_\mathrm{X}}$ ways of choosing m_X nodes out of $m_\mathrm{X}+l_\mathrm{X}$, $\mathrm{X}\in\{\mathrm{A},\mathrm{N}\}$, and $\binom{n-m_\mathrm{A}-m_\mathrm{N}}{l_\mathrm{A}+l_\mathrm{N}}$ ways of choosing $l_\mathrm{A}+l_\mathrm{N}$ nodes among the $n-m_\mathrm{A}-m_\mathrm{N}$ remaining ones, we get

$$p^{(n)}_{l_\mathrm{A},l_\mathrm{N}|m_\mathrm{A},m_\mathrm{N}}=\frac{\binom{m_\mathrm{A}+l_\mathrm{A}}{m_\mathrm{A}}\binom{m_\mathrm{N}+l_\mathrm{N}}{m_\mathrm{N}}}{\binom{n-m_\mathrm{A}-m_\mathrm{N}}{l_\mathrm{A}+l_\mathrm{N}}}\frac{g^{(n)}_{m_\mathrm{A}+l_\mathrm{A},m_\mathrm{N}+l_\mathrm{N}}}{g^{(n)}_{m_\mathrm{A},m_\mathrm{N}}}\,. \tag{4.2}$$

In particular, $p^{(n)}_{l_\mathrm{A},l_\mathrm{N}|1,0}$ ($p^{(n)}_{l_\mathrm{A},l_\mathrm{N}|0,1}$) is the sought probability that an A-type (N-type) node takes part in a group of size n which includes other l_A A-types' and l_N N-types'.

Therefore, indicated with $k_{1,0}^{(n)}$ ($k_{0,1}^{(n)}$) the expected n-degree of an A-type (N-type) node, i.e., the expected number of groups of size n to which it takes part in (per time step), and given $l_{\mathrm{A}} + l_{\mathrm{N}} = n - 1$, then $\kappa_{l_{\mathrm{A}},l_{\mathrm{N}}|1,0}^{(n)} = p_{l_{\mathrm{A}},l_{\mathrm{N}}|1,0}^{(n)} k_{1,0}^{(n)}$ ($\kappa_{l_{\mathrm{A}},l_{\mathrm{N}}|0,1}^{(n)} = p_{l_{\mathrm{A}},l_{\mathrm{N}}|0,1}^{(n)} k_{0,1}^{(n)}$) is the expected number of groups of size n in which an A-type (N-type) takes part with other l_{A} A-type's and l_{N} N-type's. These numbers, beside N and ρ, specify completely the annealed interaction structure.

To compute them, we need an expression for $g_{n_{\mathrm{A}},n_{\mathrm{N}}}^{(n)}$. Consider then a subset $\{\mathrm{A}_1,\ldots,\mathrm{A}_{n_{\mathrm{A}}},\mathrm{N}_1,\ldots,\mathrm{N}_{n_{\mathrm{N}}}\}$ of $n_{\mathrm{A}} \geqslant 1$ A-type nodes and $n_{\mathrm{N}} \geqslant 1$ N-type ones. We want to calculate the probability of finding the other $n_{\mathrm{A}} + n_{\mathrm{N}} - 1$ nodes in the subset conditioned on a single node that we can choose to be, for instance, N_1, that is, $P\left(\{\mathrm{A}_1,\ldots,\mathrm{A}_{n_{\mathrm{A}}},\mathrm{N}_2,\ldots,\mathrm{N}_{n_{\mathrm{N}}}\}\,|\mathrm{N}_1\right)$. There are $(n_{\mathrm{A}} + n_{\mathrm{N}} - 1)!$ different ways of writing this probability, one for each possible order in which the $n_{\mathrm{A}} + n_{\mathrm{N}} - 1$ nodes can be found. We can thus choose one particular ordering, say $(\mathrm{N}_2,\ldots,\mathrm{N}_{n_{\mathrm{N}}},\mathrm{A}_1,\ldots,\mathrm{A}_{n_{\mathrm{A}}})$, and apply recursively the definition of conditional probability (i.e., $P(\{x,y\}) = P(y)P(x|y)$ and $P(\{x,y\}|z) = P(y|z)P(x|\{y,z\})$ to get

$$\begin{aligned}
&P(\{\mathrm{A}_1,\ldots,\mathrm{A}_{n_{\mathrm{A}}},\mathrm{N}_2,\ldots,\mathrm{N}_{n_{\mathrm{N}}}\}|\mathrm{N}_1)\\
&\quad= (n_{\mathrm{A}} + n_{\mathrm{N}} - 1)!\\
&\qquad\times P\left(\mathrm{N}_2|\mathrm{N}_1\right)\\
&\qquad\quad\vdots\\
&\qquad\times P\left(\mathrm{N}_{n_{\mathrm{N}}}|\left\{\mathrm{N}_1,\ldots,\mathrm{N}_{n_{\mathrm{N}}-1}\right\}\right)\\
&\qquad\times P\left(\mathrm{A}_1|\left\{\mathrm{N}_1,\ldots,\mathrm{N}_{n_{\mathrm{N}}}\right\}\right)\\
&\qquad\quad\vdots\\
&\qquad\times P\left(\mathrm{A}_{n_{\mathrm{A}}}|\left\{\mathrm{N}_1,\ldots,\mathrm{N}_{n_{\mathrm{N}}},\mathrm{A}_1,\ldots,\mathrm{A}_{n_{\mathrm{A}}-1}\right\}\right)\,.
\end{aligned} \tag{4.3}$$

Given that we are not interested on the labels of the nodes but exclusively on their types, using the notation $n\mathrm{X}$ to denote a set of n X-type nodes, Eq. (4.3) becomes

$$\begin{aligned}
P(n_{\mathrm{A}}\mathrm{A},(n_{\mathrm{N}}-1)\,\mathrm{N}|\mathrm{N}) = &\binom{n_{\mathrm{A}}+n_{\mathrm{N}}-1}{n_{\mathrm{A}}}\\
&\times P\left(\mathrm{N}|\mathrm{N}\right)P\left(\mathrm{N}|2\mathrm{N}\right)\cdots P\left(\mathrm{N}|\left(n_{\mathrm{N}}-1\right)\mathrm{N}\right)\\
&\times P\left(\mathrm{A}|n_{\mathrm{N}}\mathrm{N}\right)\cdots P\left(\mathrm{A}|n_{\mathrm{N}}\mathrm{N},\left(n_{\mathrm{A}}-1\right)\mathrm{A}\right)\,,
\end{aligned} \tag{4.4}$$

where we divided by $(n_{\mathrm{N}} - 1)!$ and $n_{\mathrm{A}}!$, which are the number of permutations of the $n_{\mathrm{N}} - 1$ N-type and n_{A} A-type nodes in the sequence, respectively, equally contributing to $P(n_{\mathrm{A}}\mathrm{A},(n_{\mathrm{N}}-1)\,\mathrm{N}|\mathrm{N})$.

If the $n_{\mathrm{A}} + n_{\mathrm{N}} - 1$ nodes are drawn from a set of $n - 1 \geqslant n_{\mathrm{A}} + n_{\mathrm{N}} - 1$ nodes, there are $\binom{n-1}{n_{\mathrm{A}}+n_{\mathrm{N}}-1}$ distinct ways of doing it for any sequence chosen for the $n_{\mathrm{A}} + n_{\mathrm{N}} - 1$ nodes. Then, if there is a total of N_{N} N-type nodes in the hypergraph and, on average, each one is included in $k_{\mathrm{N}}^{(n)}$ n-edges (per time step), the expected number $g_{n_{\mathrm{A}},n_{\mathrm{N}}}^{(n)}$ of

$(n_A + n_N)$-edges composed of n_A A-type nodes and n_N N-type ones, subsets of n-edges, can be written as

$$\begin{aligned} g^{(n)}_{n_A,n_N} = & \frac{N_N k^{(n)}_N}{n_N} \binom{n-1}{n_A+n_N-1} \binom{n_A+n_N-1}{n_A} \\ & \times P(\mathrm{N}|\mathrm{N}) P(\mathrm{N}|2\mathrm{N}) \cdots P(\mathrm{N}|(n_N-1)\mathrm{N}) \\ & \times P(\mathrm{A}|n_N\mathrm{N}) \cdots P(\mathrm{A}|n_N\mathrm{N},(n_A-1)\mathrm{A}) \ , \end{aligned} \tag{4.5}$$

where the term $1/n_N$ accounts for the fact that there are n_N N-type nodes to condition upon in the set. Switching back to the more compact notation used in Eqs. (4.1) and (4.2), Eq. (4.5) becomes

$$\begin{aligned} g^{(n)}_{n_A,n_N} = & \frac{N_{0,1} k^{(n)}_{0,1}}{n_N} \binom{n-1}{n_A+n_N-1} \binom{n_A+n_N-1}{n_A} \\ & \times p^{(n)}_{0,1|0,1} p^{(n)}_{0,1|0,2} \cdots p^{(n)}_{0,1|0,n_N-1} \\ & \times p^{(n)}_{1,0|0,n_N} p^{(n)}_{1,0|1,n_N} \cdots p^{(n)}_{1,0|n_A-1,n_N} \ , \end{aligned} \tag{4.6}$$

For uniform N-type and A-type subsets, we have

$$g^{(n)}_{0,n_N} = \frac{N_{0,1} k^{(n)}_{0,1}}{n_N} \binom{n-1}{n_N-1} p^{(n)}_{0,1|0,1} p^{(n)}_{0,1|0,2} \cdots p^{(n)}_{0,1|0,n_N-1} \ , \tag{4.7}$$

$$g^{(n)}_{n_A,0} = \frac{N_{1,0} k^{(n)}_{1,0}}{n_A} \binom{n-1}{n_A-1} p^{(n)}_{1,0|1,0} p^{(n)}_{1,0|2,0} \cdots p^{(n)}_{1,0|n_A-1,0} \ . \tag{4.8}$$

as simply obtained by putting $n_A = 0$ in Eq. (4.6) and by symmetry under the exchange of the two types.

Inserting Eqs. (4.6) to (4.8) into Eq. (4.2) we are thus expressing $p^{(n)}_{l_A,l_N|m_A,m_N}$ only in terms of conditional probabilities of finding a single node of given type, i.e., of the form $p^{(n)}_{1,0|m_A,m_N}$ and $p^{(n)}_{0,1|m_A,m_N}$. The latter are the only free variables left. To close the theory, given $m_A + m_N = m \leqslant n-1$, we parametrize them as follows:

$$p^{(n)}_{1,0|m_A,m_N} = \alpha^{(n)}_{m,m_A} \rho \ , \qquad \text{if } m_A \leqslant m_N \ , \tag{4.9}$$

$$p^{(n)}_{0,1|m_A,m_N} = \alpha^{(n)}_{m,m_A} (1-\rho) \ , \qquad \text{if } m_A > m_N \ . \tag{4.10}$$

Their counterparts are found from normalization, e.g., $p^{(n)}_{0,1|m_A,m_N} = 1 - \alpha^{(n)}_{m,m_A}\rho$ if $m_A \leqslant m_N$. Parameter $\alpha^{(n)}_{m,m_A}$ governs the mixing probability in a $(m+1)$-edge, subset of an n-edge, conditioned on the presence of $m_A \leqslant m$ A-type nodes in it: $\alpha^{(n)}_{m,m_A} = 1$ corresponds to homogeneous mixing, in which case it is only the proportion of the types in the population to determine the expected mixing; for $m_A \neq m_N$, $\alpha^{(n)}_{m,m_A} < 1$ ($\alpha^{(n)}_{m,m_A} > 1$) indicates (dis)assortativity of the majority in the subset towards the remaining node, which is therefore more (less) likely than expected to be of the same

type of the majority; for $m_A = m_N$ (occurring only for m even), $\alpha^{(n)}_{m,m_A} < 1$ ($\alpha^{(n)}_{m,m_A} > 1$) indicates asymmetric preference towards N-type (A-type) nodes.

Since $m_A \in \{0, \dots, m\}$, there are $m+1$ parameters characterizing the mixing within $(m+1)$-edges. However, only one of them is free. Indeed, observe from Eq. (4.2) that the number of mixed groups $g^{(n)}_{m_A+l_A,m_N+l_N}$ (m_A, $m_N \geqslant 1$) can be expressed in terms of both $p^{(n)}_{l_A,l_N|m_A,m_N}$ and $p^{(n)}_{l_A-1,l_N+1|m_A+1,m_N-1}$ if $l_A = 1$ ($l_N = 0$), establishing a relation between $\alpha^{(n)}_{m,m_A}$ and $\alpha^{(n)}_{m,m_A+1}$; or, alternatively, in terms of both $p^{(n)}_{l_A,l_N|m_A,m_N}$ and $p^{(n)}_{l_A+1,l_N-1|m_A-1,m_N+1}$ if $l_N = 1$ ($l_A = 0$), providing a relation between $\alpha^{(n)}_{m,m_A}$ and $\alpha^{(n)}_{m,m_A-1}$. Since there are m distinct mixed group configurations (one for each $m_A \in \{1, \dots, m\}$), there are m constraints relating the $m+1$ mixing parameters, which can thus be expressed in terms of only one of them. Those relations can be found by simply substituting Eqs. (4.9) and (4.10) into Eqs. (4.6) to (4.8) and, as explained before, comparing related pairs of conditional probabilities[1]. Notice also that the sets $\{\alpha^{(n)}_{m,m_A}\}$ and $\{\alpha^{(n')}_{m,m_A}\}$, $m < \min\{n, n'\}$, satisfy the same set of relations, as the form of Eqs. (4.9) and (4.10) does not depend on m: $\alpha^{(n)}_{m,m_A}$ and $\alpha^{(n')}_{m,m_A}$ can only differ in their value, being they computed on different sets (n- and n'-edges, respectively). Therefore, we can find all the relations by just considering the case $m = n-1$ for each $n \in \{2, \dots, n_{\max}\}$. Having one free parameter for each $m \in \{1, \dots, n-1\}$, we get $n-1$ parameters to fix for group size n. Being $n \in \{2, \dots, n_{\max}\}$, the annealed structure is uniquely determined by at most $\sum_{n=2}^{n_{\max}}(n-1) = \binom{n_{\max}}{2}$ parameters (less if $e^{(n)} = 0$ for some $n \in \{2, \dots, n_{\max}\}$).

Lastly, let us define the pairwise assortativity (or just "assortativity"; or "homophily", as called later on) in n-edges, $h^{(n)}$, as the probability that two nodes are of the same type when interacting within a group of size n,

$$\begin{aligned} h^{(n)} &= \frac{N_{1,0}k^{(n)}_{1,0}p^{(n)}_{1,0|1,0} + N_{0,1}k^{(n)}_{0,1}p^{(n)}_{0,1|0,1}}{N_{1,0}k^{(n)}_{1,0} + N_{0,1}k^{(n)}_{0,1}} \\ &= 1 - \rho(1-\rho)\frac{\alpha^{(n)}_{1,0}k^{(n)}_{0,1} + \alpha^{(n)}_{1,1}k^{(n)}_{1,0}}{k^{(n)}} \end{aligned} \tag{4.11}$$

where $k^{(n)} = \rho k^{(n)}_{1,0} + (1-\rho)k^{(n)}_{0,1}$ is the average n-degree. Note that this is only one of the several ways in which assortativity can be quantified within groups of size $n > 2$. Indeed, one can consider a weaker notion of assortativity by defining an index for each one of the group compositions where one type is majoritarian (e.g., for groups of three nodes, in addition to the index counting configurations where the three nodes are of the same type, another one counting those where only two of them are of the same type; see Veldt et al. [336] for a deep analysis about the combinatorial possibilities and constraints of group assortativity). In this sense, assortativity in pairs is

1 Equivalently, the relations can be found by expressing $g^{(n)}_{m_A,m_N}$ using different orderings of the nodes in Eq. (4.3).

exceptional, for there is only one way of being majoritarian. We choose to focus on pairwise assortativity because the direct interactions we will consider—those mediating a spreading dynamics—are pairwise, hence knowing how frequently different pairs form is of primary importance here. Nonetheless, from the previous analysis we know that fixing $h^{(n)}$ alone is not sufficient to fully specify the group organization. Therefore, unless otherwise specified, we will fix the higher-order mixing parameters ($\alpha^{(n)}_{m,m_A}$, $m > 1$) and vary instead the pairwise ones ($\alpha^{(n)}_{1,m_A}$, $m_A = 0,1$). In particular, chosen a value for the average assortativity $h = \sum_{n=2}^{n_{max}} h^{(n)} (n-1) k^{(n)}/k$, being $k = \sum_{n=2}^{n_{max}} (n-1) k^{(n)}$ the average pairwise degree (i.e., the average degree over the 2-section of the hypergraph), we vary the type-assortativities $\left\{h^{(n)}\right\}$ and analyze how this affects the dynamics. Using Eq. (4.11), we find

$$h = 1 - (\alpha_{1,0} + \alpha_{1,1})\rho\,(1-\rho)\ , \tag{4.12}$$

where $\alpha_{1,0} = \sum_{n=2}^{n_{max}} \alpha^{(n)}_{1,0} (n-1) k^{(n)}_{0,1}/k = \sum_{n=2}^{n_{max}} \alpha^{(n)}_{1,1} (n-1) k^{(n)}_{1,0}/k = \alpha_{1,1}$ is the average pairwise mixing parameter.

In conclusion, the theory we developed in this section allows for the complete parametrization of the mixing patterns between two types of nodes in hypergraphs of any rank. On one hand, the set consisting of N, ρ, the (at most) $\binom{n_{max}}{2}$ mixing parameters, and the average degrees for each type, identifies an ensemble that one can sample to generate hypergraphs with those annealed properties. On the other, given a hypergraph and an assignment of types (or states, given the dynamics is binary) to its nodes, one can extract the set of parameters identifying the ensemble and so obtain global information about the mixing between the two types (or states) at each group size. An interesting possibility unlocked by the presented theory is to track, over time, the assortativity between nodes with different state. This would allow us to quantify the dynamical correlations at the various group sizes, which, as shown in Sec. 3, can crucially drive the dynamics.

4.1.2 *Pairs and triads*

As a minimal application of this formalism, we will consider agents interacting only within pairs (2-edges) and triads (3-edges). For $n_{max} = 3$ there are at most three free parameters. To start with, Eqs. (4.9) and (4.10) read

$$p^{(2)}_{1,0|0,1} = \alpha^{(2)}_{1,0}\rho\ , \tag{4.13}$$

$$p^{(2)}_{0,1|1,0} = \alpha^{(2)}_{1,1}(1-\rho)\ , \tag{4.14}$$

for $n = 2$, and

$$p^{(3)}_{1,0|0,1} = \alpha^{(3)}_{1,0}\rho \,, \tag{4.15}$$
$$p^{(3)}_{0,1|1,0} = \alpha^{(3)}_{1,1}(1-\rho) \,, \tag{4.16}$$
$$p^{(3)}_{1,0|0,2} = \alpha^{(3)}_{2,0}\rho \,, \tag{4.17}$$
$$p^{(3)}_{1,0|1,1} = \alpha^{(3)}_{2,1}\rho \,, \tag{4.18}$$
$$p^{(3)}_{0,1|2,0} = \alpha^{(3)}_{2,2}(1-\rho) \,, \tag{4.19}$$

for $n = 3$. Choosing then $\alpha^{(2)}_{1,0}$, $\alpha^{(3)}_{1,0}$ and $\alpha^{(3)}_{2,1}$ as the free parameters, one finds the relations

$$\alpha^{(n)}_{1,1} = \frac{k^{(n)}_{0,1}}{k^{(n)}_{1,0}}\alpha^{(n)}_{1,0} \quad (n = 2,3) \,, \tag{4.20}$$
$$\alpha^{(3)}_{2,0} = \frac{\alpha^{(3)}_{1,0}\left(1-\alpha^{(3)}_{2,1}\rho\right)}{1-\alpha^{(3)}_{1,0}\rho} \,, \tag{4.21}$$
$$\alpha^{(3)}_{2,2} = \frac{\alpha^{(3)}_{1,1}\alpha^{(3)}_{2,1}\rho}{1-\alpha^{(3)}_{1,1}(1-\rho)} \,. \tag{4.22}$$

We eventually find the following expressions for the mixing probabilities in triads conditioned on a single node,

$$p^{(3)}_{2,0|0,1} = \alpha^{(3)}_{1,0}\rho\,\alpha^{(3)}_{2,1}\rho \,, \tag{4.23}$$
$$p^{(3)}_{1,1|0,1} = 2\alpha^{(3)}_{1,0}\rho\left(1-\alpha^{(3)}_{2,1}\rho\right) \,, \tag{4.24}$$
$$p^{(3)}_{0,2|0,1} = \left(1-\alpha^{(3)}_{1,0}\rho\right)\left(1-\alpha^{(3)}_{2,0}\rho\right) \,, \tag{4.25}$$
$$p^{(3)}_{0,2|1,0} = \alpha^{(3)}_{1,1}(1-\rho)\left(1-\alpha^{(3)}_{2,1}\rho\right) \,, \tag{4.26}$$
$$p^{(3)}_{1,1|1,0} = 2\alpha^{(3)}_{1,1}(1-\rho)\,\alpha^{(3)}_{2,1}\rho \,, \tag{4.27}$$
$$p^{(3)}_{2,0|1,0} = \left[1-\alpha^{(3)}_{1,1}(1-\rho)\right]\left[1-\alpha^{(3)}_{2,2}(1-\rho)\right] \,, \tag{4.28}$$

which, together with Eqs. (4.13) to (4.16), specify completely the mixing patterns for a single node within groups of size $n = 2,3$. Note that, to ensure that all the mixing

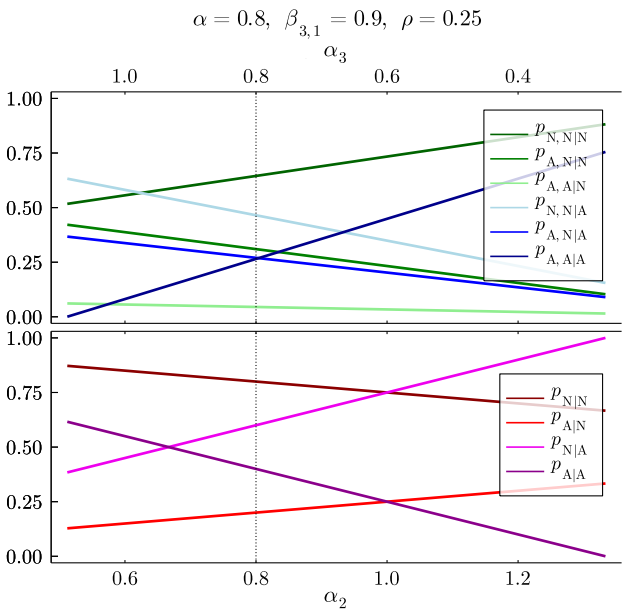


Figure 4.1: Conditional mixing probabilities for $n_{\max} = 3$, considering $\rho = 0.25$, $\beta_{3,1} = 0.9$, $k_{\mathrm{N}}^{(n)} = k_{\mathrm{A}}^{(n)}\ \forall n$, $k^{(2)} = 2k^{(3)}$ and $\alpha = 0.8$ (vertical dotted line). For each feasible value of α_2 (bottom axis), α_3 (top axis) is given by Eq. (4.32).

probabilities lie in $[0,1]$, the intervals of variation of $\alpha_{1,0}^{(2)}$, $\alpha_{1,0}^{(3)}$ and $\alpha_{2,1}^{(3)}$ must fulfil the following structural constraints:

$$\alpha_{1,0}^{(2)} \in \left[0, \min\left\{\frac{1}{\rho}, \frac{1}{1-\rho}, \frac{k_{1,0}^{(2)}}{(1-\rho)\,k_{0,1}^{(2)}}\right\}\right] , \tag{4.29}$$

$$\alpha_{1,0}^{(3)} \in \left[0, \min\left\{\frac{1}{\rho}, \frac{1}{1-\rho}, \frac{k_{1,0}^{(3)}}{(1-\rho)\,k_{0,1}^{(3)}}, \frac{k^{(3)}}{3\rho\,(1-\rho)\,k_{0,1}^{(3)}}\right\}\right] , \tag{4.30}$$

$$\alpha_{2,1}^{(3)} \in \left[\max\left\{0, \frac{2\alpha_{1,0}^{(3)}\rho - 1}{\alpha_{1,0}^{(3)}\rho^2}\right\}, \min\left\{\frac{1}{\rho}, \frac{1-\alpha_{1,1}^{(3)}\,(1-\rho)}{\alpha_{1,1}^{(3)}\rho\,(1-\rho)}\right\}\right] . \tag{4.31}$$

In particular, the forth upper bound in Eq. (4.30) guarantees the feasibility of the structure by constraining the lower bound for $\alpha_{2,1}^{(3)}$ to stay below its upper bound.

When keeping fixed the average assortativity h (hence $\alpha_{1,0}$), we get the additional relation

$$\alpha_{1,0}^{(3)} = \frac{1}{2k_{0,1}^{(3)}}\left(\alpha_{1,0}k - \alpha_{1,0}^{(2)}k_{0,1}^{(2)}\right) , \tag{4.32}$$

with $k = k^{(2)} + 2k^{(3)}$. Equation (4.32) puts further bounds on $\alpha_{1,0}^{(2)}$ in order for $\alpha_{1,0}^{(3)}$ to satisfy Eq. (4.30).

Dealing with pairs and triads only, we henceforth switch to a more explicit notation. In particular, we use $k_{\mathrm{N}}^{(n)}$ instead of $k_{0,1}^{(n)}$, $n = 2, 3$; $\kappa_{\mathrm{N|N}}$ and $\kappa_{\mathrm{A|N}}$ instead of $\kappa_{0,1|0,1}^{(2)}$ and $\kappa_{1,0|0,1}^{(2)}$, respectively; and $\kappa_{\mathrm{N,N|N}}$, $\kappa_{\mathrm{A,N|N}}$ and $\kappa_{\mathrm{A,A|N}}$ instead of $\kappa_{0,2|0,1}^{(3)}$, $\kappa_{1,1|0,1}^{(3)}$ and $\kappa_{2,0|0,1}^{(3)}$, respectively (analogously when exchanging N with A). The same notation is also used for the conditional probabilities (e.g., $p_{2,0|0,1}^{(3)}$ becomes $p_{\mathrm{A,A|N}}$). Lastly, we denote the three free parameters we have as $\alpha_2 \equiv \alpha_{1,0}^{(2)}$, $\alpha_3 \equiv \alpha_{1,0}^{(3)}$ and $\beta_{3,1} \equiv \alpha_{2,1}^{(3)}$, and the average pairwise mixing parameter as $\alpha \equiv \alpha_{1,0} = \alpha_{1,1}$. Figure 4.1 displays an example how the conditional probabilities of mixing change when varying α_2 and α_3 while keeping α fixed.

4.2 CONTAGIONS UNDER CONTEXT-INDEPENDENT BEHAVIOR

Preliminarily to the exploration of the dynamical effects that a context-dependent behavior can have on a contagion, we focus on the degenerate case in which the behavior of an individual does not depend on context, but just coincides with its type, being thus the same during each interaction. For instance, in epidemic spreading—which is the focus of this section—individuals can be split in vaccinated and not vaccinated and such division accounts for the dynamical difference between the two types of individuals. Clearly, people do not face the decision about getting vaccinated or not each time they interact with someone. Once they get a vaccine, they cannot undone it at will—they could at most wait for the vaccine efficacy to wane. Being vaccinated is thus a fixed trait. The same holds for any other prophylactic measure that cannot be arbitrarily adopted and lifted or that is usually not lifted once it has been adopted.

Assuming a context-independent behavior we can study how the contagion dynamics is exclusively affected by the way individuals with different fixed traits are distributed on the contact structure. Such distribution could be independent from those traits and thus be uniform, as if traits were assigned uniformly at random; or contacts could be more likely to occur between individuals with either similar or dissimilar traits, making the contact pattern assortative or disassortative, respectively. In fact, assortativity patterns are those usually observed in human social systems, where individuals with similar demographic, social and cultural characteristics, are more likely to interact among them than with others [337]. Such clusterization of social contacts based on individual features is known in sociology as homophily [337]. In line with this evidence, we mainly focus the next sections on homophilic patterns of prophylactic adoption.

We first consider the case in which the population is partitioned according to the vaccination status. The mixing between vaccinated and not vaccinated individuals is shown to generate three dynamical regimes where assortative mixing can be either beneficial and/or detrimental for epidemic control. We then extend these results to the case of digital proximity tracing apps and recognize the general validity of the

uncovered phenomenology, holding for other measures such as face masks usage and physical distancing.

4.2.1 *Vaccines*

Vaccines have been crucial in humanity's struggle to protect itself from infectious diseases [338]. In the 20th century, vaccines enabled to control a series of diseases [339] as well as the eradication of smallpox [340]. Nevertheless, vaccines are not uniformly adopted in the population. While on a world-wide scale a lack of access impedes equitable adoption of vaccines [341, 342], among high-income countries vaccine hesitancy is the primary barrier [343–345].

Vaccine hesitancy widely correlates with age, socio-economic status, education level or ethnicity [346–349]. Concurrently, as already observed, these same factors shape the interaction patterns in society, leading to homophily [337]. Individuals involved in a social interaction are thus likely to share similar behaviors [337], and the vaccination status is not an exception [343, 344, 350–355]. Such non-uniform, clustered vaccine adoption strongly determines how, and whether, the virus spreads in the population. A great example of this effect is provided by the recurrent measles outbreaks in high-income countries caused by clusters of vaccine hesitant individuals [343, 344, 350–352, 354].

These recurring outbreaks sparked modeling studies that analyzed the impact of homophilic vaccine adoption on the disease dynamics [91, 356–359]. Due to the high quality of vaccines against measles, these models assumed vaccine efficacy of almost 100%, and showed that clustered adoption is always detrimental, as it leads to higher final attack rates. In contrast, vaccines as the ones against influenza or variants of concern of SARS-CoV-2 have relatively low efficacy, between 20-80% [360, 361]. In the following, through the joint exploration of imperfect immunization and vaccine uptake, we offer a wider picture in which vaccination clustering is not always detrimental.

4.2.1.1 The model

We consider a standard susceptible-infected-recovered (SIR) dynamics, with transmission rate β and recovery rate μ. The fraction of people who received a vaccine is fixed as $V \in [0,1]$ (i.e., V plays the role of ρ in Sec 4.1). Upon encounter, an infected individual transmits the infection to a vaccinee at a reduced rate $\beta(1-\varepsilon)$, where $\varepsilon \in [0,1]$ represents the vaccine efficacy. That is, we consider a leaky vaccine which provides partial immunity to the vaccinee [362]. Following Sec 4.1, we label vaccinated individuals with letter A ("adopters of the vaccine") and not vaccinated's with letter N ("not adopters of the vaccine"). Since we are dealing with a simple contagion on pairwise networks, the number of contacts of an individual is specified by the 2-degree

only, $k^{(2)} \equiv k$. We assume vaccination status and degree to be uncorrelated variables and therefore, in the mean-field approximation, $k_{\mathrm{A}} = k_{\mathrm{N}} = k$.

Under homogeneous mixing and in the absence of protected individuals ($V = 0$), the basic reproduction number of the disease is $R_0 = \beta k/\mu$. If some people are protected, it becomes $R = R_0\,(1 - \varepsilon V)$. Here and in all the following sections, we refer to R as the *effective reproduction number* (or simply as the reproduction number) to emphasize that it results from the effect of some prophylactic measure which alters the "bare" growth rate R_0. Nonetheless, is should be clear that R is still a basic reproduction number, given it is computed perturbing the disease-free, inactive state. The expression for R changes under heterogeneous mixing. To see how, we use the theory of mixing developed in Sec. 4.1 for the case when all the edges have size equal to 2. From Eqs. (4.13)-(4.14) and (4.20), the number of contacts between vaccinated and not vaccinated are parametrized as follows,

$$\kappa_{\mathrm{A|N}} = \alpha V k\,, \tag{4.33}$$

$$\kappa_{\mathrm{N|N}} = (1 - \alpha V)\,k\,, \tag{4.34}$$

$$\kappa_{\mathrm{N|A}} = \alpha(1 - V)k\,, \tag{4.35}$$

$$\kappa_{\mathrm{A|A}} = [1 - \alpha(1 - V)]\,k\,. \tag{4.36}$$

where α interpolates from complete assortativity ($\alpha = 0$) to maximal disassortativity ($\alpha = \min\{1/V, 1/(1-V)\}$). From Eq. (4.12), the degree of homophily, i.e., the probability that during a contact both individuals are either vaccinated or not, thus reads

$$h = \frac{1}{k}\left[(1 - V)\kappa_{\mathrm{N|N}} + V\kappa_{\mathrm{A|A}}\right] = 1 - 2\alpha V\,(1 - V)\,. \tag{4.37}$$

Regarding the vaccine uptake against SARS-CoV-2, we show in Fig. 4.2 the value of the mixing parameter α we indirectly estimated through age-stratified data available for various regions/countries. We leveraged the correlations in the contact patterns among age groups and the levels of vaccine coverage within each age group (see Appendix D.1.1 for details). Specifically, for each region/country considered, we combined the age-stratified contact matrices [212] with the data on vaccine adoption (counting full vaccinations only), from January 2021—when first full vaccinations appeared—to September 2021 [363–366]. We find the estimated α to stay well below 1 for almost all the time, sometimes fluctuating around it only for very low levels of vaccine coverage ($V \lesssim 0.02$ for Catalonia). It must be said, however, that it would be naïve to take the temporal trends in Fig. 4.2 as ready-to-use data. Indeed, we neglect other important features beyond age such as socio-economic classes or spatial patterns [346, 349]. Including these additional features would probably reinforce the homophilic structure and thus further decrease the mixing. The results deriving from our simple estimation must be solely understood as a qualitative indication in line with existing literature reporting homophily in health behavior [353, 367, 368]. Accordingly, we focus the analysis on homophilic adoption by varying α between 0

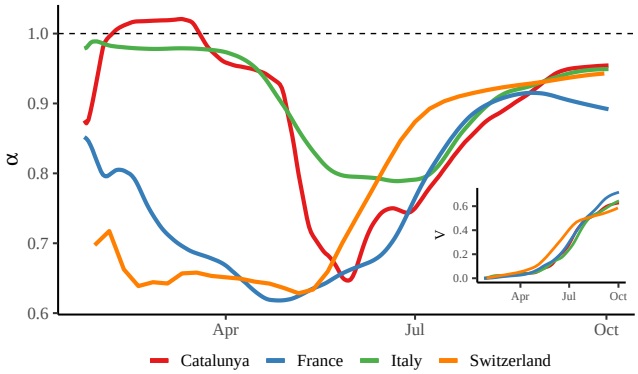


Figure 4.2: Weekly-moving average of the mixing parameter, α, as inferred for different countries/regions starting from the contact matrix among age groups and the SARS-CoV-2 vaccines uptake of each group (whose aggregate, V, is shown in the inset plot) from January to September 2021.

and 1, the latter corresponding to homogeneous mixing. Nonetheless, results under heterophilic adoption are also discussed when differing from those obtained in the homophilic case. While the results in Fig. 4.2 indicate a nontrivial evolution of mixing throughout an epidemics, studying the implications of such evolution goes beyond the scope of our work. Interpreting those implications would be hard without first understanding the effects of a constant mixing, which is our goal here.

Now, denoted with $X_Y \equiv X_Y(t)$ the fraction of vaccinated (Y = A) and not vaccinated (Y = N) people in compartment $X \in \{S, I, R\}$ at time t, the differential equations governing the dynamics read as

$$\dot{I}_N(t) = \beta \left[\kappa_{N|N} I_N(t) + \kappa_{A|N} I_A(t)\right] S_N(t) - \mu I_N(t) \,, \tag{4.38}$$

$$\dot{I}_A(t) = \beta \left[\kappa_{N|A} I_N(t) + \kappa_{A|A} I_A(t)\right] (1-\varepsilon) S_A(t) - \mu I_A(t) \,, \tag{4.39}$$

$$\dot{S}_N(t) = -\beta \left[\kappa_{N|N} I_N(t) + \kappa_{A|N} I_A(t)\right] S_N(t) \,, \tag{4.40}$$

$$\dot{S}_A(t) = -\beta \left[\kappa_{N|A} I_N(t) + \kappa_{A|A} I_A(t)\right] (1-\varepsilon) S_A(t) \,, \tag{4.41}$$

while $R_Y = 1 - S_Y - I_Y$. The final attack rate, i.e., the proportion of the population that got infected (and eventually recovered) during an entire outbreak, is given by $(1-V)R_N + VR_A$, computed once there are no infected individuals left in the population (i.e., $I_N = I_A = 0$).

4.2.1.2 Results

As the classical SIR model, this dynamical system does not allow for an explicit solution of the nontrivial, stationary state. However, we can calculate the effective reproduction number, R. The disease-free equilibrium is simply $(I_N, I_A, S_N, S_A) =$

$(0,0,1,1)$. Linearizing around it, and decomposing the Jacobian matrix $\mathbf{J}$ as in Sec. A.1.1, i.e., $\mathbf{J} = \mathbf{F} - \mathbf{V}$, we see that $\mathbf{V} = \mu\mathbb{1}$. The next-generation matrix (NGM) is therefore readily found to equal

$$\mathrm{NGM} = \frac{\beta}{\mu}\begin{pmatrix} \kappa_{\mathrm{N|N}} & \kappa_{\mathrm{A|N}} \\ (1-\varepsilon)\kappa_{\mathrm{N|A}} & (1-\varepsilon)\kappa_{\mathrm{A|A}} \end{pmatrix} . \tag{4.42}$$

Computing $\Lambda_{\max}(\mathrm{NGM})$ and inserting the expressions for the number of contacts, one gets

$$R = \frac{R_0}{2}\left[2 - \alpha + \alpha\varepsilon(1-V) - \varepsilon + \sqrt{[\alpha - \alpha\varepsilon(1-V) + \varepsilon]^2 - 4\alpha\varepsilon V}\right] . \tag{4.43}$$

Straightforward calculations prove that R/R_0 decreases monotonically with respect to all the figuring parameters. In particular, as α is heightened (i.e., homophily is lowered), the contacts among not vaccinated people decrease, leading to a smaller reproduction number. Indeed, the drivers of an outbreak are the not vaccinated individuals. Beside being intuitive, this can be checked by looking at the Jacobian matrix. Under zero mixing ($\alpha = 0$), the two sub-populations are isolated from each other (NGM is diagonal), hence the epidemics propagates with a reproduction number R_0 in the not vaccinated cluster and $R_0(1-\varepsilon) < R_0$ in the vaccinated one. For instance, $R_0 > 1$ implies that an outbreak develops in the former, but not in the latter if $\varepsilon > 1 - 1/R_0$. Increasing the mixing, part of the infections in each sub-population are now generated by individuals in the other sub-population, implying that I_{N} and I_{A} grow, respectively, slower and faster than before. In any case, since vaccinees get infected with a lower probability, $\dot{I}_{\mathrm{N}} > \dot{I}_{\mathrm{A}}$ for any mixing. Therefore, reducing R comes down to lower the growth rate associated to the not vaccinated sub-population, as increasing mixing does by reducing the contacts within the latter. As shown in Sec. 4.2.2, this monotonic decrease with α is generally lost for digital proximity tracing, as R is minimized at an intermediate value of assortativity.

As anticipated, Eq. (4.43) reduces to $R = R_0\,(1-\varepsilon V)$ for homogeneous mixing. Vaccine efficacy, ε, and coverage, V, figure only as a product in this expression. We can thus exchange them without affecting the value of R. Interestingly, this symmetry is broken under heterogenous adoption. Indeed, Eq. (4.43) contains asymmetric terms with respect to ε and V. To see how these two affect R, we vary them while keeping fixed their product, εV, which represents the probability that a randomly chosen individual is vaccinated *and* is protected during a contact. Once conveniently rewritten Eq. (4.43) to isolate the terms not proportional to the product εV, and differentiated it with respect to ε while keeping εV fixed, we get

$$\left.\frac{\partial R}{\partial \varepsilon}\right|_{\varepsilon V} = \frac{R_0}{2}(1-\alpha)\left[-1 + \frac{\varepsilon + \alpha(1+\varepsilon V - \varepsilon)}{\sqrt{[\varepsilon(1-\alpha)]^2 + \varepsilon\,[2\alpha(1-\alpha)(1+\varepsilon V)] + c(\varepsilon V)}}\right] , \tag{4.44}$$

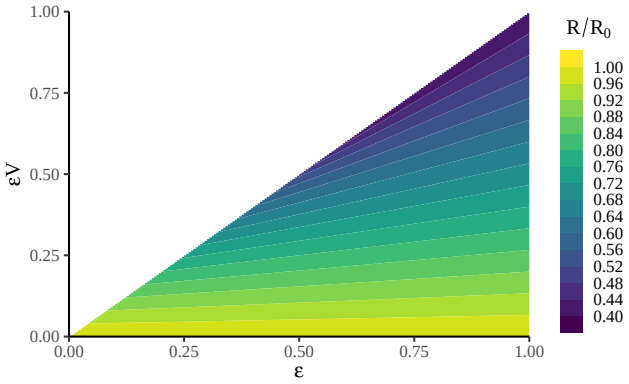


Figure 4.3: Normalized reproduction number, R/R_0, as a function of the vaccine efficacy, ε, and of εV, the product between vaccine efficacy and vaccine coverage, V. The normalized reproduction number is computed from Eq. (4.43) for $\alpha = 0.6$. For any given value of εV (a horizontal slice of the heatmap), ε can vary from that value to 1 (respectively, V varies from 1 to that value), for $\varepsilon V \leqslant \varepsilon$ (respectively, $\varepsilon V \leqslant V$).

where $c(\varepsilon V)$ collects terms which are functions of εV, hence constants. $\partial R/\partial\varepsilon|_{\varepsilon V}$ is zero for $\alpha = 1$, as we already know from $R = R_0\,(1-\varepsilon V)$. With a bit of algebra, one finds that the term in square brackets is always positive. Therefore, the sign of $\partial R/\partial\varepsilon|_{\varepsilon V}$ is simply determined by the sign of $1-\alpha$: $\partial R/\partial\varepsilon|_{\varepsilon V}$ is positive for $\alpha < 1$ and negative for $\alpha > 1$, no matter the value of εV. Said differently, under homophilic adoption ($\alpha < 1$), a population with less effective vaccines but with higher coverage is less vulnerable to develop an extensive outbreak than one with more effective vaccines but a lower coverage—given εV is the same; vice versa if adoption is heterophilic ($\alpha > 1$). Figure 4.3 illustrates the results obtained for $\alpha = 0.6$, yielding $\partial R/\partial\varepsilon|_{\varepsilon V} > 0$.

By imposing $R = 1$ and solving for α, we find the critical value of mixing, α_c, above which the disease cannot thrive, to be

$$\alpha_c = \left(1 - \frac{1}{R_0}\right) \frac{1 - R_0(1-\varepsilon)}{1 - R_0(1-\varepsilon) - \varepsilon(1-V)}\,, \tag{4.45}$$

given that $\alpha_c \geqslant [2(1-1/R_0)-\varepsilon]\,/\,[1-\varepsilon(1-V)]$ is satisfied. In particular, the condition is always met for $\varepsilon = 1$ and never met for $\varepsilon = 0$, meaning eradication is not possible in this case—unless, trivially, $R_0 < 1$. Notice that the larger is R_0, the stricter is the condition, up to the point at which it can no longer be satisfied. Imposing $\alpha_c < 1$, we find that eradication is possible in the homophilic region ($\alpha < 1$) if $\varepsilon V > 1 - 1/R_0$, otherwise some heterophily ($\alpha > 1$) is required.

Analyzed the early-stage dynamics, we now look at the final attack rate. This exhibits three dynamical regimes with respect to the mixing parameter, α. This is shown in Fig. 4.4(a), where, by increasing the basic reproduction number, R_0, the dependence of the final attack rate on α is first monotonically decreasing, then concave, and fi-

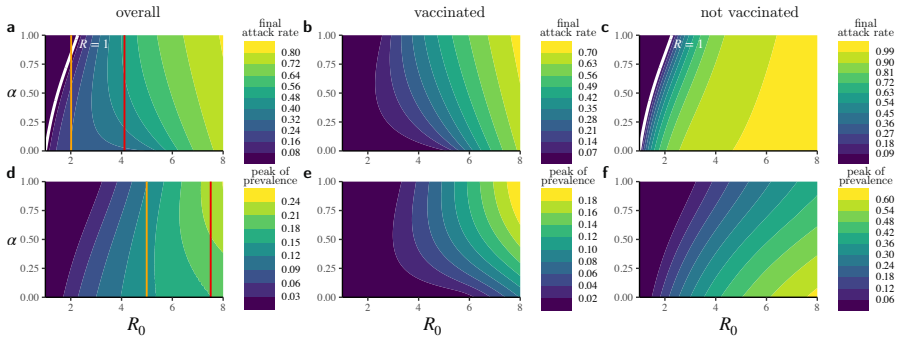


Figure 4.4: Final attack rate (a-c) and peak of prevalence (d-f) as functions of the basic reproduction number, R_0, and the mixing parameter, α, as resulting from the numerical iteration of the differential equations governing the system dynamics (Eqs. (4.38)-(4.41)). Here, the vaccine coverage is $V = 0.7$ and the vaccine efficacy $\varepsilon = 0.8$. In order to highlight the competing processes at the base of the dynamics we show, besides the results for the population overall (a, d), those for the vaccinated individuals (b, e) and for the not vaccinated ones (c, f), separately. The white solid line indicates the critical curve, $\alpha_c \equiv \alpha_c(R_0)$, at which $R = 1$, as computed from Eq. (4.45). The orange and red vertical lines, demarcate the boundaries between, respectively, the critical and the intermediate regime, and the intermediate and the saturated one, for both the final attack rate (a) and the peak of prevalence (d).

nally monotonically increasing. We refer to these three regimes as *critical*, *intermediate* and *saturated*, respectively. The former always includes part or all of the critical curve defined by $R = 1$, hence the name. The saturated regime, instead, is "saturate" of infections and does not admit eradication.

These regimes result from the competition between the dynamics associated to each sub-population, as illustrated in Figs. 4.4(b) and (c). As already observed, raising the mixing starting from zero ($\alpha = 0$), each sub-population is increasingly "diluted" with individuals of the other type. Consequently, the reciprocal protection that vaccinated people provide among them to cut infection chains is partially lost. Instead, not vaccinated individuals profit from vaccinated individuals in their vicinity and are thus subject to a lowered infection probability. Then, depending on how high the basic reproduction number, R_0, the vaccine coverage, V, and the vaccine efficacy, ε, are, the contribution from one subpopulation dominates the contribution from the other, determining in which of the three dynamical regimes the system falls overall. Figures 4.4(b) and (c), in particular, explain how the regimes are accessed by varying R_0. In (proximity of) the critical regime, the attack rate for not vaccinated individuals is very sensible to variations in mixing (i.e., α), whereas the attack rate for vaccinated individuals is much less affected, with the result that the overall attack rate decreases. This originates from the nonlinear, self-reinforcing character of the transmission dynamics. Increasing α, while reduces the contacts among not vac-

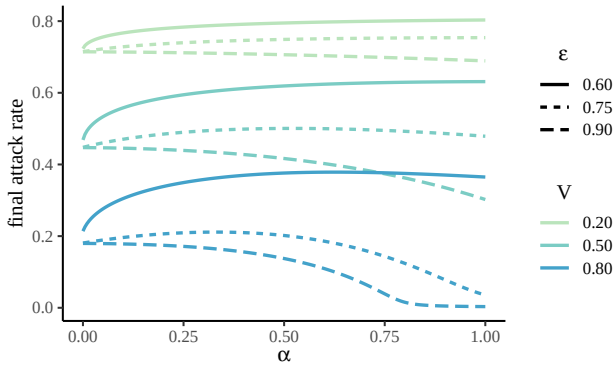


Figure 4.5: Final attack rate as a function of the mixing parameter, α, for different combinations of vaccine efficacy, ε, and coverage, V, given a basic reproduction number $R_0 = 2.5$. As a complement to Fig. 4.4, this plot illustrates how all three dynamical regimes can be also explored by varying ε and/or V.

cinated individuals linearly, it disproportionally decreases the infections among not vaccinated's—eventually benefiting vaccinated's too—, thanks to the indirect protection offered by the increased mixing with vaccinees. This holds when the system is close enough to the epidemic threshold. In the saturated regime, instead, the situation is reversed, in that mixing significantly affects the attack rate within the vaccinated subpopulation but not within the not vaccinated's. Indeed, far from the epidemic threshold, the disease is too contagious for the vaccinated individuals to provide good protection to the not vaccinated's, which experience high attack rates, whatever the mixing is. Consequently, mixing becomes strongly detrimental for vaccinees, as their attack rate rises driven by the increased contact with the highly infectious, not vaccinated subpopulation.

As Figs. 4.4(d)-(f) report, the same qualitative picture holds for the peak of prevalence, i.e., the maximum number of simultaneously infectious people during an epidemic. Note that, for each set of fixed parameters, the respective regimes of the peak of prevalence and of the final attack rate may not coincide. For example, at $R_0 = 3$, the final attack rate is in the intermediate regime while the peak of prevalence is in the critical one (compare Figs. 4.4(a) and (d)).

In Fig. 4.5, we show how the three regimes can be accessed by varying the vaccine efficacy, ε, as raising it makes the dynamics shift from the saturated to the critical regime, going through the intermediate one. Accordingly, a sufficiently high efficacy makes homogeneous mixing always beneficial, in agreement with previous studies assuming perfect vaccines ($\varepsilon = 1$) [91, 356–359]. In such case the system is always in the critical regime. Indeed, we see from Eqs. (4.38)-(4.41) that the dynamics reduces to a standard SIR model within the not vaccinated sub-population (S_N, I_N). This dynamics is solely driven by the contact rate $\kappa_{N|N}$, which is a linearly decreasing function of α. Figure 4.5 also shows the effect of varying the vaccine coverage, V. This

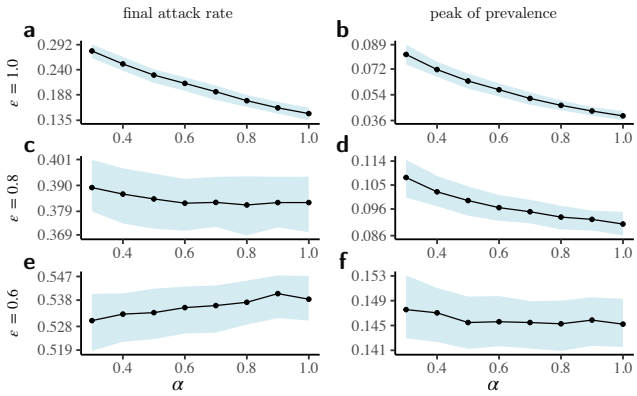


Figure 4.6: Final attack rate and peak of prevalence, respectively, for vaccine efficacy, $\varepsilon = 1.0$ (a, b), $\varepsilon = 0.8$ (c, d), and $\varepsilon = 0.6$ (e, f) as functions of the mixing parameter, α, resulting from the numerical simulations performed on top of a real-world temporal proximity network from the Copenhagen Networks Study [369] (see text for details). Dots indicate the median value, whereas the shaded area indicates first and third quartiles. Each point is obtained by averaging over 2×10^4 runs. We fixed the vaccine coverage to $V = 0.5$, the recovery rate to $\mu = 4.6 \times 10^{-4}$ (corresponding to a mean infection time of 7.5 days) and the transmission probability to $\beta = 1.142 \times 10^{-1}$, yielding $R_0 = 6$ from the estimation $R_0 = \beta(\overline{s^2}/\bar{s})/\mu$, where $\bar{s}$ and $\overline{s^2}$ are the first and second raw moment of the number of contacts (i.e., strength) per timestamp.

is not as impactful as varying efficacy and makes the system change regime only when efficacy is neither too low or too high, otherwise the system persists in the saturated or critical regime, respectively.

To corroborate the predictions made by our mean-field model, we performed Monte Carlo simulations upon a temporal proximity network estimated via Bluetooth signal exchanges in the Copenhagen Networks Study [369]. We retained from the dataset those interactions with an associated Received Signal Strength Indication (RSSI) not lower than -74 dBm, corresponding to physical distances approximately up to 2 meters [370]. In this way, we filter out those interactions which are likely to either record proximity but not face-to-face contact or be marginal for transmission. The resulting temporal network involves $N = 672$ individuals and 374884 pairwise interactions spread over 8064 timestamps, binning four weeks of recording time into five-minute intervals. In order to study different levels of assortativity in vaccine adoption, we distribute vaccines algorithmically on static aggregated network until a preset value of α is reached. The algorithm guarantees that strength (or degree) and type of nodes remain uncorrelated (see Appendix D.1.2 for details about the algorithm), allowing us to isolate the dynamical effect of mixing, consistently with what assumed in the mean-field model by taking $k_N = k_A$.

The results for the final attack rate and the peak of prevalence, respectively reported in Figs. 4.6(a) and (b) for $\epsilon = 1.0$, (c) and (d) for $\epsilon = 0.8$, and (e) and (f) for $\epsilon = 0.6$, confirm the existence of the three dynamical regimes identified by our model. Accordingly, the beneficial effect of homogeneous mixing is lost when the vaccine does not provide sufficient protection. In the saturated regime, mixing is slightly detrimental. Notice that the structural constraints of the aggregated network do not allow—our algorithm, at least—to reach arbitrarily small values of α, making the saturated regime less evident. Lastly, as in the mean-field case, we observe that final attack rate and peak of prevalence can be in different dynamical regimes (compare panels (e) and (f)).

4.2.1.3 Discussion

Depending on the need and the available information, the proposed model can be applied to different specific populations and at different scales. Currently, many epidemic models consider the mixing between not vaccinated and vaccinated individuals only implicitly, through stratification according to age (as for the results shown in Fig. 4.2) [371, 372] or socio-economic status [373]. Our findings suggest that subgroup-specific levels of mixing may have a notable impact on the spreading, thus calling for more refined empirical analysis of the vaccine uptake. Based on this, more realistic models would enable a better interpretation of epidemiological data and could act as useful guides for policy makers to design more sophisticated and precise interventions.

For instance, during the COVID-19 pandemic, various countries at some point required a 'green pass' (a proof of vaccination) to enter restaurants and workplaces. The intention behind green passes is to nudge individuals to get vaccinated, while reducing the number of physical contacts of not vaccinated individuals [374]. Consequently, green passes also reshape social contacts [375], effectively reducing the mixing rate between vaccinated and not vaccinated individuals. In light of our results, due to the associated reduction in the mixing rate, models that do not include homophily in vaccine adoption may actually overestimate the impact of green passes and measures alike, especially for low reproduction numbers. Overall, our results show that more detailed information on the correlations between social interactions and health behaviour would lead to a more comprehensive analysis.

While we considered a leaky vaccine—providing immunity with probability ε upon contact—in a mean-field model and in a real-world temporal proximity network, Hiraoka et al. [376] considered an all-or-nothing vaccine [362]—where ε is the fraction of vaccinated people acquiring perfect immunity, the remaining ones gaining no immunity at all—in random networks with different structural properties. The results they found are qualitatively equivalent to ours, which demonstrates the robustness of this phenomenology across interaction structures and its origin in an imperfect immunization. Additionally, Fefferman et al. [377, 378] showed that the presence of

homophily adds interesting phenomenology regarding the temporal trajectory of the epidemics beyond the three dynamical regimes discussed here.

4.2.2 *Digital proximity tracing*

One of the techniques employed to fight epidemics is contact tracing [379], which consists in identifying and eventually testing (and isolating) the contacts of infected individuals, in order to reconstruct and stop infection chains. In the past, contact tracing has been used against smallpox [380], tuberculosis [381], HIV [382] and Ebola [383], to name a few. The tracing is traditionally performed by public health agency officers who interview individuals who tested positive. Each individual names their potential infectious contacts and these contacts are identified and tested. The procedure can reconstruct transmissions either forward or backward in the infection chains, by asking to name the contacts which the infection potentially went to or came from, respectively. Great efforts have been devoted to analyze the efficacy of this kind of tracing [380, 384–392], called manual contact tracing. Lately, because of the COVID-19 pandemic, most of us became familiar with digital proximity tracing (DPT)—also known as digital contact tracing—, in particular with DPT apps. These leverage device-to-device communication (usually via Bluetooth Low Energy signals) to detect physical proximity between individuals, so that a notification of close contact is sent to all the app users that where detected to be close, within a certain time window in the past, to an user confirmed to be positive. The notified users are required to immediately undergo a test and, if necessary, isolate themselves. DPT ideally overcomes fundamental limitations of manual tracing. The latter is indeed labour-intensive and relies on the ability of people to remember and identify their contacts. In fact, in public contexts such as transportation or shops, we are very often in close contact with strangers that we are unable to identify. DPT is instead potentially scalable and automatic, and may therefore allow for a more prompt and precise tracing. The large-scale implementation of DPT apps to prevent the spread of SARS-COV-2 [393–395], made possible by the deep and wide penetration of smartphones, sparked numerous studies that analyzed how this novel technology may reduce the disease propagation [368, 396–403]. A variety of more theoretical studies also unveiled the physics behind the impact of DPT [404–407].

A question which has not been addressed is how the efficacy of DPT apps depends on how its adoption is distributed within a population. As for vaccines and other prophylactic tools, empirical studies suggest that the voluntary adoption of DPT apps is strongly variable and clustered. It was shown to correlate with age, income and nationality [368, 408, 409], being particularly low among socio-economically marginalized classes [408]. Accordingly, due to homophily [337], app adoption is much more probable for contacts of an app user (adopter) compared to a random individual. As a matter of fact, in Switzerland around 70% of the contacts of users were found to be

users too, while average, national adoption was only around 20% [394]. This discrepancy is in line with experimental studies that indicate how homophily importantly affects health behavior [180, 367]. Our goal is thus to characterize through a minimal model how the homophilic adoption of DPT apps affects the propagation dynamics.

4.2.2.1 The model

We build upon a model recently introduced by Bianconi et al. [405]. The model is simple enough to be analytically tractable, but retains the essential ingredients of the dynamics. It considers a discrete-time susceptible-infected-recovered (SIR) dynamics and treats it as a link percolation problem where the link-activation probability maps to the transmission probability, here denoted by β. We fix the fraction of app adopters as $T \in [0,1]$, playing—as V in Sec. 4.2.1—the role of ρ in Sec. 4.1. We assume there is no correlation between adoption status and number of contacts, therefore $k_{\mathrm{A}} = k_{\mathrm{N}} = k$, where letter A labels adopters and letter N not adopters of the app. The mixing between adopters and not adopters is again parametrized via the mixing parameter α as in Eqs. (4.33)-(4.36). Accordingly, the adoption homophily, h, reads $h = 1 - 2\alpha T\,(1-T)$.

As in Bianconi et. al [405], the model assumes that app users infected by other app users do not further transmit the disease. This corresponds to the ideal situation which assumes that the users are immediately notified and, upon notification, immediately self-isolate, thus avoid producing new infections. Nonetheless, as proved in Appendix D.1.3, the phenomenology is qualitatively unchanged when these assumptions are relaxed. In any case, we need additional variables to follow the dynamical evolution in comparison to the standard SIR model. Namely, we define $I_{\mathrm{AA}}(t)$ and $I_{\mathrm{AN}}(t)$ to be the fraction of infected individuals among adopters at time t which were infected, respectively, by adopters and not adopters. In the ideal scenario, individuals in compartment I_{AN} can transmit, those in I_{AA} do not. For not adopters such distinction is not necessary, therefore $I_{\mathrm{N}}(t)$ simply refers to the fraction of them infected at time t. Lastly, $S_{\mathrm{A}}(t)$ ($R_{\mathrm{A}}(t)$) and $S_{\mathrm{N}}(t)$ ($R_{\mathrm{N}}(t)$), respectively denote the fraction of susceptible (recovered) individuals among adopters and not adopters at time t. Defined these compartments, their dynamics is governed by the following discrete-time equations,

$$I_{\mathrm{N}}(t+1) = I_{\mathrm{N}}(t)(1-\mu) + \beta\left[\kappa_{\mathrm{N|N}} I_{\mathrm{N}}(t) + \kappa_{\mathrm{A|N}} I_{\mathrm{AN}}(t)\right] S_{\mathrm{N}}(t)\,, \tag{4.46}$$

$$I_{\mathrm{AN}}(t+1) = I_{\mathrm{AN}}(t)(1-\mu) + \beta\kappa_{\mathrm{N|A}} I_{\mathrm{N}}(t) S_{\mathrm{A}}(t)\,, \tag{4.47}$$

$$I_{\mathrm{AA}}(t+1) = I_{\mathrm{AA}}(t)(1-\mu) + \beta\kappa_{\mathrm{A|A}} I_{\mathrm{AN}}(t) S_{\mathrm{A}}(t)\,, \tag{4.48}$$

$$S_{\mathrm{N}}(t+1) = S_{\mathrm{N}}(t) - \beta\left[\kappa_{\mathrm{N|N}} I_{\mathrm{N}}(t) + \kappa_{\mathrm{A|N}} I_{\mathrm{AN}}(t)\right] S_{\mathrm{N}}(t)\,, \tag{4.49}$$

$$S_{\mathrm{A}}(t+1) = S_{\mathrm{A}}(t) - \beta\left[\kappa_{\mathrm{N|A}} I_{\mathrm{N}}(t) + \kappa_{\mathrm{A|A}} I_{\mathrm{AN}}(t)\right] S_{\mathrm{A}}(t)\,, \tag{4.50}$$

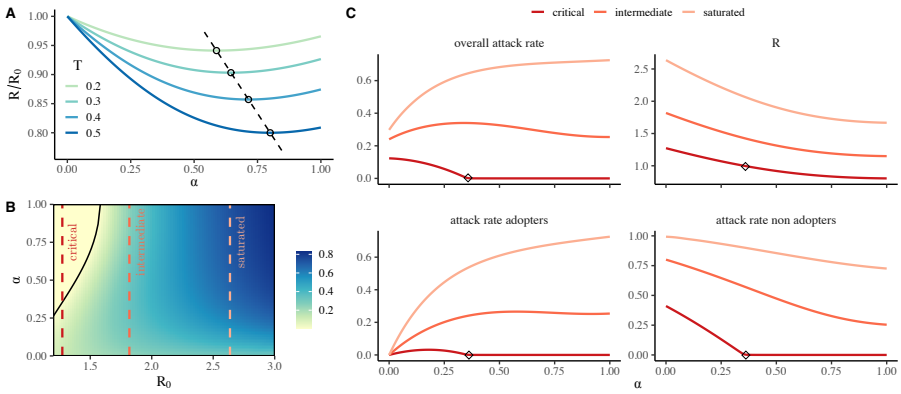


Figure 4.7: General phenomenology from the homophilic adoption of DPT apps. (a) Reproduction number, R, normalized with respect to the basic reproduction number, R_0, for different values of adoption, T, below T^*. Dots indicate the minimum at α^*, while the dashed line shows its variation for $T \in [0.1, 0.55]$. (b) Numerical solution of the dynamics for the attack rate as a function of α and R_0. The solid line indicates the threshold α_c^- for which $R = 1$. Colored, dashed lines denote the dynamical regimes: critical, intermediate and saturated. Adoption was fixed as $T = 0.7$. For these parameter values we have $\alpha_c^+ > 1$. (c) Top panels show the attack rate and the reproduction number for the different regimes defined in B. The specific attack rates for adopters and not adopters are reported in the bottom panels. Black diamonds indicate α_c^-, at which $R = 1$.

while $R_N(t) = 1 - I_N(t) - S_N(t)$ and $R_A(t) = 1 - I_{AN}(t) - I_{AA}(t) - S_A(t)$. The final attack rate is computed as $(1-T)R_N + TR_A$, once no more infected individuals are left in the population (i.e., $I_N = I_{AN} = I_{AA} = 0$).

4.2.2.2 Results

We start by linearizing Eqs. (4.46)-(4.50) around the disease-free equilibrium, $(I_N, I_{AN}, I_{AA}, S_N, S_A) = (0, 0, 0, 1, 1)$, to compute the effective reproduction number, R. The associated NGM matrix reads,

$$\text{NGM} = \frac{\beta}{\mu} \begin{pmatrix} \kappa_{N|N} & \kappa_{A|N} & 0 \\ \kappa_{N|A} & 0 & 0 \\ 0 & \kappa_{A|A} & 0 \end{pmatrix} . \tag{4.51}$$

Inserting the explicit expressions of the numbers of contacts in $\Lambda_{\max}(\text{NGM})$, we find

$$R = \frac{R_0}{2} \left[1 - \alpha T + \sqrt{(1-\alpha T)^2 + 4\alpha^2 T(1-T)} \right] . \tag{4.52}$$

As expected, R/R_0 decreases with the adoption level, T. Surprisingly, instead, it exhibits a non-monotonic dependence on the mixing parameter, α. Precisely, solving $\frac{dR}{d\alpha} = 0$ with respect to α yields

$$\alpha^* = 1 - \frac{\frac{2}{3} - T}{\frac{4}{3} - T} \,. \tag{4.53}$$

Straightforward calculations show that $\frac{dR}{d\alpha}|_{\alpha=0} < 0$ is always met, thus R is minimized at α^*. Moreover, α^* increases with T and thus, in the limit $T \to 0$, approaches its smallest value, 0.5. Imposing $\alpha^* < 1$, we find that the it lies in the homophilic region for T smaller than $T^* = 2/3$. Otherwise, the minimum either is in the heterophilic region ($\alpha^* > 1$) or is inaccessible. It is easy to prove that α^* becomes inaccessible when $T > 4/5$, for it becomes larger than its maximum allowed value, i.e., $\min\{1/T, 1/(1-T)\} = 1/T = 1.25$. Therefore, the reproduction number, R, presents a minimum (at $\alpha^* \in [0.5, 1.25]$) if $T < 4/5$, otherwise it just decreases with α. Fig. 4.7(a) shows R as a function of α for adoption $T < T^*$. The dependence of α^* on T is also indicated.

Interestingly, the non-monotonic dependence of R on the mixing parameter, α, distinguishes digital contact tracing from classical immunization strategies such as vaccination [91], where R always decreases with α, as seen in Sec. 4.2.1. In fact, differently from those individual-based strategies—that can be mapped to a site percolation problem—, DPT apps require both individuals in contact to be adopters in order to avoid further transmission—a sort of edge percolation based on end-nodes' features [405]. Under either full assortativity ($\alpha = 0$) or full disassortativity ($\alpha = \min\{1/T, 1/(1-T)\}$), R attains indeed a local maximum. In the first case, not adopters are disconnected from adopters and therefore $R = R_0$ for the sub-population of not adopters, hence for the whole population, as seen from Eq. (4.52). In the second case, adopters and not adopters tend to alternate each other along a contact chain, making DPT apps useless. Specifically, under full disassortativity, we find that $R = R_0$ for $T \leqslant 0.5$ ($\alpha = 1/(1-T)$), and $R = R_0\sqrt{1/T - 1}$ for $T > 0.5$ ($\alpha = 1/T$), which, although being smaller than R_0 given that not adopters are too few to completely avoid user-user contacts, it is still a local maximum. Then, the mixing α^* minimizes R by optimally satisfying the two opposite needs of, on one hand, increasing user-nonuser contacts to provide indirect protection to not adopters, and, on the other hand, increasing user-user contacts to let the tracing mechanism work.

If α^* takes an admissible value and $R(\alpha^*) < 1$, then there exists a range for α which eradicates the disease. Imposing $R < 1$, such range is given by $\alpha \in (\alpha_c^-, \alpha_c^+)$, where

$$\alpha_c^{\pm} = \frac{1}{2R_0(1-T)} \left[1 \pm \sqrt{1 - 4\frac{1-T}{T}(R_0 - 1)} \right] \,. \tag{4.54}$$

The existence of two physical solutions, $\alpha_c^{\pm}$, implies that increasing mixing can not only push the system below the epidemic threshold and thus eradicate the disease, but can also push it above the threshold and cause its resurgence. Imposing $\alpha_c^- < \alpha_c^+$

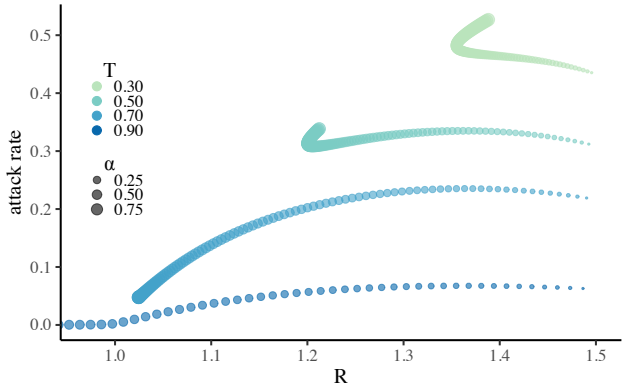


Figure 4.8: Attack rate versus reproduction number for different values of adoption, T, and mixing parameter, α. The size of the points encodes homophilic mixing, $\alpha \in [0,1]$. The basic reproduction number is fixed as $R_0 = 1.5$, which yields $T_{\min} = 2/3$ as the necessary adoption needed for eradication. The minimum for R is observed for both $T = 0.3$ and $T = 0.5$, being $T < T^*$. There is also a minimum for R when $T = 0.7$, but it is not displayed because occurring under heterophilic mixing, namely, at $\alpha \approx 1.05$. Additionally, being $T_{\min} < 0.7$, $R < 1$ around such minimum.

yields the minimum adoption, $T_{\min} = 4(R_0-1)/(1+4(R_0-1))$, needed for the range (α_c^-, α_c^+) to be finite and thus for eradication to be possible. In conclusion, if $T > T_{\min}$, then the disease is eradicated for $\alpha_c^- < \alpha < \min\{\alpha_c^+, 1/T\}$ (notice $\alpha_c^- < 1 \leqslant 1/T$), where $\min\{\alpha_c^+, 1/T\}$ equals α_c^+ for $T < {R_0}^2/(1+{R_0}^2)$ and $1/T$ otherwise. In the latter case, resurgence is impossible, for α_c^+ is inaccessible.

Following with the analysis, we see that varying the mixing between adopters and not adopters generates a very similar phenomenology as the one we already found in Sec. 4.2.1 for vaccines. Namely, the existence of three dynamical regimes for the final attack rate: the critical regime, where increasing mixing (α) lowers the attack rate; the intermediate regime, where the attack rate depends non-monotonically on mixing; and finally the saturated regime, where increasing mixing raises the attack rate. Figures 4.7(b) and (c), in particular, show how the regimes can be accessed by varying the basic reproduction number, R_0. The regimes stem from the competition between the dynamics associated to each subpopulation. On the one hand, as mixing increases, adopters provide protection to not adopters, so the attack rate among the latter decreases (bottom-right panel in Fig. 4.7(c)). On the other hand, protection vanishes inside the adoption cluster, making the attack rate raise among adopters (bottom-left panel in Fig. 4.7(c)). Which of the two processes holds the upper hand—and so determines the overall regime—depends on the infectivity, i.e., R_0, as explained by the exact same logic used in Sec. 4.2.1 for vaccines.

The complete picture is nicely resumed by putting together the final attack rate and the effective reproduction number. We do this in Fig. 4.8, where the mixing α is

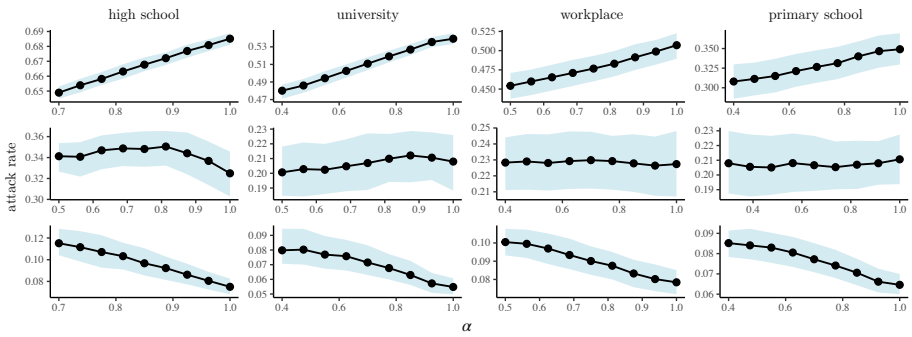


Figure 4.9: Dependence of the attack rate on α as resulting from Monte Carlo simulations performed in the four networks we generated from real-world proximity data. Each column refers to a different network, while rows indicate critical, intermediate and saturated regime respectively, from top to bottom. Note that different values of α can be accessed depending on adoption T and structural constraints of the network. Dots indicate the median value, whereas the ribbon indicates first and third quartiles. Each point is obtained by averaging over 4×10^4 runs. From left to right, the networks have 780, 636, 212 and 226 nodes. For the saturated and critical regimes we fixed $T = 0.30$ and $T = 0.90$, respectively. In the intermediate regime, T was fixed as 0.75, 0.72, 0.67 and 0.69 for the different networks (left to right). Specific to each network we set β as 0.030, 0.085, 0.110 and 0.090.

varied while considering different adoption levels T. Notably, we observe that, due to their non-monotonic dependence on mixing, attack rate and reproduction number do not always change in the same direction, in that one may increase while the other decreases. This also implies that we can observe two or three different values of the attack rate for the same value of the reproduction number, breaking the one-to-one, proportional relation the two have in a standard SIR model. Finally, Figure 4.8 illustrates how the systems moves from the saturated regime towards the critical one as adoption is increased. In fact, for low app coverage, homophilic adoption is preferable overall, for random mixing does not provide sufficient protection—neither to not adopters nor to adopters. In contrast, for high coverage, a more distributed presence of adopters effectively acts as a firewall for not adopters and enables to better protect the population, immunizing it eventually.

We corroborate our theory through Monte Carlo simulations on networks generated from real-world proximity [369, 410] and face-to-face contact [411, 412] data, collected in a diverse set of contexts. These are temporally-resolved data that we first aggregate into static, weighted networks. The latter are extremely dense, so we partially sparsify them by removing edges with a weight smaller than a threshold, ending up with the unweighted networks upon which the spreading is simulated. For each network, the threshold is chosen as approximately the highest value for which the network is still connected. To distribute the DPT apps, we use the same algorithm used

in Sec. 4.2.1 (see Appendix D.1.2 for details), which allows us to control the value of α. Figure 4.9 reports the results of the Monte Carlo simulations we performed. The three dynamical regimes of the epidemics, i.e., saturated, intermediate and critical, are observed for all the four networks, confirming our predictions. For the smallest networks—workplace and primary school—, the little dependence on α found for the intermediate regime in the analytical model, is likely to be hidden by the stochastic fluctuations. Nevertheless, the three regimes are easily recognized.

4.2.2.3 Discussion

The disease propagation fundamentally depends on how adoption of DPT apps is distributed in a population. As for vaccines (Sec 4.2.1), we proved the existence of three different dynamical regimes, originating from a nontrivial dependence of the transmission dynamics on the mixing between adopters and not adopters. Moreover, we showed that the reproduction number is in this case a convex function of the mixing, implying the existence of an optimal mixing that, if accessible, minimizes the reproduction number. Accordingly, increasing the mixing can, not only lead to eradication by pushing the system below the epidemic threshold, but also cause the reemergence of the disease. Interestingly, the different regimes in the attack rate can arise even when the minimum for the reproduction number is not accessible.

We verified the robustness of the results against the introduction of imperfections in notification and isolation (see Appendix D.1.3). In this sense, the key ingredient for the physics of DPT is the necessity for both individuals in contact to be adopters in order for DPT to work and thus potentially avoid further transmission. Consistently with what found for vaccines, where the observation of the three dynamical regimes is conditional on the vaccine being imperfect, the phenomenology for DPT apps is emphasized by the imperfections of the tracing process. Importantly, as shown in Appendix D.1.4, the phenomenology remains unchanged even inserting additional heterogeneities in the population as long as they do not correlate with app adoption. One of these is degree heterogeneity. Accordingly, the dynamical regimes in the case of vaccine adoption has been found by Hiraoka et al. [376] across different connectivity structures. At the cost of model tractability, the framework presented here may be extended to analyze the role of possible correlations with the adoption of other prophylactic measures, such as physical distancing and face masks. In particular, if the aim was to empirically quantify the impact of homophily, such extensions as well as explicitly accounting for the temporal nature of social contacts, would be necessary.

Switching our focus to the real world, adoption of DPT apps is generally very low, between 20% and 40% [394]. Accordingly, it is very improbable to control the epidemic with such low adoption. In this sense, our results indicate that homophilic adoption is beneficial to this point. However, if health authorities desire to actually contain the spread of pathogens like SARS-COV-2 through large scale adoption, i.e., to reach the

critical regime, overcoming homophily in health behavior—in particular, with respect to different socio-economic classes—is indispensable.

The phenomenology presented here, besides vaccines, has been extended to other prophylactic measures such as the use of face masks [413] and physical distancing [414]. In fact, the existence of the different dynamical regimes can be understood within a common underlying mechanism, enclosing a broad class of prophylactic measures. The essential ingredient is the lack of (sufficient) individual, independent protection offered by the prophylactic tool to the adopter. Accordingly, as seen for vaccines, only the critical regime is observed when individual protection is complete. Individual protection is absent for DPT apps by construction, as adopting them does not in itself protect the adopter; adoption from both individuals in contact is required for tracing to work. Similarly, if the vaccine is imperfect, isolated adoption may not substantially reduce the individual infection probability. More generally, any prophylactic tool that reduces the transmission probability without zeroing it out, cannot work optimally under isolated adoption; instead, shared adoption between contacts is needed. Indeed, if the adoption coverage is low in comparison to the epidemic pressure, only an assortative/clustered adoption can provide mutual protection, in which case not adopters cannot be protected (saturated regime). On the contrary, if coverage is high, adopters are anyway close to each other in the contact network and thus generate strong mutual protection. Consequently, by increasingly mixing adopters and not adopters, also the latter can be protected through a sort of herd immunity, and mixing becomes beneficial (critical regime). In between, we find the intermediate regime, where an intermediate mixing neither sufficiently protects adopters nor provides enough indirect protection to not adopters, leading a disease to spread maximally overall.

4.2.3 *Summary*

In the preceding sections, we demonstrated the nontrivial impact of homophily—the tendency of humans to interact with those who share similar characteristics and behaviors—on the effectiveness of prophylactic tools. At odd with previous findings, our results revealed that when efficacy is less than 100%, homophily may not always be detrimental; it can alternatively prove advantageous or exhibit a non-monotonic influence on epidemic control. This observed phenomenology is applicable across a diverse spectrum of prophylactic measures, including mask usage and physical distancing, in addition to vaccines and DPT.

While our models were conceived specifically for epidemic spreading, their minimalistic character strongly suggests that the phenomenology found here may be relevant for contagion processes other than biological ones. We just assumed a SIR dynamics and implemented the behavioral effect on spreading via a mere rescaling of the transmission probability. Moreover, analogous results can be shown to hold

for a SIS dynamics too (in terms of the equilibrium prevalence). If we were to model some social contagion, we would then expect to observe a similar phenomenology as a consequence of the assortative distribution of socio-cultural traits in the population. In fact, opinions, ideas, innovations and behaviors, show generally heterogeneous degrees of penetration as a function of those traits.

4.3 CONTAGIONS UNDER CONTEXT-DEPENDENT BEHAVIOR

As discussed at the beginning of this chapter, existing studies coupling infectious spreading and behavior are not able to distinguish whether a potentially infectious contact occurs in isolation or in the co-presence of other individuals. However, even though each transmission concerns only two people, the transmission is indirectly affected by the way contacts are organized within groups, for the adoption of a behavior (e.g., prophylaxis, as wearing a face-mask) depends on the level of adoption an individual observes in the entire group [99, 334, 335]. This implies that an appropriate description of the process needs to take into account either that contacts generally occur within groups, and that, concurrently to the infectious spreading, there is a decision-making process at the group-level that shapes the behavior of the individuals therein. Current models either consider the group structure [113, 257, 273, 321, 415] or the disease-behavior coupling [332, 333], but never both. In this section, we fill this gap presenting a general model of context-dependent spreading. This is representative of an entire class of models, each identified by the specific spreading process considered and by how agents' behavior is affected by the context. The models presented in the previous section, Sec. 4.2, can be seen as special cases where the effect of context is negligible—at least during a meaningful observation time.

To describe context-dependent behavior we divide individuals into behavioral types, distinguished by a different propensity to adopt a certain behavior. The correlation between individual characteristics and behavior has been observed in various forms, among which in the adoption of prophylactic measures [180, 367, 416]. At the same time, similar individuals not only behave similarly, they are also more likely to interact among them than with others, a phenomenon known in sociology as homophily [337]. The local dynamics—and the global one emerging from it—is thus determined by both, the behavior of the types and the way they mix [417–421]. To account for the latter, a theory of mixing in groups is required. The most basic instance of such a theory is the annealed heterogeneous mixing approximation developed in Sec. 4.1, used hereafter. In Sec. 4.3.2, we model the context-dependency of agents' behavior. Together, Secs. 4.1 and 4.3.2 provide the basic theoretical foundations for the study of context-dependent processes analytically. As an application of our theory, in Sec 4.3.3 we consider an epidemic spreading model incorporating context-dependent adoption of prophylactic behavior, then explored in Sec. 4.3.4. There, we reveal a rich phenomenology for the basic reproduction number and the endemic state in relation to the size and type-composition distributions of the groups, unveiling the important

dynamical implications of accounting for the modifications of the pairwise contagion created by a higher-order context.

4.3.1 *The model*

Let us consider a population of agents within which a certain entity (e.g., a pathogen, a rumour) can spread from one agent to another in a pairwise fashion. Such an interaction generally occurs in the presence of other agents not directly participating to it, i.e., within a group of some size (in particular, of size two if no other agent is present). If the co-present agents can affect the pairwise interaction, the group itself mediates an interaction modification. Thus, we henceforth use the term "group" or "higher-order interaction" to indicate a group of agents, defining the context within which the pairwise (direct) interactions conveying the spread take place between any two agents in the group.

Every agent, during each interaction, can either actively behave to modify (hinder or favor) the spread or not. Whether it chooses to adopt such behavior or not is determined by both the behavior it observes among the peers involved in the interaction and some intrinsic characteristics of its own. To account for the latter, suppose to partition the population into two behavioral types, labeled—as anticipated in Sec. 4.1—through letters A and N, standing for 'adoption-inclined' and 'not adoption-inclined': accordingly, type A is assigned to those agents which are intrinsically more prone to adopt an active behavior. Then, every time a group of agents forms, each of them adopts a behavior and a spreading event is let to happen within any pair of them. Note at this point the assumption of timescale separation we are implicitly making: the agents decide their behavior instantly, as soon as the group forms, to then let the—slow—spreading dynamics unfold throughout the duration of the group interaction. This is a reasonable (and convenient) hypothesis in some cases, like the one considered in this work, but could be not in others. There is no technical difficulty in relaxing it though, while leaving intact the core of the model: interaction modifications of the spreading dynamics generated by the context-dependency of behaviors.

Figure 4.10 provides an illustrative example of the model, while showing the inadequacy of a pairwise representation for context-dependent processes and the consequent need for higher-order representations. Indeed, since the success of each pairwise transmission depends on the behaviors taken on by all the agents in the group, the exact description of the whole process requires knowledge of either the interaction structure (i.e., who gathers with whom, and when) and the functional dependence of the agents' behavior on their type and on the type-composition of a group.

The minimal description given so far does not yet specify how the interactions take place. Among the many possible ones, we consider the following. At time step t, a generic agent i is involved in $k_i^{(n)}(t)$ groups of size $n \in \{2, \dots, n_{\max}\}$—or n-edges, in the language of hypergraphs. The composition of the groups formed at time t is encoded

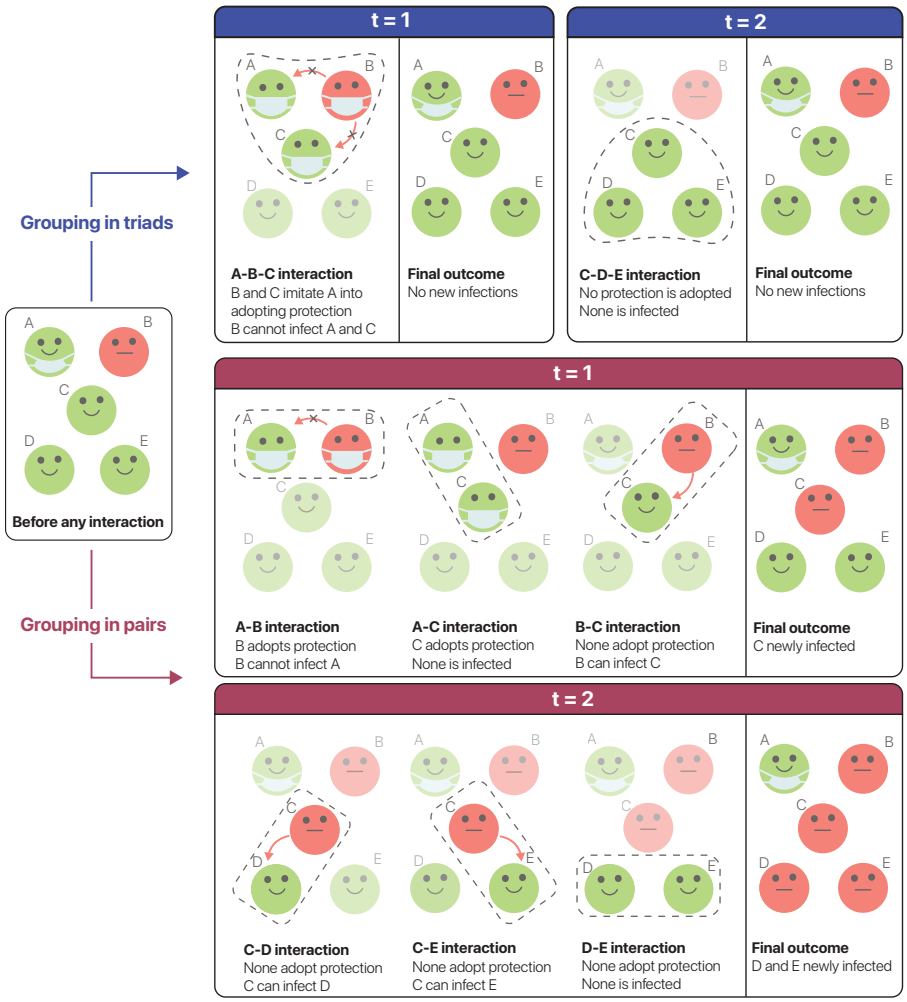


Figure 4.10: Example of a context-dependent spreading where the agents can act in order to avoid it (in either directions) by adopting a prophylactic behavior, represented as a worn face-mask. At each time t, some group interactions (dash-outlined) take place. During each interaction, an infection can occur from any infected (red) agent to any susceptible (green) one. In this example, A-type nodes (represented with a ready-to-use, lowered face-mask while no interacting) are assumed to behave actively no matter what the others do, while N-types' (those without face-mask while no interacting) do it only when they observe another node doing it, i.e., when at least an A-type's is in the group. In this setup, any transmission may be avoided when interacting in groups of three, whereas all the N-type agents may be infected when interacting in pairs. Neglecting the group organization of contacts can thus lead to radically different outcomes, proving the need for a higher-order representation to account for context.

in the adjacency tensors $\left\{A_t{}^{(n)}\right\}$, where $A_{t\,i_1,\ldots,i_n}^{(n)} = 1$ if the group $\{i_1,\ldots,i_n\}$ exists at time t, and 0 otherwise. The exact specification of the interaction structure requires however a huge amount of detailed information. While this is largely unavailable in most of the cases, we are anyway forced to resort to some approximate description if we want to gain clear knowledge from analytical results. To this end, we refer to the annealed approximation presented in Sec. 4.1, in which the value of $A_{t\,i_1,\ldots,i_n}^{(n)}$, at any t, is taken equal to its average value over the ensemble of which the hypergraph defined by $\left\{A_t{}^{(n)}\right\}$ is a concrete realization. We recall that the ensemble is specified by the number of nodes N, the fraction of A-type nodes ρ, $\binom{n_{\max}}{2}$ mixing parameters (at most), and the average degrees of the two types.

4.3.2 *Context-dependent behavior*

The next step in the definition of our model is to establish the behavioral difference between the two types of agents. The probability of actively modifying the spread can be thought of as consisting of a type-specific, context-independent part (e.g., prosociality [422]) and a context-dependent one, i.e., relying on the observed behavior of the others in the group [99]. The probability of being active is thus a dynamic object. In this regard, a simplifying assumption (seemingly realistic for some applications) is that such probability converges to some value during a timescale which is enough shorter than the duration of a group interaction. For example, the decision of wearing a face-mask may be considered to be made up in the very first moments of a gathering. With this timescale separation, the probability of being active is assumed to instantly attain its equilibrium value, effectively becoming a time-independent quantity.

Now, given a group of size n, let $q_{\mathrm{X}}^{(n_{\mathrm{X}}-1,n-1)}$ be the probability that a X-type agent adopts an active behavior when $n_{\mathrm{X}}-1$ out of the other $n-1$ agents in the group are of X-type's as well. In accordance to the meaning we associated with the labels we require that (i) the probability of behaving actively, within a mixed group, is always higher for an A-type agent than for a N-type's; and (ii) such probability increases with the number of A-type's present in the group. Taken together they imply the following chain of inequalities,

$$q_{\mathrm{A}}^{(n-1,n-1)} \geqslant q_{\mathrm{A}}^{(n_{\mathrm{A}}-1,n-1)} \geqslant q_{\mathrm{N}}^{(n_{\mathrm{N}}-1,n-1)} \geqslant q_{\mathrm{N}}^{(n-1,n-1)} \,, \tag{4.55}$$

for $1 \leqslant n_{\mathrm{A}} \leqslant n-1$, $n_{\mathrm{N}} = n - n_{\mathrm{A}}$. Equation (4.55) defines the qualitative behavioral difference observable between the types.

Having access to empirical measurements of the q_{X}s, one could simply fit them in the dynamic equations. In the absence of empirical information, we may suppose different functional forms for them, satisfying Eq. (4.55). One possibility is to let the q_{X}s emerge dynamically as a result of a behavioral adaptation, as shown in Sec. 4.3.2.1. Another one, presented in Sec. 4.3.2.2, is to postulate binary functional forms for the q_{X}s, where agents of each type are active or inactive in a deterministic way depending

on the composition of the groups they take part in. The simplicity of the binary forms will allow us to obtain explicit analytical results, gaining essential insights to interpret more complex, middle grounded scenarios.

To notice that any of the behavioral dynamics proposed here can be easily generalized to accommodate additional potential factors, such as external sources of influence (e.g., mass media), seen as a background context modifying the baseline probability of adoption, or some coupling with the spreading dynamics (e.g., awareness raised by information on an epidemics).

4.3.2.1 Social contagion model

Here we let agents' behavior emerge through a process of (simple) social contagion (e.g., via peer pressure) [180, 423, 424] modelled as a susceptible-infectious-susceptible (SIS) dynamics, where here 'S' and 'I' represent nonadoption and adoption of the active behavior, respectively, with additional endogenous transitions moving agents from state S to state I and vice versa at rates depending solely on their type. These rates quantify some cost of adopting an active behavior (e.g., the discomfort of wearing a face-mask) and the will to bear that cost (e.g., due to prosociality or vulnerability to a disease). Denoting with c_X the I-to-S rate and with b_X the S-to-I rate, $X \in \{A, N\}$, we describe the behavioral dynamics via the following system (see Appendix D.2.1 for derivation)

$$\dot{q}_A = \frac{n_N}{n-1}(q_N - q_A) + b_A(1 - q_A) - c_A q_A \,, \tag{4.56}$$

$$\dot{q}_N = \frac{n_A}{n-1}(q_A - q_N) + b_N(1 - q_N) - c_N q_N \,. \tag{4.57}$$

where we denoted $q_X^{(n_X-1,n-1)}(t)$ simply as q_X for a lighter notation. The first term of each of the two equations contains the context-dependent dynamics, which has the form of a linear consensus formation process [425]; the remaining terms contain the context-independent, type-specific dynamics.

The system admits a unique solution for each possible composition of a group. For mixed groups (i.e., $1 \leqslant n_A \leqslant n-1$), it takes the form $(q_A, q_N) = (C_1 + C_2 b_A c_N, C_1 + C_2 b_N c_A)$, where C_1 and C_2 are constants depending on all the present parameters (see Appendix D.2.1 for their explicit expressions). For uniform groups of type X (i.e., $n_X = n$), the solution simply reads $q_X = b_X/(b_X + c_X)$, $\forall n$. Interestingly, for a wide range of values of the transition rates (bs and cs) producing a significant behavioral difference between the types—which is what really motivates accounting for them—, the equilibrium probability of adoption shows a marked nonlinear dependence on the number of A-type (or N-type) agents in the group. This is especially true for q_N (q_A is always high), growing nonlinearly from low values in a uniform N-type group ($n_A = 0$) to medium-high values when all the others in the group are A-type's ($n_A = n-1$). In other words, the equilibrium solution stemming from the simple (linear) contagion mechanism takes the form of a complex (nonlinear)

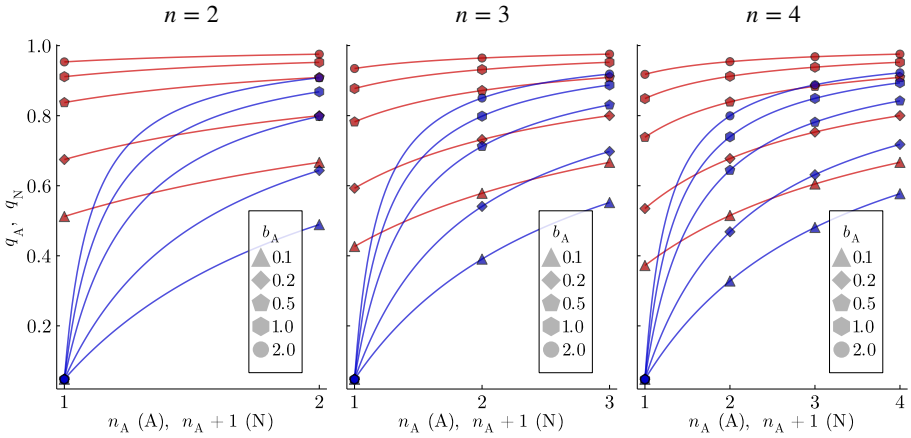


Figure 4.11: Equilibrium adoption probability for A-type individuals, q_A (red points), and for N-type ones, q_N (blue points), against the number of A-type's in the group (from 1 to n for an A-type's, from 0 to $n-1$ for a N-type's). Different symbols denote different adoption rates for A-type's, b_A, while the other rates are fixed as $c_A = c_N = 0.05$ and $b_N = c_N/20$. In Sec. 4.3.4 we use $b_A = 1.0$ (hexagons). Left plot is for 2-edges ($n = 2$), middle plot for 3-edges ($n = 3$) and right plot for 4-edges ($n = 4$). To better appreciate the nonlinear dependence on n_A, we also depicts through solid lines the equilibrium solutions as if n_A was a continuous variable.

contagion mechanism. Thanks to the timescale separation, the outcome of the simple contagion effectively acts as a complex contagion, typical of behavioral dynamics.[2]

Figure Fig. 4.11 shows the nonlinear dependence of the solutions q_A and q_N on the composition of a group, for different values of the A-type's adoption rate, b_A, and different group sizes. The other rates are fixed as used in Sec. 4.3.4, i.e., $c_A = c_N = 0.05$ and $b_N = c_N/20$. There, we will take $b_A = 1.0$ (hexagon-shaped points in Fig. 4.11), which yields a probability of adoption of around 95% and 5% in an uniform A-type and N-type group, respectively. In a $\{A, N\}$ 2-edge, the adoption probability jumps to around 87% for the N-type agent and decreases to around 91% for the A-type's. In a $\{A, N, N\}$ 3-edge, it goes to around 80% for the N-type's and to 88% for the A-type's, while in a $\{A, A, N\}$ one, to around 89% and 93%, respectively.

4.3.2.2 Binary models

We define here the binary models we call of 'easy adaptation' and 'hard adaptation'. In both of them $q_A = 1$ always, i.e., A-type agents are assumed to behave according to their intrinsic propensity and independently of the context observed in a group. We then have the following:

(i) In easy adaptation, q_N is 1 when at least an A-type is present (specifically, for

2 This observation actually raises the question of whether certain social contagions are genuinely complex, or if they only appear as such due to the temporal resolution at which they are studied.

rank-3 hypergraphs, in {A,N}, {A,N,N} and {A,A,N} groups)) and 0 otherwise (that is, in uniform N-type groups). N-type agents are easily (maximally) driven by A-type's in this scenario;

(ii) In hard adaptation, q_N is 1 when all the other individuals in the group are A-type's ({A,N} and {A,A,N} groups) and is 0 otherwise., i.e., N-type's are hardly (minimally) driven by A-type's in this case.

The only and crucial difference between the two behavioral models is, for rank-3 hypergraphs, the N-type agents' behavior in {A,N,N} groups: the presence of the A-type agent induces active behavior in the two N-type's in the easy adaptation model, but not in the hard adaptation one. As shown in the next sections, this difference modifies the spreading process by making it strongly dependent on the behavioral dynamics taking place at the group-level, hence on the properties of the higher-order interaction structure.

4.3.3 *Spreading dynamics*

Once the interaction structure and the behavior's context-dependency have been modelled, we are ready to investigate how they jointly affect a spreading process taking place upon that structure. For the sake of simplicity, we assume the spreading process to be represented by a SIS dynamics. Let thus β be the transmission rate and μ the recovery rate. We denote with $Y_X \equiv Y_X(t)$ the fraction of agents of type $X \in \{A, N\}$ in compartment $Y \in \{S, I\}$ at time t. In particular, I_X is the type-specific prevalence for type X, while $I = \rho I_A + (1-\rho) I_N$ is the prevalence overall. The dynamics can be modelled through the following system of differential equations

$$\dot{I}_A = -\mu I_A + \beta S_A \left(I_A \theta_{A\to A} + I_N \theta_{N\to A} \right) , \tag{4.58}$$

$$\dot{I}_N = -\mu I_N + \beta S_N \left(I_A \theta_{A\to N} + I_N \theta_{N\to N} \right) , \tag{4.59}$$

with $S_X = 1 - I_X$. Here $\beta\theta_{X\to Z}$ is the total transmission rate from a X-type to a Z-type, reading

$$\theta_{A\to A} = \kappa_{A|A} r_A^{(1,1)} s_A^{(1,1)} + \kappa_{A,N|A} r_A^{(1,2)} s_A^{(1,2)} + 2\kappa_{A,A|A} r_A^{(2,2)} s_A^{(2,2)} , \tag{4.60}$$

$$\theta_{N\to A} = \kappa_{N|A} r_N^{(0,1)} s_A^{(0,1)} + \kappa_{A,N|A} r_N^{(0,2)} s_A^{(1,2)} + 2\kappa_{N,N|A} r_N^{(1,2)} s_A^{(0,2)} , \tag{4.61}$$

$$\theta_{A\to N} = \kappa_{A|N} r_A^{(0,1)} s_N^{(0,1)} + \kappa_{A,N|N} r_A^{(0,2)} s_N^{(1,2)} + 2\kappa_{A,A|N} r_A^{(1,2)} s_N^{(0,2)} , \tag{4.62}$$

$$\theta_{N\to N} = \kappa_{N|N} r_N^{(1,1)} s_N^{(1,1)} + \kappa_{A,N|N} r_N^{(1,2)} s_N^{(1,2)} + 2\kappa_{N,N|N} r_N^{(2,2)} s_N^{(2,2)} , \tag{4.63}$$

where $\beta r_X^{(n_X-1,n-1)} s_Z^{(n_Z-1,n-1)}$ is the effective transmission rate for an interaction occurring within a n-edge composed of n_X X-type and n_Z Z-type agents. This is written as the product of an out-going (i.e., depending on the infected, a X-type) transmission probability $r_X^{(n_X-1,n-1)}$ and an in-going (i.e., depending on the susceptible, a Z-type) transmission probability $s_Z^{(n_Z-1,n-1)}$. With the same notation used in Sec. 4.3.2, the

superscript $(n_X - 1, n - 1)$ indicates how many of the other $n - 1$ nodes in the group are of type X (subscript), given at least an X-type is present. In other words, r_X and s_X encode the context within which a direct interaction occurs from the perspective of a X-type agent. Their form is dictated by the specific mechanism through which transmission can be modified and, via their dependence on the agents' probability to adopt active behavior, by how the mechanism is affected by the size and the type-composition of a group. For prophylactic mechanisms, such as face-masks, considered in the following, r_X and s_X act as reduction factors of the transmission probability.

Linearizing Eqs. (4.58) and (4.59) around the disease-free equilibrium, $(I_A, I_N, S_A, S_N) = (0, 0, 1, 1)$, the associated NGM matrix reads

$$\mathrm{NGM} = \frac{\beta}{\mu} \begin{pmatrix} \theta_{A\to A} & \theta_{N\to A} \\ \theta_{A\to N} & \theta_{N\to N} \end{pmatrix} . \tag{4.64}$$

Denoted with $\boldsymbol{\Theta}$ the matrix with entries $\{\theta_{X\to Z}\}$, the effective reproduction number, R, reads $R = \beta k_{\mathrm{eff}}/\mu$, with

$$k_{\mathrm{eff}} = \frac{1}{2}\left[\theta_{A\to A} + \theta_{N\to N} + \sqrt{(\theta_{A\to A} - \theta_{N\to N})^2 + 4\theta_{N\to A}\theta_{A\to N}}\right] , \tag{4.65}$$

being the largest eigenvalue of $\boldsymbol{\Theta}$, representing the effective—for spreading—average pairwise degree. Imposing $R = 1$ provides the epidemic threshold $\beta_c = \mu/k_{\mathrm{eff}}$, above which the epidemic-free equilibrium is unstable and an endemic state is reached. If there is no modification of the transmission (i.e., $r_X = s_X = 1$, $\forall X$) and there is no correlation between type and degree ($k^{(n)}_{0,1} = k^{(n)}_{1,0} = k^{(n)}$, $\forall n$), all agents are equivalent and mixing is irrelevant, hence $k_{\mathrm{eff}} = k$ and $R = R_0 \equiv \beta k/\mu$, as the SIS model on networks is recovered.

Notice that it is possible to get a closed form for the endemic solution of Eqs. (4.58) and (4.59). However, the complicated form of the solution does not allow to draw any noteworthy conclusion directly from it, hence we do not report its expression here. Nonetheless, an approximate solution that preserves the qualitative behavior of the system can be found under high prophylactic efficacy (see Appendix D.2.3). Moreover, no qualitative changes are observed in the results presented below when, instead of a SIS, a SIR is considered. In that case, the results refer to the final attack rate rather than to the equilibrium prevalence.

4.3.3.1 Application to face-masks adoption

To describe an epidemic spreading in the presence of face-masks, we need to specify a suitable form for the probabilities r_X and s_X appearing in Eqs. (4.60) to (4.63). To this end, let $\varepsilon_{\mathrm{out}}$, $\varepsilon_{\mathrm{in}} \in [0, 1]$ be the out- and in-going efficacy of a face-mask, respectively. The quantities $\varepsilon_{\mathrm{out}} q_X^{(n_X-1,n-1)}$ and $\varepsilon_{\mathrm{in}} q_X^{(n_X-1,n-1)}$ are then the probabilities that a X-type individual wears a mask and that the mask avoids, respectively, out-going

transmission when the X-type is infectious, and in-going transmission when the X-type is exposed to an infection. In more general terms, they quantify how strongly the spreading process can be affected by an agent's behavior: for perfect efficacy, the success of a spreading event is decided in a deterministic way by the eventually adopted behavior; on the contrary, for null efficacy, spreading and behavior fully decouple. Thus we have

$$r_{\mathrm{X}}^{(n_{\mathrm{X}}-1,n-1)} = 1 - \varepsilon_{\mathrm{out}} q_{\mathrm{X}}^{(n_{\mathrm{X}}-1,n-1)} , \tag{4.66}$$

$$s_{\mathrm{X}}^{(n_{\mathrm{X}}-1,n-1)} = 1 - \varepsilon_{\mathrm{in}} q_{\mathrm{X}}^{(n_{\mathrm{X}}-1,n-1)} . \tag{4.67}$$

In the following we assume $\varepsilon_{\mathrm{out}} = 0.9$ and $\varepsilon_{\mathrm{in}} = 0.5$. The effect of varying each efficacy will be discussed whenever significant.

4.3.4 *Results*

To study how the group structure of the contacts affects the spreading dynamics, we vary the interaction structure along two directions. For given average values of degree, k, and assortativity, α, we ask what is the effect of distributing (i) the degree ($k_{\mathrm{A}}^{(n)}$ and $k_{\mathrm{N}}^{(n)}$) and (ii) the assortativity ($\alpha_{1,0}^{(n)}$ and $\alpha_{1,1}^{(n)}$) differently at the varying of the group size n. To isolate the role of the mixing, we assume that type and degree are hereafter independent random variables. Therefore, a priori, $k_{\mathrm{A}}^{(n)} = k_{\mathrm{N}}^{(n)} = k^{(n)}$, $\forall n$, and consequently $\alpha_{1,1}^{(n)} = \alpha_{1,0}^{(n)} \equiv \alpha_n$. We consider individuals gathering in pairs and triads, i.e., $n = 2, 3$.

In the following, we display and analyze the results for the social contagion model of Sec. 4.3.2.1. To this end, we leverage the basic understanding provided by the binary models of Sec. 4.3.2.2. As anticipated in Sec. 4.3.2.1, we fix the adoption probabilities by taking $c_{\mathrm{A}} = c_{\mathrm{N}} = 0.05$, $b_{\mathrm{A}} = 20 c_{\mathrm{A}} = 1.0$ and $b_{\mathrm{N}} = c_{\mathrm{N}}/20 = 0.0025$. Apart from the numerical details, what ultimately matters is that the A-type agents act as modifiers of the infectious interactions in a group by lowering the transmission probability through the prophylaxis induced in the N-types'.

4.3.4.1 Varying the degree among group sizes

Suppose that the mixing parameters for each group size, α_2, α_3 and $\beta_{3,1}$, are given. We then vary $k^{(2)}$ and $k^{(3)}$ respecting the constraint $k = k^{(2)} + 2k^{(3)}$ and study how this affects the dynamics. Specifically, we ask: in order to avoid an epidemic (i.e., keep $R < 1$) or to reduce its size (i.e., minimize prevalence), should the agents interact more in pairs or in triads?

BASIC REPRODUCTION NUMBER. Let us first consider the effective reproduction number, R. We find the behavior of R to generally depend on both the mixing pattern (α_2, α_3, $\beta_{3,1}$) and the type-composition of the population (ρ), but in a different

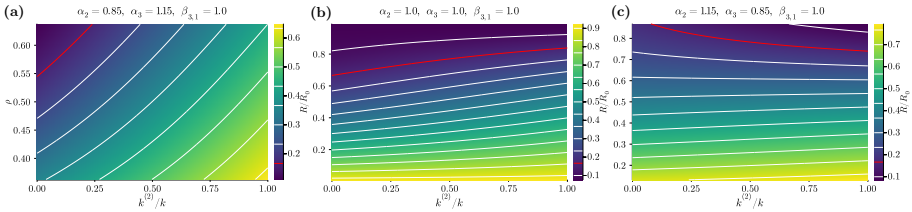


Figure 4.12: Normalized reproduction number, $R/R_0 = k_{\text{eff}}/k$, computed from Eq. (4.65), as a function of the proportion of 2-edges, $k^{(2)}/k$, and of the A-type proportion in the population, ρ, for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and (a) $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, (b) $\alpha_2 = \alpha_3 = 1.0$, **(c)** $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. Since $k = k^{(2)} + 2k^{(3)}$, $k^{(2)}/k = 0$ means nodes interacts only within triads, while $k^{(2)}/k = 1$ means they do it just in pairs. White curves report the levels indicated in the respective colorbar. The red one indicates the critical curve $R = 1$ ($R/R_0 = 1/6$), above which the epidemic-free state is stable.

way depending on the system being closer to the easy or the hard adaptation scenario. This is observed, for example, for the intermediate scenario we used, generated by the social contagion model; see Fig. 4.12. This scenario is closer to the easy adaptation limit, thus let us first analyze the latter. Assuming an ideal perfect protection in at least one direction (i.e., $\varepsilon_{\text{out}} = 1$ and/or $\varepsilon_{\text{in}} = 1$), we prove in Appendix D.2.2.1 that, under easy adaptation, there exists a threshold value

$$\tilde{\rho} = \frac{2\alpha_3 - \alpha_2}{\alpha_3 \beta_{3,1}} , \tag{4.68}$$

such that R increases or decreases with $k^{(2)}$ (for fixed k) depending on whether ρ is lower or higher than $\tilde{\rho}$. At $\rho = \tilde{\rho}$, a N-type individual has the same probability of finding at least an A-type's—and so, under easy adaptation, be induced to adoption—in a pair or in a triad. When $\rho < \tilde{\rho}$, such probability becomes lower in a pair than in a triad, leading R to increase with $k^{(2)}$; vice versa for $\rho > \tilde{\rho}$. As a consequence, expressed the critical threshold as the value $\rho = \rho_c$ at which $R = 1$ (red curves in Fig. 4.12), the threshold moves up or down with $k^{(2)}$ depending on whether $\rho_c < \tilde{\rho}$ or $\rho_c > \tilde{\rho}$. Notice from Eq. (4.68) that $\tilde{\rho} \in [0, 1]$ requires $\alpha_3 \in [\alpha_2/2, \alpha_2/(2 - \beta_{3,1})]$. When $\alpha_3 < \alpha_2/2$, $\tilde{\rho} < 0$ and R thus always decreases with $k^{(2)}$, for it is easier for a N-type's to find at least an A-type's in a pair than in a triad, whatever the fraction ρ of A-type's in the population. Said differently, the induced prophylaxis is less frequent in triads than in pairs because the former are comparatively too much assortative, therefore R is minimized (the epidemic threshold is maximized) by interacting in pairs only, i.e., for $k^{(2)} = k$ and $k^{(3)} = 0$. In the opposite case in which $\alpha_3 > \alpha_2/(2 - \beta_{3,1})$, $\tilde{\rho} > 1$ and R thus always increases with $k^{(2)}$, for a N-type's is more likely to find at least an A-type's in a triad than in a pair, no matter the value of ρ. Among the groups containing type A, while a {A,A,N} triad and {A,N} pair generate equivalent

contexts, a {A,N,N} triad has the advantage of weakening the N-N contact thanks to the prophylaxis induced in the two N-type agents, that would be absent if the two interacted as a pair. This results in R being minimized by interacting in triads only, i.e., for $k^{(2)} = 0$ and $k^{(3)} = k/2$.

Going back to the intermediate scenario considered in Fig. 4.12, treating it as if was an easy adaptation scenario with perfect unilateral protection, Eq. (4.68) yields $\tilde{\rho} \geqslant 1$ for $\alpha_2 \leqslant \alpha_3$ (panels (a) and (b)), and $\tilde{\rho} \approx 0.65$ for $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$, which is anyway not far from the true threshold, $\rho \approx 0.60$ (panel (c)). This consistency with the results reported in Fig. 4.12 confirms the explanatory power of the simplified dynamics assumed to derive Eq. (4.68).

In the hard adaptation limit, again assuming perfect unilateral protection, the shape of R is instead always monotonic, increasing or decreasing with $k^{(2)}$ solely depending on whether α_2 is smaller or larger than α_3, respectively (see Appendix D.2.2.1 for proof). Indeed, since prophylaxis is not induced in the {A,N,N} triads in this case, reducing R simply amounts to rise disassortativity (i.e., increase α) and in turn increase the frequency of adoption for the N-type's.

All in all, this analysis proves that, depending on the behavioral properties of the agents, solely changing the (first moment of the) group-size distribution can move the system from the subcritical (epidemic-free) to the supercritical (endemic) phase, and vice versa.

PREVALENCE. Looking now at the levels of prevalence at equilibrium, we can appreciate how differently the two types are affected by the interaction structure. We can first observe that I_N qualitatively behaves like R (compare panels (c), (f) and (i) in Fig. 4.13 with panels (a), (b) and (c) in Fig. 4.12, respectively), therefore the analysis following Eq. (4.68) approximately applies to I_N too. This correspondence between I_N and R is actually expected whenever there is a substantial behavioral difference between the two types and prophylactic efficacy is high. In such case (see Appendix D.2.3 for proof) the epidemic pressure mainly comes from type N and the dynamics for the latter is well approximated by a standard, one-type population SIS, where both I_N and R are largely determined by $\theta_{N\to N}$. Since, instead, $\theta_{N\to A}$ can show a different and opposite dependence on the parameters compared with $\theta_{N\to N}$, I_A might not generally follow R, as seen comparing panels (b), (e) and (h) in Fig. 4.13 with panels (a), (b) and (c) in Fig. 4.12, respectively.

The results for I_A thus require a more detailed inspection. For the intermediate scenario considered in Fig. 4.13, they broadly overlap with those obtained in the easy adaptation limit (compare with Fig. E.1). A clear exception is the case $\alpha_2 = \alpha_3$ (panel (e)), for which I_A first increases and then decreases with $k^{(2)}$ as ρ is decreased, whereas, in easy adaptation, I_A monotonically increases with $k^{(2)}$. The discrepancy is explained by the quantitative differences that make the intermediate scenario indeed depart from the easy adaptation limit and get a slight contribution from its hard adaptation "component". On one hand, in easy adaptation, the increase of I_A with

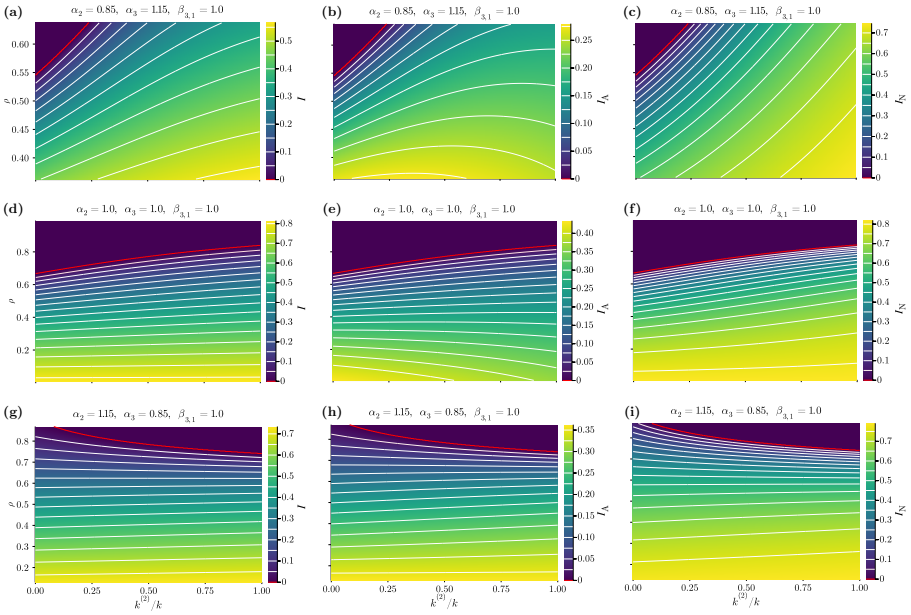


Figure 4.13: Equilibrium endemic state (fixed point of the system of Eqs. (4.58) and (4.59)) as a function of the proportion of 2-edges, $k^{(2)}/k$, and of the A-type proportion in the population, ρ, for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and (a–c) $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, **(d–f)** $\alpha_2 = \alpha_3 = 1.0$, (g–i) $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. (a, d, g) Overall prevalence, I, (b, e, h) prevalence for type A, I_A, and (c, f, i) prevalence for type N, I_N. Since $k = k^{(2)} + 2k^{(3)}$, $k^{(2)}/k = 0$ means nodes interacts only within triads, while $k^{(2)}/k = 1$ means they do it just in pairs. White curves report the levels indicated in the respective color bar. The red one indicates the critical curve $R = 1$, above which the epidemic-free state is stable.

$k^{(2)}$ gradually vanishes when ρ approaches small values: with a low proportion of A-type individuals in the population, the modification mechanism weakening the N-N interactions in $\{A, N, N\}$ triads which makes I_N, and consequently I_A, smaller, is largely compensated by the high epidemic pressure that anyway emerges via the more frequent N-type uniform groups. On the contrary, in hard adaptation (see Fig. E.3), I_A decreases faster and faster as ρ gets smaller: while I_N is not directly affected by group size (as there is no modification of the N-N interactions), I_A decreases by trading $\{A, N, N\}$ triads for $\{A, N\}$ pairs thanks to the gained bilateral protection, and $\{A, N\}$ and $\{A, N, N\}$ are those groups in which A-type individuals interact the most when ρ is small. Therefore, combining the fact that in the intermediate scenario q_N in a $\{A, N, N\}$ triad is not as high as in a $\{A, N\}$ pair, differently from the easy adaptation limit in which $q_N = 1$ in both cases, with the fact that by decreasing ρ the effect of $k^{(2)}$ on I_A gets weaker in easy adaptation but stronger in hard adaptation, results in the decrease of I_A seen in Fig. 4.13(e) for low enough ρ.

Such sort of competition between easy and hard adaptation, determining what we observe in an intermediate scenario, is certainly present also for $\alpha_2 \neq \alpha_3$. However, it has no qualitative impact when α_2 and α_3 are not too close to each other. In such case, the considered intermediate scenario has the same qualitative behavior than the easy adaptation one.

Accordingly, for $\alpha_2 < \alpha_3$ (and $\beta_{3,1} = 1$), increasing $k^{(2)}$, not only makes the weakening of the N-N interactions rarer as less interactions occur in triads, but also increases the rate at which N-N interactions take place, for pairs are more assortative than triads. This results in a rapid increase of I_{N} (as we already know from R) which, in turn, pushes I_{A} up. At the same time, especially for lower ρ, type A benefits from the increase in homophily provided by a higher $k^{(2)}$, as it isolates it from the more infectious type N. Eventually, when ρ is not too high and the increase of I_{N} gets slower (see Fig. 4.13(c)), the two contributions to I_{A} become comparable, giving the latter the nonmonotonic shape observed in Fig. 4.13(b). This holds whenever efficacy is high at least in one way, so that there is a substantial difference between the epidemic pressures of the two types; otherwise, homophily looses importance and I_{A} just grows driven by I_{N}.

Finally, for $\alpha_2 > \alpha_3$ (see Figs. 4.13(h)-(i)), I_{A} mainly behaves as I_{N} (hence as R). Differently from the previous case, pairs are now less assortative than triads and so type A cannot benefit from mixing to compensate for the larger exposure to type N implied by a higher $k^{(2)}$. I_{A} thus just follows I_{N}.

All in all, it is I_{N} to be the main driver of the observed phenomenology in the easy adaptation scenario. This leads to the conclusion that, by exploiting the indirect weakening of the N-N contacts, the system (made few exceptions) benefits from interacting more in triads than in pairs. In the limit of hard adaptation that role is instead played by I_{A} (see Appendix E.1 for a detailed discussion about the results in this scenario). As observed before, while type N is not directly affected by group size in this case, type A benefits from the protection induced in $\{\mathrm{A},\mathrm{N}\}$ pairs, absent in $\{\mathrm{A},\mathrm{N},\mathrm{N}\}$ triads, making pairs preferable to triads for the system overall. What is found for any intermediate scenario, as the one we derived from the social contagion model, eventually depends on how close this is to either one limit or the other.

At last, we considered quenched contact structures (see Appendix D.2.4 for the numerical implementation of the model) to test whether the presence of topological correlations changed the phenomenology predicted by the mean-field model, Eqs. (4.58)-(4.59). Specifically, keeping fixed the mixing parameters to some values, we varied $k^{(2)}/k$ and looked at the equilibrium endemic state. As Fig. 4.14 shows, the qualitative behavior is accurately reproduced. Note that the systematic overestimation of the prevalence is actually expected for approximation schemes that ignore dynamical correlations when the system is enough above the epidemic threshold [282].

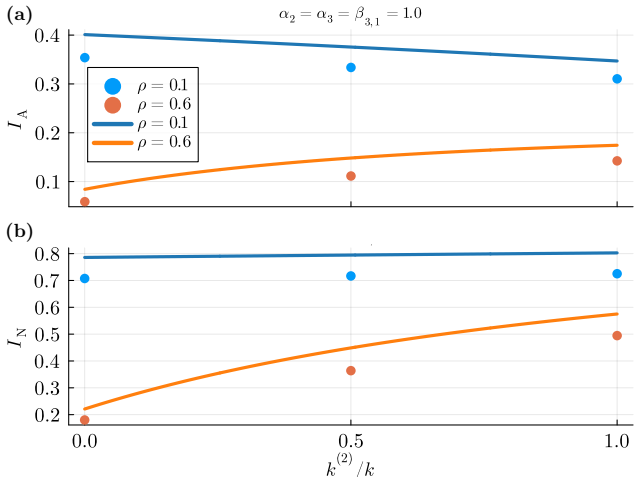


Figure 4.14: Monte Carlo results for the equilibrium endemic state as a function of the proportion of 2-edges, $k^{(2)}/k$, for regular rank-3 hypergraphs of $N = 1000$ nodes and pairwise degree $k = k^{(2)} + 2k^{(3)} = 12$. Here, $\beta = 0.1$, $\mu = 0.2$ and $\alpha_2 = \alpha_3 = \beta_{3,1} = 1.0$ (homogeneous mixing). (a) Prevalence for type A, I_A, and (b) for type N, I_N. Each point is obtained by averaging over 2500 runs with random initial conditions (infectious state and type assignment). Error bars, representing standard deviations, are smaller than point size. Solid lines represent the results predicted by the mean-field model.

4.3.4.2 Varying the type-assortativity among group sizes

We briefly summarize here the results obtained when distributing the assortativity differently between 2- and 3-edges. That is, we fix α and vary α_2, with $\alpha_3 = \alpha + (\alpha - \alpha_2)\, k^{(2)}/(2k^{(3)})$ (from Eq. (4.32)).

As proved in Appendix D.2.2.2, the reproduction number, R, increases with α_2 in the easy adaptation scenario, whereas it decreases in the hard adaptation one.

The same holds also for the type-specific prevalences I_N and I_A, hence for the prevalence overall (see Figs. E.2 and E.4), and is explained as follows. In the easy adaptation limit, a higher (lower) α_2 (α_3) means a lower rate at which $\{A, N, N\}$ triads form, implying that less N-N interactions are weakened. As a consequence, the N-type's prevalence, I_N, increases. At the same time, since in this scenario the A-N interactions are dynamically equivalent whether they occur isolated or within a triad, the A-type's probability of infection is not directly affected by the assortativity distribution. However, since I_N increases, I_A increases as well (unless $\varepsilon_{in} = 1$), although in a milder way being the A-N contacts always bilaterally protected. In hard adaptation, instead, there is no indirect mechanism modifying the N-N interactions, hence I_N is not directly affected by the assortativity distribution. On the other hand, rising (lowering) α_2 (α_3) means increasing the rate at which the A-N interactions are bilaterally

protected, implying a lower I_A. In any case, infection from an A-type's is unlikely (unless ε_{in} and ε_{out} are both small), therefore I_N is only slightly reduced by α_2.

As already observed in Sec. 4.3.4.1, what to expect in an intermediate behavioral scenario depends then on how close it is to either one of the two limit scenarios, the results being generally an interpolation between the two (see Appendix E.2 for the results found in the intermediate scenario used in Sec. 4.3.4.1). We can generally say that lowering the assortativity within triads (i.e., raising α_3) makes the endemic prevalence decrease for type N, and increase or decrease for type A depending on whether the share of A-type individuals, ρ, stays below or above a certain value: the easy adaptation "component" dominates for high ρ (I_A decreases), the hard adaptation one for low ρ (I_A increases).

To close the analysis, note that the same qualitative effects of increasing (decreasing) α_2 (α_3) are also implied by increasing the higher-order mixing parameter $\beta_{3,1}$, for the frequency of the $\{A, N, N\}$ triads is either way reduced (see Appendix E.2). While this should not be surprising, as $\beta_{3,1}$ determines the structure as much as any pairwise parameter (α_2 and α_3), it remarks the fact that exclusively relying on pairwise information does not guarantee an accurate description of the system.

4.3.5 *Summary and discussion*

We have defined a minimal model of context-dependent spreading where, during an interaction, an agent either actively behave to alter the diffusion or not depending on either the behavior it observes among the co-present peers and its intrinsic inclination to take on the active behavior.

Referring to an epidemic spreading model where context-dependency concerns the adoption of prophylactic behaviors like face-mask wearing, we have shown that accounting for the behavioral dynamics unfolding at the higher-order level of organization of the interactions, can lead to important deviations from what can be expected based on pairwise information alone. We have revealed that the direction and the magnitude of those deviations primarily depend on the properties of the behavioral dynamics; more specifically, on how effective are the (non) adoption-inclined agents in inducing (inhibiting) adoption of active behavior on the others. Then, depending also on the proportion of adoption-inclined agents in the population and the way type-assortativity distributes among the group sizes, and, secondarily, on the prophylactic efficacy, gathering more often in pairs than in triads (larger groups) can either facilitate or impede the spreading. We have proven this analytically for the basic reproduction number and shown how exclusively changing the group-size distribution can determine whether an outbreak will be subcritical—and eventually vanish—or supercritical—leading to endemicity.

More specifically, either for the reproduction number and the prevalence, we can conclude that in general (made few exceptions) when prophylactic behavior is easy to

induce, then the system generally benefits by interacting in larger groups, for many otherwise unprotected contacts would be now protected thanks to the elicited adoption. Moreover, the benefit increases when the smaller (larger) groups are the more (dis)assortative. On the contrary, when adoption is hard to induce (e.g., when some majority rule applies), smaller (and disassortative) groups become preferable.

Looking at each type-specific prevalence, we have seen that the type of an agent, not only—as expected—strongly affects its infection risk but also changes the qualitative dependence of that risk on the parameters characterizing the contact structure. Observing this becomes especially important when, for instance, adoption is mostly driven by the vulnerability to the spreading disease. For example, if elderly and young people have very different chances of suffering from severe symptoms—as it is for COVID-19 [426] or influenza [427]—, they may also have a dissimilar propensity to protect themselves and the others [428–431]. In such cases, either the overall prevalence and the reproduction number might not be the most useful indicators, for the prevalence within the vulnerable (and probably adoption-inclined) class does not generally follow the same phenomenology.

Our model stimulates a new direction in the discourse on the interplay between behavior and epidemic spreading. It helps to understand the role of contextual information in decisions to reduce the impact of infectious spreading and provides theoretical support for epidemic control policies based on public gathering restrictions.

In the end, the mean-field approximation provided here opens new possibilities in the modeling of higher-order systems, specifically serving as a basic theory to investigate virtually any process for which mutable and contextual factors can be relevant. Inspired by growing evidence in ecology, we have studied one of such processes and suggested a few more (see also Sun [320]), but perhaps many others have not yet been recognized.

4.4 ON A GENERAL FRAMEWORK FOR CONTEXT-DEPENDENT INTERACTIONS

To close this chapter, we outline a way to build a general framework for systems including context-dependent interactions. We expand over two directions. First, in Sec. 4.3, we considered all the direct interactions as pairwise, modified by a third unit. However, the direct interactions themselves can involve more than two units, as seen in Sec. 3, thus defining a hypergraph; and, moreover, we may think of a modification as caused not by just one unit, but concurrently by more units. The direct interactions can be cast in the factor graph associated to the hypergraph, which is a bipartite graph. Besides, the interaction modifications can be described by a tripartite graph where edges go from the nodes representing the units of the system, to the nodes representing the modifications, to those representing the direct interactions. The set of nodes pointing to the same interaction-modification node constitute the group of units producing that modification; the direct-interaction nodes pointed by an interaction-modification node are those on which the modification acts on. Notice

that we could squeeze the tripartite graph into a bipartite one where units directly point to the interactions they modify, but this would come with a loss of information. Indeed, it would be impossible to distinguish whether a modification is the result of a combination of pairwise modifications from different units or, instead, it is produced by an entangled action, involving many units, that cannot be separated into pairwise contributions. Representing the modifications explicitly as nodes—through the "middle" partition—allows us to keep this fundamental distinction in our framework. In truth, it is also a key step in extending the framework in a second direction.

In fact, an interaction modification can in turn be itself modified. Evidence for this comes again from ecology [432]. For instance, in coffee agroecosystems [433, 434]—whose preservation is of worldwide interest given the magnitudes of the coffee industry and market—, Azteca ants (Azteca instabilis) limit the ability of female lady beetles (Azya orbigera) to locate good sites where posit their eggs. This negative effect is understood as an interaction modification (caused by the ants) of the positive interaction between the adult beetles and the larvae they produce. But there is more. The presence of a parasitic fly (in the Phoridae family) induces defensive behavior in the ants, consequently reducing the ability of the latter to harass female beetles. That is, the modification perpetrated by the ants is in turn modified by the presence of the flies. These two modifications are a first-order and a second-order interaction modification, respectively. The tripartite representation suggested above naturally describe interaction modifications of first, second or any higher order. While we are currently not aware of third-order modifications in any kind of system, we would not be too surprised to learn that they have been observed in systems of mesmerizing complexity as ecological and social systems.

The resulting structure, encompassing both direct interactions and modifications, maps to a multilayer, multipartite (bi- and tri-partite) network, as Fig. 4.15 illustrates. Providing a novel perspective on systems with multi-order context-dependent interactions, we believe that this network representation may provide important insights about the dynamics of such systems. An interesting question is whether network measures tailored for this representation will provide already new understanding. A first application would be the analysis of ecological systems, for which a suitable representation is still lacking [432].

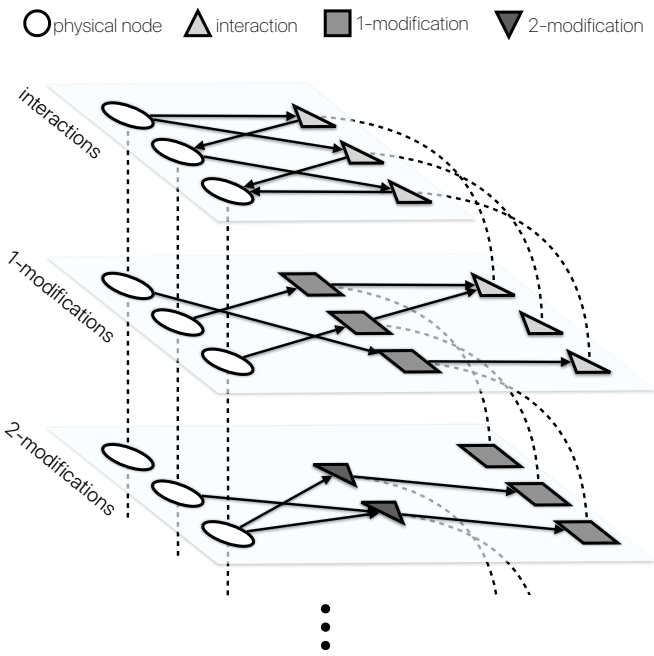


Figure 4.15: Representation of a generalized context-dependent interaction system as a multilayer, multipartite network. In top-down order, the first layer is a directed bipartite network (also a factor graph) containing the units/physical nodes of the system (circles) and their direct interactions (up-pointing triangles). The second layer is a directed tripartite network where the first-order modifications (squares) represent the action of the units on the direct interactions. The third layer is also a directed tripartite network, where the second-order modifications (down-pointing triangles) represent the action of the units on the first-order ones. The following layers are all directed tripartite networks representing the action of the units on the modifications of order n through modifications of order $(n+1)$, $n \geqslant 2$. The sub-network formed by the units is thus a multiplex intersecting all the layers, while any interaction or modification intersects two layers, the one in which it acts and the subsequent one, where it is potentially modified. Using a tripartite representation allows to distinguish whether the action of multiple units on an interaction or modification is separable or not. If separable, the overall action consists of the combination of independent pairwise modifications and is represented by multiple modifications, each one pointed by a single unit, incident on a modification at the next lower order (e.g., in the second layer, the second and third units affecting the first interaction via two independent first-order modifications). Otherwise, the overall action cannot be separated, thus we represent it as a unique modification, pointed by all the units producing it (e.g., in the third layer, the second and third units affecting the third first-order modification via a unique second-order modification).

5

CONCLUSIONS

This thesis has been devoted to deepening our understanding of how group interactions affect the propagation of biological and social contagions. To this aim, we refined existing theoretical approaches and introduced novel ones. As a result, we provided more accurate predictions, characterized new phenomena, and opened the path to unexplored modeling possibilities.

We organized our developments into two parts, according to two definitions of group, higher-order interaction. In the first part, Chapter 3, we considered interactions *directly* involving more than two agents at a time as an extension of pairwise, two-body interactions. In this context, we focused on higher-order contagions, namely, complex contagion processes—characterized by reinforcement mechanisms where transmission relies on the exposition to multiple, different sources—conveyed by group interactions, a particularly relevant representation for social contagions. Our results advanced the qualitative understanding of how group interactions and their structural organization—in the form of a hypergraph—affect the contagion dynamics. A key feature common to all the models proposed in Chapter 3 is the ability to account for local dynamical correlations. This allowed us to discover effects that elude approaches neglecting those correlations.

Section 3.1 revisited the behavior of higher-order contagions, revealing, in particular, that higher-order interactions affect the invasion threshold in quenched simplicial complexes, offering accurate predictions of such dependence. There, we realized that discrete-time formulations aiming at accounting for correlations within subsets of three or more nodes require those subsets to share at most one node. This led us to introduce the notion of edge-disjoint edge clique cover, for which we proposed a heuristic algorithm then used or adapted by other researchers working with higher-order networks [435, 436]. The algorithm also provides a way to generate hypergraphs by promoting cliques of a graph while respecting the requirement that any two maximal hyperedges share no more than one node. Such hypergraphs would serve as a reference to investigate the structural and dynamical implications of sharing more nodes, a mainly uncharted territory.

Section 3.2 extended the above analysis through a mean-field triadic approximation. Leveraging the mathematical tractability of the latter, we established a fundamental link between the behavior of higher-order contagions and the degree of overlap—

zero in linear hypergraphs, maximal in simplicial complexes—between 2- and 3-body interactions. Increasing the overlap has the double-edged effect of helping the contagion to thrive—lowering the invasion threshold—while hampering it once it becomes extensive—reducing the outbreak size. To the best of our knowledge, this result represents the first evidence of how the structural arrangement of interactions of different orders essentially affects the contagion dynamics. Careful consideration is thus necessary when choosing a higher-order representation for the system.

Section 3.3 closed the chapter by introducing a generalized approximate master equations framework accounting for dynamical correlations within and partially across groups, attaining unparalleled accuracy among mean-field models. Most importantly, the framework paves the way to investigate adaptive higher-order systems where agents can *strategically* leave and join groups in response to local dynamics (e.g., trying to avoid infectious environments)—a modeling possibility beyond the scope of earlier approaches. For both simple and complex contagions, we unveiled novel phenomena exclusively ascribable to the group structure, such as detrimental rewiring, where escaping from infectious groups makes the contagion spread even more, or a non-monotonic regime, where an intermediate rewiring rate performs worse than both a slow (yet nonzero) and a fast rewiring. Moreover, we identified and characterized two general optimal strategies for agents to hinder spreading, targeting either the dynamics—leaving infectious groups—or structure—leaving large groups. These findings result from just a first exploration of the adaptive mechanisms that can modeled with our framework, which is, additionally, by no means restricted to biological and social contagions. A future research line could be, for instance, the study of the evolution of group-structured systems where a group's functioning depends on the extent of cooperativeness of its members (e.g., informal organizations) [437–439], while the latter adapt their membership in search of more cooperative and functional environments.

The second part of the thesis, Chapter 4, built on the notion of interaction modification coming from ecology—where the presence of a third species changes the functional relation between two species —recognizing it as a more general higher-order mechanism typical of other complex systems, such as social and biological systems. The observable properties and behaviors of the units inducing the modification create a *context* that alters the direct interaction among other units. Such interaction would be generally different in another context and, therefore, cannot be considered separately from it. Context-dependency is thus an *indirect* form of higher-order interaction. Chapter 4 specifically explored the implications of context-dependent prophylactic behavior on epidemic spreading.

As a preliminary step, this required developing a theory to account for heterogeneous—assortative/homophilic versus disassortative/heterophilic—mixing between nodes of different types in a hypergraph. We provided such a theory in Sec. 4.1 in the form of a mean-field approximation mapping an ensemble of hypergraphs to a set of measurable mixing parameters. It serves as a first approximation to

investigate virtually any process for which mutable and contextual factors can be relevant. In the herein-considered problem, node types correspond to behavioral types representing individuals with different inclinations to adopt prophylactic measures.

Analyzing first the degenerate case of context-independent behavior (i.e., exclusively determined by the behavioral type, hence fixed at all times) on pairwise networks, Sec. 4.2 uncovered three dynamical regimes valid for any prophylactic measure unable to provide perfect, independent protection to the adopter, such as imperfect vaccines (Sec. 4.2.1) or, by construction, digital proximity tracing apps (Sec. 4.2.2). The three regimes are identified by how the final attack rate (or the peak of prevalence) depends on homophily. At odds with the previous belief, built on the assumption of perfect protection, that homophilic adoption is always detrimental to epidemic control, we proved that the opposite can also be true. Indeed, when infectivity is high compared to the effectiveness of the prophylactic tool, a clustered, homophilic adoption is needed to break the infection chains, or even the adopters would be scarcely protected. When infectivity is relatively low instead, since prophylaxis provides sufficient protection to even isolated adopters, these should mix as much as possible with non-adopters to dilute the contacts among the latter and offer them indirect protection—a generalized herd immunity. In such a case, which includes the perfect-protection scenario, homogeneous or even heterophilic mixing is preferable overall. Interestingly, in a situation of relatively intermediate protection, the ability of an epidemic to spread shows a non-monotonic dependence on mixing, implying eventually that different outbreak sizes (infectivities) are compatible with the same infectivity (outbreak size).

While such phenomenology remains valid in higher-order networks, interacting in a group lets individuals influence (either intentionally or not) the current behavior of the other co-present individuals, consequently modifying the local spreading between any two of them. This behavioral context-dependency results in the rich dynamics disclosed in Sec. 4.3. There, by examining groups composed of two and three individuals, we unveiled the significant impact that the interplay between epidemiological and sociological factors can exert on the outcome of an epidemic. Fixing the average contact rate and the average homophily across group sizes, we found that, as a general rule, it is beneficial (detrimental) for epidemic control gathering more frequently in triads than in pairs if prophylactic behavior is easy (hard) to induce, even more so when pairs are more (less) assortative than triads. Our developments demonstrated that incorporating the contextual character of behavior, shaped at the group level, can result in notable deviations from what might be anticipated based on pairwise information alone.

We concluded the chapter by suggesting in Sec. 4.4 a flexible network representation for systems holding context-dependent interactions of virtually any order in both the hierarchy of interaction modifications—first-order modifications of direct interactions, second-order modifications of first-order modifications, and so on—and in the number of involved units. Currently, no proper network representation can do it, not even to accommodate first-order modifications. Adopting the proposed representa-

tion might lead to reconsiderations about our understanding of the interplay between structure and dynamics for those systems characterized by context-dependent interactions. Ecological systems where first- and second-order modifications are either known or expected to exist [310, 312, 313, 433, 434] will provide the first ground to test whether or not a detailed representation as ours can offer new relevant insights.

FURTHER CONTRIBUTIONS

This thesis has only dealt with a few problems where higher-order interactions are relevant [88–90]. During the past four years, we ourselves have also investigated a few other topics (not all related to contagion processes). Building on the work by Hébert-Dufresne *et al.* [424], we used group-based approximate master equations to study the co-evolution of contagions and institutions [440]—a multi-level feedback process where a contagion spreads within and between groups represented as hyperedges, and institutions grow on top of the latter trying to extinguish the contagion. Public health departments or moderation teams on social media platforms are examples of group governances established to control harmful contagions. Nonetheless, each group is not isolated but is a piece in an ecology of other groups fighting the same contagion. Facing the dilemma of eradicating the contagion while attempting to contain the economic and social costs this entails, groups undergo an evolutionary process where they adapt their institutions based on their own and the other groups' successes and failures. Incorporating these institutional changes into the contagion dynamics revealed counterintuitive effects that cannot be predicted by regarding institutions and contagions as independent. For instance, a higher transmission rate can result in smaller outbreaks, while a slower institutional adaptation generally reduces outbreak size.

Still in the realm of evolutionary processes, we explored the effect of group interactions on the evolution of cooperative behavior. Cooperation—the act of covering some cost to benefit another individual—is observed in systems ranging from genomes to human societies [441–444], and not as a marginal phenomenon: it often profoundly shapes and makes even possible the existence of those systems. However, evolutionary game theory predicts that cooperative behavior cannot survive against selfish behavior in well-mixed populations [445, 446]. Hence, numerous mechanisms have been proposed to explain its ubiquity. Among those mechanisms, the so-called network reciprocity [447–455] helps cooperation to thrive based on two factors: a limited, recurrent set of interacting game opponents—the network neighbors—and a local adaptation mechanism through which an individual can change its behavior depending on how its neighbors in the network behave and perform. Whether this kind of reciprocity is possible also under group interactions was still unknown. Regarding group interactions as instances of a public goods game [456], we actually proved them to reinforce network reciprocity [265].

Lastly, we devoted some of our work to deepening the current understanding of synchronization in the presence of 3-body interactions. Synchronization is the fascinating phenomenon for which oscillators—units following oscillatory dynamics by themselves—spontaneously synchronize their frequencies due to their mutual influence [457–459]. Even though convincing empirical evidence for higher-order interactions in synchronization phenomena is still lacking, a considerable theoretical effort has been put in place lately [460–467], proving their validity at least as effective models. Higher-order interactions appear as small corrections in the so-called phase reduction of the dynamic equations governing the behavior of weakly coupled oscillators [457, 463]. In well-mixed systems of Kuramoto phase oscillators, the critical value of the pairwise coupling strength separating the incoherent (disordered) state from a partially coherent (ordered) one—where an extensive proportion of units oscillates phase-locked at the same frequency—is independent of 3-body (or higher) coupling strengths. However, whether some dependence on the latter can develop due to a nontrivial network structure—analogous to what we discovered for higher-order contagion processes—remains to be determined. Numerical simulations considering 3-body interactions on small hypergraphs ($\sim 10^2 - 10^3$ nodes) suggest such dependence exists. However, it is still unclear whether or not it is a finite-size effect and how specific structural properties affect it. Mapping a second-order Kuramoto model to a linear model of complex-valued phase oscillators [468], and carrying out a rank-one dimensional reduction of the resulting evolution matrix, preliminary (still unpublished) analytical predictions for the critical point appear in good agreement with numerical results, at least for small 3-body coupling strengths, and point out the degree correlation between 2- and 3-degree as responsible for an increase of the dependence of the critical point on the 3-body coupling.

WRAP-UP AND FUTURE DIRECTIONS

The study of systems with higher-order interactions is still in its infancy. When this thesis project started, few works on this subject existed in the complex system literature. Specific fields such as ecology or sociology were already familiar with higher-order interdependencies, as theoretical and experimental efforts have been put in place from much earlier. However, an interdisciplinary perspective capable of integrating knowledge from different scientific domains and recognizing known and new common mechanisms was almost absent. In four years, we witnessed an explosion of works that started the process of forging such a perspective. As a theoretical study of contagion processes relevant to biological and social systems, this thesis is our contribution to this endeavor. Beyond specific aspects, we can perhaps identify two general upshots from our findings about interacting within groups: it generally leads to stronger dynamical correlations, thus retaining the latter becomes of primary importance for a satisfactory description, often even at a qualitative level; and it can involve either direct many-body interactions, indirect interaction modifications, or multi-level

feedback loops, and all these types of non-pairwise interdependencies can be equally significant for the dynamics and potentially concurrent. Despite the advancements made in these years by the complex systems community, we just scraped the surface of the complexities hidden in the higher-order organization of systems observed at disparate scales.

Based on common limitations of current methodologies, we can delineate a few directions for future theoretical research about contagions and other processes in higher-order networks. For the sake of mathematical tractability, most of the current approaches make either explicitly (as we did ourselves in Secs. 3.1 and 3.2) or implicitly (respectively, in Sec. 3.3) the assumption that maximal cliques/hyperedges share at most one node. This is likely not the case for several systems; for instance, social groups such as face-to-face interactions [369, 410–412] or online communities [469–471], typically share two or more members. When transmission requires multiple simultaneous exposures, a larger share of nodes could be decisive for the local spillover between groups, thus for the contagion overall [291]. Furthermore, since people often coordinate their decisions, accounting for shared group members may be especially relevant when adaptive behavior is allowed, as is frequently the case within social and economic systems. Despite recent advancements [291, 292], the potential dynamical implications of shared group members remain largely unknown. Preliminary analysis suggests that approximate master equations might be good candidates for this.

Furthermore, we lack a general method to tell whether a contagion is simple or complex (and in which sense) from observed time series, albeit promising steps have been recently made. Cencetti *et al.* [472] proposed a heuristic classifier based on the order in which nodes are infected and on its correlations with their local topology (2- and 3-degree). The method shows remarkable performances for SI-like processes. Nonetheless, a classifier trained on a particular ensemble of structures is not guaranteed to perform satisfactorily on other structures, as its performance is strictly related to network properties. Moreover, the method requires full knowledge of the single nodes' degrees and their order of infection. Other research, however, has pointed out that distinguishing simple and complex contagion becomes challenging, if not impossible when information about transmission settings (e.g., group size or environmental conditions) is imperfect [114] as well as when multiple contagious entities (e.g., ideas or memes) are competing [473]. The inverse problem of retrieving the structure from observed dynamics [474–478] is also quite open for higher-order networks [479–483]. To this end, one could perhaps leverage the dependence unveiled in Sec. 3.2 of the contagion dynamics on the (average) degree of interaction overlap.

Finally, the number of possible state configurations and microscopic processes admitted by a group undergoes a combinatorial explosion with increasing the group's size. Thus, a method to make a priori estimations of the importance of interactions of higher and higher order is paramount, for it would inform us of the error made when neglecting interactions above a particular order. Finding such a method is among the most critical and challenging unresolved problems in the field.

FINAL CONSIDERATIONS

The ontological nature of (social) groups—What does identify a group? Can group agency always be reduced to a collection of individual agencies?—is a fervent matter of current philosophical debate [484–489]. Nonetheless, the recent developments brought by the complex systems community, and even earlier by specialized research fields, make it hard to deny the usefulness of the concept of group to describe and comprehend several natural systems. Besides, accounting for the higher-order organization of such systems adds a thick layer—of the thickest ones—to the already high mountain of complexities we hope to climb. Dense substructures create stronger dynamical correlations while increasing their size leads to combinatorial barriers. As we have directly seen in this thesis, higher-order structures thus call for more sophisticated models. In turn, it becomes increasingly harder to analyze and even define such models and understand the mechanistic origin of their predictions.

On a personal note, I want here to emphasize that the defining purpose of science is to generate *understanding*. Tools able to provide detailed descriptions and accurate predictions are crucial to making scientific progress, but they have no intrinsic scientific value. As a human endeavor, science is and must remain a *science for humans*, interpretable and acknowledgeable by humans. Paraphrasing Heinz von Foerster, father of second-order cybernetics, complexity is a property of the relation observer-observed:[1] the simplifications we make to understand some aspects of a system—which become *simple*—also prevent us from accessing others—which remain *complex*. Ultimately, this trade-off between interpretability and accuracy implies, in my opinion, that a limit to the comprehension of the world we can achieve does exist. Based on our short journey in modeling higher-order systems, I feel that limit to be closer—but not close! Plenty of exciting problems are out there waiting to entertain us. Boring times are far off.

1 Translated from an Edgar Morin's essay [1].

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

Appendices

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

A

BASIC METHODS

A.1 LINEAR STABILITY OF THE INACTIVE STATE

A linear stability analysis is a standard technique to obtain conditions for the stability of the stationary orbits of a dynamical system. For the purposes of this thesis, we focus in particular on fixed points. We are interested on the (local) stability of the inactive state for contagion processes, here understood as the absorbing state where all the units of the system are susceptible. The critical point at which the inactive state looses stability is known—also—as the *invasion threshold*. According to linear stability analysis, the threshold is given by the value of the infection rate—or of one of the infection rates, if many—at which the spectrum of the Jacobian matrix associated to the linearized system (around the inactive state) develops at least an eigenvalue with either nonnegative real part, if in continuous time, or with unitary modulus, if in discrete time. First, we breifly summarize the stability analysis in both cases, as we used either continuous and discrete time models. Then, we reconcile the threshold condition provided by the Jacobian matrix with the one defined by the so-called *basic reproduction number*, which is one of the most relevant quantities to understand the growth dynamics of populations, of which contagion processes are a special case.

Consider a dynamical system described by the following first-order differential equation,

$$\dot{x}(t) = f(x(t)) \,, \tag{A.1}$$

defined by the smooth map $f : \mathbb{R}^n \to \mathbb{R}^n$. Suppose that x_0 is an equilibrium, i.e., $f(x_0) = 0$. Expanding Eq. (A.1) around x_0, at first order in the infinitesimal displacement $\delta x(t) = x(t) - x_0$, one gets

$$\dot{\delta x}(t) = \boldsymbol{J}\delta x(t) \,, \tag{A.2}$$

where $\boldsymbol{J}$, of elements $J_{ij} = \partial f_i / \partial x_j(x_0)$, is the Jacobian matrix. Denoted with ϕ_i the eigenvector of $\boldsymbol{J}$ of eigenvalue λ_i, ϕ_i solves Eq. (A.2) as $\phi_i(t) = \phi_i(0)e^{\lambda_i t}$, given the initial condition $\delta x(0) = x(0) - x_0$ at $t = 0$. Being Eq. (A.2) linear, any displacement $\delta x(t)$—hence trajectory $x(t)$—solution of Eq. (A.2) can be expressed as a linear combination of the eigenvectors of $\boldsymbol{J}$, therefore as $\delta x(t) = \sum_{i=1}^{n} c_i e^{\lambda_i t}$. Supposing the eigenvectors have been labeled so that $\mathrm{Re}(\lambda \geqslant_1) \geqslant \mathrm{Re}(\lambda_2) \geqslant \cdots \geqslant \mathrm{Re}(\lambda_n)$, $\delta x(t)$ is dominated by

the largest eigenvalue(s) at large t. Thus, if $\mathrm{Re}(\lambda_1) < 0$, $\delta x(t) \to 0$ for $t \to \infty$, implying that x_0 is an asymptotically stable local equilibrium. If instead $\mathrm{Re}(\lambda_1) > 0$, $\delta x(t)$ diverges and x_0 is unstable. Therefore $\mathrm{Re}(\lambda_1) = 0$ marks the critical point at which x_0 changes stability. For a contagion process, $\mathrm{Re}(\lambda_1)$ is a function of the infection rate(s). Then, given the inactive state $x_0 = 0$, solving $\mathrm{Re}(\lambda_1) = 0$ with respect to (one of) the infection rate(s) provides the invasion threshold (expressed as a value of that infection rate).

Similarly, let us consider now a system evolving according to a discrete map,

$$x(t+1) = f(x(t)) \,, \tag{A.3}$$

$t \in \mathbb{N}$; such that $f(x_0) = x_0$, with $x_0 = 0$. The linearization gives

$$x(t+1) = \boldsymbol{J} x(t) \,, \tag{A.4}$$

If ϕ_i is an eigenvector of $\boldsymbol{J}$, then $\phi_i(t+1) = \lambda_i \phi_i(t)$, which is solved by $\phi_i(t) = \phi_i(0) {\lambda_i}^t$. The generic solution $x(t)$ of Eq. (A.4) thus reads $x(t) = \sum_{i=1}^{n} c_i {\lambda_i}^t$. This goes to zero over time if and only if the eigenvalue of largest modulus lies inside the unit circle, i.e., if $|\lambda_1| < 1$. The condition $|\lambda_1| = 1$ marks the critical point at which x_0 changes stability and therefore provides the invasion threshold.

A.1.1 *Basic reproduction number*

As long as we are exclusively interested on getting a threshold condition for some parameter of the system (e.g., an infection rate) determining the stability of an equilibrium, such condition is always implicit in $\mathrm{Re}(\lambda_1) = 0$ or $|\lambda_1| = 1$. However, we could be interested in another threshold condition which, while being equivalent to the others, has a physical meaning of interest. For instance, when studying the growth dynamics of some population, as in ecology and epidemiology. A species grows through the reproduction of its individuals; an epidemic grows through the spread of a pathogen, reproducing itself from host to host. In both cases, a basic physical quantity is how many times a typical individual/host reproduces over its lifetime; or, in other words, the expected number of secondary individuals/hosts generated by a primary individual/host (the precise meaning of "typical" and "expected" is clarified later in this section). In particular, if the initial amount of cases (individuals/hosts) in the system is negligible, that number is called *basic reproduction number* and is traditionally denoted as R_0. Whether or not an initial negligible number of cases grows to occupy a finite proportion of the system depends on R_0 [195]. If $R_0 > 1$, a primary case produces more than one secondary case, giving rise to an exponential growth of cases; if $R_0 < 1$, instead, it produces less than one secondary case, leading the cases to exponentially disappear. Therefore, the condition $R_0 = 1$ decides the local stability of the inactive state – understood as the state with no cases. In those sections in which we

explicitly refer to epidemic modeling, the stability condition is always expressed in terms of R_0.

Except for very simple models, however, finding the expression for R_0 directly from $\mathrm{Re}(\lambda_1) = 0$ or $|\lambda_1| = 1$ is generally unfeasible. Incorporating few more heterogeneities (e.g., by grouping individuals with respect to a given trait) or individual states (e.g., the exposed state, where an individual is infected but still not infectious) in the system, easily prevents extracting R_0 from the condition provided by the Jacobian spectrum. In 1990, Diekmann et al. [196], conceived the *next generation matrix* method, which allows to compute R_0 for any transmission (or population dynamics) model under a set of fairly weak assumptions. In the following, we refer to the later formulation given by van der Driessche et al. [490, 491], which further developed the theory for compartmental models.

The key idea is to split the Jacobian matrix as $\boldsymbol{J} = \boldsymbol{F} - \boldsymbol{V}$, where F_{ij} gives the rate at which individuals in state i are generated through infection by individuals in state j, while V_{ij} is the rate at which individuals, initially in state j, exit state i. One of the assumptions above imposes that infections can only be produced by infected individuals within the system. Therefore, once in a state with no infected individuals (as the inactive state), the system can only transition to states with no infected individuals, meaning the set of infection-free states is invariant. A direct consequence of this is that the linearized equations for the infected compartments fully decouple from the remaining equations, and the stability of the inactive state is thus governed by the former only. We can therefore restrict $\boldsymbol{J}$—hence $\boldsymbol{F}$ and $\boldsymbol{V}$—to the m infected compartments and forget about the $n - m$ remaining ones. Then, under the aforementioned assumptions and through a series of results from linear algebra, it can be proved that $R_0 = \Lambda_{\max}(\boldsymbol{F}\boldsymbol{V}^{-1})$, being $\boldsymbol{F}\boldsymbol{V}^{-1}$ the so-called next generation matrix (NGM) and $\Lambda_{\max}(\boldsymbol{F}\boldsymbol{V}^{-1})$ its dominant eigenvalue (what before we denoted as λ_1). The meaning associated to $\boldsymbol{F}$ and $\boldsymbol{V}^{-1}$ clarifies why R_0 precisely derives from the matrix $\boldsymbol{F}\boldsymbol{V}^{-1}$. In particular, $[\boldsymbol{V}^{-1}]_{ij}$ is the average time that an individual initially in the infected state j spends in the infected state i. $[\boldsymbol{F}\boldsymbol{V}^{-1}]_{ij}$ thus gives the average number of individuals in state i that an individual initially in state j generates during its lifetime (i.e., while in any of the infected states in its path from state j to state i). Accordingly, the eigenvector associated to R_0 can be thought of as the distribution of individuals among the infected states that maximizes the number of secondary infections per generation. Being R_0 the dominant eigenvalue, any initial distribution will converge to the former over generations. R_0 thus represents the expected number of secondary cases per generation produced after enough generations, assumed the number of cases is still negligible (i.e., the linearized system is still valid).

While we refer to van der Driessche et al. [490, 491] for the general proof, which we do not report here, we give a fast proof which is sufficient for the simple models considered in this thesis. From the decomposition of the Jacobian matrix, the linearized equation reads

$$\dot{\delta x}(t) = (\boldsymbol{F} - \boldsymbol{V})\delta x(t) \,. \tag{A.5}$$

Supposed $\boldsymbol{V}$ is invertible, applying $\boldsymbol{V}^{-1}$ to both sides of the equation and assuming $\boldsymbol{F}\boldsymbol{V}^{-1} = \boldsymbol{V}^{-1}\boldsymbol{F}$, we get

$$\boldsymbol{V}^{-1}\dot{\delta x}(t) = (\boldsymbol{F}\boldsymbol{V}^{-1} - \mathbb{1})\delta x(t) \,, \tag{A.6}$$

being $\mathbb{1}$ the identity matrix. The condition $\boldsymbol{F}\boldsymbol{V}^{-1} = \boldsymbol{V}^{-1}\boldsymbol{F}$ puts strong constraints on the form of $\boldsymbol{F}$ and $\boldsymbol{V}$, hence on the complexity of the models fulfilling it. Perhaps, the only sensible case is when $\boldsymbol{V}^{-1}$ is proportional to the identity (as it is $\boldsymbol{V}$, which is indeed invertible), say, $\boldsymbol{V}^{-1} = \mu^{-1}\mathbb{1}$. The latter means assuming that (i) there are no spontaneous (i.e., not caused by infection) transitions between infected states (no off-diagonal entries) and (ii) these states are all exited at the same rate. Condition (i), in particular, implies that a given individual can go through a single infected compartment only, and therefore μ represents the recovery rate from that compartment. Conditions (i) and (ii) are satisfied by basic models such as the susceptible-infected (SI), the susceptible-infected-recovered (SIR), the susceptible-infected-susceptible (SIS), and the susceptible-infected-recovered-susceptible (SIRS), with the additional freedom of grouping individuals with respect to some fixed trait, as long as this trait affects transmission but not recovery (if any). Going back to Eq. (A.6), given $\boldsymbol{V}^{-1} = \mu^{-1}\mathbb{1}$, we can absorb μ in the time parameter by redefining t through the map $t \mapsto \mu t$. The new temporal parameter is a generational time, for a unit corresponds to the average time spent by an individual in the infection state, i.e., to the duration of a generation (in general, we can think of $\boldsymbol{V}^{-1}$ as providing a different rescaled time for each transition an individual can experience while infected, i.e., while going through an entire generation). Consequently, Eq. (A.6) becomes

$$\dot{\delta x}(t) = (\boldsymbol{F}\boldsymbol{V}^{-1} - \mathbb{1})\delta x(t) \,, \tag{A.7}$$

from which $\boldsymbol{F}\boldsymbol{V}^{-1} - \mathbb{1} = \mu^{-1}\boldsymbol{J}$. The threshold condition for the inactive state, $\Lambda_{\max}(\boldsymbol{J}) = 0$, is thus equivalent to $\Lambda_{\max}(\boldsymbol{F}\boldsymbol{V}^{-1}) = 1$, consistently with the general result. In conclusion,

$$R_0 = \Lambda_{\max}(\boldsymbol{F}\boldsymbol{V}^{-1}) = 1 + \mu^{-1}\Lambda_{\max}(\boldsymbol{J}) \tag{A.8}$$

in this case.

Following analogous steps, this is easily extended to discrete maps (Eq. (A.3)). In fact, as demonstrated by van der Driessche et al. [492], the general proof holds true also in this case, establishing $R_0 = \Lambda_{\max}(\mathrm{NGM})$, with $\mathrm{NGM} = \boldsymbol{F}(\mathbb{1} - \boldsymbol{T})^{-1}$. The matrix $\boldsymbol{T}$ is related to the matrix $\boldsymbol{V}$ of the continuous time model through $\boldsymbol{T} = \mathbb{1} - \boldsymbol{V}$. The reason is that, compared to the continuous time case, the Jacobian matrix additionally includes an identity matrix that accounts for those individuals that persist in the same state from one time step to the next. Since $\boldsymbol{F}$ exclusively accounts for the generation of new infections (that survive the time step), all other processes, including the persistence in the same state, are included in the matrix $\boldsymbol{T}$, so that $\boldsymbol{J} = \boldsymbol{F} + \boldsymbol{T}$. In particular, the case $\boldsymbol{V} = \mu\mathbb{1}$ yields $\boldsymbol{T} = (1-\mu)\mathbb{1}$, from which $(\mathbb{1} - \boldsymbol{T})^{-1} = \mu^{-1}\mathbb{1} = \boldsymbol{V}^{-1}$. From

this, it immediately follows that $R_0 = 1 + \mu^{-1}\Lambda_{\max}(\boldsymbol{J} - \mathbb{1}) = 1 + \mu^{-1}(\Lambda_{\max}(\boldsymbol{J}) - 1)$. Since switching to the discrete-time model requires adding the identity matrix to the Jacobian matrix of the respective continuous-time model, this equation is identical to Eq. (A.8) once $\boldsymbol{J}$ is written explicitly.

A.2 QUASI-STATIONARY STATE METHOD

The quasi-stationary state (QS) method [284, 285, 493] allows for the estimation in finite-sized systems of the long-time behavior of processes with an absorbing state. In a finite-sized system, the absorbing state is the unique actual equilibrium, as there is a finite probability—however small this can be—that statistical fluctuations lead to it. Therefore, even when crossed a critical point and the absorbing state is unstable, it is still reachable. Differently put, although in the active phase the expected survival time (out of the absorbing state) increases exponentially with the system size, it remains finite, unless the thermodynamic limit ($N \to \infty$) is considered. Consequently, to recover the non-trivial long-time state distribution (the QS distribution), one must refer to the so-called *surviving averages*, which are averages restricted to runs that, after the system relaxed, have not been absorbed. Indicating with $P_\sigma(t)$ the probability for the system to be in the active (i.e., nonabsorbing) state σ at time t, the QS distribution consists of the time-independent probabilities $\bar{P}_\sigma = \lim_{t\to\infty} P_\sigma(t)/P_{\mathrm{s}}(t)$, where $P_{\mathrm{s}}(t)$ is the probability that the system survived up to time t. These are probabilities for the active states conditioned on survival. Accordingly, in the long-time QS regime, $P_\sigma(t) \approx P_{\mathrm{s}}(t)\bar{P}_\sigma$. This factorization says that, at long times, the survival probability is redistributed among all the active states proportionally to their QS probabilities.

Estimating the QS distribution requires retaining only those Monte Carlo simulations that resulted in some active state, so to compute surviving averages. Unfortunately, especially near a critical point, the majority of the simulations ends up in the absorbing state either before or soon after the transient period of the dynamics, implying that a very large numbers of Monte Carlo runs must be performed to get a sufficient statistics for the surviving configurations. The factorization above, however, suggests an alternative, more efficient way to find the QS distribution. In fact, Dickmann et al. [284] showed that, given a Markov process with a finite set of states, the respective QS distribution is the stationary solution of a macroscopic equation of motion given by the master equation associated to the original process, plus a nonlinear memory term (which thus makes the equation no longer a master equation). This additional term takes all the flux of probability that would go into the absorbing state and redirects it into the active subspace, specifically proportioning to each active state a fraction of the flux equal to the probability of that state. The QS distribution of the original process is thus the stationary distribution of the modified process. This observation provides the basis for an alternative simulation strategy that provides an unbiased estimation of the QS distribution, avoiding in this way the dispendious procedure of running the process many times. Introduced by de Oliveira et al. [493] and

then extended to complex networks by Ferreira et al. [285], this strategy implements the modified version of the original stochastic process, whose macroscopic equation is solved by the QS distribution. More precisely, each time the system falls into the absorbing state, this configuration is replaced with one active configuration stored in a list of active states. This list, at each time step, is updated with a certain probability by replacing one element of the list with the current active configuration. The longer is the time the system spends in a given configuration, the larger will be the number of respective copies in the list. Accordingly, during a sufficiently large time interval, the frequency of appearance of a configuration in the list will converge to its probability of being visited in the QS state. In other words, by letting the system evolve enough time past the transient period, the list eventually converges to an unbiased sample of the QS distribution. Obtained the latter, we can then immediately calculate the QS values of any macroscopic quantity of interest, e.g., the order parameter.

To notice that not all the processes with absorbing states admit a QS distribution, but it is reasonable to assume it exists for many processes [284]. This is the case for the SIS-like processes considered in this thesis, for which the QS method has been used.

A.3 SIMPLICIAL EPIDEMIC LINK EQUATIONS

To characterize the simplicial contagion process they introduced, Iacopini *et al.* [199] used a standard homogeneous mean-field model to approximate the dynamics in simplicial 2-complexes. As we have seen in Sec. 2.2.3.2, thanks to its simplicity, the mean-field approach allows for the analytical characterization of the qualitative phenomenology. On the flip side, however, the mean-field model gives poor quantitative predictions in quenched structures [199, 282]. In an attempt to fill this gap, Matamalas *et al.* [230] extended the MMCA [222] and the ELE [229] to simplicial 2-complexes. This resulted in the so-called 'simplicial ELE', from which the 'simplicial MMCA' is simply recovered by neglecting dynamical correlations.

As the original version, the simplicial ELE preserve the dynamical correlations within pairs of neighboring nodes by following the states of the L edges (1-simplices) in the simplicial complex. Assuming the equations for the edges can be marginalized

to recover the equations for the nodes, the model describes the simplicial contagion process through the following set of $N + L$ equations,

$$P_i^I(t+1) = (1-\mu)P_i^I(t) + \left[1 - q_i(t)q_i^{(2)}(t)\right] P_i^S(t) \,, \tag{A.9}$$

$$\begin{aligned} P_{ij}^{II}(t+1) = &\left[1 - q_{i(j)}(t)q_{i(j)}^{(2)}(t)u_{ij}^{(2)}(t)\right]\left[1 - q_{j(i)}(t)q_{j(i)}^{(2)}(t)u_{ji}^{(2)}(t)\right] P_{ij}^{SS}(t) \\ &+ (1-\mu)\left[1 - (1-\beta^{(1)})q_{i(j)}(t)q_{i(j)}^{(2)}(t)\right] P_{ij}^{SI}(t) \\ &+ (1-\mu)\left[1 - (1-\beta^{(1)})q_{j(i)}(t)q_{j(i)}^{(2)}(t)\right] P_{ij}^{IS}(t) \\ &+ (1-\mu)^2 P_{ij}^{II}(t) \,, \end{aligned} \tag{A.10}$$

where $q_i(t)$ and $q_{i(j)}(t)$ are as in Eqs. (2.33) and (2.36), while

$$q_i^{(2)}(t) = \prod_{l<m}\left(1 - A_{ilm}^{(2)}\beta^{(1)}\frac{P_{ilm}^{SII}(t)}{P_i^S(t)}\right) \,, \tag{A.11}$$

$$u_{ij}^{(2)}(t) = \prod_{l}\left(1 - A_{ijl}^{(2)}\beta^{(2)}\frac{P_{ijl}^{SII}(t)}{P_{ij}^{SI}(t)}\right) \,, \tag{A.12}$$

are the probabilities that node i is not infected via edges and triangles shared with node j, respectively. Moreover, $q_{i(j)}^{(2)}(t)$ coincides with $q_i^{(2)}(t)$ except for removing the triangles including node j. The element (i,j,l) of the adjacency tensor $\boldsymbol{A^{(2)}}$ equals 1 if there is a triangle incident on nodes i, j and l, and 0 otherwise. The equations are then closed by factorizing the 3-node probabilities in a symmetric product of 2-node probabilities. Dynamical correlations among three or more nodes are thus neglected.

Independently from the closure, comparing Eqs. (A.9) and (A.10) with Eqs. (2.27) and (2.35), we see that the simplicial ELE generalize the ELE to include the 3-body interactions via the simple substitutions $q_i \to q_i q_i^{(2)}$, et similia. This apparently innocuous extension hides, however, a subtlety that makes Eqs. (A.9) and (A.10) inconsistent. This, eventually, prevents the computation of the invasion threshold. To reveal where the inconsistency lies, it is sufficient to consider a single, isolated triangle. Say i, j and l, are the three nodes involved. According to Eq. (A.9), if node i is susceptible, it gets infected with probability

$$1 - \left(1 - \beta^{(1)}P_{j|i}^{I|S}(t)\right)\left(1 - \beta^{(1)}P_{l|i}^{I|S}(t)\right)\left(1 - \beta^{(2)}P_{jl|i}^{II|S}(t)\right) \,. \tag{A.13}$$

Take now $\beta^{(1)} = 1$. Being the pairwise transmission certain in this case, the probability for i to get infected should simply coincide with the probability $1 - P_{jl|i}^{SS|S}(t)$ that at least one among j and l is infected. Instead, Eq. (A.13) becomes $1 - P_{j|i}^{S|S}(t)P_{l|i}^{S|S}(t)\left(1 - \beta^{(2)}P_{jl|i}^{II|S}(t)\right)$, which is correct only for $\beta^{(2)} = 0$, i.e., when transmission is only pairwise and the model reduces to the original ELE. For $\beta^{(2)} \neq 0$,

there is an extra term[1] which is clearly inconsistent, as it wants j and l to be susceptible and infected at the same time. The error comes from considering the two events 'node j (l) is infected *and* infects node i through an edge' and 'nodes j and l are infected *and* infect node i through a triangle' as independent. They are not, as they indeed rely on the current states of the same nodes, j and l. Supposed these are in a certain state (with the respective probability), the rest of the stochasticity exclusively comes from the transmission processes. For instance, supposing j and l are infected (which occurs with probability $P_{jl|i}^{II|S}(t)$), whether or not the infections via either the edges or the triangle occur only depends on the probabilities $\beta^{(1)}$ and $\beta^{(2)}$. Equation (A.13), instead, tacitly assumes that the state of the nodes is redrawn (with the respective probability) each time a transmission process is considered, eventually leading to the inconsistency that a node can be simultaneously found in more states. A consequence of this is that Eqs. (A.9) and (A.10) violate the marginalization condition $P_i^{\sigma_i} = P_{ij}^{\sigma_i S} + P_{ij}^{\sigma_i I}$. This, in turn, comes with two related problems. First, the invasion threshold cannot be computed. Second, contrarily to what assumed, the $3L$ equations for the state of the edges cannot be reduced to L, leading to quantitative errors. Nonetheless, the simplicial ELE manage to provide generally good predictions, as confirmed by comparison with numerical simulations [230, 282].

Going back to the example of the isolated triangle, the consistent expression for the probability for node i to get infected, reads

$$\beta^{(1)}\left(P_{jk|i}^{IS|S}+P_{jk|i}^{SI|S}\right)+\left[1-\left(1-\beta^{(1)}\right)^2\left(1-\beta^{(2)}\right)\right]P_{jk|i}^{II|S}\,. \tag{A.14}$$

This is a combination of mutually exclusive events, as it should be, and solves the issues of the simplicial ELE.

A.4 PROBABILITY GENERATING FUNCTIONS

Probability generating functions (PGFs) are a key concept for percolation-based methods and approximate master equations, discussed in Sec. 2.3.3. We make explicit use of them in Sec. 3.3. Regarding networks as combinatorial objects, describing processes on them usually boils down to counting things—nodes or edges of a certain type, paths, motifs, etc.

A PGF is a representation of the probability distribution of discrete and non-negative random variables in the form of a power series. As such, it fully encodes the distribution. Using a PGF becomes particularly useful when an algorithm to find the distribution of some quantity we want to count is unavailable—let alone an explicit formula for it. Instead, while an analytical expression for the PGF can be found in some cases, from which a closed-form expression for the probability distribution

1 Being positive, it is partially responsible for the overestimation of the prevalence made by the simplicial ELE, as shown in Sec. 3.1. The error grows with $\beta^{(2)}$.

could be derived, the distribution can be always retrieved from the PGF (see Appendix A.4.1). Therefore, at least when we just want the numerical values of the distribution, all we need to know is to how define the PGF for our counting problem correctly.

Let us suppose we have m discrete and non-negative random variables, $\{n_i\}_{i=1,\dots,m}$, with joint probability mass function $p_{n_1,\dots,n_m}$. The associated (multinomial) PGF is defined as

$$G(x_1,\dots,x_m) = \sum_{n_1,\dots,n_m=0}^{\infty} p_{n_1,\dots,n_m} x_1^{n_1}\cdots x_m^{n_m} \,. \tag{A.15}$$

That is, a PGF is a series that encodes the generic element $p_{n_1,\dots,n_m}$ of the distribution as the coefficient of the monomial $x_1^{n_1}\cdots x_m^{n_m}$. The powers in the monomial thus represent the counts. In the words of Herbert S. Wilf, "probability generating functions are a bridge between discrete mathematics, on the one hand, and continuous analysis, on the other" [494].

Let us list below a set of basic properties of PGFs. First, supposing we got an expression for the PGF, i.e., a function of the m arguments $x_1,\dots,x_m$, it is immediate to see that the coefficient (probability) $p_{n_1,\dots,n_m}$ is found through differentiation as

$$p_{n_1,\dots,n_m} = \left(\prod_{i=1}^{m} \frac{1}{n_i!}\frac{\partial^{n_i}}{\partial {x_i}^{n_i}}\right) G(x_1,\dots,x_m)\bigg|_{x_1,\dots,x_m=0} \,. \tag{A.16}$$

Second, $G(1,\dots,1) = \sum_{n_1,\dots,n_m=0}^{\infty} p_{n_1,\dots,n_m} = 1$ expresses the normalization of the probability mass function. Third, the mixed moment $\langle {n_1}^{k_1}\cdots {n_m}^{k_m}\rangle$ is given by

$$\langle {n_1}^{k_1}\cdots {n_m}^{k_m}\rangle = \prod_{i=1}^{m}\left(x_i\frac{\partial}{\partial x_i}\right)^{k_i} G(x_1,\dots,x_m)\bigg|_{x_1,\dots,x_m=1} \,. \tag{A.17}$$

Very simply, applying $\partial/\partial x_i$ provides a n_i from ${x_i}^{n_i}$, which becomes ${x_i}^{n_i-1}$. Multiplying the latter by x_i gives ${x_i}^{n_i}$ back. Repeating this operation provides any power of n_i we wish.

The two properties below are essential for our purposes. Suppose that the m random variables are counting all the same quantity, so that we can actually sum them. For example, imagine we want to know how many infected nodes there are across the m groups a node belongs to. Then, Eq. (A.15) takes the simpler form

$$G(x) = \sum_{n_1,\dots,n_m=0}^{\infty} p_{n_1,\dots,n_m} x^{\sum_{i=1}^{m} n_i} \tag{A.18a}$$

$$= \sum_{n=0}^{\infty} \underbrace{\sum_{n_1,\dots,n_m=0}^{\infty} p_{n_1,\dots,n_m}\delta_{n,\sum_{i=1}^{m} n_i}}_{q_n} x^n = \sum_{n=0}^{\infty} q_n x^n \,, \tag{A.18b}$$

being $\delta_{a,b}$ the Kronecker delta. Since we are only keeping information about the sum of the m random variables, not about each of them, the associated PGF is a function of

a single variable. Equation (A.18b) makes this fact explicit. If the m variables are statistically independent, then $p_{n_1,\dots,n_m} = \prod_{i=1}^{m} p_{n_i}$, and from Eq. (A.18a) it immediately follows that the PGF can be factorized too, i.e.,

$$G(x) = \prod_{i=1}^{m} \underbrace{\sum_{n_i=0}^{\infty} p_{n_i} x^{n_i}}_{G_i(x)} = \prod_{i=1}^{m} G_i(x) \,. \tag{A.19}$$

The first key property thus states the following: the PGF of the sum of independent random variables equals the product of the PGFs of the single variables. As a special case, if the variables are also identically distributed, then $G_i(x) = g(x)$, $\forall i$, and we obtain $G(x) = (g(x))^m$.

Assume now that also the number m of random variables is itself a random variable. Going back to the previous example, it means that we don't know how many groups a node belongs to, we just know the distribution of this number. Denoted with $p^{(m)}_{n_1,\dots,n_m}$ the joint probability mass function for the entire set of $m+1$ variables, the associated PGF reads

$$G(x) = \sum_{m=0}^{\infty} \sum_{n_1,\dots,n_m=0}^{\infty} p^{(m)}_{n_1,\dots,n_m} x^{\sum_{i=1}^{m} n_i} \,. \tag{A.20}$$

If m is statistically independent from the other variables, then $p^{(m)}_{n_1,\dots,n_m} = h_m p_{n_1,\dots,n_m}$, being $\{h_m\}$ the distribution for m. Equation (A.20) becomes

$$G(x) = \sum_{m=0}^{\infty} h_m \sum_{n_1,\dots,n_m=0}^{\infty} p_{n_1,\dots,n_m} x^{\sum_{i=1}^{m} n_i} \,, \tag{A.21}$$

which is nothing but an average of PGFs over the distribution $\{h_m\}$. In particular, if the m random variables are all independent and identically distributed (i.i.d) with PGF $g(x)$, using Eq. (A.19), Eq. (A.21) becomes

$$G(x) = \sum_{m=0}^{\infty} h_m (g(x))^m = H\left(g(x)\right) \,, \tag{A.22}$$

where $H(x)$ is the PGF associated to $\{h_m\}$. Equation (A.22) states the second key property: the PGF of the sum of a *variable* number of i.i.d random variables is found by the composition of the PGF of that number with the PGF associated to each of those variables. Consistently, if the value of the variable m was known to be k (that is, $h_m = \delta_{m,k}$), then Eq. (A.22) would reduce to $G(x) = (g(x))^k$. Consider again the previous example. Assume that the states of nodes in different groups are i.i.d. variables and that they are also independent from the number of groups a node belongs to. Then, the distribution for the total number of infected neighbors of a random node is characterized by a PGF constructed by composing the PGF of the number of groups to which a random node belongs to (say $H(x)$) with the PGF of the number of infected nodes in a random group (say $g(x)$).

A.4.1 *Retrieving the numerical values of the coefficients of a PGF*

Recovering the probability distribution by repeated differentiation as in Eq. (A.16) is impractical, as the number of terms for a partial derivative of degree n requires B_n terms, being B_n the Bell number[2]. A more efficient way to find $p_{n_1,\dots,n_m}$ is through a discrete Fourier transform, as

$$p_{n_1,\dots,n_m} = \frac{1}{\prod_{i=1}^{m} N_i} \sum_{u_1=0}^{N_1-1} \cdots \sum_{u_m=0}^{N_m-1} G(e^{2\pi i u_1/N_1}, \dots, e^{2\pi i u_m/N_m}) e^{-2\pi i \sum_{i=1}^{m} n_i u_i/N_i} \,, \quad \text{(A.23)}$$

where N_i is the maximal count we can have for the i-th variable—anything we can count in the physical world comes limited in number. Equation (A.23) is easily verified to be an identity by inserting Eq. (A.15) and using $\sum_{u=0}^{N-1} e^{2\pi i (r-s)u/N} = N\delta_{r,s}$. Computing the entire set $\{p_{n_1,\dots,n_m}\}$ of $N_1 \times \cdots \times N_m$ coefficients using Eq. (A.23) directly, takes $\mathcal{O}(N_1^2 \cdots N_m^2)$ operations, which is still computationally expensive. Using the fast Fourier transform, however, the complexity greatly reduces to $\mathcal{O}((N_1 \log N_1) \cdots (N_m \log N_m))$. We compute FFTs via the FFTW.jl Julia package [496], which provides bindings to the FFTW library [497].

2 $B_n = 1, 1, 2, 5, 15, 52, 203, 877, 4140, \dots$ for $n = 0, 1, 2, 3, 4, 5, 6, 7, 8, \dots$ [495].

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

B

METHODS FOR SEC. 3

B.1 METHODS FOR SEC. 3.1

B.1.1 *Derivation of the invasion threshold predicted by the MECLE*

In this section, we derive the invasion threshold $\beta_{\rm cr}^{(1)}$, defined as the value of $\beta^{(1)}$ at which the inactive state looses stability. Above this threshold, the system reaches an active state. In the presence of a bi-stable region, $\beta_{\rm cr}^{(1)}$ identifies the rightmost transition.

We linearize Eq. (3.5) by regarding of the same order ϵ, with $\epsilon \to 0$, all the probabilities of states with at least one infected node, i.e., $P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g} \in \mathcal{O}(\epsilon)$ if $\exists k : \sigma_{i_k} = I$. Consequently, $P^{S\dots S}_{i_1\dots i_n,g} \in 1-\mathcal{O}(\epsilon)$. Without this assumption, higher-order dynamical correlations would be lost, and the critical point would not be correctly located.

For a lighter notation, the value of the time step is omitted from now on. All the states appearing in $q^{(r)}_{i_k,g}$, see Eq. (3.10), include at least one infected node, therefore we have,

$$q^{(r)}_{i_k,g} = 1 - \sum_{\{j_1,\dots,j_r\}\in\Gamma^{(r)}_{i_k,g}} \left[\sum_{l=1}^{r} \frac{1-w^{(r)}_{l,g}}{l!\,(r-l)!} \sum_{k_1\neq\dots\neq k_r=1}^{r} P^{I\dots IS\dots SS}_{j_{k_1}\dots j_{k_l} j_{k_{l+1}}\dots j_{k_r} i_k,g} \right] + \mathcal{O}(\epsilon^2)\,, \quad \text{(B.1)}$$

where the squared brackets contain $\mathcal{O}(\epsilon)$ terms only. We want to stress here that, if in the considered clique there was at least another node $i_{\tilde{k}}$ forming with i_k an edge included in some other maximal clique, let us say a (g,r)-clique, then $q^{(r)}_{i_k,g}$ could not be linearized. Indeed, the closure in Eq. (3.3) would now have $P^{S\sigma_{i_{\tilde{k}}}}_{i_k i_{\tilde{k}}}$ in the denominator, since we must condition on the state of the shared edge $\{i_k, i_{\tilde{k}}\}$. The sum over the states of the (g,r)-clique in Eq. (3.10) would contain terms with $\sigma_{i_{\tilde{k}}} = I$. Like the numerator, $P_{i_k i_{\tilde{k}}} \in \mathcal{O}(\epsilon)$, and those terms would be $\mathcal{O}(1)$, preventing linearization. Please, notice that this is not a specific issue of our discrete-time formulation; rather, it is of any model willing to preserve dynamical correlations in structures with non-edge-disjoint subgraphs (see for instance Wu et al. [498], where a continuous-time pair-approximation model is presented to study epidemic spreading in clustered networks).

Going back to the derivation, since $P^{S\dots S}_{i_1\dots i_n,g}$ is fixed by normalization,

$$P^{S\dots S}_{i_1\dots i_n,g} = 1 - \sum_{\{\sigma_1,\dots,\sigma_n\}\neq\{S,\dots,S\}} P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g} \,, \tag{B.2}$$

we only need the linearized equations for arrival states with at least one infected node, i.e., $\{\sigma'_{i_1},\dots,\sigma'_{i_n}\}\neq\{S,\dots,S\}$, which are 2^n-1 in number. Retaining the $\mathcal{O}(\epsilon)$ terms in $\phi_{i_k,g}$, the MECLE (Eq. (3.5)) become

$$\begin{aligned} P^{\sigma'_{i_1}\dots\sigma'_{i_n}}_{i_1\dots i_n,g} = & \sum_{\substack{\{\sigma_{i_1},\dots,\sigma_{i_n}\}\\ \neq\{S,\dots,S\}}} P^{\sigma_{i_1}\dots\sigma_{i_n}}_{i_1\dots i_n,g} \left[(1-\mu)^{N_{I|I}}\ \mu^{N_{S|I}}\left(1-w^{(n-1)}_{N_I,g}\right)^{N_{I|S}}\left(w^{(n-1)}_{N_I,g}\right)^{N_{S|S}}\right] \\ & + \mathbf{1}_{(\sigma_{i_k}=S\ \forall k,\ \exists!k:\sigma'_{i_k}=I)} \sum_{(g,r)}{}^{*}\left(1-q^{(r)}_{i_k,g}\right) + \mathcal{O}(\epsilon^2) \,, \end{aligned} \tag{B.3}$$

where $N_{\sigma'|\sigma} = \left|\{i_{k=1,\dots,n}|\sigma_{i_k}=\sigma,\sigma'_{i_k}=\sigma'\}\right|$ is the number of nodes going from state σ to state σ'; and the asterisk (*) means that the focal (g,n)-clique, $\{i_1,\dots,i_n\}$, is excluded from the sum in Eq. (B.1). The second term in Eq. (B.3) contains only those transitions bringing the focal clique from the state $\{S,\dots,S\}$ to a state with exactly one infected node, due to infections coming from external cliques. This is the only transition the latter can cause. This fact is a direct consequence of the used approximation closure, Eq. (3.3), as it treats the states of adjacent maximal cliques as conditionally independent variables with respect to the node they share. Therefore, the joint probability for the the focal clique and any other adjacent clique to be both in an infected state is of order ϵ^2, hence negligible. Accordingly, external cliques appear in only n of the 2^n-1 equations, corresponding to the n arrival states containing just one infected node. The remaining 2^n-n-1 equations, instead, only contain state probabilities of the focal clique. As a special case of Eq. (B.3), the dynamic equation for a single node (Eq. (3.11)), reads

$$P^I_i = P^I_i(1-\mu) + \sum_{(g,r)}\left(1-q^{(r)}_{i_k,g}\right) + \mathcal{O}(\epsilon^2) \,, \tag{B.4}$$

with $q^{(r)}_{i_k,g}$ from Eq. (B.1).

Given $C^{(n)}$ is the number of maximal $(\cdot,n)$-cliques in the system, by using Eq. (B.3) only, we would have to solve a linear system of $\sum_{n=2}^{m}(2^n-1)\,C^{(n)}$ equations. As we know, we can reduce this number to $n_{\text{eq}} = N + n^{(\text{cli})}_{\text{eq}}$, being $n^{(\text{cli})}_{\text{eq}} = \sum_{n=2}^{m}(2^n-n-1)\,C^{(n)}$ the reduced number of equations for the maximal cliques. We do this by leveraging the consistency relations $P^I_{i_k} = \sum_{\{\sigma_{\neg i_k}\}} P^{I\{\sigma_{\neg i_k}\}}_{i_k\{\neg i_k\},g}$, $\forall k\in\{1,\dots,n\}$, to express n of the state probabilities of each clique in terms of the n state probabilities of its nodes; and then using Eq. (B.4). Conveniently, we could choose the n state probabilities of a clique to replace to be those in which the clique contains a single infected node (notice that we already made this kind of choice in Sec. 2.3.2.2, when

we reduced the linearized equations of the ELE model to Eqs. (2.38)-(2.39)). In this way, the $2^n - n - 1$ equations for a clique would only contain state probabilities of the clique itself and of the nodes it includes, not of external cliques. With this choice, suitably ordering the state variables, the $n_{\text{eq}}^{(\text{cli})} \times n_{\text{eq}}^{(\text{cli})}$ sub-matrix $\boldsymbol{J}^{(\text{cli})}$ of the $n_{\text{eq}} \times n_{\text{eq}}$ Jacobian matrix $\boldsymbol{J}$, would be block diagonal, each $(2^n - n - 1) \times (2^n - n - 1)$ block being associated to one of the $C^{(n)}$ maximal cliques (in the case of the ELE model, since $n = 2$, those blocks are just scalars; see Eq. (2.39)). At this point, one could compute $\Lambda_{\max}(\boldsymbol{J})$ numerically and look for the smallest value of $\beta^{(1)}$ for which $|\Lambda_{\max}(\boldsymbol{J})|$ crosses 1. The block structure would allow to speed up the computation, especially for dense structures, for which $N \ll n_{\text{eq}}^{(\text{cli})}$.

In truth, we can further reduce the number of linear equations to consider to just N. We have seen an analogous reduction in Sec. 2.3.2.2, when going from $N + C^{(2)}$ to N equations. Here, we extend the method for cliques of any size and for more general contagion dynamics. The key observation is that, in the large time limit, the state of the system aligns with the dominant eigenvector of the original Jacobian matrix. At the threshold, $\Lambda_{\max}(\boldsymbol{J}) = 1$, meaning the dominant eigenvector is a fixed point of the linearized dynamics. The state of the system, converging to it, thus approaches stationarity. Imposing $P_{i_1 \dots i_n, g}^{\sigma'_{i_1} \dots \sigma'_{i_n}} = P_{i_1 \dots i_n, g}^{\sigma_{i_1} \dots \sigma_{i_n}}$ in Eq. (B.3), we get $2^n - n - 1$ constraints for each of the $C^{(n)}$ maximal $(\cdot, n)$-cliques. With such constraints ($n_{\text{eq}}^{(\text{cli})}$ in total), we can write the state probabilities of every $(\cdot, n)$-clique as a linear combinations of the marginal probabilities of its nodes, $P_{i_1}^I, \dots, P_{i_n}^I$. Given a (g, n)-clique $\{i_1, \dots, i_n\}$, its generic state $P_{i_1 \dots i_n, g}^{\sigma_{i_1} \dots \sigma_{i_n}}$ decomposes as,

$$P_{i_1 \dots i_n, g}^{\sigma_{i_1} \dots \sigma_{i_n}} = X_{N_I, g}^{(n-1)} \sum_{k : \sigma_{i_k} = I} P_{i_k}^I + Y_{N_I, g}^{(n-1)} \sum_{k : \sigma_{i_k} = S} P_{i_k}^I + \mathcal{O}(\epsilon^2) \,, \tag{B.5}$$

where we can take $Y_{n, g}^{(n-1)} = 0$, as there are no nodes in state S for $N_I = n$. Equation (B.5) is the most general linear combination which respects the existing symmetry between nodes in the same state in the clique. The linear coefficients, which are functions of all the m_1 microscopic parameters of the model ($\{\beta^{(s)}\}_{s=1,\dots,m_1-1}$ and μ), connect the state probability of a clique to the state probabilities of its single nodes when the system is in the critical region around the invasion threshold. The coefficients are found by eliminating the clique state probabilities from the $n_{\text{eq}}^{(\text{cli})}$ equations of the form of Eq. (B.3), after imposing stationarity. Equivalently, one can directly insert Eqs. (B.5) into the $n_{\text{eq}}^{(\text{cli})}$ stationary equations and get a determined system of equations for the linear coefficients in Eq. (B.5). We have to find one coefficient ($X_{n, g}^{(n-1)}$) for $N_I = n$ and two coefficients ($X_{N_I, g}^{(n-1)}$ and $Y_{N_I, g}^{(n-1)}$) for each $N_I \in \{1, \dots, n-1\}$, hence $2n - 1$ of them in total. Summing over every n from 2 to m_0 for $g = 0$, and from 3 to m_1 for $g = 1$, we get a maximum of ${m_0}^2 - 1$ and ${m_1}^2 - 3$ linear coefficients to fix, respectively.

We can show, however, that only one coefficient is left to be determined for each size n (given g). The number of equations of the system for the linear coefficients can be

thus reduced from ${m_0}^2-1$ to m_0-1 for $g=0$, and from ${m_1}^2-5$ to m_1-2 for $g=1$. To see this, observe that summing the state probabilities of a $(\cdot,n+1)$-clique over all the possible states of a single node in it, we must recover the state probability of the $(\cdot,n)$-clique containing the other n nodes. In other words, the state of the sub-clique must be recovered by marginalizing the state of any clique including it. Let us consider a (g,n)-clique $\{i_1,\dots,i_n\}$ in a state with l nodes in state I. We can conveniently order the indices of the nodes so that those l nodes occupy the first l positions. Then, given the $(g,n+1)$-clique $\{i_1,\dots,i_n,i_{n+1}\}$, the relation $P^{I\dots IS\dots S}_{i_1,\dots,i_l,i_{l+1},\dots,i_n,g} = P^{I\dots IS\dots SS}_{i_1,\dots,i_l,i_{l+1},\dots,i_n,i_{n+1},g} + P^{I\dots IS\dots SI}_{i_1,\dots,i_l,i_{l+1},\dots,i_n,i_{n+1},g}$, must hold. The insertion of Eq. (B.5) in the latter, imposes the following constrains,

$$\text{(C1)}\quad X^{(n)}_{l+1,g}+Y^{(n)}_{l,g}=0\ ;$$
$$\text{(C2)}\quad X^{(n)}_{l+1,g}+X^{(n)}_{l,g}=X^{(n-1)}_{l,g}\ ;$$
$$\text{(C3)}\quad Y^{(n)}_{l+1,g}+Y^{(n)}_{l,g}=Y^{(n-1)}_{l,g}\ ;$$

with $l\in\{1,\dots,n\}$. Directly from Eq. (B.5), $P^{I\dots I}_{i_1,\dots,i_{n+1},g}>0$ implies $X^{(n)}_{n+1,g}>0$. (C1) thus yields $Y^{(n)}_{n,g}<0$. The latter, coupled to $P^{I\dots IS}_{i_1,\dots,i_n,i_{n+1},g}>0$, implies $X^{(n)}_{n,g}>0$; in turn, (C1) imposes $Y^{(n)}_{n-1,g}<0$. We can proceed in this way down to $l=1$. Thus, for any l, $X^{(n)}_{l,g}>0$ and $Y^{(n)}_{l,g}=-X^{(n)}_{l+1,g}<0$. As a consequence, (C3) is identical to (C2). The latter allows us to find the quantities $\{X^{(n)}_{l,g}\}_{l=1,\dots,n}$ recursively. Starting from $n=1$, (C2) yields $X^{(1)}_{2,g}+X^{(1)}_{1,g}=1$, for $X^{(0)}_{1,g}\equiv X^{(0)}_1$ must be 1 to trivially satisfy $P^I_i=X^{(0)}_1P^I_i$; thus $X^{(1)}_{1,g}=1-X^{(1)}_{2,g}$. Given $Y^{(1)}_{1,g}<0$, $P^{IS}_{i_1,i_2,g}>0$ requires $X^{(1)}_{1,g}>0$, and thus $X^{(1)}_{2,g}\in(0,1)$. Going to $n=2$, (C2) implies $X^{(2)}_{3,g}+X^{(2)}_{2,g}=X^{(1)}_{2,g}$ for $l=2$, hence $X^{(2)}_{2,g}=X^{(1)}_{2,g}-X^{(2)}_{3,g}$; and $X^{(2)}_{2,g}+X^{(2)}_{1,g}=X^{(1)}_{1,g}$ for $l=1$, from which $X^{(2)}_{1,g}=X^{(1)}_{1,g}-X^{(2)}_{2,g}=1-2X^{(1)}_{2,g}+X^{(2)}_{3,g}$. Again, being $Y^{(2)}_{2,g}<0$, $P^{IIS}_{i_1,i_2,i_3,g}>0$ requires $X^{(2)}_{2,g}>0$, therefore $X^{(2)}_{3,g}<X^{(1)}_{2,g}<1$; and given $Y^{(2)}_{1,g}<0$, $P^{ISS}_{i_1,i_2,i_3,g}>0$ imposes $X^{(2)}_{1,g}>0$, yielding $X^{(2)}_{2,g}<X^{(1)}_{1,g}=1-X^{(1)}_{2,g}$. More generally, for any n, we start using (C2) from $l=n$ and find $X^{(n)}_{n,g}=X^{(n)}_{n+1,g}-X^{(n-1)}_{n,g}$. Then at $l=n-1$, we get $X^{(n)}_{n-1,g}=X^{(n)}_{n,g}-X^{(n-1)}_{n-1,g}$, where $X^{(n)}_{n,g}$ has just been found and $X^{(n-1)}_{n-1,g}$ is already known from the same procedure performed at $n-1$ – building up from $n=2$. Proceeding in this way, one finds that the generic $X^{(n)}_{l,g}$ with $l\in\{1,\dots,n\}$ can be expressed as a linear combination of the n quantities $\{X^{(l)}_{l+1,g}\}_{l=1,\dots,n}$. Besides, the conditions $P^{I\dots IS\dots S}_{i_1,\dots,i_l,i_{l+1},\dots,i_n,g}>0$, coupled to $Y^{(n)}_{l,g}<0$, impose a chain of inequalities that bound any $X^{(n)}_{l,g}$ (hence $Y^{(n)}_{l,g}$) between 0 and 1. In particular, we can show that, for any n, it holds $(n+1)X^{(n)}_{n+1,g}<nX^{(n-1)}_{n,g}$. Let us denote with $\langle P^I\rangle_n$ the average of the state probabilities $\{P^I_j\}_{j=i_1,\dots,i_n}$ associated to the nodes of a clique $\{i_1,\dots,i_n\}$, so that $P^{I\dots I}_{i_1,\dots,i_n}=nX^{(n-1)}_n\langle P^I\rangle_n$. Similarly, for the clique $\{i_1,\dots,i_n,i_{n+1}\}$, superset of the former, we have $P^{I\dots II}_{i_1,\dots,i_n,i_{n+1}}=(n+1)X^{(n)}_{n+1}\langle P^I\rangle_{n+1}$. Then, since (i) $P^{I\dots I}_{i_1,\dots,i_n}=P^{I\dots II}_{i_1,\dots,i_n,i_{n+1}}+P^{I\dots IS}_{i_1,\dots,i_n,i_{n+1}}>P^{I\dots II}_{i_1,\dots,i_n,i_{n+1}}$, and (ii) we can always choose i_{n+1}

to be the node with the highest infection probability among the $n+1$ nodes, so that $\langle P^I \rangle_n \leqslant \langle P^I \rangle_{n+1}$; the above inequality follows. Summing up, $\forall (g,n)$, $l \in \{1,\dots,n\}$, we proved the following properties:

(P1) $Y^{(n)}_{l,g} = -X^{(n)}_{l+1,g} < 0$;

(P2) $X^{(n)}_{n+1,g} < \frac{n}{n+1} X^{(n-1)}_{n,g} < X^{(n-1)}_{n,g} < 1$;

(P3) $X^{(n)}_{l,g} \in (0,1)$;

(P4) $X^{(n)}_{l,g} = \sum_{l'=1}^{n} c^{(n)}_{l',g} X^{(l')}_{l'+1,g}$, with $c^{(n)}_{l',g} \in \mathbb{Z}$.

From (P4) it follows that the only coefficients left to be determined are the $m_0 - 1$ coefficients $\{X^{(n-1)}_{n,g}\}_{n=2,\dots,m_0}$ and the $m_1 - 2$ coefficients $\{X^{(n-1)}_{n,g}\}_{n=3,\dots,m_1}$. As explained in the previous paragraph, these are found from the stationary equations.

Finally, plugging Eq. (B.5) into Eq. (B.4), we are left with a closed system of N linear equations for the N node state probabilities $\{P^I_i\}_{i=1,\dots,N}$. Before presenting the associated reduced Jacobian matrix, $\boldsymbol{J'}$, let us define the generalized adjacency matrices $\left\{\boldsymbol{A^{(0,r)}}\right\}_{r\in\{1,\dots,m_0-1\}}$ and $\left\{\boldsymbol{A^{(1,r)}}\right\}_{r\in\{2,\dots,m_1-1\}}$, such that $A^{(g,r)}_{ij}$ equals 1 if there is a $(g,r+1)$-clique in $\mathcal{D}(\mathcal{K})$ including both nodes i and j, and 0 otherwise. Notice that these matrices depend on the decomposition $\mathcal{D}(\mathcal{K})$. Moreover, being the $\mathcal{D}(\mathcal{K})$ edge-disjoint, only one of these matrices can have a non-zero element in the position corresponding to a given pair of nodes. Consequently, the ($\mathcal{D}(\mathcal{K})$-independent) adjacency matrix of $\mathcal{K}^{(1)}$, the underlying graph of $\mathcal{K}$, is simply obtained by the sum of all those generalized adjacency matrices. From the latter, we can define the generalized (g,r)-degree of node i, $k^{(g,r)}_i$, as the number of $(g,r+1)$-cliques incident on node i in $\mathcal{D}(\mathcal{K})$, computed as $k^{(g,r)}_i = \sum_{j=1}^{N} A^{(g,r)}_{ij}$. Besides, the ($\mathcal{D}(\mathcal{K})$-independent) pairwise degree k_i of node i in $\mathcal{K}^{(1)}$ can be computed as $k_i = k^{(0)}_i + k^{(1)}_i$, where $k^{(0)}_i = \sum_{r=1}^{m_0} r k^{(0,r)}_i$ and $k^{(1)}_i = \sum_{r=2}^{m_1} r k^{(1,r)}_i$ are the total number of neighbors of i within, respectively, $(0,\cdot)$-cliques and $(1,\cdot)$-cliques.

Endowed with these definitions and substituted Eq. (B.5) in Eq. (B.4), we eventually get

$$P^I(t+1) = \boldsymbol{J'} P^I(t) \,, \tag{B.6}$$

with

$$\boldsymbol{J'} = \boldsymbol{M} - \boldsymbol{D} + (1-\mu)\mathbb{1} \,, \tag{B.7}$$

having defined the matrices $\boldsymbol{M}$ and $\boldsymbol{D}$, of elements

$$M_{ij} = \sum_{(g,r)} A^{(g,r)}_{ij} \sum_{l=1}^{r} \left(1 - w^{(r)}_{l,g}\right) \left[\binom{r-1}{l-1} X^{(r)}_{l,g} + \binom{r-1}{l} Y^{(r)}_{l,g} \right] \,, \tag{B.8}$$

$$D_{ij} = -\delta_{ij} \left[\sum_{(g,r)} k^{(g,r)}_i \sum_{l=1}^{r} \left(1 - w^{(r)}_{l,g}\right) \binom{r}{l} Y^{(r)}_{l,g} \right] \,. \tag{B.9}$$

The matrices $\boldsymbol{M}$ and $\boldsymbol{D}$ account for the transition from state S to state I, and the form they take is understood in this way. The sum in Eq. (B.1) over all the state configurations of a $(g, r+1)$-clique, using Eq. (B.5), yields a term $X^{(r)}_{l,g} P^I_j$ or $Y^{(r)}_{l,g} P^I_j$ every time a node j included in the clique is in state I or S, respectively. Consider i as the focal node – the one that could transition from S to I – and j one of its neighbors through a $(g, r+1)$-clique, i.e., $A^{(g,r)}_{ij} = 1$ in Eq. (B.8). Then, if j is in state I, there are $\binom{r-1}{l-1}$ $(= \binom{r-1}{r-l})$ ways in which the other $l-1$ $(r-l)$ nodes in state I (S) can be chosen among the $r-1$ remaining ones in the clique; if in state S, instead, there are $\binom{r-1}{r-l-1}$ $(= \binom{r-1}{l})$ ways in which the other $r-l-1$ (l) nodes in state S (I) can be chosen among the $r-1$ remaining ones. Regarding the focal node, this will contribute a term $Y^{(r)}_{l,g} P^I_i$ – stored in Eq. (B.9) – for each of the state configurations in which the other r nodes can be found. These configurations are $\binom{r}{l}$ $(=\binom{r}{r-l})$, which is the number of ways in which l $(r-l)$ out of r nodes can be chosen to be in state I (S).

We could, at this point, require $\Lambda_{\max}(\boldsymbol{J'}) = 1$ and find the invasion threshold as the smallest $\beta^{(1)}$ satisfying that condition. Nonetheless, we can make the threshold condition more explicit. One way to do it, is imposing $P^I(t+1) = P^I(t)$ and converting Eq. (B.6) into the eigenvalue equation $\boldsymbol{B}P^I = (\mu/\beta)P^I$, once defined $\boldsymbol{B} = (1/\beta)(\boldsymbol{M} - \boldsymbol{D})$. This has the same form of Eq. (2.44), found in Sec. 2.3.2.2 for the ELE. Since the coefficients $X^{(r)}_{l,g}$ (hence $Y^{(r)}_{l,g}$) lie in $(0,1)$, the entries of $\boldsymbol{B}$ are limited functions of the microscopic parameters $\{\beta^{(s)}\}_{s=1,\dots,m_1-1}$ and μ. We can thus use the same reasoning presented after Eq. (2.44) to prove that $\beta^{(1)}_{\rm cr}$ can be found from the implicit equation $\beta^{(1)} = \mu/\Lambda_{\max}(\boldsymbol{B})$. An alternative, more robust and elegant way – outlined at the end of Sec. 2.3.2.2 – is the following. Let us first prove that $\boldsymbol{M}$ is a non-negative and irreducible matrix, so that the Perron-Frobenius theorem can be applied. In fact, the non-negativity comes directly from the physical meaning of its entries. They quantify the by-definition positive contribution of neighboring nodes to the infection probability of a focal node. Nevertheless, we can prove the non-negativity even without this information. The probability $w^{(r)}_{l,g}$, being a decreasing function of l, takes its minimum for $l = r$ and its maximum for $l = 1$. Consequently, the value of the sum $\sum_{l=1}^{r}(1 - w^{(r)}_{l,g})[\cdots]$ in Eq. (B.8), is limited by $(1 - w^{(r)}_{1,g})\sum_{l=1}^{r}[\cdots]$ from below and by $(1 - w^{(r)}_{r,g})\sum_{l=1}^{r}[\cdots]$ from above; with

$$\begin{aligned}\sum_{l=1}^{r}[\cdots] &\equiv \sum_{l=1}^{r}\left[\binom{r-1}{l-1}X^{(r)}_{l,g} + \binom{r-1}{l}Y^{(r)}_{l,g}\right] \\ &= \sum_{l=1}^{r}\binom{r-1}{l-1}X^{(r)}_{l,g} + \sum_{l=2}^{r+1}\binom{r-1}{l-1}Y^{(r)}_{l-1,g} \\ &= \sum_{l=2}^{r}\binom{r-1}{l-1}\left(X^{(r)}_{l,g} + Y^{(r)}_{l-1,g}\right) + X^{(r)}_{1,g} = X^{(r)}_{1,g} > 0\,,\end{aligned}$$

where we used $Y^{(r)}_{l-1,g} = -X^{(r)}_{l,g}$ and $\binom{r-1}{r+1} = 0$. Therefore, $0 < (1 - w^{(r)}_{1,g})X^{(r)}_{1,g} < \sum_{l=1}^{r}(1 - w^{(r)}_{l,g})[\cdots] < (1 - w^{(r)}_{r,g})X^{(r)}_{1,g} < 1$, which proves that $\boldsymbol{M}$ is non-negative (and

bounded). Its irreducibility simply derives from assuming $\mathcal{K}$ (or equivalently, its underlying graph $\mathcal{K}^{(1)}$) to be connected. Connectedness also implies that $\boldsymbol{D}$ has only nonzero elements in the diagonal. Given $Y^{(r)}_{l,g} < 0$, these elements are all positive, and this obviously holds true if μ is added to them. Therefore, the matrix $\widetilde{\boldsymbol{D}} = \boldsymbol{D} + \mu\mathbb{1}$ is invertible, and the inverse, $\widetilde{\boldsymbol{D}}^{-1}$, has elements $[\widetilde{D}^{-1}]_{ij} = \delta_{ij}(D_{ij} + \mu)^{-1}$. In the end, by imposing $P^I(t+1) = P^I(t)$, we get

$$\boldsymbol{M}P^I = \widetilde{\boldsymbol{D}}P^I \,, \tag{B.10}$$

which, applying $\widetilde{\boldsymbol{D}}^{-1}$ to both sides, becomes

$$\widetilde{\boldsymbol{M}}P^I = P^I \,, \tag{B.11}$$

where we defined $\widetilde{\boldsymbol{M}} = \widetilde{\boldsymbol{D}}^{-1}\boldsymbol{M}$, of elements $\widetilde{M}_{ij} = M_{ij}(D_{ii} + \mu)^{-1}$. As $\boldsymbol{M}$, $\widetilde{\boldsymbol{M}}$ is a non-negative and irreducible matrix. By the Perron-Frobenius theorem, $\widetilde{\boldsymbol{M}}$ admits a unique dominant eigenvector $P^I_\star$ with all entries real and positive. This is associated to $\Lambda_{\max}(\widetilde{\boldsymbol{M}})$, which is real and positive too. Since P^I is positive, $P^I = P^I_\star$ is the only acceptable solution of Eq. (B.11). In the limit $\epsilon \to 0$, and up to a $\mathcal{O}(1)$ global rescaling, $P^I = P^I_\star$ represents the expected state of the system right at the invasion threshold. In conclusion, fixed the values of the recovery probability, μ, and of the higher-order infection probabilities, $\{\beta^{(s)}\}_{s=2,\dots,m_1}$, the invasion threshold $\beta^{(1)}_{\rm cr}$ is implicitly found as the smallest value of $\beta^{(1)}$ such that

$$\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = 1 \,. \tag{B.12}$$

More generally, Eq. (B.12) imposes a constraint among the m_1 parameters of the model, identifying one or more surfaces in the m_1-dimensional parameter space. Among them, there is the critical surface, the one marking the loss of stability for the inactive state, from which we can isolate the critical value of any of the m_1 parameters, e.g., $\beta^{(1)}$.

B.1.1.1 Node-based mean-field approximation

The node-based, homogeneous mean-field (MF) approximation of Eq. (3.11) is found by neglecting both the state correlations and the local structural heterogeneity among the nodes, i.e., regarding every node as the "average node" in the structure [499]. Given any (g,r)-clique $\{i_1,\dots,i_r\}$, it follows $P^{I\dots IS\dots S}_{i_1\dots i_l i_{l+1}\dots i_r,g} = \rho^l\,(1-\rho)^{r-l}$; and, for any node i, $k_i^{(g,r)} = \bar{k}^{(g,r)}$, where $\bar{k}^{(g,r)}$ is the average value of the (g,r)-degree of the nodes in the structure. Thus, Eq. (3.11) becomes

$$\rho\,(t+1) = \rho\,(t)\,(1-\mu) + (1-\rho\,(t))\left(1 - \prod_{(g,r)} \bar{q}_g^{(r)}\right) \,, \tag{B.13}$$

where

$$\bar{q}_g^{(r)} = \left[1 - \sum_{l=1}^{r} \binom{r}{l} \left(1 - w_{l,g}^{(r)}\right) \rho^l \left(1-\rho\right)^{r-l}\right]^{\bar{k}^{(g,r)}} . \tag{B.14}$$

The stationary solution is then got imposing $\rho\left(t+1\right) = \rho\left(t\right)$ in Eq. (B.13). The linearization around the epidemic-free state is then implemented by taking $\rho = \epsilon \to 0$. Looking at Eq. (B.14), the only $\mathcal{O}(\epsilon)$ terms are given by $l = 1$, whatever the couple (g, r). That is, only the pair-wise probability $\beta^{(1)}$ contributes in the MF approximation. With few algebra, one gets the renowned formula

$$\beta_{\rm cr}^{(1)} = \frac{\mu}{\langle k \rangle} , \tag{B.15}$$

where $\langle k \rangle = \sum_{i=1}^{N} k_i / N$, is the average degree of a node in $\mathcal{K}^{(1)}$. Equation (B.15) holds identical also for the MF approximation of the simplicial ELE (MMCA) [230].

More generally, following Ref. [499], it is straightforward to prove that under the node-based, heterogeneous mean-field approximation of either the MECLE and the simplicial ELE, the threshold reads

$$\beta_{\rm cr}^{(1)} = \mu \frac{\langle k \rangle}{\langle k^2 \rangle} , \tag{B.16}$$

being $\langle k^2 \rangle = \sum_{i=1}^{N} {k_i}^2 / N$. Equation (B.16) is also found in continuous time [275].

Finally, any quenched model treating the nodes states as independent, thus including the MMCA approximation of the MECLE and of the simplicial ELE, yields the same prediction made by the MMCA in pairwise networks, Eq. (2.31). Accordingly, Eqs. (B.15) and (B.16) are recovered as the homogeneous and heterogeneous mean-filed approximations of the MMCA. All these approaches, whether quenched or annealed, share one fundamental assumption, namely, the absence of dynamical correlations. This causes their inability to predict that $\beta_{\rm cr}^{(1)}$ does depend on $\beta^{(2)}$, or on any other higher-order parameter.

B.2 METHODS FOR SEC. 3.3

B.2.1 *General A-GAME*

For a general adaptive hypergraph model, we allow both susceptible and infectious nodes to rewire away from groups, respectively at rates γ_S and γ_I. This leads to the following additional adaptive terms (tagged with the 'a' superscript), schematized in Fig. B.1, to the static GAME framework,

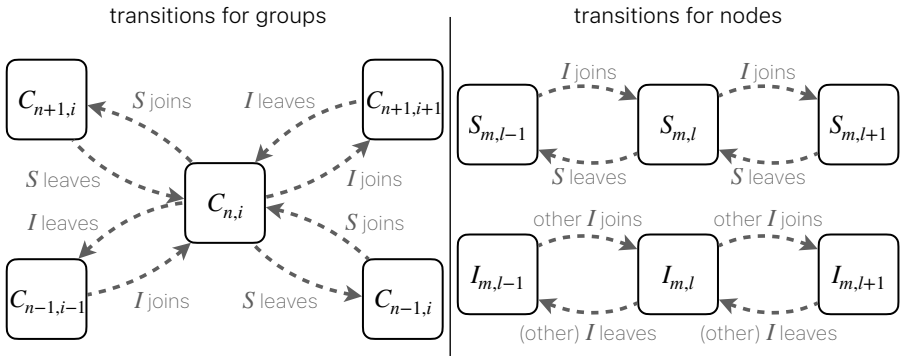


Figure B.1: States and transitions for the general version of the A-GAME, where also infected nodes can rewire. This is a graphical representation of Eqs. (B.17). While susceptible nodes rewiring from group to group only affect the sizes of the latter (hence the group size distribution and the pairwise degree distribution of nodes), the migration of infected nodes can also affect the activity state of the groups, thus the active membership (and degree) of the nodes therein, giving rise to a more complex dynamics.

$$\begin{aligned}
\dot{C}^{\mathrm{a}}_{n,i} &= \gamma_S \mathbf{1}_{i\geqslant\bar{i}} \left[(n+1-i)C_{n+1,i} - (n-i)C_{n,i}\right] \\
&\quad + \gamma_I \left[\mathbf{1}_{i\geqslant\bar{i}}(i+1)C_{n+1,i+1} - \mathbf{1}_{i>\bar{i}} i C_{n,i}\right] \\
&\quad + \left(\frac{\eta \mathbf{1}_{i<\bar{i}}}{C_{i<\bar{i}}} + 1 - \eta\right) \left[\gamma_S \Omega_{S|i\geqslant\bar{i}} C_{n-1,i} - \left(\gamma_S \Omega_{S|i\geqslant\bar{i}} + \gamma_I \Omega_{I|i>\bar{i}}\right) C_{n,i}\right] \\
&\quad + \left(\frac{\eta \mathbf{1}_{i\leqslant\bar{i}}}{C_{i<\bar{i}}} + 1 - \eta\right) \gamma_I \Omega_{I|i>\bar{i}} C_{n-1,i-1} \,, \qquad \text{(B.17a)} \\
\dot{S}^{\mathrm{a}}_{m,l} &= \gamma_S \left[\eta + (1-\eta)C_{i<\bar{i}}\right] \left[(l+1)S_{m,l+1} - l S_{m,l}\right] \\
&\quad + \gamma_I \Omega_{I|i>\bar{i}} \left(\frac{\eta}{\Omega_{S|i<\bar{i}}} + \frac{1-\eta}{\Omega_S}\right) \\
&\qquad \times \Omega_{S|i=\bar{i}-1} \left[(m-l+1)S_{m,l-1} - (m-l)S_{m,l}\right] \,, \qquad \text{(B.17b)} \\
\dot{I}^{\mathrm{a}}_{m,l} &= \gamma_I \left[\eta + (1-\eta)C_{i<\bar{i}}\right] \left[(l+1)I_{m,l+1} - l I_{m,l}\right] \\
&\quad + \gamma_I \frac{\Gamma_{I|i=\bar{i}+1}}{\Omega_{I|i>\bar{i}}} \left[(l+1)I_{m,l+1} - l I_{m,l}\right] \\
&\quad + \gamma_I \Omega_{I|i>\bar{i}} (1-\eta) \frac{\Omega_{I|i=\bar{i}}}{\Omega_I} \left[(m-l+1)I_{m,l-1} - (m-l)I_{m,l}\right] \,. \qquad \text{(B.17c)}
\end{aligned}$$

These equations are simply added to Eqs. (3.22). We defined $C_{i<\bar{i}} = \sum_{n,i<\bar{i}} C_{n,i}$, $\Omega_S = \sum_{n,i}(n-i)C_{n,i}$, $\Omega_{S|i\geqslant\bar{i}} = \sum_{n,i\geqslant\bar{i}}(n-i)C_{n,i}$, $\Omega_{S|i<\bar{i}} = \sum_{n,i<\bar{i}}(n-i)C_{n,i}$, $\Omega_{S|i=\bar{i}-1} = \sum_n (n-\bar{i}+1)C_{n,\bar{i}-1}$, $\Omega_I = \sum_{n,i} i C_{n,i}$, $\Omega_{I|i>\bar{i}} = \sum_{n,i>\bar{i}} i C_{n,i}$, $\Omega_{I|i=\bar{i}} = \sum_n \bar{i} C_{n,\bar{i}}$, and $\Gamma_{I|i=\bar{i}+1} = \sum_n \bar{i}(\bar{i}+1)C_{n,\bar{i}+1}$. The first two terms in Eq. (B.17a) account for S and I nodes leaving groups, the third and the forth for S and I nodes joining groups. The

first term in Eq. (B.17b) accounts for S nodes rewiring away, while the second term for I nodes rewiring to groups with $i = \bar{i} - 1$, making them active for the S nodes therein. Lastly, the first term in Eq. (B.17c) considers I nodes rewiring away, while the second (third) term accounts for I nodes leaving (joining) groups with $i = \bar{i} + 1$ ($i = \bar{i}$), making them inactive (active) for the I nodes therein. The simpler version where only S nodes rewire, Eqs. (3.45), is recovered by setting $\gamma_I = 0$ ($\gamma_S \equiv \gamma$).

C

SUPPLEMENTARY INFORMATION FOR SEC. 3

C.1 SUPPLEMENTARY INFORMATION FOR SEC. 3.1

C.1.1 *Robustness of the MECLE under variable estimation of the EECC*

Here we assess the dependence of the predictions made by the MECLE on the use of different EECCs found for a given SC, $\mathcal{K}$. If the maximal cliques of the underlying graph $\mathcal{K}^{(1)}$ of the latter are all mutually edge-disjoint, then there is a unique EECC. Otherwise, two cases must be distinguished: (i) $\mathcal{K}$ is strictly 0-connected, so the only non-edge-disjoint cliques in $\mathcal{K}^{(1)}$ are cliques also in $\mathcal{K}$ (i.e., they are $(0,\cdot)$-cliques); (ii) $\mathcal{K}$ is non-strictly 0-connected, so the non-edge-disjoint cliques in $\mathcal{K}^{(1)}$ can be either cliques or simplices in $\mathcal{K}$ (i.e., $(0,\cdot)$- or $(1,\cdot)$-cliques, respectively). The MECLE model strictly applies only to the case (i) but, as shown in Sec. 3.1, it keeps working well when the strictly 0-connectedness is only slightly broken. In order to prove how robust is the model in both cases, we consider structures possessing high proportions of edges shared by multiple cliques. Indeed, the higher is the number of times edges are shared, the higher is the probability that the proposed heuristic returns (proportionally) different EECCs. We recall that the algorithm is not deterministic only when, at some point, among the cliques with the minimum score, there are multiple ones having maximum order, among which one is randomly chosen (step 3.b of the proposed heuristic).

Let us first focus on case (i). The best way to prove the robustness of the MECLE model is, in this case, considering a graph (i.e., a 1-dimensional SC), so that the number of non-edge-disjoint maximal $(0,\cdot)$-cliques is maximized; otherwise, some of them would be $(1,\cdot)$-cliques, which in this case are supposed strictly 0-connected and hence not affecting the edge covering. In Fig. C.1 we show and discuss the results obtained computing several EECCs of a graph generated by the Dorogovtsev-Mendes model [147]. There are no tree-like regions in such a graph. Rather, 3-cliques percolate it entirely through shared edges. Despite this high rate of edge overlap, the prediction made by the MECLE is substantially independent from the used covering.

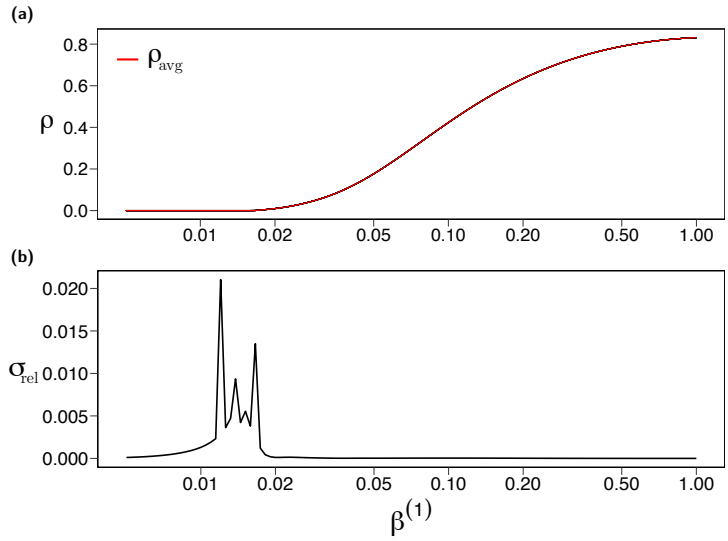


Figure C.1: The robustness of the MECLE under different EECCs is assessed for a graph generated from the Dorogovtsev-Mendes model, forming a 0-connected simplicial complex. This boasts a very high rate of edge overlap and, therefore, potentially different EECCs could lead the MECLE to make notably dissimilar predictions. Indeed, about 34% of the edges in the network are shared at least by two maximal cliques. Specifically, within that 34%, edges are shared by about 2.5 maximal cliques on average, with a standard deviation of 0.85 and a skewness of 2.3. The recovery probability is fixed as $\mu = 0.2$. (a) 20 prevalence curves obtained by 20 different EECCs are shown in black, while in red is reported their average curve $\rho_{\text{avg}}(\beta^{(1)})$, defined by the average value of ρ at each value of $\beta^{(1)}$. Since the deviations are below the width of the red line, the black lines are not visible. (b) The ratio σ_{rel} between the sample standard deviation σ and ρ_{avg} for each $\beta^{(1)}$. The value of σ_{rel} vanishes everywhere, expect for a small region around the invasion threshold. This is below 2.5% for values of $\rho_{\text{avg}} \sim 10^{-4} \sim N^{-1}$, proving the remarkable robustness of the model.

Moving to case (ii), we consider several EECCs of the clique complex of the Dorogovtsev-Mendes network used in (i). The clique complex is obtained by converting all the 3-cliques of the network in 2-faces, so it is 1-connected. To force strict 0-connectedness, many 2-faces must be sacrificed, implying that the resulting EECCs might substantially differ from one another. Nonetheless, the results reported in Fig. C.2 show the low variability among the predictions made by the model. Accordingly, such variability becomes negligible when the simplicial complex only slightly breaks strict 0-connectedness, so that we may just use the first EECC we estimate. This is the case for the simplicial complexes considered in Fig. 3.5, for which the model is still reliable.

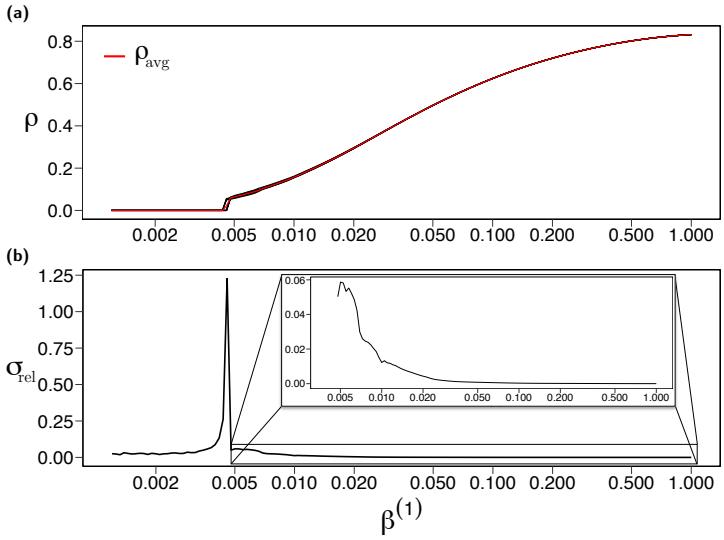


Figure C.2: The robustness of the MECLE under different EECCs is assessed for a 1-connected simplicial complex generated from the Dorogovtsev-Mendes network used in Fig. C.1. The recovery probability is $\mu = 0.2$. (a) 20 prevalence curves obtained by 20 different EECCs are shown in black, while in red is reported their average curve $\rho_{\text{avg}}(\beta^{(1)})$, defined by the average value of ρ at each value of $\beta^{(1)}$. (b) The ratio σ_{rel} between the sample standard deviation σ and ρ_{avg} for each $\beta^{(1)}$. The pick at about $\beta^{(1)} = 0.0047$ is due to some curves transitioning at slightly different values of $\beta^{(1)}$. In fact, the uncertainty about the location of the critical point is of only about 0.0002, corresponding to a relative uncertainty of less than 5%. The inset plot shows a zoom of σ_{rel} to the right of the transition: σ_{rel} stays below 6% next to the transition, while rapidly decreasing towards zero for larger values of $\beta^{(1)}$, hence enlightening the robustness of the model.

C.1.2 *MECLE for simplicial 2-complexes*

We here illustrate the form taken by the MECLE when the interaction structure is a simplicial 2-complex, i.e., the $(3,3)$ implementation of the model. Cliques and faces can only have order $n = 2, 3$.

The evolution of the probability P_i^I for node i being infected is governed by Eq. (3.11), which now takes the form

$$P_i^I(t+1) = P_i^I(t)(1-\mu) + P_i^S(t)\left(1 - q_{i,0}^{(1)} q_{i,0}^{(2)} q_{i,1}^{(2)}\right), \tag{C.1}$$

where

$$q_{i,0}^{(1)} = \prod_{j \in \Gamma_{i,0}^{(1)}} \left(1 - \beta^{(1)} P_{j|i,0}^{I|S}\right) , \tag{C.2}$$

$$q_{i,g}^{(2)} = \prod_{\{j,k\} \in \Gamma_{i,g}^{(2)}} \left\{1 - \beta^{(1)} \left(P_{jk|i,g}^{IS|S} + P_{jk|i,g}^{SI|S}\right)\right. \tag{C.3}$$

$$\left. - \left[1 - \left(1 - \beta^{(1)}\right)^2 \left(1 - g\beta^{(2)}\right)\right] P_{jk|i,g}^{II|S} \right\} . \tag{C.4}$$

According to Eq. (3.5), the state of a $(0,1)$-clique $\{i,j\}$ is governed by the following equations,

$$\begin{aligned} P_{ij,0}^{II}(t+1) = {} & P_{ij,0}^{SS}(t) \left(1 - q_{i(j),0}^{(1)} q_{i,0}^{(2)} q_{i,1}^{(2)}\right) \left(1 - q_{j(i),0}^{(1)} q_{j,0}^{(2)} q_{j,1}^{(2)}\right) \\ & + P_{ij,0}^{SI}(t) \left(1 - (1-\beta^{(1)}) q_{i(j),0}^{(1)} q_{i,0}^{(2)} q_{i,1}^{(2)}\right) (1-\mu) \\ & + P_{ij,0}^{IS}(t) (1-\mu) \left(1 - (1-\beta^{(1)}) q_{j(i),0}^{(1)} q_{j,0}^{(2)} q_{j,1}^{(2)}\right) \\ & + P_{ij,0}^{II}(t) (1-\mu)^2 , \end{aligned} \tag{C.5a}$$

$$\begin{aligned} P_{ij,0}^{IS}(t+1) = {} & P_{ij,0}^{SS}(t) \left(1 - q_{i(j),0}^{(1)} q_{i,0}^{(2)} q_{i,1}^{(2)}\right) \left(q_{j(i),0}^{(1)} q_{j,0}^{(2)} q_{j,1}^{(2)}\right) \\ & + P_{ij,0}^{SI}(t) \left(1 - (1-\beta^{(1)}) q_{i(j),0}^{(1)} q_{i,0}^{(2)} q_{i,1}^{(2)}\right) \mu \\ & + P_{ij,0}^{IS}(t) (1-\mu) \left((1-\beta^{(1)}) q_{j(i),0}^{(1)} q_{j,0}^{(2)} q_{j,1}^{(2)}\right) \\ & + P_{ij,0}^{II}(t) (1-\mu) \mu . \end{aligned} \tag{C.5b}$$

where $q_{i(j),0}^{(1)}$ coincides with $q_{i,0}^{(1)}$ except for excluding the $(0,1)$-clique $\{i,j\}$ from the product, and analogously for the other similar terms.

The state of a $(0,2)$-clique $\{i,j,k\}$ follows the equations,

$$
\begin{aligned}
P^{III}_{ijk,0}(t+1) = {} & P^{SSS}_{ijk,0}(t)\left(1-q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(1-q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(1-q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SSI}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)(1-\mu) \\
& + P^{SIS}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)(1-\mu)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{ISS}_{ijk,0}(t)(1-\mu)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SII}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)(1-\mu)^2 \\
& + P^{ISI}_{ijk,0}(t)(1-\mu)\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)(1-\mu) \\
& + P^{IIS}_{ijk,0}(t)(1-\mu)^2\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{III}_{ijk,0}(t)(1-\mu)^3 \,, \qquad \text{(C.6a)} \\
P^{IIS}_{ijk,0}(t+1) = {} & P^{SSS}_{ijk,0}(t)\left(1-q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(1-q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SSI}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\mu \\
& + P^{SIS}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)(1-\mu)\left(\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{ISS}_{ijk,0}(t)(1-\mu)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SII}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)(1-\mu)\,\mu \\
& + P^{ISI}_{ijk,0}(t)(1-\mu)\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\mu \\
& + P^{IIS}_{ijk,0}(t)(1-\mu)^2\left(\left(1-\beta^{(1)}\right)^2 q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{III}_{ijk,0}(t)(1-\mu)^2\,\mu \,, \qquad \text{(C.6b)} \\
P^{ISS}_{ijk,0}(t+1) = {} & P^{SSS}_{ijk,0}(t)\left(1-q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SSI}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\left(\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\mu \\
& + P^{SIS}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\mu\left(\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{ISS}_{ijk,0}(t)(1-\mu)\left(\left(1-\beta^{(1)}\right)q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\left(\left(1-\beta^{(1)}\right)q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{SII}_{ijk,0}(t)\left(1-\left(1-\beta^{(1)}\right)^2 q^{(1)}_{i,0}q^{(2)}_{i(jk),0}q^{(2)}_{i,1}\right)\mu^2 \\
& + P^{ISI}_{ijk,0}(t)(1-\mu)\left(\left(1-\beta^{(1)}\right)^2 q^{(1)}_{j,0}q^{(2)}_{j(ik),0}q^{(2)}_{j,1}\right)\mu \\
& + P^{IIS}_{ijk,0}(t)(1-\mu)\,\mu\left(\left(1-\beta^{(1)}\right)^2 q^{(1)}_{k,0}q^{(2)}_{k(ij),0}q^{(2)}_{k,1}\right) \\
& + P^{III}_{ijk,0}(t)(1-\mu)\,\mu^2 \,. \qquad \text{(C.6c)}
\end{aligned}
$$

where $q^{(2)}_{i(jk),0}$ coincides with $q^{(2)}_{i,0}$ except for excluding the $(0,2)$-clique $\{i,j,k\}$ from the product, and analogously for the other similar terms.

Finally, for a $(1,2)$-clique $\{i,j,k\}$, we get the following equations,

$$
\begin{aligned}
P^{III}_{ijk,1}(t+1) = {} & P^{SSS}_{ijk,1}(t) \left(1 - q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(1 - q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \left(1 - q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{SSI}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) (1-\mu) \\
& + P^{SIS}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) (1-\mu) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{ISS}_{ijk,1}(t) (1-\mu) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{SII}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) (1-\mu)^2 \\
& + P^{ISI}_{ijk,1}(t) (1-\mu) \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) (1-\mu) \\
& + P^{IIS}_{ijk,1}(t) (1-\mu)^2 \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{III}_{ijk,1}(t) (1-\mu)^3 \ , \qquad \text{(C.7a)} \\
P^{IIS}_{ijk,1}(t+1) = {} & P^{SSS}_{ijk,1}(t) \left(1 - q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(1 - q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \left(q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{SSI}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \mu \\
& + P^{SIS}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) (1-\mu) \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right)\right) \\
& + P^{ISS}_{ijk,1}(t) (1-\mu) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right)\right) \\
& + P^{SII}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) (1-\mu)\, \mu \\
& + P^{ISI}_{ijk,1}(t) (1-\mu) \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \mu \\
& + P^{IIS}_{ijk,1}(t) (1-\mu)^2 \left(\left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{III}_{ijk,1}(t) (1-\mu)^2 \mu \ , \qquad \text{(C.7b)} \\
P^{ISS}_{ijk,1}(t+1) = {} & P^{SSS}_{ijk,1}(t) \left(1 - q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \left(q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{SSI}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right)\right) \mu \\
& + P^{SIS}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \mu \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right)\right) \\
& + P^{ISS}_{ijk,1}(t) (1-\mu) \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right)\right) \left(\left(1 - \left(1-\beta^{(1)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right)\right) \\
& + P^{SII}_{ijk,1}(t) \left(1 - \left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{i,0} q^{(2)}_{i,0} q^{(2)}_{i(jk),1}\right) \mu^2 \\
& + P^{ISI}_{ijk,1}(t) (1-\mu) \left(\left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{j,0} q^{(2)}_{j,0} q^{(2)}_{j(ik),1}\right) \mu \\
& + P^{IIS}_{ijk,1}(t) (1-\mu)\, \mu \left(\left(1-\beta^{(1)}\right)^2 \left(1-\beta^{(2)}\right) q^{(1)}_{k,0} q^{(2)}_{k,0} q^{(2)}_{k(ij),1}\right) \\
& + P^{III}_{ijk,1}(t) (1-\mu)\, \mu^2 \ . \qquad \text{(C.7c)}
\end{aligned}
$$

where $q^{(2)}_{i(jk),1}$ coincides with $q^{(2)}_{i,1}$ except for excluding the $(1,2)$-clique $\{i,j,k\}$ from the product, and analogously for the other similar terms.

C.1.3 *Invasion threshold for highly symmetrical structures*

Here we make predictions for some particularly symmetrical structures, for which Eq. (B.12) can be recast in a more explicit form to compute $\beta^{(1)}_{\rm cr}$. As a particular case, we show the monotonous decrease of $\beta^{(1)}_{\rm cr}$ with respect to $\beta^{(2)}$ in the herein examined 2-dimensional SCs. We stress that such dependence, as like as any other one, of $\beta^{(1)}_{\rm cr}$ on the higher-order couplings, is completely overlooked by uncorrelated-state models.

C.1.3.1 Regular SCs

For a simplicial complex such that $k^{(g,r)}_i = k^{(g,r)}$, $\forall i \in V$, $\forall (g,r)$, since all the non-zero elements of $\widetilde{\boldsymbol{M}}$ are equal to the same constant, from the Collatz–Wielandt formula [223], it follows,

$$\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = \sum_{j=1}^{N} \widetilde{M}_{ij} = \frac{\sum_{(g,r)} r k^{(g,r)} \sum_{l=1}^{r} (1-w^{(r)}_{l,g}) \left[\binom{r-1}{l-1} X^{(r)}_{l,g} + \binom{r-1}{l} Y^{(r)}_{l,g} \right]}{\mu - \sum_{(g,r)} k^{(g,r)} \sum_{l=1}^{r} (1-w^{(r)}_{l,g}) \binom{r}{l} Y^{(r)}_{l,g}} \,, \tag{C.8}$$

where both equalities hold for any chosen $i \in V$. Using Eq. (B.12), $\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = 1$, one can solve with respect to $\beta^{(1)}$ and express $\beta^{(1)}_{\rm cr}$ in terms of all the other parameters.

In particular, for a simplicial 2-complex, using the properties (P1) and (P4) proved in Sec. B.1.1, $\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = 1$ becomes,

$$\begin{aligned} &\beta^{(1)2} \left[k^{(0,2)} \left(3X^{(2)}_{3,0} - 2X^{(1)}_{2,0} \right) + k^{(1,2)} (1-\beta^{(2)}) \left(3X^{(2)}_{3,1} - 2X^{(1)}_{2,1} \right) \right] \\ &+ \beta^{(1)} \left\{ k^{(0,1)} \left(1 - 2X^{(1)}_{2,0} \right) + 2k^{(0,2)} \left(1 - 2X^{(1)}_{2,0} - 2X^{(2)}_{3,0} \right) \right. \\ &\qquad \left. + 2k^{(1,2)} \left[1 - 2X^{(1)}_{2,1} - 2X^{(2)}_{3,1} + \beta^{(2)} \left(3X^{(2)}_{3,1} - 2X^{(1)}_{2,1} \right) \right] \right\} \\ &+ 2k^{(1,2)} \beta^{(2)} \left(X^{(1)}_{2,1} - 2X^{(2)}_{3,1} \right) - \mu = 0 \,, \end{aligned} \tag{C.9}$$

where $X^{(1)}_{2,g}$ and $X^{(2)}_{3,g}$ are themselves functions of $\beta^{(1)}$, $\beta^{(2)}$ and μ. More precisely, only those with $g = 1$ depend on $\beta^{(2)}$ – taking $k^{(1,2)} = 0$, $\beta^{(2)}$ must disappear from the equation.

We can observe a few things from Eq. (C.9). First, if only $k^{(0,1)}$ was nonzero – meaning that the simplicial complex would be just a graph – and we neglected $2X^{(1)}_{2,0}$, we would recover the node-based solution $\beta^{(1)}_{\rm cr} = \mu / k^{(0,1)}$. Therefore, $-2X^{(1)}_{2,0}$ (but also

$-2X_{2,1}^{(1)}$) is the correction due to 2-nodes correlations. By not neglecting it, we indeed recover the prediction made by the ELE model for a regular graph, which is

$$\beta_{\rm cr}^{(1)} = \frac{\mu}{k^{(0,1)} - \frac{1-\mu}{1-\mu/2}}, \quad \text{given } k^{(0,1)} > 0,\ k^{(g,2)} = 0\,. \tag{C.10}$$

Since $(1-\mu)/(1-\mu/2) > 0$ for $\mu \in [0,1)$, we recover the well-known fact that the pair-based threshold is higher than the node-based one. This is because a newly infected node cannot infect back the neighbor from which it was infected, for the neighbor will take, on average, $1/\mu > 1$ time steps to recover. Consistently, the two coincide for $\mu = 1$, since in this case the neighbor will be already recovered at the end of the time step; in other words, state correlations took one time step to disappear ($X_{2,g}^{(1)}$ vanishes). Besides, the fact that $(1-\mu)/(1-\mu/2) < 1$ implies that the threshold above, in discrete time, is smaller than the one predicted in continuous time, namely, $\beta_{\rm cr}^{(1)} = \mu/(k^{(0,1)}-1)$. The two converge only for $\mu \to 0$, for a node infecting another node has, almost surely, not recovered after one time step (in continuous time, the next event, almost surely, is not recovery).

Second, the rest of the terms in Eq. (C.9) yield the correction due to 3-clique correlations, which obviously depends on how many 3-cliques and 2-simplices the nodes participate in, and on the intensity $\beta^{(2)}$ of the 3-body interactions. As it should be, Eq. (C.9) becomes symmetric with respect to 3-cliques and 2-simplices when $\beta^{(2)} = 0$. Moreover, the same argument regarding the vanishing of the correlations for $\mu = 1$, given for the pair approximation, holds identical in the presence of higher-order correlations. Indeed, property (P2) in Sec. B.1.1 implies that $X_{3,g}^{(2)}$ vanishes if $X_{2,g}^{(1)}$ does. In the general case, a closed-form solution for Eq. (C.9) can be found, but its rather complicated expression makes it not very informative. Therefore, we do not show it here; we report its predictions. Figure C.3 shows how $\beta_{\rm cr}^{(1)}$ decreases with $\beta^{(2)}$ for two classes of regular simplicial 2-complexes. In particular, the periodic triangular clique complex, considered in Sec. 3.1.4, falls in one of this classes.

C.1.3.2 Friendship SCs

As an opposite case, we consider now extremely heterogeneous SCs. A Friendship graph F_n is a Windmill graph $\mathrm{Wd}(m,n)$ whose "sails" are cliques of order $m = 3$. It consists of $N = 2n+1$ nodes, where n is the number of 3-cliques incident on the central node. Starting from F_n, we convert a fraction $p_\triangle$ of the n 3-cliques in 2-faces. The central node has $k^{(0,2)} = (1-p_\triangle)n \equiv n_0$ and $k^{(1,2)} = p_\triangle n \equiv n_1$. Then, $2n_0$ of the peripheral nodes have each $k^{(0,2)} = 1$ and $k^{(1,2)} = 0$, while the remaining $2n_1$ have $k^{(1,2)} = 1$ and $k^{(0,2)} = 0$. The greater N (hence, n), the higher the degree disparity between the central node and the peripheral ones. In the large N limit, the average number of neighbors $\bar{k} = 3\frac{N-1}{N}$ tends to 3, whereas any higher m-moment diverges as N^{m-1}. Accordingly, we expect the invasion threshold to vanish in that limit.

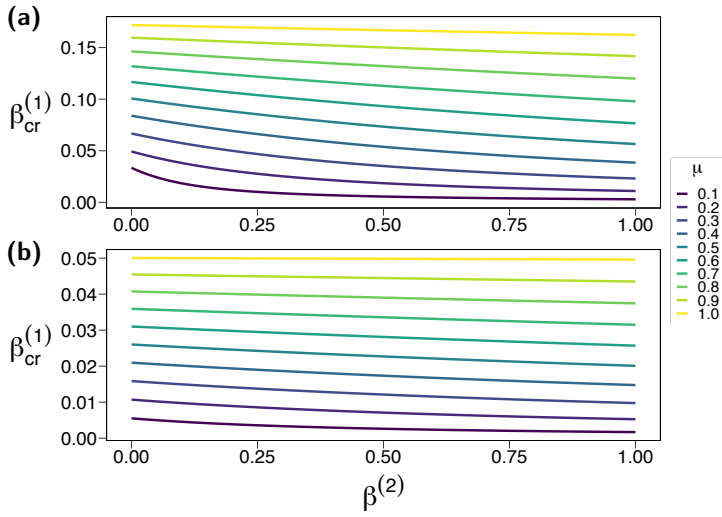


Figure C.3: The value of the invasion threshold $\beta_{cr}^{(1)}$, as computed from Eq. (C.9), is shown against $\beta^{(2)}$ for (a) any regular clique 2-complex with $k^{(0,1)} = k^{(0)} = 0$ and $k^{(1,2)} = k^{(1)} = 3$ (notice that the periodic triangular clique complex used in Sec. 3.1.4 is one of them); (b) any regular clique 2-complex with $k^{(0,1)} = k^{(0)} = 12$ and $k^{(1,2)} = k^{(1)} = 4$ (a proxy for large random SCs). Note that the HoMF approximation, Eq. (B.15), wrongly predicts $\beta_{cr}^{(1)} = \mu/k, \forall \beta^{(2)}$, where $k = 6$ in (a), $k = 20$ in (b).

In order to find a closed expression for the invasion threshold $\beta_{cr}^{(1)}$, we take advantage of the nearly block structure featured by matrix $\widetilde{\boldsymbol{M}}$. It can be partitioned as

$$\widetilde{\boldsymbol{M}} = \begin{pmatrix} \boldsymbol{B} & \boldsymbol{P} \\ \boldsymbol{C} & 0 \end{pmatrix} , \tag{C.11}$$

where $\boldsymbol{B}$ is a $(N-1)\times(N-1)$ block diagonal matrix, where each 2×2 block corresponds to the peripheral edge of a $(0,3)$- or a $(1,3)$-clique. That is, $\boldsymbol{B}$ can be put in the form

$$\boldsymbol{B} = \mathrm{diag}(\underbrace{\boldsymbol{B_0},\dots,\boldsymbol{B_0}}_{n_0},\underbrace{\boldsymbol{B_1},\dots,\boldsymbol{B_1}}_{n_1}) , \tag{C.12}$$

where

$$\boldsymbol{B_g} = \begin{pmatrix} 0 & \widetilde{M}_{(P,g)} \\ \widetilde{M}_{(P,g)} & 0 \end{pmatrix} , \tag{C.13}$$

and

$$\widetilde{M}_{(P,g)} = \frac{\beta^{(1)}\left(X^{(2)}_{1,g} + Y^{(2)}_{1,g}\right) + \left(1 - \left(1-\beta^{(1)}\right)^2\left(1 - g\beta^{(2)}\right)\right) X^{(2)}_{2,g}}{\mu - 2\beta^{(1)} Y^{(2)}_{1,g} - \left(1 - \left(1-\beta^{(1)}\right)^2\left(1 - g\beta^{(2)}\right)\right) Y^{(2)}_{2,g}} , \tag{C.14}$$

for $g = 0, 1$; where $X^{(2)}_{1,g} = 1 - 2X^{(1)}_{2,g} + X^{(2)}_{3,g}$, $X^{(2)}_{2,g} = X^{(1)}_{2,g} - X^{(2)}_{3,g}$, $Y^{(2)}_{1,g} = -X^{(2)}_{2,g}$, and $Y^{(2)}_{2,g} = -X^{(2)}_{3,g}$. Then, P is a $(N-1)\times 1$ matrix whose elements equal $\widetilde{M}_{(P,0)}$ or $\widetilde{M}_{(P,1)}$ depending on whether the peripheral node (hence the 'P') corresponding to the considered row participates, respectively, in a $(0,3)$- or a $(1,3)$-clique. Similarly, C is a $1\times(N-1)$ matrix whose elements equal $\widetilde{M}_{(C,0)}$ or $\widetilde{M}_{(C,1)}$ depending on whether the central node (hence the 'C') participates, respectively, in a $(0,3)$- or a $(1,3)$-clique with the peripheral node corresponding to the considered column; where

$$\widetilde{M}_{(C,g)} = \frac{\beta^{(1)}\left(X^{(2)}_{1,g} + Y^{(2)}_{1,g}\right) + \left(1 - \left(1-\beta^{(1)}\right)^2\left(1 - g\beta^{(2)}\right)\right) X^{(2)}_{2,g}}{\mu - \sum_{g=0,1} n_g \left[2\beta^{(1)} Y^{(2)}_{1,g} + \left(1 - \left(1-\beta^{(1)}\right)^2\left(1 - g\beta^{(2)}\right)\right) Y^{(2)}_{2,g}\right]} , \tag{C.15}$$

for $g = 0, 1$.

We now compute the determinant of

$$\widetilde{\boldsymbol{M}} - \lambda \mathbb{1}_N = \begin{pmatrix} \boldsymbol{B} - \lambda \mathbb{1}_{N-1} & \boldsymbol{P} \\ \boldsymbol{C} & -\lambda \end{pmatrix} , \tag{C.16}$$

where $\mathbb{1}_N$ is the $N\times N$ identity matrix. Thanks to the Schur complement formula [223], we can compute it as

$$\det(\widetilde{\boldsymbol{M}} - \lambda \mathbb{1}_N) = \det(\boldsymbol{B} - \lambda \mathbb{1}_{N-1}) \left[-\lambda - \boldsymbol{C}\left(\boldsymbol{B} - \lambda \mathbb{1}_{N-1}\right)^{-1} \boldsymbol{P}\right] . \tag{C.17}$$

Using the properties of block diagonal matrices [223],

$$\det(\boldsymbol{B} - \lambda \mathbb{1}_{N-1}) = \left(\lambda^2 - \widetilde{M}^2_{(P,0)}\right)^{n_0} \left(\lambda^2 - \widetilde{M}^2_{(P,1)}\right)^{n_1} , \tag{C.18}$$

implying that $\lambda_{(P,g)} \equiv \widetilde{M}_{(P,g)}$, $\forall g \in \{0,1\}$, solves $\det(\widetilde{M} - \lambda \mathbb{1}_N) = 0$; and therefore one between $\lambda_{(P,0)}$ and $\lambda_{(P,1)}$ is the dominant eigenvalue of $\boldsymbol{B}$. However, receiving contributions from peripheral nodes only, the latter can be shown to never coincide with the dominant eigenvalue of $\widetilde{\boldsymbol{M}}$. In particular, when both $n_0 > 1$ and $n_1 > 1$, we already know this is true, for the dominant eigenvalue is simple [223]. Therefore, let us suppose $\lambda \neq \lambda_{(P,g)}$, $\forall g \in \{0,1\}$, and look for $\Lambda_{\max}(\widetilde{\boldsymbol{M}})$ in the other factor, the one containing the contribution coming also from the central node. It is easily found that

$$\left(\boldsymbol{B} - \lambda \mathbb{1}_{N-1}\right)^{-1} = \mathrm{diag}(\underbrace{\widetilde{\boldsymbol{B}}_{\boldsymbol{0}}^{-1}, \ldots, \widetilde{\boldsymbol{B}}_{\boldsymbol{0}}^{-1}}_{n_0}, \underbrace{\widetilde{\boldsymbol{B}}_{\boldsymbol{1}}^{-1}, \ldots, \widetilde{\boldsymbol{B}}_{\boldsymbol{1}}^{-1}}_{n_1}) , \tag{C.19}$$

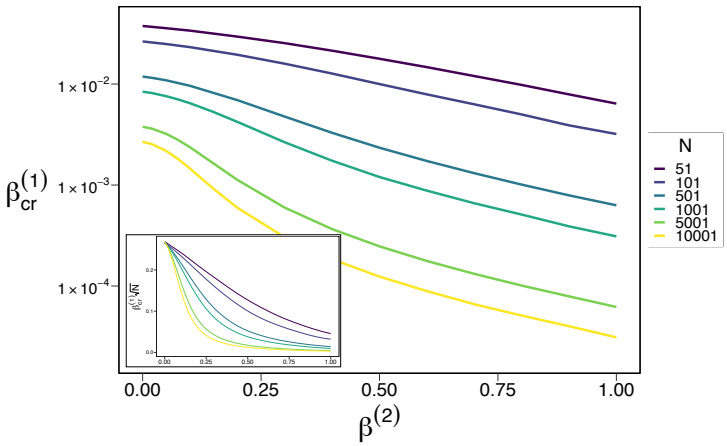


Figure C.4: The value of the invasion threshold $\beta_{cr}^{(1)}$ is shown against $\beta^{(2)}$ for simplicial 2-complexes constructed from the Friendship graph F_n, a proxy for extremely heterogeneous structures, where $n = (N-1)/2$. Here $p_\triangle = 0.5$, thus $k^{(0,2)} = k^{(1,2)} = n/2$ for the central node, $k^{(1,2)} = 1$ and $k^{(0,2)} = 0$ for half of the peripheral nodes, and $k^{(0,2)} = 0$ and $k^{(1,2)} = 1$ for the other half. The recovery probability is $\mu = 0.2$. The mean-field approximation, Eq. (B.15), wrongly predicts $\beta_{cr}^{(1)} = \mu/\bar{k}$, $\forall\beta^{(2)}$, where $\bar{k} = 3(N-1)/N$. The inset plot shows the curves $\beta_{cr}^{(1)}\sqrt{N}$ vs. $\beta^{(2)}$. These all collapse to the same value $x^* \approx 0.27$ at $\beta^{(2)} = 0$, implying that $\beta_{cr}^{(1)} \to 0$ for $N \to \infty$ as $x^*/\sqrt{N}$, in accordance with previous findings [227]. Besides, the larger N – hence the degree disparity –, the faster $\beta_{cr}^{(1)}$ decreases with $\beta^{(2)}$.

being

$$\widetilde{\boldsymbol{B}}_g^{-1} = \frac{1}{\lambda^2 - \widetilde{M}_{(P,g)}^2}\begin{pmatrix} -\lambda & -\widetilde{M}_{(P,g)} \\ -\widetilde{M}_{(P,g)} & -\lambda \end{pmatrix}, \tag{C.20}$$

the inverse matrix of $\widetilde{\boldsymbol{B}}_g = \boldsymbol{B}_g - \lambda\mathbb{1}_2$, $g = 0, 1$. With a few algebra, it follows

$$\boldsymbol{C}\left(\boldsymbol{B} - \lambda\mathbb{1}_{N-1}\right)^{-1}\boldsymbol{P} = -2\left(n_0\frac{\widetilde{M}_{(P,0)}\widetilde{M}_{(C,0)}}{\lambda - \widetilde{M}_{(P,0)}} + n_1\frac{\widetilde{M}_{(P,1)}\widetilde{M}_{(C,1)}}{\lambda - \widetilde{M}_{(P,1)}}\right). \tag{C.21}$$

We now impose $\det(\widetilde{\boldsymbol{M}} - \lambda\mathbb{1}_N) = 0$ which, using the previous result $\lambda \neq \lambda_{(P,g)}$, $\forall g \in \{0,1\}$, reduces to $-\lambda - \boldsymbol{C}\left(\boldsymbol{B} - \lambda\mathbb{1}_{N-1}\right)^{-1}\boldsymbol{P} = 0$. This can be rearranged in the form

$$\begin{aligned}&\lambda^3 - \lambda^2\left(\widetilde{M}_{(P,0)} + \widetilde{M}_{(P,1)}\right)\\ &+ \lambda\left(\widetilde{M}_{(P,0)}\widetilde{M}_{(P,1)} - 2n_0\widetilde{M}_{(P,0)}\widetilde{M}_{(C,0)} - 2n_1\widetilde{M}_{(P,1)}\widetilde{M}_{(C,1)}\right)\\ &+ 2\widetilde{M}_{(P,0)}\widetilde{M}_{(P,1)}\left(n_0\widetilde{M}_{(C,0)} + n_1\widetilde{M}_{(C,1)}\right) = 0\,.\end{aligned} \tag{C.22}$$

The largest of the solutions of Eq. (C.22) is $\lambda = \Lambda_{\max}(\widetilde{\boldsymbol{M}})$. Imposing $\Lambda_{\max}(\widetilde{\boldsymbol{M}}) = 1$ and solving with respect to $\beta^{(1)}$, we finally obtain $\beta^{(1)}_{\mathrm{cr}}$ as a function of the microscopic parameters, $\beta^{(2)}$ and μ, and of N and $p_{\triangle}$. Results are shown in Fig. C.4, where several values of N are explored for $p_{\triangle} = 0.5$. There, it is shown that the invasion threshold vanishes in the limit of large N, as we expect, given the second moment of the link-degree distribution diverges with N. For fixed N, the threshold decreases with $\beta^{(2)}$, and this dependence gets stronger with N, hence with the degree disparity between the central node and the peripheral ones. This, coupled to the similar dependence found for Dorogovtsev-Mendes simplicial complexes (Fig. 3.4) and the weaker one found for regular structures (Fig. C.3), indicates that the dependence of $\beta^{(1)}_{\mathrm{cr}}$ on $\beta^{(2)}$ grows with the triangle-degree heterogeneity. In fact, also the triangle-degree distribution has a diverging second moment with N. Despite their simplicity, Friendship simplicial complexes capture some basic dynamical implications of more complex heterogeneous structures.

c.1.4 *Continuous-time model*

Let us first sketch how to derive the continuous-time limit of the MECLE. This requires making the following substitutions,

$$\mu \longrightarrow \mu\,\Delta t\ ,$$
$$\beta^{(s)} \longrightarrow \beta^{(s)}\,\Delta t\ ,$$

where now μ and $\beta^{(s)}$ represent rates, i.e., probabilities per unit time. Once multiplied by the duration of the time interval Δt, they give the probabilities of the respective processes to occur during that interval. Then, we take $\Delta t \to 0$. Consequently, all those terms appearing in the equations as second or greater powers of Δt becomes negligible. In other words, only single-node state changes need to be considered, as they are the only $\mathcal{O}(\Delta t)$ processes. The resulting system of ODEs is thus quite simpler than the respective difference equations in discrete-time.

The linearized equations around the inactive state, as Eq. (B.3), include all the possible infection channels within a focal facet, hence any infection rate up to $\beta^{(n)}$, supposed n is the dimension of the facet. Consequently, all such rates up to $\beta^{(n)}$ do show up in the threshold condition, Eq. (B.12), for the invasion threshold, $\beta^{(1)}_{\mathrm{cr}}$. Focusing on simplicial 2-complexes, Fig. C.5 reports how $\beta^{(1)}_{\mathrm{cr}}$ depends on $\beta^{(2)}$ as computed from the linearized equations for a Dorogovtsev-Mendes SC.

Since the MECLE are by construction defined over strictly 0-connected SCs, their continuous-time model is too. To obtain a continuous-time model valid for simplicial complexes of any q-connectedness, we rather define the model by considering time as continuous from the beginning. We provide equations for simplicial 2-complexes. The extension to any dimension is straightforward.

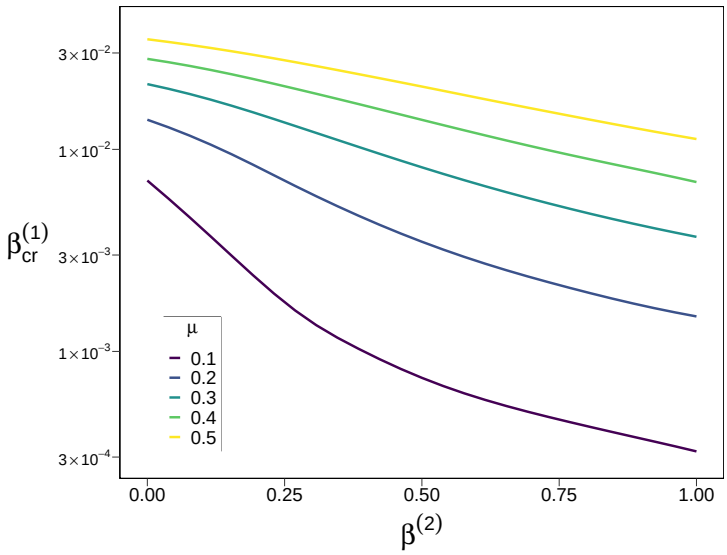


Figure C.5: The value of the invasion threshold $\beta_{cr}^{(1)}$, as predicted in the continuous-time limit of the MECLE, is shown against $\beta^{(2)}$ for a Dorogovtsev-Mendes simplicial complex with $\bar{k}^{(0,1)} = 1.10$ and $\bar{k}^{(1,2)} = 1.45$. The results are very similar to their continuous-time counterparts, shown in Fig. 3.4.

Let us define the following overall infection rates,

$$v_i = \beta^{(1)} \sum_{l\in\Gamma_i^{(1)}} P_{il}^{SI}\,, \quad w_i = \beta^{(2)} \sum_{l,m\in\Gamma_i^{(2)}} P_{ilm}^{SII}\,,$$

$$v_{i(j)}^{\sigma_j} = \beta^{(1)} \sum_{l\in\Gamma_i^{(1)}\,j} P_{ijl}^{S\sigma_j I}\,, \quad w_{i(j)}^{\sigma_j} = \beta^{(2)} \sum_{\{l,m\}\in\Gamma_i^{(2)}\setminus j} P_{ijlm}^{S\sigma_j II}\,, \quad w_{i,j} = \beta^{(2)} \sum_{l\in\Gamma_{i,j}} P_{ijl}^{SII}\,,$$

$$v_{i(jk)}^{\sigma_j\sigma_k} = \beta^{(1)} \sum_{l\in\Gamma_i^{(1)}\setminus\{j,k\}} P_{ijkl}^{S\sigma_j\sigma_k I}\,, \quad w_{i(jk)}^{\sigma_j\sigma_k} = \beta^{(2)} \sum_{\{l,m\}\in\Gamma_i^{(2)}\setminus\{j,k\}} P_{ijklm}^{S\sigma_j\sigma_k II}\,, \quad w_{i,j(k)}^{\sigma_k} = \beta^{(2)} \sum_{l\in\Gamma_{i,j}\setminus k} P_{ijkl}^{SI\sigma_k I}\,,$$

with $w_{i,j} = w^{\sigma_k}_{i,j(k)} = 0$, $\forall k$, if the 2-clique $\{i,j\}$ is not face of any 2-simplex. The dynamic equations then read,

$$\dot{P}^I_i = -\mu P^I_i + v_i + w_i\,, \tag{C.23a}$$

$$\dot{P}^{IS}_{ij} = -\left(\mu+\beta^{(1)}\right)P^{IS}_{ij} - v^I_{j(i)} - w_{j,i} - w^I_{j(i)} + \mu P^{II}_{ij} + v^S_{i(j)} + w^S_{i(j)}\,, \tag{C.23b}$$

$$\dot{P}^{II}_{ij} = -2\mu P^{II}_{ij} + \beta^{(1)}P^{IS}_{ij} + v^I_{j(i)} + w_{j,i} + w^I_{j(i)} + \beta^{(1)}P^{SI}_{ij} + v^I_{i(j)} + w_{i,j} + w^I_{i(j)}\,, \tag{C.23c}$$

$$\begin{aligned}\dot{P}^{ISS}_{ijk,g} =& -\left(\mu+2\beta^{(1)}\right)P^{ISS}_{ijk,g} - v^{IS}_{k(ij)} - w^S_{k,i(j)} - w^{IS}_{k(ij)}\\ &+\mu P^{ISI}_{ijk,g} - v^{IS}_{j(ik)} - w^S_{j,i(k)} - w^{IS}_{j(ik)} + v^{SS}_{i(jk)} + w^{SS}_{i(jk)}\,,\end{aligned} \tag{C.23d}$$

$$\begin{aligned}\dot{P}^{IIS}_{ijk,g} =& -\left(2\mu+2\beta^{(1)}+g\beta^{(2)}\right)P^{IIS}_{ijk,g} - v^{II}_{k(ij)} - w^I_{k,i(j)} - w^I_{k,j(i)} - w^{II}_{k(ij)} + \mu P^{III}_{ijk,g}\\ &+\beta^{(1)}P^{ISS}_{ijk} + v^{IS}_{j(ik)} + w^S_{j,i(k)} + w^{IS}_{j(ik)}\\ &+\beta^{(1)}P^{SIS}_{ijk,g} + v^{IS}_{i(jk)} + w^S_{i,j(k)} + w^{IS}_{i(jk)}\,,\end{aligned} \tag{C.23e}$$

$$\begin{aligned}\dot{P}^{III}_{ijk,g} =& -3\mu P^{III}_{ijk,g} + 2\left(\beta^{(1)}+g\beta^{(2)}\right)\left(P^{IIS}_{ijk,g}+P^{ISI}_{ijk,g}+P^{SII}_{ijk,g}\right)\\ &+v^{II}_{k(ij)} + w^I_{k,i(j)} + w^I_{k,j(i)} + w^{II}_{k(ij)}\\ &+v^{II}_{j(ik)} + w^I_{j,i(k)} + w^I_{j,k(i)} + w^{II}_{j(ik)}\\ &+v^{II}_{i(jk)} + w^I_{i,k(j)} + w^I_{i,j(k)} + w^{II}_{i(jk)}\,.\end{aligned} \tag{C.23f}$$

These equations are then closed by specifying a moment closure for the 4-node and 5-node state probabilities in terms of the 2-node and 3-node ones.

C.2 SUPPLEMENTARY INFORMATION FOR SEC. 3.2

C.2.1 *Generation and properties of the hypergraphs used in Fig. 3.8*

The hypergraphs used in Fig. 3.8 are derived from two real-world datasets. One dataset contains face-to-face interactions recorded during a conference [293], the other one proximity data recorded within a university campus [294]. The procedure we use to generate the hypergraphs is the same for both datasets.

These consist of time-resolved interactions (each representing a face-to-face interaction or proximity between two people) which, once aggregated, yields very dense networks [293]. Therefore, we first build a static pairwise network where each edge is assigned a weight equal to its number of appearances (i.e., how many times the interaction between the two agents has been detected throughout the entire observation time) and then threshold it.

Starting from an empty network with only the nodes in the dataset ($N = 219$ for the conference's and $N = 672$ for the university campus'), edges are listed in decreasing order of weight and added to the network starting from the first one. Since some nodes only participate to edges with very low weight, waiting until all nodes are included

would yield a network identical to the original one, except for just few missing edges. To avoid this, we stop including edges when the 95% of the nodes has been connected to some other node (notice that disconnected components may still exist at this point), indeed thresholding the original network[1]. The remaining degree-0 nodes are then connected to the other nodes at random. If the network is disconnected, the connected components are connected to the largest connected component by adding an edge at random between each of them and the latter. In practice, at the moment in which the 95% of the nodes is reached, there exists a component containing almost all nodes and very few other components of very few nodes. Consequently, the few edges added to connect the network do not affect the properties of the thresholded network.

The binary network obtained in this way represents the pairwise backbone to which we add three-body interactions in order to get rank-3 hypergraphs. To do this, we first list all the 3-cliques in the network[2]. Then, to each 3-clique, we add a three-body interaction (i.e., a 3-edge containing the three nodes) with probability h, such that, if the addition occurs, a 2-simplex (triangle) is formed. Otherwise (occurring with probability $1-h$), the 3-edge is added to three unconnected nodes chosen at random, so that a three-body interaction not overlapped with two-body interactions is formed. Notice that the total number of three-body interactions added is independent from h; only their distribution over the system changes with it. Setting $h=0$ yields a linear hypergraph. Increasing h, more and more frequently three-body interactions overlap with two-body interactions. At $h=1$, the structure becomes a simplicial 2-complex.

Some of the basic properties of the generated hypergraphs are listed below. Figure C.6 reports their 2-degree and 3-degree distributions, which are well reproduced by exponential and gaussian distributions, respectively.

CONFERENCE'S DATASET. The hypergraphs consist of $N=219$ nodes. The 2-degree $\kappa^{(1)}$ (number of 2-edges incident on a node) is distributed heterogeneously. The first and the second raw moments of the 2-degree distribution are $\langle\kappa^{(1)}\rangle\approx 33.05$ and $\langle{\kappa^{(1)}}^2\rangle\approx 2034.16$, giving a high variance of $\mathrm{var}(\kappa^{(1)})\approx 941.84$. The structure also shows 2-degree assortativity (coefficient $r=0.1$ [500]). At last, the first and the second raw moments of the 3-degree (number of 3-edges incident on a node) distribution are $\langle\kappa^{(2)}\rangle\approx 16.22$ and $\langle{\kappa^{(2)}}^2\rangle\approx 280.03$, giving a low variance of $\mathrm{var}(\kappa^{(2)})\approx 16.97$.

1 The original networks include many large cliques. Since our model assumes cliques of up to 3 nodes, when an edge is included as above we check whether a 4-clique formed, in which case the edge is ignored. Including 4-cliques or larger ones does not change qualitatively the results. On the other hand, avoiding them makes the network less dense and the phenomenology easier to appreciate.

2 Notice that some 3-cliques share two nodes with other 3-cliques, meaning that those 2-cliques which are part of more than one 3-clique appear multiple times in the list. Each appearance is considered as a different interaction. An alternative method that would avoid repeated 2-cliques would consist in finding an edge-disjoint edge clique cover of the network, as done in Sec. 3.1. This would change the number of 2-cliques and 3-cliques in the network, but the results would be qualitatively unaffected.

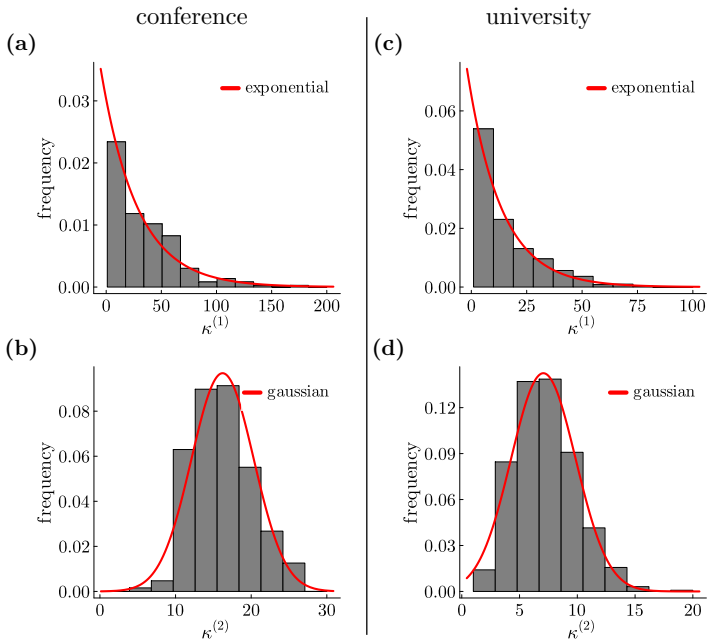


Figure C.6: Histograms of the 2-degree and 3-degree distributions of the hypergraphs generated from the conference's dataset ((a) and (b)) and the university campus' dataset ((c) and (d)). These are well reproduced by exponential (mean: $1/\langle\kappa^{(1)}\rangle$) and gaussian (mean: $\langle\kappa^{(2)}\rangle$; variance: $\mathrm{var}(\kappa^{(2)})$) distributions, respectively (red curves).

UNIVERSITY CAMPUS'S DATASET. The hypergraphs consist of $N = 672$ nodes. The 2-degree $\kappa^{(1)}$ is distributed heterogeneously. The first and the second raw moments of the 2-degree distribution are $\langle\kappa^{(1)}\rangle \approx 15.3$ and $\langle{\kappa^{(1)}}^2\rangle \approx 479.67$, giving a high variance of $\mathrm{var}(\kappa^{(1)}) \approx 245.56$. The structure also shows 2-degree assortativity (coefficient $r = 0.19$ [500]). At last, the first and the second raw moments of the 3-degree distribution are $\langle\kappa^{(2)}\rangle \approx 7.06$ and $\langle{\kappa^{(2)}}^2\rangle \approx 57.68$, giving a low variance of $\mathrm{var}(\kappa^{(2)}) \approx 7.84$.

C.2.2 *Results for the simplicial SIR model*

We show here the results of numerical simulations we performed using a higher-order susceptible-infectious-recovered (SIR) contagion model. This generalization of the SIR model has been very recently analyzed in detail by Lv et al. [501]. The authors showed that, as in the SIS model, sufficiently high values of the three-body infection rate, $\beta^{(2)}$, make the phase transition for the final attack rate (R_∞) discontinuous in both homogeneous and heterogeneous simplicial complexes. Using however

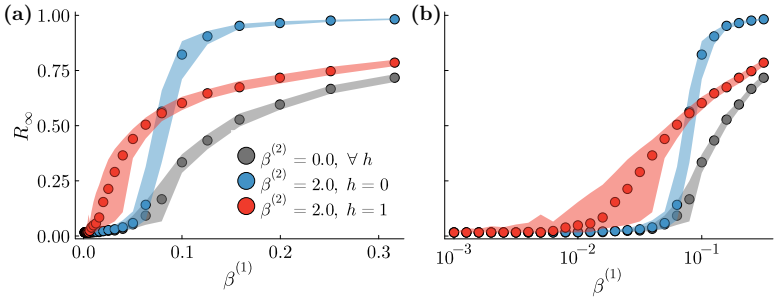


Figure C.7: Numerical simulations performed on the hypergraphs constructed from the university campus' [294] datasets when the contagion dynamics is a SIR process. (a) Final attack rate, R_∞, versus the two-body infection rate, $\beta^{(1)}$. Points denote medians computed over 100 random initializations; ribbons cover from the 25-th to the 75-th percentile. The added three-body interactions form triangles with probability $h = 0$ (linear hypergraph; blue curve) and $h = 1$ (simplicial complex; red curve). Clearly, turning off the three-body interactions ($\beta^{(2)} = 0$), varying h has no effect on the dynamics (grey curve). (b) Same as panel (a), but with logarithmic abscissa to stretch the low-$\beta^{(1)}$ interval and better appreciate the shift of the invasion threshold.

a node-based approximation, their approach is insensible to the way in which two- and three-body interactions are arranged in the structure, hence to their degree of overlap. Consequently, it is not possible to discern a simplicial complex from a linear hypergraph (or any other intermediate structure).

In Fig. C.7, we show the results of the numerical simulations performed on hypergraphs generated from the university campus' dataset. These confirm the generality of the phenomenology our mean-field model revealed: (i) three-body interactions affect the invasion threshold only if they overlap with two-body interactions ($h > 0$); (ii) a larger overlap (higher h) implies lower invasion thresholds but also smaller outbreaks; and (iii) varying exclusively the degree of overlap can change the nature of the phase transition. About the last point, it should be noted that the discontinuity of the transition for $\beta^{(2)} = 2$ and $h = 0$ is blurred by strong finite-size effects.

C.3 SUPPLEMENTARY INFORMATION FOR SEC. 3.3

C.3.1 *A-GAME: Additional results*

We report here few additional findings for the adaptive hypergraph model studied in Sec. 3.3.2.1, which results are displayed in Fig. 3.11. Figure C.8(a) helps distinguish the different regimes determined by the rewiring accuracy, η. Figures C.8(b) and (c) then mimic Figs. 3.11(b) and (c) but for complex contagion ($\nu = 2$), showing that increasing either the speed (γ) or the accuracy (η) of rewiring heightens the persistence threshold,

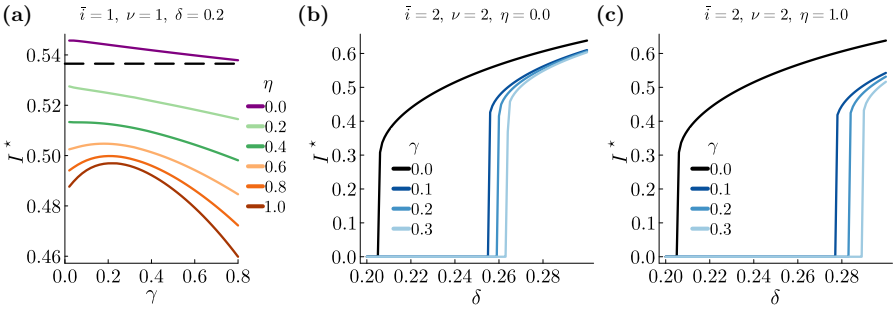


Figure C.8: (a) Horizontal slices of the phase diagram in Fig. 3.11(a) ($\bar{i} = \nu = 1$, $\delta = 0.2$). The purple curve corresponds to detrimental rewiring, where the equilibrium fraction of infected nodes, $I^\star$, is larger than in absence of rewiring (dashed line); green curves represent increasingly beneficial rewiring; orange curves denote non-monotonic beneficial rewiring, leading to a local maximum in $I^\star$. (b and c) As panels (b) and (c) in Fig. 3.11, but for complex contagion ($\nu = 2$). Lower (solid) and upper (dashed) branches of the prevalence $I^\star$ versus the per-node infection rate δ using different rewiring rates γ for, respectively, random ($\eta = 0.0$) and perfectly targeted ($\eta = 1.0$) rewiring. While there is no invasion threshold for this dynamics, the persistence threshold, as for simple contagion, increases due to either a faster or more accurate rewiring.

as for simple contagion. Differently from the latter (see Fig. 3.11(d)), however, since the threshold-like complex contagion does not admit an invasion threshold by construction, there exists no value of $\bar{i}$ able to make the transition continuous. Lastly, Fig. C.9 shows how the degree evolves differently for susceptible and infected nodes. In particular, the reduced ability of susceptible nodes to avoid infectious groups when the rewiring accuracy is lower, explains why the contagion persists more (i.e., the persistence threshold diminishes) under a poorer accuracy (η closer to 0).

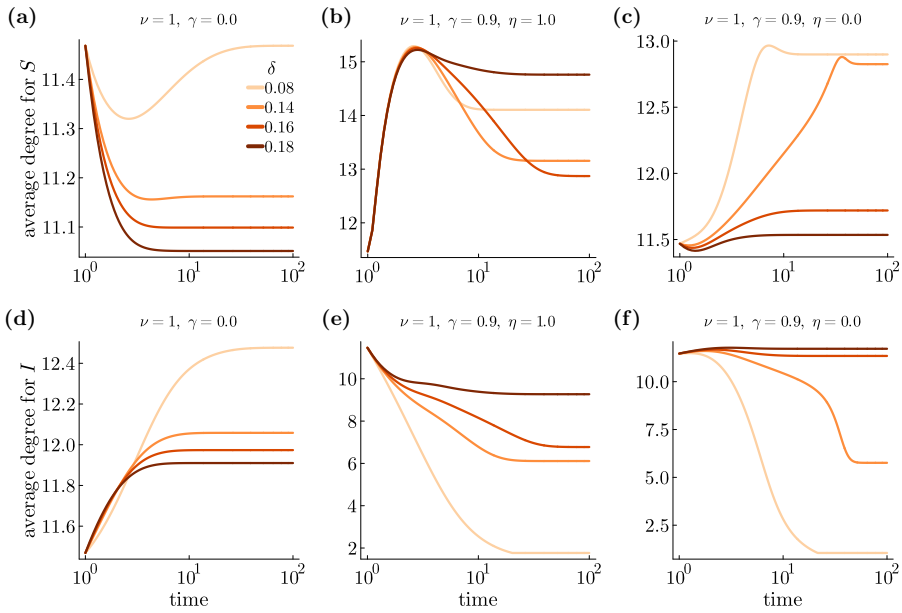


Figure C.9: Analysis of the evolution of the degree of susceptible ((a)-(c)) and infected ((d)-(f)) nodes for the adaptive hypergraphs in Figs. 3.11(b) and (c). ((a) and (d)) In absence of rewiring ($\gamma = 0.0$), as in pairwise networks, infections localize in higher degree nodes, so a positive (negative) correlation between being infected (susceptible) and degree emerges. In the case $\delta = 0.08$, the only one in which $I^\star = 0$, the degree of susceptible nodes gets back to the initial (uncorrelated) value given every node is now susceptible and the average degree is conserved under no rewiring. ((b) and (e)) When rewiring is sufficiently accurate (here is perfect, $\eta = 1.0$) we find, except for a different transient that alters the group size distribution and thus the average degree, the same positive (negative) correlation between being infected (susceptible) and degree. Notice that $I^\star = 0$ except for $\delta = 0.18$. ((c) and (f)) When rewiring is sufficiently inaccurate (here is uniformly random, $\eta = 0.0$), instead, while the susceptible nodes still mature higher degrees, the connectivity of the infected nodes can remain unaffected or even slightly increase. In this example, this happens for $\delta = 0.16$ and 0.18, for which the system is still above the persistence threshold ($I^\star > 0$). Decreasing accuracy prevents susceptible nodes from effectively isolating infected nodes, thus producing a lower persistence threshold.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

D

METHODS FOR SEC. 4

D.1 METHODS FOR SECTION 4.2

D.1.1 *Mixing parameter estimation from vaccination data*

We infer the mixing parameter α of SARS-CoV-2 vaccine adoption through an age-stratified approach for different regions/countries. We make use of an empirical contact matrix $\mathbf{C}$ [212] and the vaccination coverage $v_i(t)$ in age-strata $i = 1, 2, \dots, M$ at time t [363–366]. The entry C_{ij} of the contact matrix represents the average number of daily contacts of an individual in age-strata i with an individual in age-strata j. Vaccination data is stratified into intervals of 10 years $(0-9, 10-19, \dots, 70-79, 80+)$, while the contact matrices use 5 year bins. Therefore, as a first step, we average the empirical contact matrices to transform them into 10 year bins according to the census data.

Generally, the empirical contact matrices do not respect the balance equation $C_{ij}N_i = C_{ji}N_j$, where N_i represents the number of individuals in age-strata i. There are various approaches to fix this issue [502]. Here, we defined the matrix $\mathbf{C}$ as

$$C_{ij} = \frac{N_i \widetilde{C}_{ij} + N_j \widetilde{C}_{ji}}{N_i} \,, \tag{D.1}$$

such that the balance equation is met, being $\widetilde{\mathbf{C}}$ the empirical matrix. To infer $\alpha(t)$ we leverage its relation with the homophily, $h(t)$; i.e., $\alpha(t) = (1 - h(t))/[2V(1-V)]$ (Eq. (4.37)), where V is the fraction of vaccinated individuals in the overall population. The homophily is computed as

$$h(t) = \frac{\sum_{i,j=1}^{M} N_i \left[v_i(t) v_j(t) + (1 - v_i(t))(1 - v_j(t)) \right]}{\sum_{i,j=1}^{M} C_{ij} N_i} \,. \tag{D.2}$$

D.1.2 *Procedure for the distribution of binary types in a network*

Given a (weighted) network of N nodes, we want to assign the types '0' and '1' to the nodes in order to reach a desired value for the homophily, h. This is computed as the sum of the weights over the homophilic edges (0-0 and 1-1), normalized by the sum over all the edges. Denoted with $a_i \in \{0,1\}$ the type of node i and with w_{ij} the weight of the edge between nodes i and j, the homophily reads

$$h = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\left[a_i a_j + (1-a_i)(1-a_j)\right]}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \,. \tag{D.3}$$

The mixing parameter, α, is then obtained from Eq. (4.37) as $\alpha = (1-h)/[2\rho(1-\rho)]$, being $\rho = \sum_{i=1}^{N} a_i/N$ the proportion of nodes of type 1 in the network.

To reach a desired value for α, we initially distribute the vaccine at random ($\alpha \approx 1$) and then iteratively swap the type of two randomly selected neighboring nodes whenever this leads to a decrease in α, until the latter attains the preset value (± 0.01). Additionally, if the average strength is higher (lower) for type 1 nodes than for type 0's, a swap is only allowed if it increases (decreases) the strength of the latter. Without this condition, the algorithm induces a spurious correlation between type and the strength of a node. Given the role that more connected nodes play in driving the spreading dynamics, such correlation can importantly affect the results. Therefore, only innocuous, not correlating swaps are carried out.

We apply the algorithm for the case of vaccination in Sec. 4.2.1, where type 1 is assigned to vaccinated individuals, and the case of digital proximity tracing apps in Sec. 4.2.2, where type 1 is assigned to individuals adopting the app. In both cases, the type assignment is made on top of weighted networks resulting from the aggregation of temporal networks generated by a five-minutes binning of the data. Accordingly, the weight w_{ij} measures the total amount of time nodes i and j interacted, understood as the number of five-minutes bins this occurred (e.g., $w_{ij} = 3$ means that the edge appeared in 3 bins during the entire observation time).

D.1.3 *Efficacy reduction of contact tracing*

The analysis presented in Sec. 4.2.2 is based on the assumption that app users infected by other app users do not further transmit the disease, meaning that the contact tracing app is fully effective. Clearly, in real scenarios contact tracing may lose some degree of efficacy due to different factors. For example, efficacy may be reduced by individuals that do not to self-isolate after notification. Furthermore, since also presymptomatic individuals could be infectious, substantial time may pass between infection

and notification, leaving the door open for the disease to spread further. Notification delays might be also due to technological/organizational reasons.

Here we introduce both these factors by generalizing Eqs. (4.46)-(4.48). Specifically, we indicate with $\epsilon \in [0,1]$ the probability that an app user, once notified about a contact with an infectious app user, does not self-isolate (and, therefore, potentially extend the infection chain). Then, with $\gamma \in [0,1]$, we indicate the reduction of the infectious period (in principle, $1/\mu$) of an infected app user due to self-isolation *after* notification. The larger γ, the larger the delay between infection and notification, hence longer the period during which the infectious adopter can infect. Combined the two factors, an individual in compartment I_{AA} has an effective infectious period rescaled by $\xi = \epsilon + \gamma(1-\epsilon)$. Indeed, such an user is infectious if (i) it does not isolate (with probability ϵ), no matter how late it receives the notification; (ii) it isolates (with probability $1-\epsilon$), but getting to the user takes the notification a time γ, in units of infectious period. Rescaling the infectious period in compartment I_{AA} is here equivalent to rescale the associated infection rate as $\beta \to \xi\beta$. Perfect tracing is thus recovered for $\epsilon = \gamma = 0$, while a completely ineffective tracing corresponds to $\epsilon = \gamma = 1$. In this setup, the number of infected individuals in the compartments I_{N} and I_{AA} can also grow through the interaction with the ones in I_{AA}. In the end, Eqs. (4.46)-(4.48) take the new form,

$$I_{\mathrm{N}}(t+1) = I_{\mathrm{N}}(t)(1-\mu) + \beta\left[\kappa_{\mathrm{N|N}} I_{\mathrm{N}}(t) + \kappa_{\mathrm{A|N}}\left(I_{\mathrm{AN}}(t) + \xi I_{\mathrm{AA}}(t)\right)\right] S_{\mathrm{N}}(t)\,, \tag{D.4}$$

$$I_{\mathrm{AN}}(t+1) = I_{\mathrm{AN}}(t)(1-\mu) + \beta\kappa_{\mathrm{N|A}} I_{\mathrm{N}}(t) S_{\mathrm{A}}(t)\,, \tag{D.5}$$

$$I_{\mathrm{AA}}(t+1) = I_{\mathrm{AA}}(t)(1-\mu) + \beta\kappa_{\mathrm{A|A}}\left[I_{\mathrm{AN}}(t) + \xi I_{\mathrm{AA}}(t)\right] S_{\mathrm{A}}(t)\,, \tag{D.6}$$

Consistently, Eqs. (4.46)-(4.48) as well as the phenomenology they generate are recovered for $\xi = 0$. The linearization of Eqs. (D.4)-(D.6) yields the following next-generation matrix,

$$\mathrm{NGM} = \frac{\beta}{\mu}\begin{pmatrix} \kappa_{\mathrm{N|N}} & \kappa_{\mathrm{A|N}} & \xi\kappa_{\mathrm{A|N}} \\ \kappa_{\mathrm{N|A}} & 0 & 0 \\ 0 & \kappa_{\mathrm{A|A}} & \xi\kappa_{\mathrm{A|A}} \end{pmatrix}. \tag{D.7}$$

The effective reproduction number, $R = \Lambda_{\max}(\mathrm{NGM})$, then reads

$$R = \frac{R_0}{2}\Big[1 - \alpha T + \xi(1-\alpha(1-T)) + \sqrt{\left[1-\alpha T - \xi\left(1-\alpha\left(1-T\right)\right)\right]^2 + 4\alpha^2 T(1-T)}\Big]\,. \tag{D.8}$$

This expression correctly reduces to $R = R_0$ for $\xi = 1$, as a standard SIR is recovered. We observe that the functional form of R with respect to α is independent of $\xi \in [0,1)$. Following the same analysis performed in the manuscript, we can thus show that the phenomenology predicted in the fully effective case is still valid.

Firstly, we impose $\frac{dR}{d\alpha} = 0$ to prove the existence of $\alpha = \alpha^*$ at which R reaches its minimum value, for any given values of T and ξ. As before, it yields a second-order equation having a single positive solution $\alpha = \alpha^*$, reading

$$\alpha^* = \frac{T - \xi\left[T + 1 - \xi\,(2 - T) + \xi^2\,(1 - T)\right] + \sqrt{(T + \xi - \xi T)^2\,(1 - \xi)^3}}{(1 - \xi)\left[(4 - 3T)\,T - 2\xi T\,(1 - T) + \xi^2\,(1 - T)^2\right]}\,. \tag{D.9}$$

By differentiation, one sees that α^* is a decreasing function of ξ. Then, imposing $\alpha^* \leqslant 1$, we find the upper bound T^* for the app coverage T, below which α^* lies in the homophilic range, going from fully assortativeness ($\alpha = 0$) to homogeneous mixing ($\alpha = 1$). It reads,

$$T^* = \frac{1}{3 + \xi}\left(1 + \frac{1 + \xi}{\sqrt{1 - \xi}}\right)\,, \tag{D.10}$$

reducing to $T^* = 2/3$ for $\xi = 0$. Additionally, T^* takes physical values, that is $T^* \leqslant 1$, if and only if $\xi \leqslant \left(\sqrt{5} - 1\right)/2 \approx 0.618$ [1]. Consequently, for greater values of ξ, $\alpha^* < 1$ for any $T \in [0, 1]$, meaning the minimum for R is always in the homophilic region when tracing is poor enough.

Secondly, we find that eradication (i.e., $R < 1$) is possible for $\alpha \in (\alpha_c^-, \alpha_c^+)$, being

$$\begin{aligned}\alpha_c^{\pm} =& \frac{1}{2R_0 T\,(1 - T)\,(1 - \xi)}\,\big[T + \xi\,(1 - T) - \xi R_0 \\ &\pm \sqrt{\left[T + \xi\,(1 - T - R_0)\right]^2 - 4T\,(1 - T)\,(1 - \xi)\left[R_0\,(1 + \xi) - 1 - \xi {R_0}^2\right]}\big]\end{aligned} \tag{D.11}$$

the two roots of the equation $R = 1$.

Alike Fig. 4.7, Fig. D.1 condenses the outcome of the above analysis; precisely, for the dynamics obtained with $\gamma = 0.30$ and $\epsilon = 0.20$ ($\xi = 0.44$). Overall, we observe that the reduction in efficacy of the tracing process emphasizes the phenomenology found in the fully effective case.

D.1.4 *Annealed network formulation*

The model presented in Sec. 4.2.2 has been derived under the homogeneous mean field approximation, which assumes that all the individuals in the population have the same number of contacts k. In that formulation we see, from Eq. (4.52), that the dependence on k in the effective reproduction number factors out. In what follows we show that the same happens when the heterogeneity in the number of contacts (degree heterogeneity) is taken into account, given the latter do not correlate with app

1 As a matter of curiosity, this value exactly equals $1/\varphi$, being φ the golden ratio.

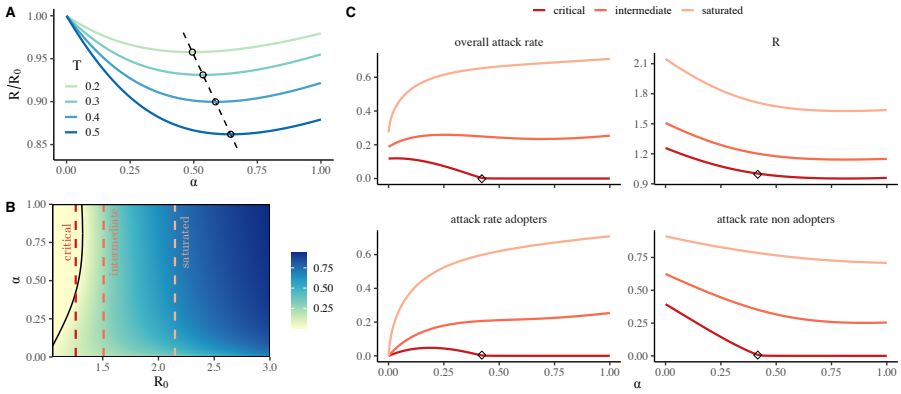


Figure D.1: Analogous illustration to Fig. 4.7, but for $\gamma = 0.30$ and $\epsilon = 0.20$, giving $\xi = 0.44$. **A:** Normalized reproduction number, R/R_0, for different values of adoption, T, below $T^* \approx 0.85$. Dots indicate the minimum at α^*, while the dashed line shows its variation for $T \in [0.1, 0.55]$. **B:** Attack rate as a function of α and R_0. The solid line indicates the threshold $\alpha_c^{\pm}$ for which $R = 1$. Colored, dashed lines denote the dynamical regimes: critical, intermediate and saturated. Adoption was fixed as $T = 0.7$. **C:** Top panels show the attack rate and the reproduction number for the different regimes defined in B. The specific attack rates for adopters and non adopters are reported in the bottom panels. Black diamonds indicate α_c^-, at which $R = 1$.

adoption. Thus, in the annealed network formalism [80, 209], Eqs. (4.46)-(4.50) are generalized to take the form,

$$I_{\mathrm{N}}^{k}(t+1) = I_{\mathrm{N}}^{k}(t)(1-\mu) + \frac{\beta}{k}\sum_{k'}\mathbf{C}_{kk'}\left[\kappa_{\mathrm{N}|\mathrm{N}} I_{\mathrm{N}}^{k'}(t) + \kappa_{\mathrm{A}|\mathrm{N}} I_{\mathrm{AN}}^{k'}(t)\right] S_{\mathrm{N}}^{k}(t)\,, \tag{D.12}$$

$$I_{\mathrm{AN}}^{k}(t+1) = I_{\mathrm{AN}}^{k}(t)(1-\mu) + \frac{\beta}{k}\sum_{k'}\mathbf{C}_{kk'}\kappa_{\mathrm{N}|\mathrm{A}} I_{\mathrm{N}}^{k'}(t) S_{\mathrm{A}}^{k}(t)\,, \tag{D.13}$$

$$I_{\mathrm{AA}}^{k}(t+1) = I_{\mathrm{AA}}^{k}(t)(1-\mu) + \frac{\beta}{k}\sum_{k'}\mathbf{C}_{kk'}\kappa_{\mathrm{A}|\mathrm{A}} I_{\mathrm{AN}}^{k'}(t) S_{\mathrm{A}}^{k}(t)\,, \tag{D.14}$$

$$S_{\mathrm{N}}^{k}(t+1) = S_{\mathrm{N}}^{k}(t) - \frac{\beta}{k}\sum_{k'}\mathbf{C}_{kk'}\left[\kappa_{\mathrm{N}|\mathrm{N}} I_{\mathrm{N}}^{k'}(t) + \kappa_{\mathrm{A}|\mathrm{N}} I_{\mathrm{AN}}^{k'}(t)\right] S_{\mathrm{N}}^{k}(t)\,, \tag{D.15}$$

$$S_{\mathrm{A}}^{k}(t+1) = S_{\mathrm{A}}^{k}(t) - \frac{\beta}{k}\sum_{k'}\mathbf{C}_{kk'}\left[\kappa_{\mathrm{N}|\mathrm{A}} I_{\mathrm{N}}^{k'}(t) + \kappa_{\mathrm{A}|\mathrm{A}} I_{\mathrm{AN}}^{k'}(t)\right] S_{\mathrm{A}}^{k}(t)\,, \tag{D.16}$$

where $\mathbf{C}_{kk'} = kP(k'|k)$ are the entries of the *connectivity matrix* $\mathbf{C}$, as defined in [80]. $P(k'|k)$ is the probability that a node of degree k connects to a node of degree k',

and thus accounts for potential degree-degree correlations in the network. The next-generation matrix is now given by the tensor

$$\mathrm{NGM} = \frac{\beta}{\mu k}\begin{pmatrix} \kappa_{\mathrm{N|N}}\mathbf{C}_{kk'} & \kappa_{\mathrm{A|N}}\mathbf{C}_{kk'} & 0 \\ \kappa_{\mathrm{N|A}}\mathbf{C}_{kk'} & 0 & 0 \\ 0 & \kappa_{\mathrm{A|A}}\mathbf{C}_{kk'} & 0 \end{pmatrix} = \frac{\beta}{\mu k}\begin{pmatrix} \kappa_{\mathrm{N|N}} & \kappa_{\mathrm{A|N}} & 0 \\ \kappa_{\mathrm{N|A}} & 0 & 0 \\ 0 & \kappa_{\mathrm{A|A}} & 0 \end{pmatrix} \otimes \mathbf{C}_{kk'} \,. \tag{D.17}$$

This splits into the tensor product of two matrices that encode the effect of mixing in the app adoption and the network structure, respectively. As a consequence, its spectral radius factorizes as well, reading

$$R = \frac{\beta\Lambda(\mathbf{C})}{2\mu k}\left[1 - \alpha T + \sqrt{(1-\alpha T)^2 + 4\alpha^2 T(1-T)}\right] , \tag{D.18}$$

where $\Lambda(\mathbf{C})$ is the spectral radius of the connectivity matrix $\mathbf{C}_{kk'}$. This allows us to redefine the basic reproduction number as $R_0 = \beta\Lambda(\mathbf{C})/\mu$, which simply reduces to $\beta k/\mu$ in the homogeneous mean field approximation. Meanwhile, the contribution of mixing in the app adoption remains unaffected. Accordingly, the effective reproduction number, R, is still given by Eq. (4.52), meaning that the phenomenology presented in the manuscript is qualitatively unchanged. The same argument also holds (i) when including the effects presented in Sec. D.1.3, in which case R would still be given by Eq. (D.8), but with $R_0 = \beta\Lambda(\mathbf{C})/\mu$; (ii) for any other type of heterogeneity – encoded in a matrix, like $\mathbf{C}$ encodes degree information – that is introduced into the dynamics, as long as it is does not correlate with the app adoption; and (iii) considering other prophylactic measures, as this could only change the dimension and the entries of the transmission part of next-generation matrix – for instance, for vaccine adoption, that part would be given by Eq. (4.42).

D.2 METHODS FOR SECTION 4.3

D.2.1 *Derivation of the social contagion model*

We detail here the derivation of the social contagion model presented in Sec. 4.3.2.1. Consider a group of size n, composed of n_{A} and n_{N} agents of type A and N, respectively, and denote with $q_{\mathrm{X}}(t)$ the probability of adopting an active behavior for a X-type's at time t. An A-type agent in the adopter state is induced to switch to the nonadopter state by the current nonadopters in the group. These are, on average, $(1-q_{\mathrm{A}}(t))\,(n_{\mathrm{A}}-1) + (1-q_{\mathrm{N}}(t))\,n_{\mathrm{N}}$. If, instead, the A-type's is in the nonadopter state, it is induced to switch to the adopter one by the $q_{\mathrm{A}}(t)\,(n_{\mathrm{A}}-1) + q_{\mathrm{N}}(t)n_{\mathrm{N}}$ adopters that, on average, it currently finds in the group. The same holds for a N-type agent given $n_{\mathrm{A}}-1$ and n_{N} in the expressions above are replaced by n_{A} and $n_{\mathrm{N}}-1$, respectively. Additionally, agents can spontaneously move between the two behavioral

compartments at rates depending solely on their type. These rates quantify some cost of adopting an active behavior and the will to bear that cost. Denoting with c_X the I-to-S rate and with b_X the S-to-I rate, for type X, the adaptive behavioral dynamics is described by the following system of two differential equations

$$\begin{aligned}\dot{q}_A = &(1-q_A)\left[q_A\frac{n_A-1}{n-1}+q_N\frac{n_N}{n-1}\right]\\ &-q_A\left[(1-q_A)\frac{n_A-1}{n-1}+(1-q_N)\frac{n_N}{n-1}\right]\\ &+b_A(1-q_A)-c_Aq_A\,,\end{aligned} \tag{D.19}$$

$$\begin{aligned}\dot{q}_N = &(1-q_N)\left[q_N\frac{n_N-1}{n-1}+q_A\frac{n_A}{n-1}\right]\\ &-q_N\left[(1-q_N)\frac{n_N-1}{n-1}+(1-q_A)\frac{n_A}{n-1}\right]\\ &+b_N(1-q_N)-c_Nq_N\,,\end{aligned} \tag{D.20}$$

where $q_X \equiv q_X^{(n_X-1,n-1)}(t)$ for a lighter notation. The $n-1$ in the denominator comes from assuming that each of the $n-1$ sources weights the same in affecting the behavior of a focal agent. It is immediate to see that the terms involving only one type ($\propto (1-q_X)\,q_X$), as well as the mixed quadratic terms ($\propto q_Aq_N$), cancel out. Therefore, we are left with the linear system defined by Eqs. (4.56) and (4.57), and the solutions described after it.

In the solution for mixed groups (i.e., $1 \leqslant n_A \leqslant n-1$), the constants C_1 and C_2, which depend on all the figuring parameters, read

$$C_1 = \frac{(n-1)\,b_Ab_N+\sum_{X=A,N}n_Xb_X}{(n-1)\,(b_A+c_A)\,(b_N+c_N)+\sum_{X=A,N}n_X\,(b_X+c_X)}\,, \tag{D.21}$$

$$C_2 = \frac{n-1}{(n-1)\,(b_A+c_A)\,(b_N+c_N)+\sum_{X=A,N}n_X\,(b_X+c_X)}\,. \tag{D.22}$$

The difference $q_A-q_N = C_2\,(b_Ac_N-b_Nc_A)$ requires $b_Ac_N > b_Nc_A$ in order for A-type agents to be actually more inclined to be adopters than N-type's are. We can thus choose $b_A > c_A$ (or $b_A \gg c_A$) and $b_N < c_N$ (or $b_N \ll c_N$). Also, it requires C_2 to be large enough and, from Eq. (D.22), we see this means considering not too large values for the rates (specifically for b_A and c_N, when $b_A \gg c_A$ and $b_N \ll c_N$).

D.2.2 *Dependence of the basic reproduction number on the structural parameters*

D.2.2.1 Degree

The basic reproduction number, $R = \beta k_{\text{eff}}/\mu$, is here studied as a function of the 2-degree, $k^{(2)}$, while keeping fixed the total pairwise degree, $k = k^{(2)} + 2k^{(3)}$. We assume type and degree to be uncorrelated, i.e., $k_{\text{A}}^{(n)} = k_{\text{N}}^{(n)} = k^{(n)}$. For any fixed k, we want to establish the sign of $\partial R/\partial k^{(2)}|_k = (\beta/\mu)\, \partial k_{\text{eff}}/\partial k^{(2)}|_k$. To this end, given the complicated expression of k_{eff} (Eq. (4.65)), we simplify it by referring to the binary behavioral scenarios presented in Sec. 4.3.2.2. In both of them, it holds $r_{\text{A}} = 1 - \varepsilon_{\text{out}}$ and $s_{\text{A}} = 1 - \varepsilon_{\text{in}}$. Moreover, we assume perfect protection in at least one way, that is, $\varepsilon_{\text{in}} = 1$ and/or $\varepsilon_{\text{out}} = 1$.

EASY ADAPTATION. Here, $r_{\text{N}} = 1 - \varepsilon_{\text{out}}$ and $s_{\text{N}} = 1 - \varepsilon_{\text{in}}$ for $1 \leqslant n_{\text{N}} \leqslant n-1$, with $\varepsilon_{\text{in}} = 1$ and/or $\varepsilon_{\text{out}} = 1$, and $r_{\text{N}} = s_{\text{N}} = 1$ for $n_{\text{N}} = n$. Only $\theta_{\text{N}\to\text{N}}$ is nonzero, hence $k_{\text{eff}} = \theta_{\text{N}\to\text{N}}$ and is given by

$$k_{\text{eff}} = k^{(2)}\,(1 - \alpha_2\rho) + \left(k - k^{(2)}\right)\left[1 - \alpha_3\rho\,(2 - \beta_{3,1}\rho)\right] \,, \tag{D.23}$$

where the two terms multiplying the degrees account for the probability that a N-type individual takes part to a type-N uniform group of size 2 and 3, respectively, for these are the groups where it is not an adopter. Then we obtain

$$\left.\frac{\partial k_{\text{eff}}}{\partial k^{(2)}}\right|_k = \rho\left[\alpha_3\,(2 - \beta_{3,1}\rho) - \alpha_2\right] \,, \tag{D.24}$$

and equaling it to zero, we get the solutions $\rho = 0$ and $\rho = \tilde{\rho}$, with

$$\tilde{\rho} = \frac{2\alpha_3 - \alpha_2}{\alpha_3\beta_{3,1}} \,, \tag{D.25}$$

given $\alpha_3\beta_{3,1} \neq 0$. Also, note that $\tilde{\rho} \in [0,1]$ requires $\alpha_3 \in [\alpha_2/2, \alpha_2/\,(2 - \beta_{3,1})]$. Technically, $\tilde{\rho}$ results from the fact that the frequency of the {A,A,N} and {A,N,N} triads are both quadratic (respectively, increasing and decreasing) functions of ρ^2, whereas the frequency of the {A,N} pairs increases linearly with ρ. Finally, differentiating Eq. (D.24) with respect to ρ and computing it at $\rho = 0$ and $\rho = \tilde{\rho}$, one finds

$$\frac{\partial}{\partial\rho}\left.\frac{\partial k_{\text{eff}}}{\partial k^{(2)}}\right|_k = \begin{cases} 2\alpha_3 - \alpha_2 & \text{at } \rho = 0 \,, \\ \alpha_2 - 2\alpha_3 & \text{at } \rho = \tilde{\rho} \,. \end{cases} \tag{D.26}$$

Thus, it follows that $\partial k_{\text{eff}}/\partial k^{(2)}|_k$ is positive for $\rho \in (0, \tilde{\rho})$ and negative for $\rho \in (\tilde{\rho}, 1]$. Rephrasing it, k_{eff} (hence R) is an increasing or decreasing function of $k^{(2)}$ depending on whether ρ is lower or higher than $\tilde{\rho}$.

HARD ADAPTATION. In this scenario, $r_N = 1 - \varepsilon_{out}$ and $s_N = 1 - \varepsilon_{in}$ for $n_N = 1$, with $\varepsilon_{in} = 1$ and/or $\varepsilon_{out} = 1$, and $r_N = s_N = 1$ otherwise. We get

$$k_{eff} = k^{(2)} (1 - \alpha_2 \rho) + \left(k - k^{(2)}\right) (1 - \alpha_3 \rho) \ , \tag{D.27}$$

the two terms – notice, both linear in ρ – accounting for the probability that a N-type individual finds at least another N-type's in a group of size 2 and 3, respectively, not being an adopter in such cases. We obtain

$$\left.\frac{\partial k_{eff}}{\partial k^{(2)}}\right|_k = \rho (\alpha_3 - \alpha_2) \ , \tag{D.28}$$

therefore k_{eff} (R) increases or decreases with $k^{(2)}$ solely depending on whether α_3 is larger or smaller than α_2, respectively.

D.2.2.2 Assortativity

Here we study how the basic reproduction number is affected by the way in which assortativity is distributed between 2- and 3-edges. No correlation is assumed between type and degree. Fixed the average pairwise mixing parameter, $\alpha = \left(\alpha_2 k^{(2)} + 2\alpha_3 k^{(3)}\right)/k$, with $k = k^{(2)} + 2k^{(3)}$, we want to establish the sign of $\partial R/\partial \alpha_2|_\alpha = (\beta/\mu)\, \partial k_{eff}/\partial \alpha_2|_\alpha$.

EASY ADAPTATION. Here, $r_N = 1 - \varepsilon_{out}$ and $s_N = 1 - \varepsilon_{in}$ for $1 \leqslant n_N \leqslant n - 1$, and $r_N = s_N = 1$ for $n_N = n$. After some algebra, one gets to

$$\begin{aligned}
\left.\frac{\partial k_{eff}}{\partial \alpha_2}\right|_\alpha &= \frac{k^{(2)}}{2} \rho (1 - \beta_{3,1}\rho) (\varepsilon_{out} + \varepsilon_{in} - \varepsilon_{out}\varepsilon_{in}) \\
&\quad \times \frac{C + \left[C^2 + [2k\alpha (1 - \varepsilon_{out})(1 - \varepsilon_{in})]^2 \rho (1 - \rho)\right]^{\frac{1}{2}}}{\left[C^2 + [2k\alpha (1 - \varepsilon_{out})(1 - \varepsilon_{in})]^2 \rho (1 - \rho)\right]^{\frac{1}{2}}} \qquad \text{(D.29)} \\
&\geqslant \frac{k^{(2)}}{2} \rho (1 - \beta_{3,1}\rho) (\varepsilon_{out} + \varepsilon_{in} - \varepsilon_{out}\varepsilon_{in}) \\
&\quad \times \frac{C + |C|}{\left[C^2 + [2k\alpha (1 - \varepsilon_{out})(1 - \varepsilon_{in})]^2 \rho (1 - \rho)\right]^{\frac{1}{2}}} \geqslant 0 \ . \qquad \text{(D.30)}
\end{aligned}$$

Whatever the value of C, which is a function of the various parameters, since the first term in Eq. (D.29) is nonnegative (recall the condition $\beta_{3,1} \leqslant 1/\rho$, Eq. (4.31)) and the second term within the square root as well, immediately follow both inequalities. From Eq. (D.29) we see that $\partial k_{eff}/\partial \alpha_2|_\alpha$ is strictly positive when at least one between ε_{out} and ε_{in} is nonzero, i.e., trivially, whenever there is protection. Additionally, we note that the dependence of k_{eff} on α_2 is reduced by making $1 - \beta_{3,1}\rho$ smaller. From any equation among Eqs. (4.18), (4.24) and (4.26), we see that that means lowering the

frequency of $\{A, N, N\}$ triads, confirming them as the responsible for that dependence.

HARD ADAPTATION. In this scenario, $r_N = 1 - \varepsilon_{out}$ and $s_N = 1 - \varepsilon_{in}$ for $n_N = 1$, and $r_N = s_N = 1$ otherwise. With a bit of algebra, one finds

$$\begin{aligned}\left.\frac{\partial k_{eff}}{\partial \alpha_2}\right|_{\alpha} &\propto -k^{(2)}\rho\,(1-\rho)\,(1-\beta_{3,1}\rho)\,(1-\varepsilon_{out})\,(1-\varepsilon_{in}) \\ &\quad\times\left[\varepsilon_{out}+\varepsilon_{in}-2\varepsilon_{out}\varepsilon_{in}\right. \\ &\quad\left.\times\left(\beta_{3,1}\rho+(1-\beta_{3,1}\rho)\,\frac{\alpha_2 k^{(2)}}{\alpha k}\right)\right]\alpha k \qquad \text{(D.31)} \\ &\leqslant -k^{(2)}\rho\,(1-\rho)\,(1-\beta_{3,1}\rho)\,(1-\varepsilon_{out})\,(1-\varepsilon_{in}) \\ &\quad\times(\varepsilon_{out}+\varepsilon_{in}-2\varepsilon_{out}\varepsilon_{in})\,\alpha k \leqslant 0\,, \qquad \text{(D.32)}\end{aligned}$$

where the constant of proportionality in Eq. (D.31) is positive. The first inequality comes from the constraints $\alpha_2 \leqslant \alpha k/k^{(2)}$, ensuring $\alpha_3 \geqslant 0$ (combining Eqs. (4.30) and (4.32)), and $\beta_{3,1} \leqslant 1/\rho$ (Eq. (4.31)), so that the term in square brackets is a decreasing function of α_2 and thus takes its minimum value at $\alpha_2 = \alpha k/k^{(2)}$. Since the remaining terms in Eq. (D.32) are all nonnegative, we finally have $\partial k_{eff}/\partial\alpha_2|_\alpha \leqslant 0$. Combining this result with the previous one, we see that the sign of $\partial k_{eff}/\partial\alpha_2|_\alpha$ is thus dictated by whether the adaptation by the N-type individuals is easy or hard. From Eq. (D.31) we note that the derivative goes to zero for either full out-going ($\varepsilon_{out} = 1$) or in-going ($\varepsilon_{in} = 1$) protection (besides $\varepsilon_{out} = \varepsilon_{in} = 0$). In such cases, secondary infections are exclusively generated by N-types', but in absence of indirect modifications, their state is unaffected by how assortativity is distributed among group sizes. In the end, as before, lowering $1 - \beta_{3,1}\rho$ reduces the dependence of k_{eff} on α_2, corroborating the role of the $\{A, N, N\}$ triads.

D.2.3 *Approximate dynamics for high prophylactic efficacy*

Let us assume that the probability of adoption is low in type-N uniform groups and high for type A in any group. Considering first the limit of high out-going prophylactic efficacy, i.e., $1 - \varepsilon_{out} \ll 1$, we have $\theta_{A\to A}, \theta_{A\to N} \in O(1 - \varepsilon_{out})$, hence $\theta_{A\to A} \ll \theta_{N\to A}$ and $\theta_{A\to N} \ll \theta_{N\to N}$. In such regime, Eqs. (4.58) and (4.59) approximate to

$$\dot{I}_A \approx -\mu I_A + \beta\,(1 - I_A)\,I_N\theta_{N\to A}\,, \qquad \text{(D.33)}$$

$$\dot{I}_N \approx -\mu I_N + \beta\,(1 - I_N)\,I_N\theta_{N\to N}\,. \qquad \text{(D.34)}$$

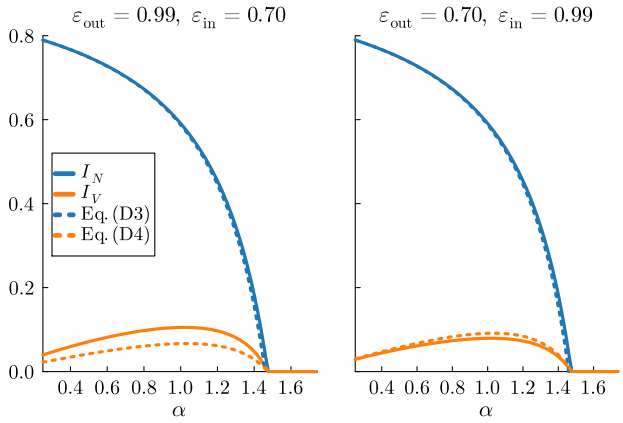


Figure D.2: Exact versus approximated prevalence as given by Eqs. (D.35) and (D.36) (dotted). Here, $\alpha_2 = \alpha_3 = \alpha$, $\beta_{3,1} = 1.0$, $\rho = 0.45$, $k = k^{(2)} + 2k^{(3)} = 6$ and $\beta/\mu = 1$.

Equation (D.34) is a standard SIS dynamics for I_{N}, hence its nonzero fixed point (stable for $\beta\theta_{\text{N}\to\text{N}} \geqslant \mu$) reads

$$I_{\text{N}}{}^{\star} = 1 - \frac{\mu}{\beta\theta_{\text{N}\to\text{N}}} \approx 1 - \frac{1}{R}\,, \tag{D.35}$$

where the approximation comes from $k_{\text{eff}} \approx \theta_{\text{N}\to\text{N}}$ (see Eq. (4.65)). As a consequence, $I_{\text{N}}{}^{\star}$ and R share the same dependence on the various parameters. The nonzero fixed point for I_{A} is then

$$I_{\text{A}}{}^{\star} = \frac{1}{1 + \frac{\mu}{\beta\theta_{\text{N}\to\text{A}} I_{\text{N}}{}^{\star}}} \approx \frac{\beta}{\mu}\theta_{\text{N}\to\text{A}}. \tag{D.36}$$

In particular, if also the in-going efficacy is high, then Eq. (D.36) implies $I_{\text{A}}^{\star} \ll I_{\text{N}}^{\star}$. From Eqs. (D.35) and (D.36) is easy to see how $I_{\text{N}}{}^{\star}$ and $I_{\text{A}}{}^{\star}$ can show different shapes. For instance, everything else left unchanged, an overall increased homophily yields a higher $\theta_{\text{N}\to\text{N}}$ but a lower $\theta_{\text{N}\to\text{A}}$, therefore $I_{\text{N}}{}^{\star}$ grows (as R does) but $I_{\text{A}}{}^{\star}$ can either increase or decrease depending on whether the rate at which $\theta_{\text{N}\to\text{A}}$ decreases is lower or higher than that at which $I_{\text{N}}{}^{\star}$ (growing with $\theta_{\text{N}\to\text{N}}$) increases. In other words, the reproduction number may not be informative about the endemic equilibrium for type A – not even qualitatively –, although it generally may for type N.

Considering now the limit of high in-going prophylactic efficacy, i.e., $1 - \varepsilon_{\text{in}} \ll 1$, we get $\theta_{\text{A}\to\text{A}}, \theta_{\text{N}\to\text{A}} \in O(1 - \varepsilon_{\text{in}})$, implying that $I_{\text{A}}^{\star} \in O(1 - \varepsilon_{\text{in}})$ is the only solution. In turn, the latter yields $I_{\text{N}}^{\star}$ as given – again – by Eq. (D.35), the solution of Eq. (D.34).

Figure D.2 gives an example of how the approximation performs for either high out-going or in-going efficacy. It is less accurate in the former case for type A because, not being the in-going efficacy high enough, the A-to-A infection rate is not so small as assumed. In any case, the approximation preserves the qualitative behavior.

D.2.4 *Numerical simulation of the microscopic spreading dynamics*

The microscopic SIS spreading dynamics is simulated as a discrete-time Markov process. Letting $\Delta t \ll 1$ be the time step duration, the local state of the nodes is updated synchronously at the discrete times $t = 0, \Delta t, 2\Delta t, \dots$, up to a maximum time step. At each step, every node interacts within each of the $k_i^{(n)}$ n-edges including it. The contact structure is encoded in the adjacency tensors $\left\{A^{(n)}\right\}$ such that $A^{(n)}_{i_1,\dots,i_n} = 1$ if nodes $i_1, \dots, i_n$ form a n-edge and 0 otherwise. In the following we consider $n = 2, 3$. Denoting with $\sigma_i(t) \in \{0, 1\}$ the binary variable representing the state of node $i \in \{1, \dots, N\}$ at time t ($\sigma_i(t) = 0$ if susceptible, $\sigma_i(t) = 1$ if infected), we need to compute the transition probabilities of infection, $P\left(\sigma_i(t + \Delta t) = 1 | \sigma_i(t) = 0\right)$, and recovery, $P\left(\sigma_i(t + \Delta t) = 0 | \sigma_i(t) = 1\right)$. Finally, let $\mathrm{X}_i \in \{\mathrm{A}, \mathrm{N}\}$ be the type of node i.

If the interaction occurs in a 2-edge, we have four possible values for the transmission probability, depending on the types of the two involved nodes, say i and j. If i is the susceptible node and j the infected one, then the transmission rate is $\beta r^{(1,1)}_{\mathrm{X}_j} s^{(1,1)}_{\mathrm{X}_i}$ if $\mathrm{X}_i = \mathrm{X}_j$ and $\beta r^{(0,1)}_{\mathrm{X}_j} s^{(0,1)}_{\mathrm{X}_i}$ if $\mathrm{X}_i \neq \mathrm{X}_j$. Accordingly, the transmission probability from j to i during the time interval $[t, t + \Delta t]$ is, respectively, $\beta r^{(1,1)}_{\mathrm{X}_j} s^{(1,1)}_{\mathrm{X}_i} \Delta t$ and $\beta r^{(0,1)}_{\mathrm{X}_j} s^{(0,1)}_{\mathrm{X}_i} \Delta t$. The out-going and in-going probabilities, r and s, are given in Eqs. (4.66) and (4.67) and are computed at the equilibrium of the adoption dynamics (i.e., using the fixed point of Eqs. (4.56)-(4.57)). The probability $w_i^{(2)}(t)$ that i does not get infected via any of the 2-edges incident on it, thus reads

$$w_i^{(2)}(t) = \prod_{j=1}^{N} A^{(2)}_{i,j} \left[1 - \sigma_j(t) \beta \tau^{(2)}_{i,j} \Delta t\right] , \tag{D.37}$$

where

$$\tau^{(2)}_{i,j} \equiv \tau^{(2)}\left(\mathrm{X}_i, \mathrm{X}_j\right) = \delta_{\mathrm{X}_i,\mathrm{X}_j} r^{(1,1)}_{\mathrm{X}_j} s^{(1,1)}_{\mathrm{X}_i} + \left(1 - \delta_{\mathrm{X}_i,\mathrm{X}_j}\right) r^{(0,1)}_{\mathrm{X}_j} s^{(0,1)}_{\mathrm{X}_i} , \tag{D.38}$$

being $\delta_{x,y}$ a Kronecker delta.

If the interaction is part of a 3-edge, the transmission probability can take eight different values, depending on the type of all the three nodes, say i, j and l. Focusing on the interaction between i and j, if the former is susceptible and the other one is infected, we have: $\beta r^{(2,2)}_{\mathrm{X}_j} s^{(2,2)}_{\mathrm{X}_i}$ if $\mathrm{X}_i = \mathrm{X}_j = \mathrm{X}_l$; $\beta r^{(1,2)}_{\mathrm{X}_j} s^{(1,2)}_{\mathrm{X}_i}$ if $\mathrm{X}_i = \mathrm{X}_j \neq \mathrm{X}_l$; $\beta r^{(1,2)}_{\mathrm{X}_j} s^{(0,2)}_{\mathrm{X}_i}$ if $\mathrm{X}_i \neq \mathrm{X}_j$ and $\mathrm{X}_j = \mathrm{X}_l$; and $\beta r^{(0,2)}_{\mathrm{X}_j} s^{(1,2)}_{\mathrm{X}_i}$ if $\mathrm{X}_i \neq \mathrm{X}_j$ and $\mathrm{X}_i = \mathrm{X}_l$. Then, the probability $w_i^{(3)}(t)$ that i does not get infected within any of the 3-edges it takes part to, reads

$$w_i^{(3)}(t) = \prod_{j=1}^{N} \prod_{l>j}^{N} A^{(3)}_{i,j,l} \left[1 - \sigma_j(t) \beta \tau^{(3)}_{i,j,l} \Delta t\right] \left[1 - \sigma_l(t) \beta \tau^{(3)}_{i,l,j} \Delta t\right] , \tag{D.39}$$

where

$$\begin{aligned}\tau^{(3)}_{i,j;l} &\equiv \tau^{(3)}\left(X_i, X_j; X_l\right) \\ &= \delta_{X_i,X_j}\left[\delta_{X_i,X_l} r^{(2,2)}_{X_j} s^{(2,2)}_{X_i} + \left(1-\delta_{X_i,X_l}\right) r^{(1,2)}_{X_j} s^{(1,2)}_{X_i}\right] \\ &\quad + \left(1-\delta_{X_i,X_j}\right)\left[\delta_{X_i,X_l} r^{(0,2)}_{X_j} s^{(1,2)}_{X_i} + \left(1-\delta_{X_i,X_l}\right) r^{(1,2)}_{X_j} s^{(0,2)}_{X_i}\right] . \end{aligned} \tag{D.40}$$

All in all, the probability that a susceptible node i gets infected via at least one of its contacts is

$$P\left(\sigma_i(t+\Delta t)=1|\sigma_i(t)=0\right) = 1 - w^{(2)}_i(t) w^{(3)}_i(t) , \tag{D.41}$$

while its probability of recovery if infected is simply

$$P\left(\sigma_i(t+\Delta t)=0|\sigma_i(t)=1\right) = \mu\Delta t . \tag{D.42}$$

The stochastic process is implemented by drawing a value for $u_i \sim \text{Uniform}(0,1)$ independently for each node i, and at each time step. If node i is susceptible, it gets infected if $P\left(\sigma_i(t+\Delta t)=1|\sigma_i(t)=0\right) > u_i$; if infected, it recovers if $P\left(\sigma_i(t+\Delta t)=0|\sigma_i(t)=1\right) > u_i$.

The continuous dynamics corresponds to the limit $\Delta t \to 0$, as only the terms linear in Δt survive. In the simulation, we took $\Delta t = 0.05$. Using the quasistationary state (QS) method (with 50 stored active states and a 25% probability of update)[285], we let the system run for a transient of 1000 time steps and then averaged over the last 500 ones.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

E

SUPPLEMENTARY INFORMATION FOR SEC. 4.3

E.1 ENDEMIC STATE FOR THE TWO BINARY MODELS

In this section we show the results for the endemic equilibrium in the two limits of easy and hard adaptation, useful to properly interpret the results obtained for intermediate scenarios like the ones entailed by the social contagion model of Sec. 4.3.2.1. We either vary (i) the 2- and 3-degree for a fixed overall degree $k = k^{(2)} + 2k^{(3)}$, for given values of the mixing parameters (α_2, α_3 and $\beta_{3,1}$); (ii) the pairwise assortativity distribution between pairs (α_2) and triads (α_3) for a fixed average pairwise assortativity $\alpha = (\alpha_2 k^{(2)} + 2\alpha_3 k^{(3)})/(k^{(2)} + 2k^{(3)})$, assumed $k_A^{(n)} = k_N^{(n)} = k^{(n)}$, $n = 2, 3$. In both cases we use $\varepsilon_{out} = 0.9$ and $\varepsilon_{in} = 0.5$. Results are shown in Figs. E.1-E.4. The results found in the hard adaptation limit when varying the degree distribution among group sizes are analysed below, while all the others are already discussed in Sec. 4.3.4.

E.1.1 *Varying the degree distribution among group sizes: hard adaptation*

Here we provide a detailed analysis of the results obtained in the limit of hard adaptation when varying the distribution of the degree between the two group sizes (see Fig. E.3). In this behavioral scenario, I_A is the main driver of the observed phenomenology. For $\alpha_2 = \alpha_3$ (recall $\beta_{3,1} = 1$), I_A decreases with $k^{(2)}$, for the A-N interactions in which the N-type individual is also an adopter becomes more frequent (panel (e)). Since the N-type individuals self-protect more often, also I_N decreases, but only slightly (panel (f)), as the epidemic pressure from the type A is small (unless efficacy is low in both ways). For $\alpha_2 < \alpha_3$, the effect on I_N produced by the more frequent N-N interactions, as implied by a raised $k^{(2)}$, overcomes the little benefit coming from the protection in the fewer A-N interactions, hence I_N increases (similarly to R; panel (c)). This growth is faster for high values of ρ, for the $\{A, N, N\}$ triads become scarce and only few protected A-N interactions are gained interacting within pairs. Moreover, for high enough ρ, the growing pressure coming from type N is able to compensate the benefit for type A entailed by a lower mixing with type N. When the efficacy is very high (and hence mixing is important), I_A can even show a non-monotonous shape,

with a maximum at small values of $k^{(2)}$; otherwise it decreases monotonically with the latter (panel (b)). On the other hand, for low ρ (meaning A-type's mainly find themselves in the highly infectious $\{A, N, N\}$ configuration when interacting in triads) and poor protection in both ways, the relative decrease of I_A is sufficiently rapid to even induce a slight decrease of I_N.

Considering $\alpha_2 > \alpha_3$, I_N (like R) decreases with $k^{(2)}$ (panel (i)) as a consequence of the N-N interactions becoming sparser and, secondarily, of the A-N interactions being more often bilaterally protected. The decrease is faster when ρ is high as the many A-type individuals can better decrease the epidemic pressure among the N-type's. The increased rate at which A-N interactions are fully protected generally makes I_A decrease as well (panel (h)). An exception to this exists for low values of ρ when the in-going and out-going efficacy are, respectively, high and low. In such situation, on one hand, since the efficacy is high in at least one way (in-going), the two types have sensibly different epidemic pressures, therefore the mixing matters and for type A is convenient to reduce the contacts with type N; on the other, the transmission coming from type N is only poorly stopped (ε_{out} low), hence the more frequent adoption from type N, as entailed by a larger $k^{(2)}$, benefits type A only marginally. Altogether, type A finds it convenient to mix less with type N, as implied by a smaller $k^{(2)}$, but only if ρ is not too high, otherwise the rapid decrease of I_N with $k^{(2)}$ drives I_A to decrease as well.

E.2 VARYING THE TYPE-ASSORTATIVITY DISTRIBUTION AMONG GROUP SIZES: INTERMEDIATE SCENARIO (SOCIAL CONTAGION)

We here focus on the effect of distributing the assortativity differently between 2- and 3-edges considering the intermediate scenario derived from the social contagion model of Sec. 4.3.2.1.

BASIC REPRODUCTION NUMBER. As Fig. E.5(a) shows, the basic reproduction number, R, increases (decreases) with α_2 (α_3), whatever is the proportion ρ of A-type individuals in the population. In other words, for a given overall assortativity, the minimal (maximal) (dis)assortativity within triads is the optimal mixing scheme to minimize the early growth rate of an outbreak. This is in accordance with the fact that in easy adaptation, which is a proxy for the considered scenario, R decreases monotonically with α_2 under unilateral perfect protection, as proved in Appendix D.2.2.2. In fact, in the considered scenario, q_N jumps from around 0.05 to 0.80 when going from a $\{N, N, N\}$ triad to a $\{A, N, N\}$ one, thus a single A-type suffices to notably reduce the transmission probability of all the direct interactions in the triad, including the N-N one. Minimizing the assortativity within triads implies maximizing the frequency of the $\{A, N, N\}$ triads (see any equation among Eqs. (4.17), (4.24) and (4.26)), and therefore the rate at which N-N direct interactions are indirectly weakened (the

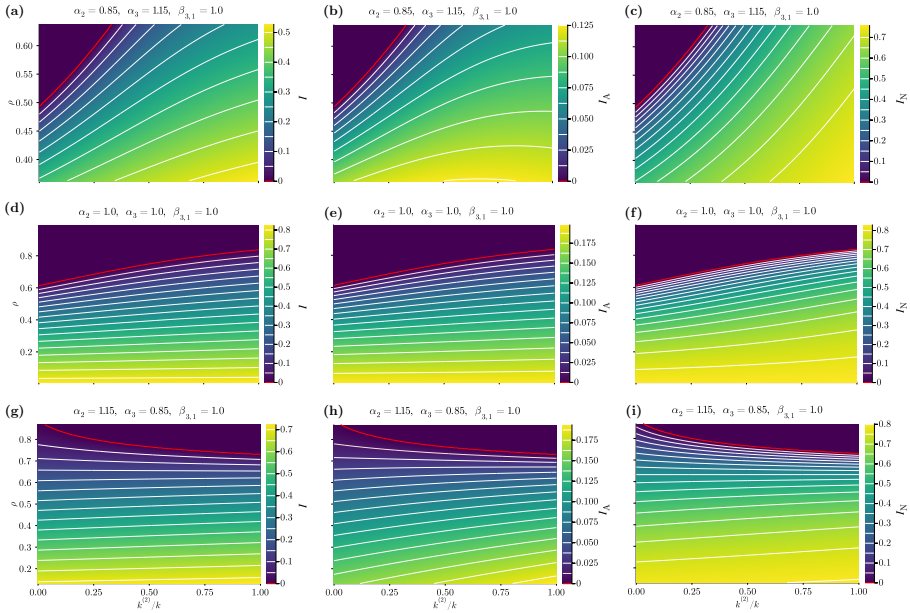


Figure E.1: **Easy adaptation.** Equilibrium endemic state as a function of $k^{(2)}/k$ and ρ for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and **(a–c)** $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, **(d–f)** $\alpha_2 = \alpha_3 = 1.0$, **(g–i)** $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. **(a, d, g)** Overall prevalence, I, **(b, e, h)** prevalence for type A, I_A, and **(c, f, i)** prevalence for type N, I_N. White curves report the levels indicated in the respective colorbar. The red one indicates the critical curve $R = 1$, above which the epidemic-free state is stable.

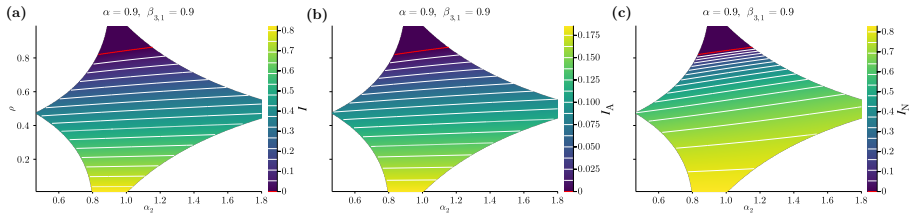


Figure E.2: **Easy adaptation.** Equilibrium endemic state as a function of α_2 and ρ for $\alpha = 0.9$, $\beta_{3,1} = 0.9$, $k^{(2)} = 2k^{(3)}$ and $R_0 = k = 6$. **(a)** Overall prevalence, I, **(b)** prevalence for type A, I_A, and **(c)** prevalence for type N, I_N. White curves report the levels indicated in the respective colorbar. The red one indicates the critical curve $R = 1$, above which the epidemic-free state is stable.

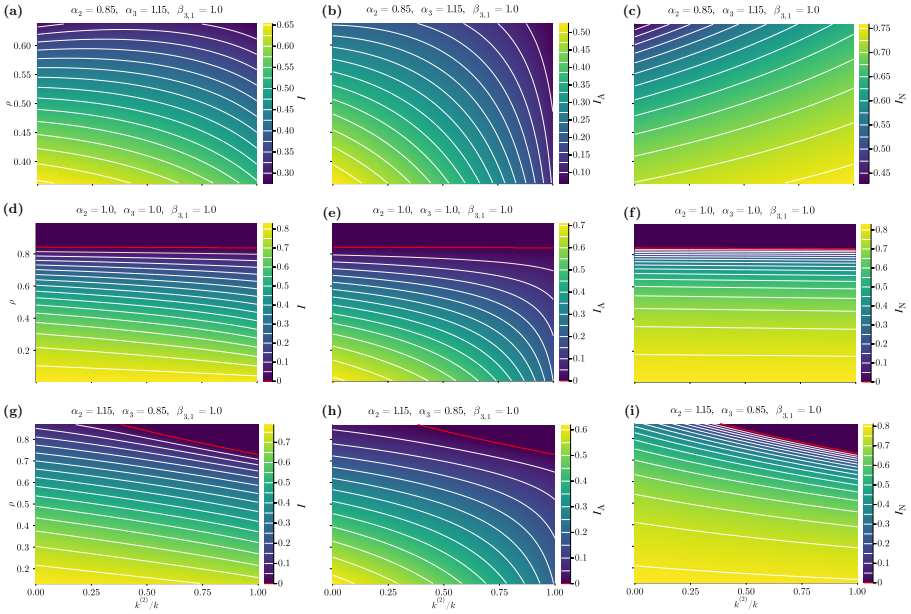


Figure E.3: **Hard adaptation.** Equilibrium endemic state as a function of $k^{(2)}/k$ and ρ for $\beta_{3,1} = 1.0$, $R_0 = k = 6$, and **(a–c)** $\alpha_2 = 0.85$ and $\alpha_3 = 1.15$, **(d–f)** $\alpha_2 = \alpha_3 = 1.0$, **(g–i)** $\alpha_2 = 1.15$ and $\alpha_3 = 0.85$. **(a, d, g)** Overall prevalence, I, **(b, e, h)** prevalence for type A, I_A, and **(c, f, i)** prevalence for type N, I_N. White curves report the levels indicated in the respective colorbar. The red one indicates the critical curve $R = 1$, above which the epidemic-free state is stable.

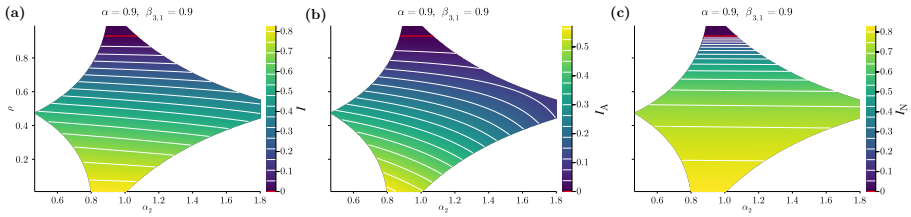


Figure E.4: **Hard adaptation.** Equilibrium endemic state as a function of α_2 and ρ for $\alpha = 0.9$, $\beta_{3,1} = 0.9$, $k^{(2)} = 2k^{(3)}$ and $R_0 = k = 6$. **(a)** Overall prevalence, I, **(b)** prevalence for type A, I_A, and **(c)** prevalence for type N, I_N. White curves report the levels indicated in the respective colorbar. The red one indicates the critical curve $R = 1$, above which the epidemic-free state is stable.

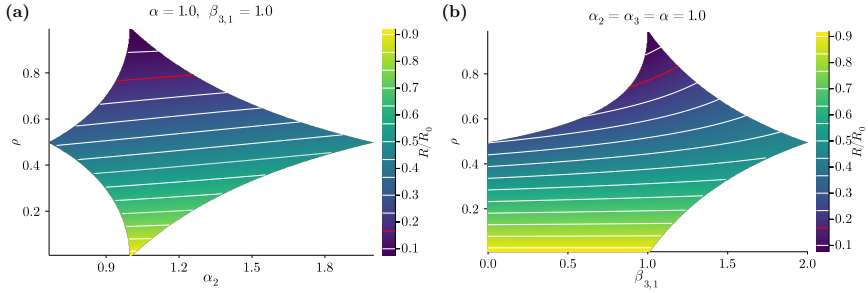


Figure E.5: **Intermediate scenario.** Normalized reproduction number, $R/R_0 = k_{\text{eff}}/k$, as a function of ρ and of **(a)** α_2 ($\alpha = 1.0$, $\beta_{3,1} = 1.0$), and **(b)** $\beta_{3,1}$ ($\alpha_2 = \alpha_3 = \alpha = 1.0$). We fixed $k^{(2)} = 2k^{(3)}$ and $R_0 = k = 6$. White curves report the levels indicated in the colorbar. The red one indicates the curve $R = 1$ ($R/R_0 = 1/6$), above which the epidemic-free state is stable.

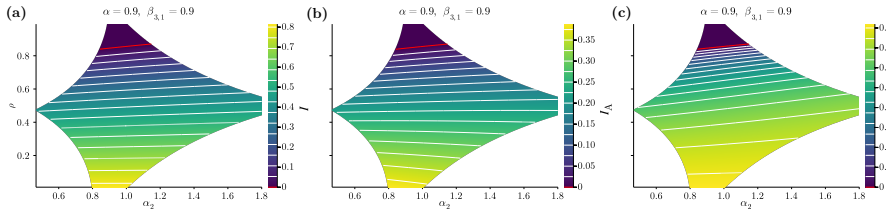


Figure E.6: **Intermediate scenario.** Equilibrium endemic state as a function of α_2 and ρ for $\alpha = 0.9$, $\beta_{3,1} = 0.9$, $k^{(2)} = 2k^{(3)}$ and $R_0 = k = 6$. **(a)** Overall prevalence, I, **(b)** prevalence for type A, I_{A}, and **(c)** prevalence for type N, I_{N}. White curves report the levels indicated in the respective colorbar. The red one indicates the curve $R = 1$, above which the epidemic-free state is stable.

A-N ones are anyway weak due to the A-type). The frequency of the $\{\text{A},\text{N},\text{N}\}$ triads also grows raising $\beta_{3,1}$, leading again R to increase with α_2 (see Fig. E.5(b)).

PREVALENCE. The results for the equilibrium prevalence are reported in Fig. E.6. We assume a moderate overall assortativity by taking $\alpha = 0.9$, and we fix $\beta_{3,1} = 0.9$. The phenomenology is qualitatively unaltered by choosing other values for these parameters, whose effect is mainly changing the ranges within which α_2 and α_3 can vary (in particular, $\alpha \neq 1$ shifts the ranges, while $\beta_{3,1} \neq 1$ makes them asymmetric with respect to the exchange $\rho \leftrightarrow 1 - \rho$). The results are explained by the same logic discussed in Sec. 4.3.4. For any intermediate scenario, we can generally say that lowering the assortativity within triads makes the endemic prevalence decrease (or, at least, not increase) for type N, and increase or decrease for type A depending on whether the share of A-type individuals, ρ, stays below or above a certain value: the easy adaptation "component" dominates for high ρ, the hard adaptation one for low ρ. Lastly, as already seen for the basic reproduction number, increasing $\beta_{3,1}$ has the same qualitative effect of increasing α_2.

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

BIBLIOGRAPHY

1. Bocchi, G. & Ceruti, M. *La sfida della complessità* (Mondadori Bruno, 2007).
2. Anderson, P. W. More Is Different: Broken symmetry and the Nature of the hierarchical structure of science. *Science* **177,** 393–396 (1972).
3. Izyumov, Y. A. & Syromyatnikov, V. N. *Phase transitions and crystal symmetry* (Springer Science & Business Media, 2012).
4. Alberts, B. *Molecular biology of the cell* (Garland science, 2017).
5. Peskin, C. & of Mathematical Sciences, C. I. *Mathematical Aspects of Heart Physiology* (Courant Institute of Mathematical Sciences, New York University, 1975).
6. Néda, Z., Ravasz, E., Vicsek, T., Brechet, Y. & Barabási, A.-L. Physics of the rhythmic applause. *Physical Review E* **61,** 6987 (2000).
7. Wilczek, F. Quantum time crystals. *Physical Review Letters* **109,** 160401 (2012).
8. Shapere, A. & Wilczek, F. Classical time crystals. *Physical Review Letters* **109,** 160402 (2012).
9. Zhang, J., Hess, P. W., Kyprianidis, A, Becker, P., Lee, A, Smith, J, Pagano, G., Potirniche, I.-D., Potter, A. C., Vishwanath, A. *et al.* Observation of a discrete time crystal. *Nature* **543,** 217–220 (2017).
10. Choi, S., Choi, J., Landig, R., Kucsko, G., Zhou, H., Isoya, J., Jelezko, F., Onoda, S., Sumiya, H., Khemani, V. *et al.* Observation of discrete time-crystalline order in a disordered dipolar many-body system. *Nature* **543,** 221–225 (2017).
11. Simon, H. A. The architecture of complexity. *Proceedings of the American Philosophical Society* **106,** 467–482 (1962).
12. Wimsatt, W. C. The Ontology of Complex Systems: Levels of Organization, Perspectives, and Causal Thickets1. *Canadian Journal of Philosophy Supplementary Volume* **20,** 207–274 (1994).
13. Koffka, K. *Principles of Gestalt Psychology* 176 (Harcourt, Brace, 1935).
14. Laughlin, R. B. & Pines, D. The theory of everything. *Proceedings of the National Academy of Sciences* **97,** 28–31 (2000).
15. Geroch, R. & Hartle, J. B. Computability and physical theories. *Foundations of Physics* **16,** 533–550 (1986).
16. Moore, C. Unpredictability and undecidability in dynamical systems. *Physical Review Letters* **64,** 2354 (1990).
17. Da Costa, N. C. & Doria, F. A. Undecidability and incompleteness in classical mechanics. *International Journal of Theoretical Physics* **30,** 1041–1073 (1991).
18. Cubitt, T. S., Perez-Garcia, D. & Wolf, M. M. Undecidability of the spectral gap. *Nature* **528,** 207–211 (2015).
19. Bausch, J., Cubitt, T. S., Lucia, A. & Perez-Garcia, D. Undecidability of the spectral gap in one dimension. *Physical Review X* **10,** 031038 (2020).
20. Cubitt, T. S. A Note on the Second Spectral Gap Incompleteness Theorem. *arXiv preprint arXiv:2105.09854* (2021).
21. Turing, A. M. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* **s2-42,** 230–265 (1937).
22. Gödel, K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik* **38,** 173–198 (1931).

23. Fortunato, S., Bergstrom, C. T., Börner, K., Evans, J. A., Helbing, D., Milojević, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B. *et al.* Science of science. *Science* **359,** eaao0185 (2018).

24. Huang, K. *Statistical mechanics* (John Wiley & Sons, 2008).

25. Weaver, W. Science and complexity. *American Scientist* **36,** 536–544 (1948).

26. Wasserman, S. & Faust, K. in *Social Network Analysis: Methods and Applications* 3–27 (Cambridge University Press, 1994).

27. Strogatz, S. H. Exploring complex networks. *Nature* **410,** 268–276 (2001).

28. Newman, M. *Networks* (Oxford University Press, 2018).

29. Estrada, E. *The structure of complex networks: theory and applications* (Oxford University Press, USA, 2012).

30. Barabási, A.-L. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371,** 20120375 (2013).

31. Latora, V., Nicosia, V. & Russo, G. *Complex networks: principles, methods and applications* (Cambridge University Press, 2017).

32. Moreno, J. L. *Who shall survive? A new approach to the problem of human interrelations* (Nervous and Mental Disease Publishing, 1934).

33. Moreno, J. L. & Jennings, H. H. Statistics of Social Configurations. *Sociometry* **1,** 342–374 (1938).

34. Costa, L. d. F., Oliveira Jr, O. N., Travieso, G., Rodrigues, F. A., Villas Boas, P. R., Antiqueira, L., Viana, M. P. & Correa Rocha, L. E. Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics* **60,** 329–412 (2011).

35. Börner, K., Sanyal, S., Vespignani, A. *et al.* Network science. *Annual Review of Information Science and Technology* **41,** 537–607 (2007).

36. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y. & Porter, M. A. Multilayer networks. *Journal of Complex Networks* **2,** 203–271 (2014).

37. Stauffer, D. & Aharony, A. *Introduction to percolation theory* 2nd ed. (Taylor & Francis, 1992).

38. Marro, J. & Dickman, R. *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, 1999).

39. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Physics Reports* **424,** 175–308 (2006).

40. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical processes on complex networks* (Cambridge University Press, 2008).

41. Vespignani, A. Modelling dynamical processes in complex socio-technical systems. *Nature Physics* **8,** 32–39 (2012).

42. Porter, M. A. & Gleeson, J. P. Dynamical systems on networks. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials* **4** (2016).

43. Gross, T. & Sayama, H. *Adaptive networks* (Springer, 2009).

44. Sayed, A. H. Adaptive networks. *Proceedings of the IEEE* **102,** 460–497 (2014).

45. Holme, P. & Saramäki, J. Temporal networks. *Physics Reports* **519,** 97–125 (2012).

46. Nitschke, J. R. Molecular networks come of age. *Nature* **462,** 736–738 (2009).

47. Ashkenasy, G., Hermans, T. M., Otto, S. & Taylor, A. F. Systems chemistry. *Chemical Society Reviews* **46,** 2543–2554 (2017).

48. Alm, E. & Arkin, A. P. Biological networks. *Current Opinion in Structural Biology* **13,** 193–202 (2003).

49. Barabasi, A.-L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics* **5,** 101–113 (2004).

50. Barabási, A.-L., Gulbahce, N. & Loscalzo, J. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics* **12,** 56–68 (2011).

51. Chan, S. Y. & Loscalzo, J. The emerging paradigm of network medicine in the study of human disease. *Circulation Research* **111,** 359–374 (2012).

52. Ings, T. C., Montoya, J. M., Bascompte, J., Blüthgen, N., Brown, L., Dormann, C. F., Edwards, F., Figueroa, D., Jacob, U., Jones, J. I. *et al.* Ecological networks–beyond food webs. *Journal of Animal Ecology* **78,** 253–269 (2009).

53. Guimaraes Jr, P. R. The structure of ecological networks across levels of organization. *Annual Review of Ecology, Evolution, and Systematics* **51,** 433–460 (2020).

54. Solé, R. V., Corominas-Murtra, B., Valverde, S. & Steels, L. Language networks: Their structure, function, and evolution. *Complexity* **15,** 20–26 (2010).

55. Goyal, S. *Connections: an introduction to the economics of networks* (Princeton University Press, 2012).

56. Bramoullé, Y., Galeotti, A. & Rogers, B. W. *The Oxford Handbook of the Economics of Networks* (Oxford University Press, 2016).

57. Gale, D. M. & Kariv, S. Financial networks. *American Economic Review* **97,** 99–103 (2007).

58. Bardoscia, M., Barucca, P., Battiston, S., Caccioli, F., Cimini, G., Garlaschelli, D., Saracco, F., Squartini, T. & Caldarelli, G. The physics of financial networks. *Nature Reviews Physics* **3,** 490–507 (2021).

59. Lazer, D. Networks in political science: Back to the future. *PS: Political Science & Politics* **44,** 61–68 (2011).

60. Macy, M. W. & Willer, R. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* **28,** 143–166 (2002).

61. Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network analysis in the social sciences. *Science* **323,** 892–895 (2009).

62. Knoke, D. & Yang, S. *Social network analysis* (SAGE publications, 2019).

63. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. *Reviews of Modern Physics* **81,** 591 (2009).

64. Hinrichsen, H. Non-equilibrium critical phenomena and phase transitions into absorbing states. *Advances in Physics* **49,** 815–958 (2000).

65. Broadbent, S. R. & Hammersley, J. M. *Percolation processes: I. Crystals and mazes* in *Mathematical proceedings of the Cambridge philosophical society* **53** (1957), 629–641.

66. Frisch, H. & Hammersley, J. Percolation processes and related topics. *Journal of the Society for Industrial and Applied Mathematics* **11,** 894–918 (1963).

67. Harris, T. E. Contact interactions on a lattice. *The Annals of Probability* **2,** 969–988 (1974).

68. Mollison, D. Spatial contact models for ecological and epidemic spread. *Journal of the Royal Statistical Society: Series B (Methodological)* **39,** 283–313 (1977).

69. Grassberger, P. On the critical behavior of the general epidemic process and dynamical percolation. *Mathematical Biosciences* **63,** 157–172 (1983).

70. Cardy, J. L. & Grassberger, P. Epidemic models and percolation. *Journal of Physics A: Mathematical and General* **18,** L267 (1985).

71. Janssen, H. K. Renormalized field theory of dynamical percolation. *Zeitschrift für Physik B Condensed Matter* **58,** 311–317 (1985).

72. Callaway, D. S., Newman, M. E., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: Percolation on random graphs. *Physical Review Letters* **85,** 5468 (2000).

73. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Resilience of the internet to random breakdowns. *Physical Review Letters* **85,** 4626 (2000).

74. Moore, C. Epidemics and percolation in small-world networks. *Physical Review E* **61,** 5678 (2000).

75. Newman, M. E. Spread of epidemic disease on networks. *Physical Review E* **66,** 016128 (2002).

76. Newman, M. E. Properties of highly clustered networks. *Physical Review E* **68,** 026121 (2003).

77. Kenah, E. & Robins, J. M. Second look at the spread of epidemics on networks. *Physical Review E* **76,** 036113 (2007).

78. Miller, J. C. Percolation and epidemics in random clustered networks. *Physical Review E* **80,** 020901 (2009).

79. Critical phenomena in complex networks. *Reviews of Modern Physics* **80,** 1275 (2008).

80. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Reviews of Modern Physics* **87,** 925 (2015).

81. Le Bon, G. *The crowd: A study of the popular mind* (Courier Corporation, 2002).

82. Anderson, R. M. & May, R. M. *Infectious diseases of humans: dynamics and control* (Oxford University Press, 1992).

83. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2008).

84. Kiss, I. Z., Miller, J. C. & Simon, P. L. *Mathematics of epidemics on networks* 31 (Springer, 2017).

85. Wang, W., Tang, M., Stanley, H. E. & Braunstein, L. A. Unification of theoretical approaches for epidemic spreading on complex networks. *Reports on Progress in Physics* **80,** 036603 (2017).

86. Bretto, A. *Hypergraph theory. An introduction.* (Springer, 2013).

87. Lambiotte, R., Rosvall, M. & Scholtes, I. From networks to optimal higher-order models of complex systems. *Nature Physics* **15,** 313–320 (2019).

88. Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G. & Petri, G. Networks beyond pairwise interactions: structure and dynamics. *Physics Reports* (2020).

89. Battiston, F., Amico, E., Barrat, A., Bianconi, G., Ferraz de Arruda, G., Franceschiello, B., Iacopini, I., Kéfi, S., Latora, V., Moreno, Y. *et al.* The physics of higher-order interactions in complex systems. *Nature Physics* **17,** 1093–1098 (2021).

90. Bick, C., Gross, E., Harrington, H. A. & Schaub, M. T. What are higher-order networks? *SIAM Review* **65,** 686–731 (2023).

91. Salathé, M. & Bonhoeffer, S. The effect of opinion clustering on disease outbreaks. *Journal of the Royal Society Interface* **5,** 1505–1508 (2008).

92. Salathé, M. & Jones, J. H. Dynamics and control of diseases in networks with community structure. *PLOS Computational Biology* **6,** e1000736 (2010).

93. Hébert-Dufresne, L., Noël, P.-A., Marceau, V., Allard, A. & Dubé, L. J. Propagation dynamics on networks featuring complex topologies. *Physical Review E* **82,** 036115 (2010).

94. Allard, A., Hébert-Dufresne, L., Noël, P.-A., Marceau, V. & Dubé, L. J. Bond percolation on a class of correlated and clustered random graphs. *Journal of Physics A: Mathematical and Theoretical* **45,** 405005 (2012).

95. Stegehuis, C., Van Der Hofstad, R. & Van Leeuwaarden, J. S. Epidemic spreading on complex networks with community structures. *Scientific Reports* **6,** 29748 (2016).

96. Guilbeault, D., Becker, J. & Centola, D. in *Complex spreading phenomena in social systems* 3–25 (Springer, 2018).

97. Asch, S. E. in *Organizational influence processes* 295–303 (Routledge, 2016).

98. Asch, S. E. *Social psychology* (Englewood Cliffs, NJ:Prentice Hall, 1952).

99. Asch, S. E. Opinions and social pressure. *Scientific American* **193,** 31–35 (1955).

100. Rosenberg, L. Group size, prior experience, and conformity. *The Journal of Abnormal and Social Psychology* **63,** 436 (1961).

101. Milgram, S., Bickman, L. & Berkowitz, L. Note on the drawing power of crowds of different size. *Journal of Personality and Social Psychology* **13,** 79 (1969).

102. Godwin, W. F. & Restle, F. The road to agreement: Subgroup pressures in small group consensus processes. *Journal of Personality and Social Psychology* **30,** 500 (1974).

103. Kanter, R. M. Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology* **82,** 965–990 (1977).

104. Latané, B., Nida, S. *et al.* Social impact theory and group influence: A social engineering perspective. *Psychology of Group Influence,* 3–34 (1980).

105. Latané, B. The psychology of social impact. *American Psychologist* **36,** 343 (1981).

106. Stasser, G. & Davis, J. H. Group decision making and social influence: A social interaction sequence model. *Psychological Review* **88,** 523 (1981).

107. Tanford, S. & Penrod, S. Social influence model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin* **95,** 189 (1984).

108. Wilson, E. B. & Worcester, J. The law of mass action in epidemiology. *Proceedings of the National Academy of Sciences* **31,** 24–34 (1945).

109. Wilson, E. B. & Worcester, J. The law of mass action in epidemiology: II. *Proceedings of the National Academy of Sciences* **31,** 109–116 (1945).

110. Capasso, V. & Serio, G. A generalization of the Kermack-McKendrick deterministic epidemic model. *Mathematical Biosciences* **42,** 43–61 (1978).

111. Liu, W.-m., Hethcote, H. W. & Levin, S. A. Dynamical behavior of epidemiological models with nonlinear incidence rates. *Journal of Mathematical Biology* **25,** 359–380 (1987).

112. Dodds, P. S. & Watts, D. J. Universal Behavior in a Generalized Model of Contagion. *Physical Review Letters* **92,** 218701 (21 2004).

113. St-Onge, G., Sun, H., Allard, A., Hébert-Dufresne, L. & Bianconi, G. Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Physical Review Letters* **127,** 158301 (2021).

114. St-Onge, G., Hébert-Dufresne, L. & Allard, A. Nonlinear bias toward complex contagion in uncertain transmission settings. *Proceedings of the National Academy of Sciences* **121,** e2312202121 (2024).

115. Ajelli, M., Poletti, P., Melegaro, A. & Merler, S. The role of different social contexts in shaping influenza transmission during the 2009 pandemic. *Scientific Reports* **4,** 1–7 (2014).

116. Leclerc, Q. J., Fuller, N. M., Knight, L. E., Funk, S., Knight, G. M., covid 19 Working Group, C. *et al.* What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Research* **5** (2020).

117. Thompson, H. A., Mousa, A., Dighe, A., Fu, H., Arnedo-Pena, A., Barrett, P., Bellido-Blasco, J., Bi, Q., Caputi, A., Chaw, L. *et al.* Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) setting-specific transmission rates: a systematic review and meta-analysis. *Clinical Infectious Diseases* **73,** e754–e764 (2021).

118. Allen, J. G. & Ibrahim, A. M. Indoor Air Changes and Potential Implications for SARS-CoV-2 Transmission. *JAMA: The Journal of the American Medical Association* **325,** 2112–2113 (2021).

119. Robles-Romero, J. M., Conde-Guillén, G., Safont-Montes, J. C., García-Padilla, F. M. & Romero-Martín, M. Behaviour of aerosols and their role in the transmission of SARS-CoV-2; a scoping review. *Reviews in Medical Virology* **32,** e2297 (2022).

120. Gross, J. L. & Yellen, J. *Handbook of graph theory* (CRC press, 2003).

121. Munkres, J. R. *Elements of algebraic topology* (CRC press, 2018).

122. Carlsson, G. Topology and data. *Bulletin of the American Mathematical Society* **46,** 255–308 (2009).

123. Lum, P. Y., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., Carlsson, J. & Carlsson, G. Extracting insights from the shape of complex data using topology. *Scientific Reports* **3,** 1236 (2013).

124. Giusti, C., Pastalkova, E., Curto, C. & Itskov, V. Clique topology reveals intrinsic geometric structure in neural correlations. *Proceedings of the National Academy of Sciences* **112,** 13455–13460 (2015).

125. Carlsson, G. & De Silva, V. Zigzag persistence. *Foundations of Computational Mathematics* **10,** 367–405 (2010).

126. Reininghaus, J., Huber, S., Bauer, U. & Kwitt, R. *A stable multi-scale kernel for topological machine learning* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), 4741–4748.

127. Taylor, D., Klimm, F., Harrington, H. A., Kramár, M., Mischaikow, K., Porter, M. A. & Mucha, P. J. Topological data analysis of contagion maps for examining spreading processes on networks. *Nature Communications* **6,** 7723 (2015).

128. Atkin, R. H. From cohomology in physics to q-connectivity in social science. *International Journal of Man-Machine Studies* **4,** 139–167 (1972).

129. Bollobás, B. *Random Graphs* 2nd ed. (Cambridge University Press, 2001).

130. Erdős, P., Rényi, A. *et al.* On the evolution of random graphs. *Publ. math. inst. hung. acad. sci* **5,** 17–60 (1960).

131. Broido, A. D. & Clauset, A. Scale-free networks are rare. *Nature Communications* **10,** 1017 (2019).

132. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).

133. Price, D. d. S. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science* **27,** 292–306 (1976).

134. Bollobás, B., Riordan, O., Spencer, J. & Tusnády, G. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms* **18,** 279–290 (2001).

135. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Reviews of Modern Physics* **87,** 925–979 (3 2015).

136. Redner, S. How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B* **4,** 131–134 (1998).

137. Faloutsos, M., Faloutsos, P. & Faloutsos, C. On power-law relationships of the internet topology. *ACM SIGCOMM Computer Communication Review* **29,** 251–262 (1999).

138. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabási, A.-L. The large-scale organization of metabolic networks. *Nature* **407,** 651–654 (2000).

139. Leskovec, J., Kleinberg, J. & Faloutsos, C. *Graphs over time: densification laws, shrinking diameters and possible explanations* in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (2005), 177–187.

140. Voitalov, I., Van Der Hoorn, P., Van Der Hofstad, R. & Krioukov, D. Scale-free networks well done. *Physical Review Research* **1,** 033034 (2019).

141. Holme, P. Rare and everywhere: Perspectives on scale-free networks. *Nature Communications* **10,** 1016 (2019).

142. Holland, P. W. & Leinhardt, S. Transitivity in structural models of small groups. *Comparative Group Studies* **2,** 107–124 (1971).

143. Granovetter, M. S. The strength of weak ties. *American Journal of Sociology* **78,** 1360–1380 (1973).

144. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world'networks. *Nature* **393,** 440–442 (1998).

145. Travers, J. & Milgram, S. in *Social networks* 179–197 (Elsevier, 1977).

146. Budrikis, Z. 25 years of small-world network theory. *Nature Reviews Physics* **5,** 440–440 (2023).

147. Dorogovtsev, S. N., Mendes, J. F. & Samukhin, A. N. Size-dependent degree distribution of a scale-free growing network. *Physical Review E* **63,** 062101 (2001).

148. Bender, E. A. & Canfield, E. R. The asymptotic number of labeled graphs with given degree sequences. *Journal of Combinatorial Theory, Series A* **24,** 296–307 (1978).

149. Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics* **1,** 311–316 (1980).

150. Molloy, M. & Reed, B. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, Orobability and Computing* **7,** 295–305 (1998).

151. Gallo, G., Longo, G., Pallottino, S. & Nguyen, S. Directed hypergraphs and applications. *Discrete Applied Mathematics* **42,** 177–201 (1993).

152. Young, J.-G., Petri, G., Vaccarino, F. & Patania, A. Construction of and efficient sampling from the simplicial configuration model. *Physical Review E* **96,** 032312 (2017).

153. Chodrow, P. S. Configuration models of random hypergraphs. *Journal of Complex Networks* **8,** cnaa018 (2020).

154. Chodrow, P. & Mellor, A. Annotated hypergraphs: models and applications. *Applied Network Science* **5,** 1–25 (2020).

155. Tarde, G. *The Laws of Imitation* (Henry Holt and Company, New York, 1903).

156. Djellal, F. & Gallouj, F. The laws of imitation and invention: Gabriel Tarde and the evolutionary economics of innovation. *Available at SSRN 4344822* (2023).

157. Blumer, H. in *Principles of Sociology* (ed Park, R. E.) 219–288 (New York, Barnes & Noble, 1939).

158. Katz, E. & Lazarsfeld, P. F. *Personal influence: The part played by people in the flow of mass communications* (Free Press, 1955).

159. Coleman, J., Katz, E. & Menzel, H. The diffusion of an innovation among physicians. *Sociometry* **20,** 253–270 (1957).

160. Goffman, W. & Newill, V. Generalization of epidemic theory. *Nature* **204,** 225–228 (1964).

161. Daley, D. J. & Kendall, D. G. Stochastic rumours. *IMA Journal of Applied Mathematics* **1,** 42–55 (1965).

162. Katz, E. & Lazarsfeld, P. F. *Personal Influence, The part played by people in the flow of mass communications* (Transaction Publishers, 1966).

163. Grindereng, M. P. Fashion diffusion. *Journal of Home Economics* **59,** 171–174 (1967).

164. Schelling, T. C. *et al.* Dynamic models of segregation. *Journal of Mathematical Sociology* **1,** 143–186 (1971).

165. Granovetter, M. Threshold models of collective behavior. *American Journal of Sociology* **83,** 1420–1443 (1978).

166. Centola, D., Willer, R. & Macy, M. The emperor's dilemma: A computational model of self-enforcing norms. *American Journal of Sociology* **110,** 1009–1040 (2005).

167. Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I. & Castillo-Chávez, C. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications* **364,** 513–536 (2006).

168. López-Pintado, D. Diffusion in complex social networks. *Games and Economic Behavior* **62,** 573–590 (2008).

169. Dornbusch, R., Park, Y. C. & Claessens, S. Contagion: Understanding how it spreads. *The World Bank Research Observer* **15,** 177–197 (2000).

170. Anderson, E. W. Customer satisfaction and word of mouth. *Journal of Service Research* **1,** 5–17 (1998).

171. Goldenberg, J., Libai, B. & Muller, E. Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters* **12,** 211–223 (2001).

172. Leskovec, J., Adamic, L. A. & Huberman, B. A. The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)* **1,** 5–es (2007).

173. Romero, D. M., Meeder, B. & Kleinberg, J. *Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter* in *Proceedings of the 20th international conference on World wide web* (2011), 695–704.

174. Bakshy, E., Rosenn, I., Marlow, C. & Adamic, L. *The role of social networks in information diffusion* in *Proceedings of the 21st International Conference on World Wide Web* (2012), 519–528.

175. Weng, L., Flammini, A., Vespignani, A. & Menczer, F. Competition among memes in a world with limited attention. *Scientific Reports* **2,** 335 (2012).

176. Weng, L., Menczer, F. & Ahn, Y.-Y. Virality prediction and community structure in social networks. *Scientific Reports* **3,** 1–6 (2013).

177. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *American Journal of Sociology* **113,** 702–734 (2007).

178. Mønsted, B., Sapieżyński, P., Ferrara, E. & Lehmann, S. Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PLOS ONE* **12,** e0184148 (2017).

179. Chwe, M. S.-Y. Structure and strategy in collective action. *American Journal of Sociology* **105,** 128–156 (1999).

180. Centola, D. The spread of behavior in an online social network experiment. *Science* **329,** 1194–1197 (2010).

181. Marwell, G. & Oliver, P. *The critical mass in collective action* (Cambridge University Press, 1993).

182. Bernoulli, D. & Blower, S. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Reviews in Medical Virology* **14,** 275 (2004).

183. Brownlee, J. Historical note on Farr's theory of the epidemic. *British Medical Journal* **2,** 250 (1915).

184. Snow, J. On continuous molecular changes, more particularly in their relation to epidemic diseases. 1853. *Reviews of Infectious Diseases,* 441–447 (1985).

185. The Milroy Lectures on epidemic disease in England—The evidence of variability and of persistency of type. *The Lancet* **167.** Originally published as Volume 1, Issue 4305, 569–574 (1906).

186. Ross, R. An application of the theory of probabilities to the study of a priori pathometry.–Part I. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **92,** 204–230 (1916).

187. Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry.–Part II. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **93,** 212–225 (1917).

188. Ross, R. & Hudson, H. P. An application of the theory of probabilities to the study of a priori pathometry.–Part III. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **93,** 225–240 (1917).

189. McKendrick, A. G. An application of the theory of probabilities to the study of a priori pathometry.–Part III. *Paludisme* **1,** 54–66 (1912).

190. Kermack, W. O. & McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **115,** 700–721 (1927).

191. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. II.—The problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **138,** 55–83 (1932).

192. Kermack, W. O. & McKendrick, A. G. Contributions to the mathematical theory of epidemics. III.—Further studies of the problem of endemicity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **141,** 94–122 (1933).

193. Érdi, P. & Tóth, J. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models* (Manchester University Press, 1989).

194. Heesterbeek, H. The law of mass-action in epidemiology: a historical perspective. *Ecological Paradigms Lost: Routes of Theory Change,* 81–104 (2005).

195. Diekmann, O., Heesterbeek, H. & Britton, T. *Mathematical tools for understanding infectious disease dynamics* (Princeton University Press, 2013).

196. Diekmann, O., Heesterbeek, J. A. & Metz, J. A. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* **28,** 365–382 (1990).

197. Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Holt-Saunders, 1976).

198. Castellano, C. & Pastor-Satorras, R. Thresholds for epidemic spreading in networks. *Physical Review Letters* **105,** 218701 (2010).

199. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nature Communications* **10,** 1–9 (2019).

200. Kiss, I. Z., Iacopini, I., Simon, P. L. & Georgiou, N. Insights from exact social contagion dynamics on networks with higher-order structures. *Journal of Complex Networks* **11,** cnad044 (2023).

201. Ethier, S. N. & Kurtz, T. G. *Markov processes: characterization and convergence* (John Wiley & Sons, 2009).

202. Gardiner, C. W. *et al. Handbook of stochastic methods* (Springer Berlin, 1985).

203. Oppenheim, I, Shuler, K. & Weiss, G. Stochastic and deterministic formulation of chemical rate equations. *The Journal of Chemical Physics* **50,** 460–466 (1969).

204. Kurtz, T. G. Solutions of ordinary differential equations as limits of pure jump Markov processes. *Journal of Applied Probability* **7,** 49–58 (1970).

205. Kurtz, T. G. Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. *Journal of Applied Probability* **8,** 344–356 (1971).

206. Kadanoff, L. P. More is the same; phase transitions and mean field theories. *Journal of Statistical Physics* **137,** 777–797 (2009).

207. Hethcote, H. W. & Yorke, J. A. *Gonorrhea transmission dynamics and control* (Springer, 2014).

208. Anderson, R., Medley, G., May, R. & Johnson, A. A preliminary study of the transmission dynamics of the human immunodeficiency virus (HIV), the causative agent of AIDS. *Mathematical Medicine and Biology: a Journal of the IMA* **3,** 229–263 (1986).

209. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Physical Review Letters* **86,** 3200 (2001).

210. Wallinga, J., Teunis, P. & Kretzschmar, M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *American Journal of Epidemiology* **164,** 936–944 (2006).

211. Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J. *et al.* Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine* **5,** e74 (2008).

212. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Computational Biology* **13,** e1005697 (2017).

213. Prem, K., Zandvoort, K. v., Klepac, P., Eggo, R. M., Davies, N. G., for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, C., Cook, A. R. & Jit, M. Projecting contact matrices in 177 geographical regions: an update and comparison with empirical data for the COVID-19 era. *PLOS Computational Biology* **17,** e1009098 (2021).

214. Mistry, D., Litvinova, M., Pastore y Piontti, A., Chinazzi, M., Fumanelli, L., Gomes, M. F., Haque, S. A., Liu, Q.-H., Mu, K., Xiong, X. *et al.* Inferring high-resolution human mixing patterns for disease modeling. *Nature Communications* **12,** 323 (2021).

215. Mézard, M., Parisi, G. & Virasoro, M. A. *Spin glass theory and beyond: An Introduction to the Replica Method and Its Applications* (World Scientific Publishing Company, 1987).

216. Holme, P. & Saramäki, J. *Temporal network theory* (Springer, 2019).

217. Cowen, L., Goddard, W. & Jesurum, C. E. Defective coloring revisited. *Journal of Graph Theory* **24,** 205–219 (1997).

218. Ganesh, A., Massoulié, L. & Towsley, D. *The effect of network topology on the spread of epidemics* in *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.* **2** (2005), 1455–1466.

219. Van Mieghem, P. The N-intertwined SIS epidemic network model. *Computing* **93,** 147–169 (2011).

220. Wang, Y., Chakrabarti, D., Wang, C. & Faloutsos, C. *Epidemic spreading in real networks: An eigenvalue viewpoint* in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.* (2003), 25–34.

221. Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J. & Faloutsos, C. Epidemic thresholds in real networks. *ACM Transactions on Information and System Security (TISSEC)* **10,** 1–26 (2008).

222. Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S. & Moreno, Y. Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *Europhysics Letters* **89,** 38009 (2010).

223. Meyer, C. D. *Matrix analysis and applied linear algebra* (SIAM, 2000).

224. Eames, K. T. D. & Keeling, M. J. Modeling dynamic and network heterogeneities in the spread of Sexually Transmitted Diseases. *Proceedings of the National Academy of Sciences* **99,** 13330–13335 (2002).

225. Sharkey, K. J. Deterministic epidemiological models at the individual level. *Journal of Mathematical Biology* **57,** 311–331 (2008).

226. Cator, E. & Van Mieghem, P. Second-order mean-field susceptible-infected-susceptible epidemic threshold. *Physical Review E* **85,** 056111 (2012).

227. Mata, A. S. & Ferreira, S. C. Pair quenched mean-field theory for the susceptible-infected-susceptible model on complex networks. *Europhysics Letters* **103,** 48003 (2013).

228. Sharkey, K. J., Kiss, I. Z., Wilkinson, R. R. & Simon, P. L. Exact equations for SIR epidemics on tree graphs. *Bulletin of Mathematical Biology* **77,** 614–645 (2015).

229. Matamalas, J. T., Arenas, A. & Gómez, S. Effective approach to epidemic containment using link equations in complex networks. *Science Advances* **4,** eaau4212 (2018).

230. Matamalas, J. T., Gómez, S. & Arenas, A. Abrupt phase transition of epidemic spreading in simplicial complexes. *Physical Review Research* **2,** 012049 (2020).

231. Kiss, I. Z., Morris, C. G., Sélley, F., Simon, P. L. & Wilkinson, R. R. Exact deterministic representation of Markovian SIR epidemics on networks with and without loops. *Journal of Mathematical Biology* **70,** 437–464 (2015).

232. Watts, D. J. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* **99,** 5766–5771 (2002).

233. Miller, J. C. Epidemic size and probability in populations with heterogeneous infectivity and susceptibility. *Physical Review E* **76,** 010101 (2007).

234. Gleeson, J. P. Cascades on correlated and modular random networks. *Physical Review E* **77,** 046117 (2008).

235. Karrer, B. & Newman, M. E. Message passing approach for general epidemic models. *Physical Review E* **82,** 016101 (2010).

236. Wilkinson, R. R. & Sharkey, K. J. Message passing and moment closure for susceptible-infected-recovered epidemics on finite networks. *Physical Review E* **89,** 022808 (2014).

237. Gleeson, J. P. & Porter, M. A. in *Complex Spreading Phenomena in Social Systems: Influence and Contagion in Real-World Social Networks* (eds Lehmann, S. & Ahn, Y.-Y.) 81–95 (Springer International Publishing, Cham, 2018).

238. Allard, A. & Hébert-Dufresne, L. Percolation and the effective structure of complex networks. *Physical Review X* **9,** 011023 (2019).

239. Shrestha, M., Scarpino, S. V. & Moore, C. Message-passing approach for recurrent-state epidemic models on networks. *Physical Review E* **92,** 022821 (2015).

240. Radicchi, F. & Castellano, C. Beyond the locally treelike approximation for percolation on real networks. *Physical Review E* **93,** 030302 (2016).

241. Eames, K., Bansal, S., Frost, S. & Riley, S. Six challenges in measuring contact networks for use in modelling. *Epidemics* **10,** 72–77 (2015).

242. Gleeson, J. P. Binary-state dynamics on complex networks: Pair approximation and beyond. *Physical Review X* **3,** 021004 (2013).

243. Dickman, R. Kinetic phase transitions in a surface-reaction model: mean-field theory. *Physical Review A* **34,** 4246 (1986).

244. Dietz, K. & Hadeler, K. Epidemiological models for Sexually Transmitted Diseases. *Journal of Mathematical Biology* **26,** 1–25 (1988).

245. Matsuda, H., Ogita, N., Sasaki, A. & Satō, K. Statistical mechanics of population: the lattice Lotka-Volterra model. *Progress of Theoretical Physics* **88,** 1035–1049 (1992).

246. Satō, K., Matsuda, H. & Sasaki, A. Pathogen invasion and host extinction in lattice structured populations. *Journal of Mathematical Biology* **32,** 251–268 (1994).

247. Levin, S. A. & Durrett, R. From individuals to epidemics. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **351,** 1615–1621 (1996).

248. Rand, D. Correlation equations and pair approximations for spatial ecologies. *Advanced Ecological Theory: Principles and Applications,* 100–142 (1999).

249. Keeling, M. J. The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society B* **266,** 859–867 (1999).

250. Bauch, C. T. The spread of infectious diseases in spatially structured populations: an invasory pair approximation. *Mathematical Biosciences* **198,** 217–237 (2005).

251. House, T. & Keeling, M. J. The impact of contact tracing in clustered populations. *PLOS Computational Biology* **6,** e1000721 (2010).

252. Bauch, C. A versatile ODE approximation to a network model for the spread of Sexually Transmitted Diseases. *Journal of Mathematical Biology* **45,** 375–395 (2002).

253. House, T., Davies, G., Danon, L. & Keeling, M. J. A motif-based approach to network epidemics. *Bulletin of Mathematical Biology* **71,** 1693–1706 (2009).

254. Marceau, V., Noël, P.-A., Hébert-Dufresne, L., Allard, A. & Dubé, L. J. Adaptive networks: Coevolution of disease and topology. *Physical Review E* **82,** 036116 (2010).

255. Fennell, P. G. & Gleeson, J. P. Multistate dynamical processes on networks: analysis through degree-based approximation frameworks. *SIAM Review* **61,** 92–118 (2019).

256. O'Sullivan, D. J., O'Keeffe, G. J., Fennell, P. G. & Gleeson, J. P. Mathematical modeling of complex contagion on clustered networks. *Frontiers in Physics* **3,** 71 (2015).

257. St-Onge, G., Thibeault, V., Allard, A., Dubé, L. J. & Hébert-Dufresne, L. Master equation analysis of mesoscopic localization in contagion dynamics on higher-order networks. *Physical Review E* **103,** 032301 (2021).

258. Keating, L. A., Gleeson, J. P. & O'Sullivan, D. J. Multitype branching process method for modeling complex contagion on clustered networks. *Physical Review E* **105,** 034306 (2022).

259. Valente, T. W. *Network models of the diffusion of innovations* **303.484 V3** (Hampton Press, 1995).

260. Cowan, R. & Jonard, N. Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control* **28,** 1557–1575 (2004).

261. Watts, D. J. & Dodds, P. S. Influentials, networks, and public opinion formation. *Journal of Consumer Research* **34,** 441–458 (2007).

262. Bianconi, G., Darst, R. K., Iacovacci, J. & Fortunato, S. Triadic closure as a basic generating mechanism of communities in complex networks. *Physical Review E* **90,** 042806 (2014).

263. Palla, G., Derényi, I., Farkas, I. & Vicsek, T. Uncovering the overlapping community structure of complex networks in Nature and society. *Nature* **435,** 814–818 (2005).

264. Porter, M. A., Onnela, J.-P. & Mucha, P. J. Communities in networks. *Notices of the AMS* (2009).

265. Burgio, G., Matamalas, J. T., Gómez, S. & Arenas, A. Evolution of Cooperation in the Presence of Higher-Order Interactions: from Networks to Hypergraphs. *Entropy* **22,** 744 (7 2020).

266. Dai, X, Kovalenko, K, Molodyk, M, Wang, Z, Li, X, Musatov, D, Raigorodskii, A., Alfaro-Bittner, K, Cooper, G., Bianconi, G *et al.* D-dimensional oscillators in simplicial structures: Odd and even dimensions display different synchronization scenarios. *Chaos, Solitons & Fractals* **146,** 110888 (2021).

267. Ghorbanchian, R., Restrepo, J. G., Torres, J. J. & Bianconi, G. Higher-order simplicial synchronization of coupled topological signals. *Communications Physics* **4,** 120 (2021).

268. Tadić, B. & Gupte, N. Hidden geometry and dynamics of complex networks: Spin reversal in nanoassemblies with pairwise and triangle-based interactions (a). *Europhysics Letters* **132,** 60008 (2021).

269. Sun, H., Ziff, R. M. & Bianconi, G. Renormalization group theory of percolation on pseudofractal simplicial and cell complexes. *Physical Review E* **102,** 012308 (1 2020).

270. Andjelković, M., Tadić, B. & Melnik, R. The topology of higher-order complexes associated with brain hubs in human connectomes. *Scientific Reports* **10,** 1–10 (2020).

271. Ugander, J., Backstrom, L., Marlow, C. & Kleinberg, J. Structural diversity in social contagion. *Proceedings of the National Academy of Sciences* **109,** 5962–5966 (2012).

272. Rosas, F. E., Mediano, P. A., Luppi, A. I., Varley, T. F., Lizier, J. T., Stramaglia, S., Jensen, H. J. & Marinazzo, D. Disentangling high-order mechanisms and high-order behaviours in complex systems. *Nature Physics* **18,** 476–477 (2022).

273. Bodó, Á., Katona, G. Y. & Simon, P. L. SIS epidemic propagation on hypergraphs. *Bulletin of Mathematical Biology* **78,** 713–735 (2016).

274. Jhun, B., Jo, M. & Kahng, B. Simplicial SIS model in scale-free uniform hypergraph. *Journal of Statistical Mechanics: Theory and Experiment* **2019,** 123207 (2019).

275. Landry, N. W. & Restrepo, J. G. The effect of heterogeneity on hypergraph contagion models. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30,** 103117 (2020).

276. De Arruda, G. F., Petri, G. & Moreno, Y. Social contagion models on hypergraphs. *Physical Review Research* **2,** 023032 (2020).

277. Li, Z., Deng, Z., Han, Z., Alfaro-Bittner, K., Barzel, B. & Boccaletti, S. Contagion in simplicial complexes. *Chaos, Solitons & Fractals* **152,** 111307 (2021).

278. Cisneros-Velarde, P. & Bullo, F. Multigroup SIS epidemics with simplicial and higher order interactions. *IEEE Transactions on Control of Network Systems* **9,** 695–705 (2021).

279. Higham, D. J. & De Kergorlay, H.-L. Epidemics on hypergraphs: Spectral thresholds for extinction. *Proceedings of the Royal Society A* **477,** 20210232 (2021).

280. Higham, D. J. & de Kergorlay, H.-L. Mean field analysis of hypergraph contagion models. *SIAM Journal on Applied Mathematics* **82,** 1987–2007 (2022).

281. Ferraz de Arruda, G., Petri, G., Rodriguez, P. M. & Moreno, Y. Multistability, intermittency, and hybrid transitions in social contagion models on hypergraphs. *Nature Communications* **14,** 1375 (2023).

282. Burgio, G., Arenas, A., Gómez, S. & Matamalas, J. T. Network clique cover approximation to analyze complex contagions through group interactions. *Communications Physics* **4,** 111 (2021).

283. Kou, L. T., Stockmeyer, L. J. & Wong, C.-K. Covering edges by cliques with regard to keyword conflicts and intersection graphs. *Communications of the ACM* **21,** 135–139 (1978).

284. Dickman, R. & Vidigal, R. Quasi-stationary distributions for stochastic processes with an absorbing state. *Journal of Physics A: Mathematical and General* **35,** 1147 (2002).

285. Ferreira, S. C., Castellano, C. & Pastor-Satorras, R. Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results. *Physical Review E* **86,** 041125 (2012).

286. Barrat, A., Ferraz de Arruda, G., Iacopini, I. & Moreno, Y. in *Higher-order systems* 329–346 (Springer, 2022).

287. Chowdhary, S., Kumar, A., Cencetti, G., Iacopini, I. & Battiston, F. Simplicial contagion in temporal higher-order networks. *Journal of Physics: Complexity* **2,** 035019 (2021).

288. Joslyn, C. A., Aksoy, S. G., Callahan, T. J., Hunter, L. E., Jefferson, B., Praggastis, B., Purvine, E. & Tripodi, I. J. *Hypernetwork science: from multidimensional networks to computational topology* in *International conference on complex systems* (2020), 377–392.

289. LaRock, T. & Lambiotte, R. Encapsulation structure and dynamics in hypergraphs. *Journal of Physics: Complexity* **4,** 045007 (2023).

290. Landry, N. W., Young, J.-G. & Eikmeier, N. The simpliciality of higher-order networks. *EPJ Data Science* **13** (2024).

291. Kim, J.-H. & Goh, K.-I. Higher-Order Components Dictate Higher-Order Contagion Dynamics in Hypergraphs. *Physical Review Letters* **132,** 087401 (2024).

292. Kim, J., Lee, D.-S. & Goh, K.-I. Contagion dynamics on hypergraphs with nested hyperedges. *Physical Review E* **108,** 034313 (2023).

293. Génois, M. & Barrat, A. Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Science* **7,** 1–18 (2018).

294. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the copenhagen networks study. *Scientific Data* **6,** 315 (2019).

295. Zhang, Y., Lucas, M. & Battiston, F. Higher-order interactions shape collective dynamics differently in hypergraphs and simplicial complexes. *Nature Communications* **14,** 1605 (2023).

296. Feld, S. L. Why your friends have more friends than you do. *American Journal of Sociology* **96,** 1464–1477 (1991).

297. House, T. & Keeling, M. J. Deterministic epidemic models with explicit household structure. *Mathematical Biosciences* **213,** 29–39 (2008).

298. Lindquist, J., Ma, J., van den Driessche, P. & Willeboordse, F. H. Effective degree network disease models. *Journal of Mathematical Biology* **62,** 143–164 (2011).

299. Unicomb, S., Iñiguez, G., Kertész, J. & Karsai, M. Reentrant phase transitions in threshold driven contagion on multiplex networks. *Physical Review E* **100,** 040301 (4 2019).

300. Unicomb, S., Iñiguez, G., Gleeson, J. P. & Karsai, M. Dynamics of cascades on burstiness-controlled temporal networks. *Nature Communications* **12,** 133 (2021).

301. St-Onge, G., Thibeault, V., Allard, A., Dubé, L. J. & Hébert-Dufresne, L. Social confinement and mesoscopic localization of epidemics on networks. *Physical Review Letters* **126,** 098301 (2021).

302. St-Onge, G., Thibeault, V., Allard, A., Dubé, L. J. & Hébert-Dufresne, L. Master equation analysis of mesoscopic localization in contagion dynamics on higher-order networks. *Physical Review E* **103,** 032301 (3 2021).

303. St-Onge, G., Iacopini, I., Latora, V., Barrat, A., Petri, G., Allard, A. & Hébert-Dufresne, L. Influential groups for seeding and sustaining nonlinear contagion in heterogeneous hypergraphs. *Communications Physics* **5,** 1–16 (2022).

304. Gross, T., D'Lima, C. J. D. & Blasius, B. Epidemic dynamics on an adaptive network. *Physical Review Letters* **96,** 208701 (2006).

305. Zanette, D. H. & Risau-Gusmán, S. Infection spreading in a population with evolving contacts. *Mathematical Epidemiology* **34,** 135–148 (2008).

306. Shaw, L. B. & Schwartz, I. B. Fluctuating epidemics on adaptive networks. *Physical Review E* **77,** 066101 (2008).

307. Van Segbroeck, S., Santos, F. C. & Pacheco, J. M. Adaptive contact networks change effective disease infectiousness and dynamics. *PLOS Computational Biology* **6,** e1000895 (2010).

308. Tunc, I. & Shaw, L. B. Effects of community structure on epidemic spread in an adaptive network. *Physical Review E* **90,** 022801 (2014).

309. Scarpino, S. V., Allard, A. & Hébert-Dufresne, L. The effect of a prudent adaptive behaviour on disease transmission. *Nature Physics* **12,** 1042–1046 (2016).

310. Wootton, J. T. Indirect effects and habitat use in an intertidal community: interaction chains and interaction modifications. *The American Naturalist* **141,** 71–89 (1993).

311. Okuyama, T & Bolker, B. M. in *Ecology and evolution of trait-mediated indirect interactions: linking evolution, community, and ecosystem* 186–204 (Cambridge University Press, 2012).

312. Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. Beyond pairwise mechanisms of species coexistence in complex communities. *Nature* **546,** 56–64 (2017).

313. Mickalide, H. & Kuehn, S. Higher-order interaction between species inhibits bacterial invasion of a phototroph-predator microbial community. *Cell Systems* **9,** 521–533 (2019).

314. Sahasrabuddhe, R., Neuhäuser, L. & Lambiotte, R. Modelling non-linear consensus dynamics on hypergraphs. *Journal of Physics: Complexity* **2,** 025006 (2021).

315. Glaeser, E. & Sunstein, C. R. Does more speech correct falsehoods? *The Journal of Legal Studies* **43,** 65–93 (2014).

316. Solé, R. V., Corominas-Murtra, B. & Fortuny, J. Diversity, competition, extinction: the ecophysics of language change. *Journal of The Royal Society Interface* **7,** 1647–1664 (2010).

317. Wolosker, H., Dumin, E., Balan, L. & Foltyn, V. N. D-amino acids in the brain: D-serine in neurotransmission and neurodegeneration. *The FEBS Journal* **275,** 3514–3526 (2008).

318. Cho, W.-H., Barcelon, E. & Lee, S. J. Optogenetic glia manipulation: possibilities and future prospects. *Experimental Neurobiology* **25,** 197 (2016).

319. Berg, J., Tymoczko, J. & Stryer, L. *Biochemistry, Fifth Edition* (W.H. Freeman, 2002).

320. Sun, H., Radicchi, F., Kurths, J. & Bianconi, G. The dynamic Nature of percolation on networks with triadic interactions. *Nature Communications* **14,** 1308 (2023).

321. St-Onge, G., Thibeault, V., Allard, A., Dubé, L. J. & Hébert-Dufresne, L. Social confinement and mesoscopic localization of epidemics on networks. *Physical Review Letters* **126,** 098301 (2021).

322. Lai, A., Poon, C. & Cheung, A. Effectiveness of facemasks to reduce exposure hazards for airborne infections among general populations. *Journal of the Royal Society Interface* **9,** 938–948 (2012).

323. Eikenberry, S. E., Mancuso, M., Iboi, E., Phan, T., Eikenberry, K., Kuang, Y., Kostelich, E. & Gumel, A. B. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling* **5,** 293–308 (2020).

324. Worby, C. J. & Chang, H.-H. Face mask use in the general population and optimal resource allocation during the COVID-19 pandemic. *Nature Communications* **11,** 1–9 (2020).

325. Catching, A., Capponi, S., Yeh, M. T., Bianco, S. & Andino, R. Examining the interplay between face mask usage, asymptomatic transmission, and social distancing on the spread of COVID-19. *Scientific Reports* **11,** 1–11 (2021).

326. Howard, J., Huang, A., Li, Z., Tufekci, Z., Zdimal, V., Van Der Westhuizen, H.-M., Von Delft, A., Price, A., Fridman, L., Tang, L.-H. *et al.* An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences* **118,** e2014564118 (2021).

327. Leech, G., Rogers-Smith, C., Monrad, J. T., Sandbrink, J. B., Snodin, B., Zinkov, R., Rader, B., Brownstein, J. S., Gal, Y., Bhatt, S. *et al.* Mask wearing in community settings reduces SARS-CoV-2 transmission. *Proceedings of the National Academy of Sciences* **119,** e2119266119 (2022).

328. Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., El-Harakeh, A., Bognanni, A., Lotfi, T., Loeb, M. *et al.* Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *The Lancet* **395,** 1973–1987 (2020).

329. Islam, N., Sharp, S. J., Chowell, G., Shabnam, S., Kawachi, I., Lacey, B., Massaro, J. M., D'Agostino, R. B. & White, M. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *BMJ* **370** (2020).

330. Koh, W. C., Naing, L. & Wong, J. Estimating the impact of physical distancing measures in containing COVID-19: an empirical analysis. *International Journal of Infectious Diseases* **100,** 42–49 (2020).

331. Jones, N. R., Qureshi, Z. U., Temple, R. J., Larwood, J. P., Greenhalgh, T. & Bourouiba, L. Two metres or one: what is the evidence for physical distancing in covid-19? *BMJ* **370** (2020).

332. Bedson, J., Skrip, L. A., Pedi, D., Abramowitz, S., Carter, S., Jalloh, M. F., Funk, S., Gobat, N., Giles-Vernick, T., Chowell, G. *et al.* A review and agenda for integrated disease models including social and behavioural factors. *Nature Human Behaviour* **5,** 834–846 (2021).

333. Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L. & Bauch, C. T. Coupled disease–behavior dynamics on complex networks: A review. *Physics of Life Reviews* **15,** 1–29 (2015).

334. Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N. *et al.* Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour* **4,** 460–471 (2020).

335. Bokemper, S. E., Cucciniello, M., Rotesi, T., Pin, P., Malik, A. A., Willebrand, K., Paintsil, E. E., Omer, S. B., Huber, G. A. & Melegaro, A. Beliefs about Mask Efficacy and the Effect of Social Norms on Mask Wearing Intentions for COVID-19 Risk Reduction. *medRxiv,* 2021–03 (2021).

336. Veldt, N., Benson, A. R. & Kleinberg, J. Combinatorial characterizations and impossibilities for higher-order homophily. *Science Advances* **9,** eabq3200 (2023).

337. McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: Homophily in social networks. *Annual Review of Sociology* **27,** 415–444 (2001).

338. McNeill, W. H. *Plagues and Peoples* (Anchor Books, New York, USA, 1976).

339. Pollard, A. J. & Bijker, E. M. A guide to vaccinology: from basic principles to new developments. *Nature Reviews Immunology* **21,** 83–100 (2021).

340. Smith, G. L. & McFadden, G. Smallpox: anything to declare? *Nature Reviews Immunology* **2,** 521–527 (2002).

341. Smith, J., Lipsitch, M. & Almond, J. W. Vaccine production, distribution, access, and uptake. *The Lancet* **378,** 428–438 (2011).

342. Wouters, O. J., Shadlen, K. C., Salcher-Konrad, M., Pollard, A. J., Larson, H. J., Teerawattananon, Y. & Jit, M. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *The Lancet* **397,** 1023–1034 (2021).

343. May, T. & Silverman, R. D. "Clustering of exemptions" as a collective action threat to herd immunity. *Vaccine* **21,** 1048–1051 (2003).

344. Parker, A. A., Staggs, W., Dayan, G. H., Ortega-Sánchez, I. R., Rota, P. A., Lowe, L., Boardman, P., Teclaw, R., Graves, C. & LeBaron, C. W. Implications of a 2005 measles outbreak in Indiana for sustained elimination of measles in the United States. *The New England Journal of Medicine* **355,** 447–455 (2006).

345. Editorial. Vaccine hesitancy: a generation at risk. *The Lancet Child & Adolescent Health* **3,** 281 (2019).

346. Robertson, E., Reeve, K. S., Niedzwiedz, C. L., Moore, J., Blake, M., Green, M., Katikireddi, S. V. & Benzeval, M. J. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. *Brain, Behavior, and Immunity* **94,** 41–50 (2021).

347. Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M. D. & Paterson, P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007-2012. *Vaccine* **32,** 2150–2159 (2014).

348. MacDonald, N. E. & on Vaccine Hesitancy, S. W. G. Vaccine hesitancy: Definition, scope and determinants. *Vaccine* **33,** 4161–4164 (2015).

349. Cascini, F., Pantovic, A., Al-Ajlouni, Y., Failla, G. & Ricciardi, W. Attitudes, acceptance and hesitancy among the general population worldwide to receive the COVID-19 vaccines and their contributing factors: A systematic review. *eClinicalMedicine* **40** (2021).

350. Richard, J. l, Masserey-Spicher, V., Santibanez, S. & Mankertz, A. Measles outbreak in Switzerland–an update relevant for the European football championship (EURO 2008). *Eurosurveillance* **13,** 8043 (2008).

351. Omer, S. B., Enger, K. S., Moulton, L. H., Halsey, N. A., Stokley, S. & Salmon, D. A. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. *American Journal of Epidemiology* **168,** 1389–1396 (2008).

352. Atwell, J. E., Van Otterloo, J., Zipprich, J., Winter, K., Harriman, K., Salmon, D. A., Halsey, N. A. & Omer, S. B. Nonmedical vaccine exemptions and pertussis in California, 2010. *Pediatrics* **132,** 624–630 (2013).

353. Barclay, V. C., Smieszek, T., He, J., Cao, G., Rainey, J. J., Gao, H., Uzicanin, A. & Salathé, M. Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity. *PLOS ONE* **9,** e87042 (2014).

354. Lieu, T. A., Ray, G. T., Klein, N. P., Chung, C. & Kulldorff, M. Geographic clusters in underimmunization and vaccine refusal. *Pediatrics* **135,** 280–289 (2015).

355. Edge, R., Keegan, T., Isba, R. & Diggle, P. Observational study to assess the effects of social networks on the seasonal influenza vaccine uptake by early career doctors. *BMJ Open* **9,** e026997 (2019).

356. Mbah, M. L. N., Liu, J., Bauch, C. T., Tekel, Y. I., Medlock, J., Meyers, L. A. & Galvani, A. P. The impact of imitation on vaccination behavior in social contact networks. *PLOS Computational Biology* **8,** e1002469 (2012).

357. Liu, F., Enanoria, W. T. A., Zipprich, J., Blumberg, S., Harriman, K., Ackley, S. F., Wheaton, W. D., Allpress, J. L. & Porco, T. C. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for California. *BMC Public Health* **15,** 447 (2015).

358. Glasser, J. W., Feng, Z., Omer, S. B., Smith, P. J. & Rodewald, L. E. The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study. *The Lancet Infectious Diseases* **16,** 599–605 (2016).

359. Kuylen, E., Willem, L., Broeckhove, J., Beutels, P. & Hens, N. Clustering of susceptible individuals within households can drive measles outbreaks: an individual-based model exploration. *Scientific Reports* **10,** 19645 (2020).

360. Centers for Disease Control and Prevention. *CDC Seasonal Flu Vaccine Effectiveness Studies* https://www.cdc.gov/flu/vaccines-work/effectiveness-studies.htm. 2021.

361. Pouwels, K. B. *et al.* Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nature Medicine* (2021).

362. Shim, E. & Galvani, A. P. Distinguishing vaccine efficacy and effectiveness. *Vaccine* **30,** 6700–6705 (2012).

363. *Catalunya's COVID-19 vaccine uptake data* https://dadescovid.cat/descarregues.

364. *France's COVID-19 vaccine uptake data* https://www.data.gouv.fr/fr/datasets.

365. *Italy's COVID-19 vaccine uptake data* https://github.com/italia/covid19-opendata-vaccini/tree/master/dati.

366. *Switzerland's COVID-19 vaccine uptake data* https://www.covid19.admin.ch/en/vaccination/persons.

367. Centola, D. An experimental study of homophily in the adoption of health behavior. *Science* **334,** 1269–1272 (2011).

368. Moreno López, J. A., García, B. A., Bentkowski, P., Bioglio, L., Pinotti, F., Boëlle, P.-Y., Barrat, A., Colizza, V. & Poletto, C. Anatomy of digital contact tracing: Role of age, transmission setting, adoption and case detection. *Science Advances* (2021).

369. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the Copenhagen Networks Study. *Scientific Data* **6,** 315 (2019).

370. Sekara, V. & Lehmann, S. The strength of friendship ties in proximity sensor data. *PLOS ONE* **9,** e100915 (2014).

371. Bubar, K. M., Reinholt, K., Kissler, S. M., Lipsitch, M., Cobey, S., Grad, Y. H. & Larremore, D. B. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science* **371,** 916–921 (2021).

372. Sonabend, R. *et al.* Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *The Lancet* **398,** 1825–1835 (2021).

373. Brand, S. P. C. *et al.* COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science* **374,** 989–994 (2021).

374. Wilf-Miron, R., Myers, V. & Saban, M. Incentivizing Vaccination Uptake: The "Green Pass" Proposal in Israel. *JAMA: The Journal of the American Medical Association* **325,** 1503–1504 (2021).

375. Dada, S., Battles, H., Pilbeam, C., Singh, B., Solomon, T. & Gobat, N. Learning from the past and present: social science implications for COVID-19 immunity-based documentation. *Humanities and Social Sciences Communications* **8,** 219 (2021).

376. Hiraoka, T., Rizi, A. K., Kivelä, M. & Saramäki, J. Herd immunity and epidemic size in networks with vaccination homophily. *Physical Review E* **105,** L052301 (2022).

377. Fefferman, N. H., Silk, M. J., Pasquale, D. K. & Moody, J. Homophily in risk and behavior complicate understanding the COVID-19 epidemic curve. *medRxiv* (2021).

378. Young, M. J., Silk, M. J., Pritchard, A. J. & Fefferman, N. H. *Diversity in Valuing Social Contact and Risk Tolerance Lead to the Emergence of Homophily in Populations Facing Infectious Threats* 2021.

379. Rohani, P. & Keeling, M. *Modeling infectious diseases in humans and animals* 1–368 (Princeton University Press, 2011).

380. Eichner, M. Case isolation and contact tracing can prevent the spread of smallpox. *American Journal of Epidemiology* **158,** 118–128 (2003).

381. Fox, G. J., Barry, S. E., Britton, W. J. & Marks, G. B. *Contact investigation for tuberculosis: A systematic review and meta-analysis* 2013.

382. Ramstedt, K., Hallhagen, G., Lundin, B.-I., Hakansson, C., Johannisson, G., LÖWhagen, G.-B., Norkrans, G. & Giesecke, J. Contact tracing for human immunodeficiency virus (HIV) infection. *Sexually Transmitted Diseases* **17,** 37–41 (1990).

383. Swanson, K. C., Altare, C., Wesseh, C. S., Nyenswah, T., Ahmed, T., Eyal, N., Hamblion, E. L., Lessler, J., Peters, D. H. & Altmann, M. Contact tracing performance during the Ebola epidemic in Liberia, 2014-2015. *PLoS Neglected Tropical Diseases* **12** (2018).

384. Howell, M. R., Kassler, W. J. & Haddix, A. Partner Notification to Prevent Pelvic Inflammatory Disease in Women. *Sexually Transmitted Diseases* **24,** 287–292 (1997).

385. Kretzschmar, M., Van Duynhoven, Y. T. & Severijnen, A. J. Modeling prevention strategies for gonorrhea and chlamydia using stochastic network simulations. *American Journal of Epidemiology* **144,** 306–317 (1996).

386. Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* **101,** 6146–6151 (2004).

387. Klinkenberg, D., Fraser, C. & Heesterbeek, H. The effectiveness of contact tracing in emerging epidemics. *PLOS ONE* **1** (2006).

388. Porco, T. C., Holbrook, K., Fernyak, S. E., Portnoy, D., Reiter, R. & Aragón, T. J. Logistics of community smallpox control through contact tracing and ring vaccination: A stochastic network model. *BMC Public Health* **4** (2004).

389. Müller, J., Kretzschmar, M. & Dietz, K. Contact tracing in stochastic and deterministic epidemic models. *Mathematical Biosciences* **164,** 39–64 (2000).

390. Kiss, I. Z., Green, D. M. & Kao, R. R. Infectious disease control using contact tracing in random and scale-free networks. *Journal of the Royal Society Interface* **3,** 55–62 (2006).

391. Eames, K. T. & Keeling, M. J. Contact tracing and disease control. *Proceedings of the Royal Society B* **270,** 2565–2571 (2003).

392. Becker, N. G., Glass, K., Li, Z. & Aldis, G. K. Controlling emerging infectious diseases like SARS. *Mathematical Biosciences* **193,** 205–221 (2005).

393. Braithwaite, I., Callender, T., Bullock, M. & Aldridge, R. W. *Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19* 2020.

394. Salathé, M., Althaus, C. L., Anderegg, N., Antonioli, D., Ballouz, T., Bugnion, E., Capkun, S., Jackson, D., Kim, S.-I., Larus, J. *et al.* Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. *medRxiv* (2020).

395. Rodríguez, P. *et al.* A population-based controlled experiment assessing the epidemiological impact of digital contact tracing. *Nature Communications* **12,** 1–6 (2021).

396. Kucharski, A. J., Klepac, P., Conlan, A. J., Kissler, S. M., Tang, M. L., Fry, H., Gog, J. R., Edmunds, W. J., Emery, J. C., Medley, G. *et al.* Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *The Lancet Infectious Diseases* **20,** 1151–1160 (2020).

397. Aleta, A., Martin-Corral, D., y Piontti, A. P., Ajelli, M., Litvinova, M., Chinazzi, M., Dean, N. E., Halloran, M. E., Longini Jr, I. M., Merler, S. *et al.* Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4,** 964–971 (2020).

398. Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D. & Fraser, C. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368** (2020).

399. Barrat, A., Cattuto, C., Kivelä, M., Lehmann, S. & Saramäki, J. Effect of manual and digital contact tracing on COVID-19 outbreaks: a study on empirical contact data. *Journal of The Royal Society Interface* **18,** 20201000 (2021).

400. Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health* **8,** e488–e496 (2020).

401. Lorch, L., Kremer, H., Trouleau, W., Tsirtsis, S., Szanto, A., Schölkopf, B. & Gomez-Rodriguez, M. Quantifying the effects of contact tracing, testing, and containment measures in the presence of infection hotspots. *ACM Transactions on Spatial Algorithms and Systems* **8,** 1–28 (2022).

402. Bradshaw, W. J., Alley, E. C., Huggins, J. H., Lloyd, A. L. & Esvelt, K. M. Bidirectional contact tracing could dramatically improve COVID-19 control. *Nature Communications* **12** (2021).

403. Cencetti, G., Santin, G., Longa, A., Pigani, E., Barrat, A., Cattuto, C., Lehmann, S., Salathé, M. & Lepri, B. Digital proximity tracing on empirical contact networks for pandemic control. *Nature Communications* **12,** 1655 (2021).

404. Reyna-Lara, A., Soriano-Paños, D., Gómez, S., Granell, C., Matamalas, J. T., Steinegger, B., Arenas, A. & Gómez-Gardeñes, J. Virus spread versus contact tracing: Two competing contagion processes. *Physical Review Research* **3,** 013163 (1 2021).

405. Bianconi, G., Sun, H., Rapisardi, G. & Arenas, A. Message-passing approach to epidemic tracing and mitigation with apps. *Physical Review Research* **3,** L012014 (1 2021).

406. Kryven, I. & Stegehuis, C. Contact tracing in configuration models. *Journal of Physics: Complexity* **2,** 025004 (2021).

407. Kojaku, S., Hébert-Dufresne, L., Mones, E., Lehmann, S. & Ahn, Y.-Y. The effectiveness of backward contact tracing in networks. *Nature Physics,* 1–7 (2021).

408. Von Wyl, V., Höglinger, M., Sieber, C., Kaufmann, M., Moser, A., Serra-Burriel, M., Ballouz, T., Menges, D., Frei, A. & Puhan, M. A. Drivers of Acceptance of COVID-19 Proximity Tracing Apps in Switzerland: Panel Survey Analysis. *JMIR Public Health Surveill* **7,** e25701 (2021).

409. Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F. & Lowe, W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nature Human Behaviour* (2021).

410. Salathé, M., Kazandjieva, M., Lee, J. W., Levis, P., Feldman, M. W. & Jones, J. H. A high-resolution human contact network for infectious disease transmission. *Proceedings of the National Academy of Sciences* **107,** 22020–22025 (2010).

411. Stehlé, J. *et al.* High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE* **6,** e23176 (2011).

412. Génois, M. & Barrat, A. Can co-location be used as a proxy for face-to-face contacts? *EPJ Data Science* **7,** 11 (2018).

413. Watanabe, H. & Hasegawa, T. Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks. *Physica A: Statistical Mechanics and its Applications* **603,** 127760 (2022).

414. Kadelka, C. & McCombs, A. Effect of homophily and correlation of beliefs on COVID-19 and general infectious disease outbreaks. *PLOS ONE* **16,** 1–20 (2021).

415. Burgio, G., St-Onge, G. & Hébert-Dufresne, L. Adaptive hypergraphs and the characteristic scale of higher-order contagions using generalized approximate master equations. *arXiv preprint arXiv:2307.11268* (2023).

416. Barceló, J. & Sheen, G. C.-H. Voluntary adoption of social welfare-enhancing behavior: Mask-wearing in Spain during the COVID-19 outbreak. *PLOS ONE* **15,** e0242764 (2020).

417. Burgio, G., Steinegger, B., Rapisardi, G. & Arenas, A. Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics. *Physical Review Research* **3,** 033128 (2021).

418. Burgio, G., Steinegger, B. & Arenas, A. Homophily impacts the success of vaccine roll-outs. *Communications Physics* **5,** 1–7 (2022).

419. Watanabe, H. & Hasegawa, T. Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks. *Physica A: Statistical Mechanics and its Applications,* 127760 (2022).

420. Rizi, A. K., Faqeeh, A., Badie-Modiri, A. & Kivelä, M. Epidemic spreading and digital contact tracing: Effects of heterogeneous mixing and quarantine failures. *Physical Review E* **105,** 044313 (2022).

421. Hiraoka, T., Rizi, A. K., Kivelä, M. & Saramäki, J. Herd immunity and epidemic size in networks with vaccination homophily. *Physical Review E* **105,** L052301 (2022).

422. Eisenberg, N., Fabes, R. A. & Spinrad, T. L. *Prosocial development.* (John Wiley & Sons, Inc., 2006).

423. Hill, A. L., Rand, D. G., Nowak, M. A. & Christakis, N. A. Infectious disease modeling of social contagion in networks. *PLOS Computational Biology* **6,** e1000968 (2010).

424. Hébert-Dufresne, L., Waring, T. M., St-Onge, G., Niles, M. T., Kati Corlew, L., Dube, M. P., Miller, S. J., Gotelli, N. J. & McGill, B. J. Source-sink behavioural dynamics limit institutional evolution in a group-structured society. *Royal Society Open Science* **9,** 211743 (2022).

425. DeGroot, M. H. Reaching a consensus. *Journal of the American Statistical association* **69,** 118–121 (1974).

426. Sorensen, R. J., Barber, R. M., Pigott, D. M., Carter, A., Spencer, C. N., Ostroff, S. M., Reiner Jr, R. C., Abbafati, C., Adolph, C., Allorant, A. *et al.* Variation in the COVID-19 infection-fatality ratio by age, time, and geography during the pre-vaccine era: A systematic analysis. *The Lancet* **399,** 1469–1488 (2022).

427. Cozza, V., Campbell, H., Chang, H. H., Iuliano, A. D., Paget, J., Patel, N. N., Reiner, R. C., Troeger, C., Viboud, C., Bresee, J. S. *et al.* Global seasonal influenza mortality estimates: A comparison of 3 different approaches. *American Journal of Epidemiology* **190,** 718–727 (2021).

428. Badillo-Goicoechea, E., Chang, T.-H., Kim, E., LaRocca, S., Morris, K., Deng, X., Chiu, S., Bradford, A., Garcia, A., Kern, C. *et al.* Global trends and predictors of face mask usage during the COVID-19 pandemic. *BMC Public Health* **21,** 1–12 (2021).

429. Haischer, M. H., Beilfuss, R., Hart, M. R., Opielinski, L., Wrucke, D., Zirgaitis, G., Uhrich, T. D. & Hunter, S. K. Who is wearing a mask? Gender-, age-, and location-related differences during the COVID-19 pandemic. *PLOS ONE* **15,** e0240785 (2020).

430. Sánchez-Arenas, R., Doubova, S. V., González-Pérez, M. A. & Pérez-Cuevas, R. Factors associated with COVID-19 preventive health behaviors among the general public in Mexico City and the State of Mexico. *PLOS ONE* **16,** e0254435 (2021).

431. Kaewpan, W., Rojpaisarnkit, K., Pengpid, S. & Peltzer, K. Factors affecting face mask-wearing behaviors to prevent COVID-19 among Thai people: A binary logistic regression model. *Frontiers in Psychology* **13,** 996189 (2022).

432. Golubski, A. J., Westlund, E. E., Vandermeer, J. & Pascual, M. Ecological networks over the edge: hypergraph trait-mediated indirect interaction (TMII) structure. *Trends in Ecology & Evolution* **31,** 344–354 (2016).

433. Vandermeer, J., Perfecto, I. & Philpott, S. Ecological complexity and pest control in organic coffee production: uncovering an autonomous ecosystem service. *BioScience* **60,** 527–537 (2010).

434. Perfecto, I., Vandermeer, J. & Philpott, S. M. Complex ecological interactions in the coffee agroecosystem. *Annual Review of Ecology, Evolution, and Systematics* **45,** 137–158 (2014).

435. Mann, P., Smith, V. A., Mitchell, J. B. & Dobson, S. Degree correlations in graphs with clique clustering. *Physical Review E* **105,** 044314 (2022).

436. Keating, L. A., Gleeson, J. P. & O'Sullivan, D. J. A generating-function approach to modelling complex contagion on clustered networks with multi-type branching processes. *Journal of Complex Networks* **11,** cnad042 (2023).

437. Dooley, K. J. A complex adaptive systems model of organization change. *Nonlinear Dynamics, Psychology, and Life Sciences* **1,** 69–97 (1997).

438. Cordes, C., Richerson, P. J., McElreath, R. & Strimling, P. A naturalistic approach to the theory of the firm: The role of cooperation and cultural evolution. *Journal of Economic Behavior & Organization* **68,** 125–139 (2008).

439. Cordes, C., Richerson, P., McElreath, R. & Strimling, P. How does opportunistic behavior influence firm size? An evolutionary approach to organizational behavior. *Journal of Institutional Economics* **7,** 1–21 (2011).

440. St-Onge, J., Burgio, G., Rosenblatt, S. F., Waring, T. M. & Hébert-Dufresne, L. Paradoxes in the co-evolution of contagions and institutions. *arXiv preprint arXiv:2310.03672* (2023).

441. Dawkins, R. *The selfish gene* (Oxford University Press, 2016).

442. Pennisi, E. *On the origin of cooperation* 2009.

443. Dugatkin, L. A. *Cooperation among animals: an evolutionary perspective* (Oxford University Press, USA, 1997).

444. Nowak, M. & Highfield, R. *Supercooperators: Altruism, evolution, and why we need each other to succeed* (Simon and Schuster, 2011).

445. Roca, C. P., Cuesta, J. A. & Sánchez, A. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews* **6,** 208–249 (2009).

446. Szabó, G., Vukov, J. & Szolnoki, A. Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices. *Physical Review E* **72,** 047107 (2005).

447. Nowak, M. A. & May, R. M. Evolutionary games and spatial chaos. *Nature* **359,** 826–829 (1992).

448. Ohtsuki, H., Hauert, C., Lieberman, E. & Nowak, M. A. A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441,** 502–505 (2006).

449. Killingback, T., Doebeli, M. & Knowlton, N. Variable investment, the continuous prisoner's dilemma, and the origin of cooperation. *Proceedings of the Royal Society B* **266,** 1723–1728 (1999).

450. Santos, F. C., Rodrigues, J. & Pacheco, J. M. Graph topology plays a determinant role in the evolution of cooperation. *Proceedings of the Royal Society B* **273,** 51–55 (2006).

451. Gómez-Gardenes, J., Campillo, M., Floría, L. M. & Moreno, Y. Dynamical organization of cooperation in complex topologies. *Physical Review Letters* **98,** 108103 (2007).

452. Santos, F. C., Santos, M. D. & Pacheco, J. M. Social diversity promotes the emergence of cooperation in public goods games. *Nature* **454,** 213–216 (2008).

453. Gomez-Gardenes, J., Romance, M., Criado, R., Vilone, D. & Sánchez, A. Evolutionary games defined at the network mesoscale: The public goods game. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **21** (2011).

454. Tanimoto, J. *Fundamentals of evolutionary game theory and its applications* (Springer, 2015).

455. Traulsen, A. & Claussen, J. C. Similarity-based cooperation and spatial segregation. *Physical Review E* **70,** 046128 (2004).

456. Olson Jr, M. *The Logic of Collective Action: Public Goods and the Theory of Groups, with a new preface and appendix* (Harvard University Press, 1971).

457. Kuramoto, Y. *Chemical oscillations, waves, and turbulence* (Dover Publications, 2003).

458. Strogatz, S. H. *Sync: The Emerging Science of Spontaneous Order* (Hyperion, 2003).

459. Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y. & Zhou, C. Synchronization in complex networks. *Physics Reports* **469,** 93–153 (2008).

460. Tanaka, T. & Aoyagi, T. Multistable attractors in a network of phase oscillators with three-body interactions. *Physical Review Letters* **106,** 224101 (2011).

461. Bick, C., Ashwin, P. & Rodrigues, A. Chaos in generically coupled phase oscillator networks with nonpairwise interactions. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **26** (2016).

462. Skardal, P. S. & Arenas, A. Abrupt desynchronization and extensive multistability in globally coupled oscillator simplexes. *Physical Review Letters* **122,** 248301 (2019).

463. León, I. & Pazó, D. Phase reduction beyond the first order: The case of the mean-field complex Ginzburg-Landau equation. *Physical Review E* **100,** 012211 (2019).

464. Skardal, P. S. & Arenas, A. Higher order interactions in complex networks of phase oscillators promote abrupt synchronization switching. *Communications Physics* **3,** 218 (2020).

465. Millán, A. P., Torres, J. J. & Bianconi, G. Explosive higher-order Kuramoto dynamics on simplicial complexes. *Physical Review Letters* **124,** 218301 (2020).

466. Skardal, P. S., Arola-Fernández, L., Taylor, D. & Arenas, A. Higher-order interactions can better optimize network synchronization. *Physical Review Research* **3,** 043193 (2021).

467. Gallo, L., Muolo, R., Gambuzza, L. V., Latora, V., Frasca, M. & Carletti, T. Synchronization induced by directed higher-order interactions. *Communications Physics* **5,** 263 (2022).

468. Arola-Fernández, L. *Synchronization in Complex Networks Under Uncertainty* (Universitat Rovira i Virgili, 2022).

469. Benson, A. R., Abebe, R., Schaub, M. T., Jadbabaie, A. & Kleinberg, J. Simplicial closure and higher-order link prediction. *Proceedings of the National Academy of Sciences* **115,** E11221–E11230 (2018).

470. Benson, A. R., Kumar, R. & Tomkins, A. *Sequences of sets* in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (2018), 1148–1157.

471. Lee, G., Choe, M. & Shin, K. *How do hyperedges overlap in real-world hypergraphs?-patterns, measures, and generators* in *Proceedings of the web conference 2021* (2021), 3396–3407.

472. Cencetti, G., Contreras, D. A., Mancastroppa, M. & Barrat, A. Distinguishing simple and complex contagion processes on networks. *Physical Review Letters* **130,** 247401 (2023).

473. Hébert-Dufresne, L., Scarpino, S. V. & Young, J.-G. Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement. *Nature Physics* **16,** 426–431 (2020).

474. Timme, M. Revealing network connectivity from response dynamics. *Physical Review Letters* **98,** 224101 (2007).

475. Napoletani, D. & Sauer, T. D. Reconstructing the topology of sparsely connected dynamical networks. *Physical Review E* **77,** 026103 (2008).

476. Wang, W.-X., Chen, Q., Huang, L., Lai, Y.-C. & Harrison, M. A. F. Scaling of noisy fluctuations in complex networks and applications to network prediction. *Physical Review E* **80,** 016116 (2009).

477. Ren, J., Wang, W.-X., Li, B. & Lai, Y.-C. Noise bridges dynamical correlation and topology in coupled oscillator networks. *Physical Review Letters* **104,** 058701 (2010).

478. Wu, X., Zhou, C., Chen, G. & Lu, J.-a. Detecting the topologies of complex networks with stochastic perturbations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **21** (2011).

479. Wang, H., Ma, C., Chen, H.-S., Lai, Y.-C. & Zhang, H.-F. Full reconstruction of simplicial complexes from binary contagion and Ising data. *Nature Communications* **13,** 3043 (2022).

480. Lizotte, S., Young, J.-G. & Allard, A. Hypergraph reconstruction from noisy pairwise observations. *arXiv preprint arXiv:2208.06503* (2022).

481. Santoro, A., Battiston, F., Petri, G. & Amico, E. Higher-order organization of multivariate time series. *Nature Physics* **19,** 221–229 (2023).

482. Malizia, F., Corso, A., Gambuzza, L. V., Russo, G., Latora, V. & Frasca, M. Reconstructing higher-order interactions in coupled dynamical systems. *arXiv preprint arXiv:2305.06609* (2023).

483. Delabays, R., De Pasquale, G., Dörfler, F. & Zhang, Y. Hypergraph reconstruction from dynamics. *arXiv preprint arXiv:2402.00078* (2024).

484. Tuomela, R. *The philosophy of sociality: The shared point of view* (Oxford University Press, 2007).

485. List, C. & Pettit, P. *Group agency: The possibility, design, and status of corporate agents* (Oxford University Press, 2011).

486. Tuomela, R. *Social ontology: Collective intentionality and group agents* (Oxford University Press, 2013).

487. Townsend, L. Social Ontology: Collective Intentionality and Group Agents, written by Raimo Tuomela. *Journal of Social Ontology* **1,** 183–187 (2015).

488. Harp, R. Social Ontology: Collective Intentionality and Group Agents, written by Raimo Tuomela. *Journal of Moral Philosophy* **14,** 608–611 (2017).

489. Lackey, J. *The epistemology of groups* (Oxford University Press, USA, 2021).

490. Van den Driessche, P. & Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* **180,** 29–48 (2002).

491. Van den Driessche, P. & Watmough, J. Further notes on the basic reproduction number. *Mathematical Epidemiology,* 159–178 (2008).

492. Allen, L. J. & van den Driessche, P. The basic reproduction number in some discrete-time epidemic models. *Journal of Difference Equations and Applications* **14,** 1127–1147 (2008).

493. De Oliveira, M. M. & Dickman, R. How to simulate the quasistationary state. *Physical Review E* **71,** 016129 (2005).

494. Wilf, H. S. *generatingfunctionology* (CRC press, 2005).

495. *The On-Line Encyclopedia of Integer Sequences (OEIS)* https://oeis.org/A000110.

496. *FFTW.jl Julia package* https://github.com/JuliaMath/FFTW.jl (2023).

497. *FFTW library* http://www.fftw.org (2023).

498. Wu, Q. & Hadzibeganovic, T. An individual-based modeling framework for infectious disease spreading in clustered complex networks. *Applied Mathematical Modelling* **83,** 1–12 (2020).

499. Gómez, S., Gómez-Gardenes, J., Moreno, Y. & Arenas, A. Nonperturbative heterogeneous mean-field approach to epidemic spreading in complex networks. *Physical Review E* **84,** 036105 (2011).

500. Newman, M. Assortative mixing in networks. *Physical Review Letters* **89,** 208701 (2002).

501. Lv, X., Fan, D., Li, Q., Wang, J. & Zhou, L. Simplicial SIR rumor propagation models with delay in both homogeneous and heterogeneous networks. *Physica A: Statistical Mechanics and its Applications* **627,** 129131 (2023).

502. Arregui, S., Aleta, A., Sanz, J. & Moreno, Y. Projecting social contact matrices to different demographic structures. *PLOS Computational Biology* **14,** e1006638 (2018).

UNIVERSITAT ROVIRA I VIRGILI
CONTAGION PROCESSES ON HIGHER-ORDER NETWORKS
Giulio Burgio

UNIVERSITAT
ROVIRA i VIRGILI