# Contagion à effet de seuil dans les réseaux complexes

Samuel Lee Unicomb



**THESE de DOCTORAT DE L'UNIVERSITE DE LYON**

opérée par

**l'Ecole Normale Supérieure de Lyon**

**Ecole Doctorale** N° accréditation : 512

**Ecole Doctorale** : École Doctorale en Informatique et Mathématiques de Lyon

**Spécialité de doctorat** : Systèmes Complexes

**Discipline** : Informatique

Soutenue publiquement le 14/01/2020, par :

**Samuel UNICOMB**

---

# Threshold driven contagion on complex networks

***Contagion à effet de seuil dans les réseaux complexes***

---

## Devant le jury composé de :

**Saramäki, Jari** Professor, Aalto University, Finland, Rapporteur

**Moreno, Yamir** Professor, University of Zaragoza, Spain, Rapporteur

**Burioni, Raffaella** Professor, Università di Parma, Italy, Rapporteure

**Magnien, Clémence** Directeur de recherche, Sorbonne Université, France, Examinatrice

**Avrachenkov, Konstantin** Directeur de recherche, Inria Sophia Antiplos, Examinateur

**Fleury, Eric** Professor, Inria Paris, Co-encadrant de thèse

**Karsai, Márton** Associate Professor, CEU, Hungary, Directeur de thèse

## ABSTRACT

Networks arise frequently in the study of complex systems, since interactions among the components of such systems are critical. Networks can act as a substrate for dynamical process, such as the diffusion of information or disease throughout populations. Network structure can determine the temporal evolution of a dynamical process, including the characteristics of the steady state.

The simplest representation of a complex system is an undirected, unweighted, single layer graph. In contrast, real systems exhibit heterogeneity of interaction strength and type. Such systems are frequently represented as weighted multiplex networks, and in this work we incorporate these heterogeneities into a master equation formalism in order to study their effects on spreading processes. We also carry out simulations on synthetic and empirical networks, and show that spreading dynamics, in particular the speed at which contagion spreads via threshold mechanisms, depend non-trivially on these heterogeneities. Further, we show that an important family of networks undergo reentrant phase transitions in the size and frequency of global cascades as a result of these interactions.

A challenging feature of real systems is their tendency to evolve over time, since the changing structure of the underlying network is critical to the behaviour of overlying dynamical processes. We show that one aspect of temporality, the observed "burstiness" in interaction patterns, leads to non-monotic changes in the spreading time of threshold driven contagion processes.

The above results shed light on the effects of various network heterogeneities, with respect to dynamical processes that evolve on these networks.

## RÉSUMÉ

Les interactions entre les composants des systèmes complexes font émerger différents types de réseaux. Ces réseaux peuvent jouer le rôle d'un substrat pour des processus dynamiques tels que la diffusion d'informations ou de maladies dans des populations. Les structures de ces réseaux déterminent l'évolution d'un processus dynamique, en particulier son régime transitoire, mais aussi les caractéristiques du régime permanent.

Les systèmes complexes réels manifestent des intéractions hétérogènes en type et en intensité. Ces systèmes sont représetés comme des réseaux pondérés à plusieurs couches. Dans cette thèse, nous développons une équation maîtresse afin d'intégrer ces hétérogénéités et d'étudier leurs effets sur les processus de diffusion. À l'aide de simulations mettant en jeu des réseaux réels et générés, nous montrons que les dynamiques de diffusion sont liées de manière non triviale à l'hétérogénéité de ces réseaux, en particulier la vitesse de propagation d'une contagion basée sur un effet de seuil. De plus, nous montrons que certaines classes de réseaux sont soumises à des transitions de phase réentrantes fonctions de la taille des "global cascades".

La tendance des réseaux réels à évoluer dans le temps rend difficile la modélisation des processus de diffusion. Nous montrons enfin que la durée de diffusion d'un processus de contagion basé sur un effet de seuil change de manière non-monotone du fait de la présence de "rafales" dans les motifs d'intéractions. L'ensemble de ces résultats mettent en lumière les effets de l'hétérogénéité des réseaux vis-à-vis des processus dynamiques y évoluant.

## PUBLICATIONS

1. Unicomb, S., Iñiguez, G. & Karsai, M. Threshold driven contagion on weighted networks. *Scientific reports* **8,** 3094 (2018).
2. Unicomb, S., Iñiguez, G., Kertész, J. & Karsai, M. Reentrant phase transitions in threshold driven contagion on multiplex networks. *Phys. Rev. E* **100,** 040301 (2019).
3. Unicomb, S., Gleeson, J. P., Iñiguez, G. & Karsai, M. The dynamics of cascades on bursty temporal networks. *Working paper* (2019).

# ACKNOWLEDGMENTS

I would like to thank first and foremost my supervisor, Márton Karsai, for his constant support over the course of this thesis. On both a scientific and technical level, as well a broader academic sense, Márton's guidance was absolutely indispensable. Evident in every meeting was an absolute and sincere drive for the betterment of our common work. I owe a great deal also to Gerardo Iñiguez, my second scientific advisor and supervisor, not just for his expertise and creativity, but for his constant enthusiasm for the problems at hand.

I wish to sincerely thank my co-authors James P. Gleeson and János Kertész, for their expertise shared during our common work. I'm particularly indebted to James for a tremendously formative three-month visit at the University of Limerick in Ireland.

Last but not least I would like to thank my jury, namely Jari Saramäki, Yamir Moreno, Raffaella Burioni, Clémence Magnien, Konstantin Avrachenkov and Eric Fleury. I'm especially grateful to Eric for his support at the beginning of my thesis. Without him, this thesis would likely never have gotten its feet off the ground.

# CONTENTS

# LIST OF FIGURES

## LIST OF TABLES



## ACRONYMS

AME approximate master equation
ER Erdös-Rényi
BFS breadth-first search
FFP forest-fire process
WS Watts-Strogatz

| | |
|---|---|
| BA | Barabási-Albert |
| RK | Runge-Kutta |
| MC | Monte Carlo |
| LSA | linear stability analysis |
| VFA | vector field analysis |
| SI | susceptible-infected |
| SIS | susceptible-infected-susceptible |
| SIR | susceptible-infected-recovered |
| ODE | ordinary differential equation |
| CSS | computational social science |

## NOMENCLATURE

The following is a list of commonly used symbols. As much as possible, the page number indicates gives the first page on which the symbol is used and defined. A number of symbols are only used in Part I, the introductory section, which includes Chapters 1 and 2. We indicate those symbols for which this is the case.

| Symbol | Meaning | Chapter | Page |
|---|---|---|---|
| $\zeta$ | branching number | 2 | 75 |
| $C$ | clustering coefficient | 1 | 20 |
| $d$ | graph diameter | 1 | 19 |
| $E$ | set of edges $(u, v)$ | | 13 |
| $E_i$ | edge set in layer $i$ | | 14 |
| $e_{uv}$ | edge index | | 13 |
| $G(V, E)$ | graph $G$, composed of $V$ and $E$ | | 13 |
| $l$ | path length | 1 | 10 |
| $M$ | number of multiplex layers | | 14 |
| $N$ | network size, number of nodes | | 10 |
| $u$, $v$ | vertex indices | | 13 |
| $V$ | set of vertices | | 13 |

| | | |
|---|---|---|
| $w_{uv}$ | weight of edge $(u, v)$ | 13 |
| $\text{supp}(p_k)$ | support of distribution $p_k$ | 54 |

# Part I

# BACKGROUND

In Chapter 1 we review the existing literature, outlining the systems that have been of interest to network scientists, and giving a brief review of the main random graph models. We sketch the dynamical systems to which our mathematical methods apply. Then, in Chapter 2 we provide a background to the analytic tools of which we make use throughout each of the remaining chapters. In particular, this chapter introduces a master equation formalism that we extend to edge heterogeneous systems. We conclude by comparing these methods to alternative techniques used to understand dynamical processes on networks. In Part II, we apply these tools to a number of specific systems.

# 1

# INTRODUCTION

## 1.1 INTRODUCTION

Complex systems are structures composed of many entities whose interactions lead to collective behaviour not easily inferred from the examination of individual entities. Due to the interactions between these entities, which are essential to emergent phenomena, networks have emerged as a compelling modelling framework. Abstracting a complex system as a network allows one to disregard features of the entities that are relevant to system-wide behaviour.

Complex systems have long been the focus of study, but have seen a renewal in methodology due to the emergence of so called "big data." This refers to the abundance of quantitative information about real systems due to the improving ability to observe, record, read, sensor, and output data. This increased availability of data is due in large part to maturation of technology; with the digitisation of communication, economics and finance, as well as medicine. This has triggered the emergence of interdisciplinary techniques, and indeed a new field in its own right, namely network science, complete with its own methodology, techniques and literature.

The modelling of complex systems often proceeds in the following steps. First, one defines the system itself, identifying a set of discrete entities, and the interactions among them. From these interactions emerges a network, whose nodes represent the elements of the system, and edges the interactions between these elements. The study of the structure of the resulting network is a central research direction, and has lead to fundamental results in the science of complex systems. Second, nodes and edges may be assigned a dimension or state, that is allowed to evolve over time, according to a mixture of stochastic and deterministic rules, almost always based upon the local topology of the network. This amounts to a flow of information over the network, which can be viewed as a substrate for the overlying node dynamics. We refer to the time evolution of such a phenomena as *dynamical*

*processes*, which thus amount to the second layer of a model built upon the network itself.

In this thesis we will be concerned with idealised models of dynamical processes that, despite their simplicity, provide meaningful insight into real systems. In particular, we concentrate on binary-state models of node dynamics, where the spread of information or disease is modelled by attributing to nodes one of two possible states. In so doing, we pay attention to emphasise the relationship between network structure and dynamics, since spreading processes can often be framed in terms of more fundamental, static systems. This is especially true when it comes to various forms of contagion, which denotes an important class of transmission models, and percolation, which refers to a network's connectivity.

In this introductory chapter, we first give an outline of the notions of network science that are useful in understanding the contributions of this thesis, to which subsequent chapters are dedicated. First, in Section 1.2, we discuss the real systems that have stimulated research in network science, in particular the study of dynamical processes and the spread of information. In Section 1.3, we discuss the main types of graphs studied in network science, and outline their representations. Then, in Section 1.4, we outline notions and results from random graph theory that arise recurrently throughout this thesis. In particular, we emphasise the structural properties of networks that most affect dynamical processes that are running on top of them. It will be seen that there is a close relationships between network structure and the evolution of such dynamics. We shall see that dynamical processes can sometimes be reduced to more fundamental, static problems that have been well studied from a network structure point of view. Then, in Section 1.5, we'll talk more about the different types of dynamical systems, including a simple but important class of dynamics known as binary-state processes, that will be the focus of this thesis. In particular, we will be interested in a class of model called "social contagion," which has stimulated the study of threshold models, to which we pay particular attention in Section 1.6. Finally, in Section 1.7, we introduce the computational tools that are essential to the results of this thesis, namely Monte Carlo simulation of networks and dynamics and Runge-Kutta solutions of master equation systems. Following this general introduction, in Section 1.8 we'll outline the structure of this thesis, and contextualise it within the existing literature.

## 1.2 COMPLEX SYSTEMS AS NETWORKS

It is useful to briefly outline what we mean by complex systems, networks and graphs. A complex system is seldom defined explicitly in terms of a network. A key aspect of complex systems is that they are composed of a large number of entities whose interactions are significant. Inherent in the network description of complex systems is the suppression of detail; often it is not the nature of an interaction that matters, it is the fact that an interaction is occurring that is important. From these interactions, we observe emergent phenomena such as self-organisation, synchronisation and pattern formation - behaviour that is difficult to extrapolate from the properties of individual constituents of the system. Based on these interactions, there are clearly many ways of defining a network. For example, the brain (see Section 1.2.4 for further discussion) is a classic complex system from which there are countless ways of defining network models, depending on the research question. One "microscopic" network model may be to represent neurons as nodes and axons as links. Another might be to have nodes represent entire regions, and links to represent correlations in their activity; the network itself is derivative of the complex system, and the network definition depends on the goal of the study itself.

In the following sections we outline a number of systems that are of interest to network scientists. For each, we give a brief outline of the common network formulations, which are often constrained by the nature of the available data and its collection method. Where applicable, we outline the relevant dynamical processes that occur on each network.

### 1.2.1 *Computational social science*

*Outline of system.* The emerging discipline of computational social science (CSS) is based on the study of social phenomena using large scale behavioural datasets [33, 134]. When such datasets entail the interaction between individuals, we enter the field of social network analysis. Nodes in a social network can be defined in a variety of ways, and can represent a variety of entities. Most often, nodes in a social network represent individuals, but can equally represent collectivities such as families, groups of friends, teams or communities. Most frequently, social network studies are on humans, but animal sociological studies are also common [61, 76]. The motivations behind the field of CSS are well outlined in a manifesto by Conte and coauthors [57], where "against a background of financial crises, riots and international epi-

*social network*

demics, the urgent need for a greater comprehension of the complexity of our interconnected global society is clear." Central to CSS, therefore, is not only the underlying networks, but the dynamical processes that they mediate.

Edges in a social network most commonly represent dyadic relationships between pairs of individuals. This can be generalised to so-called simplicial complex models of social networks [107], where one studies groupings of nodes and edges. We shall see that heterogeneity of edge type is fundamental property of social networks [168], and plays an important role in network formation, as well as overlying dynamical processes. In a social context, edge heterogeneities can be roughly categorised as familial, friendship, acquaintanceship, or in the weakest sense, recognition. Such heterogeneities are a dominant feature of real networks. The role of "weak" ties such as acquaintanceship is especially important in large networks, as was pointed out in the seminal work by Granovetter [96]. Whether it be in a large firm containing thousands of employees, or society as a whole, weakly interacting nodes serve to form bridges between distant parts of the network. The first clues of this effect can be traced back to Milgram's famous 1967 letter-forwarding study [141]. This experiment inferred the distribution of path lengths in an acquaintanceship network by asking participants to pass a letter to one of their acquaintances in order to reach an assigned target individual. Although most of the letters were simply lost, about one quarter reached the target individual, and on average, only required the help of six individuals in order to do so. This experiment lead to the concept of "six degrees of separation", the ramifications of which continue to be studied to this day (see Section 1.4). This idea was finally formalised in a rather beautiful one parameter model in [199], discussed in Section 1.4.3.

*six degrees of separation*

A vast array of social networks are of interest, and different social networks entail different research questions and methodologies, depending on the context. One may be interested in the psychology of the individuals in question, which usually entails smaller networks with mode detailed node and edge attributes. Here one may investigate networks with dozens of nodes, with the resulting network being relatively dense [98, 147]. This may be the case when studying the nature of intimate relationships, such as friendships and marriages. In Ref. [156] for example, authors study mother-daughter and father-son relationships via communication data, and draw conclusions relating to sex differences in the gender-bias of preferred relationships. In such a study, the global structure of the network can be largely ignored in favour of distinct, local neighbourhoods. Similarly, in Ref. [36], a

study of online dating data, the authors study the messaging strategies between individuals based on perceived attractiveness. Here as well, global network structure can be discarded. In such studies, emphasis is on human interest, rather than the implications of large-scale structure and collective phenomena. Historically, the study of small-scale social networks predate "big data," and labour-intensive surveys were conducted. Even before big data, crucial insights had been obtained from these methods, such as the observation of the small-world effect by Milgram.

The global structure of online social networks or of immense interest. A well-known 2011 study of Facebook data was carried out in [188], where the authors study the number of users and friendships, the degree distribution, path lengths, clustering, and mixing patterns. It is found that 99.91% of nodes belong to a single giant connected component. The authors confirmed the "six degrees of separation" phenomenon, showing a mean shortest-path distance of about 4.74, or an average of 3.74 intermediaries. It is perhaps unsurprising that this is lower than the value obtained in Milgram's letter-forwarding experiment, [1] given both the low cost of forming a link on the Facebook network, and the increased population size in the Facebook study. [2] Via analysis of community structure and the degeneracy of graph neighbourhoods, the authors of [188] showed that while the Facebook graph as a whole is clearly sparse, the graph neighbourhoods of users contain surprisingly dense structure. Characterising assortativity patterns by studying the basic demographic and network properties of users, clear degree assortativity. Similar structural analysis has been carried out on other social networks such as Twitter [127, 129], where in addition to studies of demographics, the problem of maximising influence [182] is of strong interest.

It is useful to note that different fields make drastically different use of social networks, where they can even go by different names and terminology. For example, the study of networks can go by the term actor-network theory in the sociological literature [133], where networks are introduced to provide a theoretical and methodological approach to social theory. In sharp contrast, in a setting such as public health, network studies are used to inform large scale intervention policies. Researchers might use a centrality analysis of a network of

1. Milgram's 1967 study gave a mean shortest-path distance of 6.4, or 5.4 intermediaries, using data from 44 complete sender-recipient paths.

2. In fact, a 2016 study of the Facebook network [27] show the mean shortest-path length to be as low as 4.57, requiring 3.57 intermediaries. The authors account for the decrease from the 2011 value by pointing out that the number of users had subsequently doubled.

drug addicts to detect good candidates for costly training in healthful practices, with the hope that these individuals would then diffuse the practices through the network. Similarly, understanding the spread of obesity [48], as well as the spread of smoking [49], can be achieved by examining social networks. In the latter study, discernible clusters of smokers and non-smokers were present in the network, and the clusters extended to three degrees of separation. Clusters of smokers remained the same across time, suggesting that whole groups of people were quitting in unison. The general relationship between social relationships and health is outlined in [55, 120].

*Dynamical processes.* Within the context of social networks, the dynamical processes studied focus on the spread of information and behaviour throughout the network. Recently, models have been introduced to understand the diffusion of memes, generically referred to as a shareable unit of information. Mobile phone call and communication network data [118, 130, 152] provide an ideal setting for studying the spread of information. This can be very hard to study directly, although network data itself is increasingly easy to come by, data pertaining to the spreading process demands an additional layer, and is therefore relatively sparse. Experiments have nevertheless been attempted, such as the famous work by Centola [45], where he studied spread of behaviour in an online social network experiment. The motivation is highly practical, and the results actionable. For many, the ultimate goal in such studies is to optimise the spread of influence through a social network, either by maximising desirable influece [122], or minimising negative influence.

*meme propogation*

### 1.2.2 *Computational economics and finance*

*Outline of system.* Economic and financial systems are build upon individuals and institutions such as banks and governments engaging in complex interactions and transactions [62]. It is immediately clear that the ensuing dependencies invite network representation. Clearly, there is an overlap here with the preceeding section, since economic and financial systems are inherently social [109, 111]. Behind every transaction is an individual, and markets of all types are intrinsically related to social networks of professionals. Behind trade, investment and borrowing markets are networks that determine the dynamics and stability of the resulting systems.

In studies of the structure of economic and financial networks [71, 110, 151], is typical to examine the formation and stochastic evolution of networks in view of the potential "payoff" from an economic or

social activity, with respect to network structure. Over time, individuals form and sever links connecting themselves to other individuals based on the improvement the resulting network offers them relative to the current network. Such a process creates a sequence of networks that are referred to as *improving paths*. Such approaches highlight the economy of networks themselves, theorising that the cost of forming and maintaining links is balanced by the benefits derived from such relationships. We shall see in Chapter 4 that broad classes of networks can be shown to result from entropy maximisation processes, driven by differences in the cost of maintaining various ties in the network. Typically, networks that form according to these principles drastically affect the behaviour of overlying dynamical processes.

*Dynamical processes on economic networks.* The bulk of studies regarding dynamics on financial networks concern risk and contagion [7, 73]. Systemic risk refers to the risk of default of a large portion of the financial system. This can be shown to depend on the network of financial exposures between connected institutions [25], typically banks. The key node-level quantities are asset size, leverage, and a financial connectivity measure given by the fraction of a financial institution's liabilities held by other financial institutions [84]. The goal is ultimately preventative, to use knowledge of network structure and criticality to inform financial policy.

*systemic risk*

### 1.2.3 *Technology and Infrastructure networks*

A major topic motivating the use of network analysis, are technological and infrastructure systems. From communication systems to public utilities, ideas from network science can be readily applied to areas of practical importance.

*Internet.* Studies of the internet often relate to its growth [94, 177], using models of preferential attachement [41] or the principles of statistical physics [158]. In [177] for example, the authors perform a study of growth mechanisms to understand the simultaneous emergence of fractality, modularity, as well as the small world effect and the scale-free property in real world complex networks. Studies of structure and topology including path-length distribution and diameter [4, 192, 205]. In the latter, the topology of the internet provides its connectivity, and consequently, the effectiveness with which users can locate information. In [159] the authors focus on the topological and dynamical properties of real Internet maps. Such studies have helped to build an accurate picture of the Internet's topology.

Studies of robustness have been a major focus of internet research. [67, 186]. In particular, in [54], a common property of many large networks, including the Internet, is that the connectivity of the various nodes follows a scale-free power-law distribution. The stability of such networks with respect to crashes are often examined, with respect to attack strategies such as random removal of sites. Such studies frequently draw upon percolation theory, aiming to derive general conditions for the critical fraction of nodes, $p_c$, that needs to be removed before the network disintegrates. A well known result is that for a class of scale-free networks, the transition never takes place, unless the network is finite. In the special case of the physical structure of the Internet, the network is impressively robust, with $p_c > 0.99$.

*Transport.* In this section we follow the well-known work by Barthelemy [21]. Here, airlines provide an important example of a spatial network, as they connect distant parts of real physical networks, in contrast to other transport networks. Nodes are identified as airports, located in two-dimensional space. A number of studies have shown that its structure displays both the small world and scale free property. In contrast, the degree distribution of metro networks within a range of cities is relatively narrow [131, 132], with the range of variation normally being of the order of one decade. Due to the relatively small range of variation in network size $N$ in most studies, the behaviour of the mean shortest path is not clear and could be fitted by a logarithm or a power law as well. Barthelemy notes that the mean shortest path is usually large compared to $\ln N$, suggesting that the behaviour of the mean shortest path might not be logarithmic with $N$ but more likely scales as $N^{1/2}$, a behaviour typical of a two-dimensional lattice. Rail also appears to be consistent with a two dimensional lattice, scaling as $\sqrt{N}$. An interesting study is [77] where the authors apply the concept of Dunbar's number to transport networks. They ask whether cognitive constraints affect the way we navigate cities.

*Road.* Still following [21], it is possible to construct various networks from the physical network of roads and streets, and as such measure spatial properties of a city by examining road networks. Edges represent roads, and nodes intersections. Except for roads that cross each other uses bridges, planar graphs are a useful approximation. This underlines that graph properties like degree distributions are constrained; it is unlikely for an intersection (node) to connect more than say 5 roads (edges). Clearly, such networks are closer to a lattice than a small world network. As such, studies generally focus on the total length $l$ of roads as a function of $N$ across a large number of cities. In contrast to

airline networks, the degree distribution of road networks thus defined are narrow.

*Utilities.* A key infrastructure network is power grids, which are clearly important when considering the impact of blackouts. In power grid networks, nodes represent power plants, as well as distributing and transmission substations, while edges correspond to transmission lines. Like road and rail networks, such networks are generally planar. Their degree distribution is usually exponential, in both Europe and the US.

*Dynamical processes.* As mentioned in the case of the Internet, the study of dynamical processes on infrastructure networks focus largely on stability to random and targeted attacks. This is particularly important from a security point of view, as the interconnectedness of communication and energy networks has been shown to increase their vulnerability [38]. Here, cascading failures have been shown to rapidly propagate throughout real networks, leading to large scale blackouts. In the case of the Internet, the spread of information is a major topic of study, with goals of optimisation of the spread of desirable information, and the suppression of undesirable content such as viruses. Finally, physical transport throughout networks, often modelled by random walkers, is a typical paradigm, drawing upon the tools of classical physics.

### 1.2.4 *Biological networks*

Biological networks refer to the large class of networks describing complex biological systems [5, 207]. This is best illustrated with the example of a cell, which can be viewed as a multiplex of at least three types of networks, describing protein-protein, protein-DNA, and protein-metabolite interactions. In other words, the nodes in such networks represent proteins, DNA and metabolites, whereas edges represent interactions. A general review outlining these dependencies can be found in [16]. Given the importance of these networks, there is increasing momentum within biological network communities to systematically catalogue all molecules and their interactions within a living cell. A type of connectome, as has been done for genomes, and partially for the mammal brain, is now needed and in the case of cell interaction networks. This arises from a need to understand how these molecules and the interactions between them determine the function of the resulting systems. Network perspectives attempt to rejuvenate traditional approaches to the study of biological systems.

The identification of biological networks gives rise to the study of their structure. For example, [81] studies the community structure in social and biological networks, including the community structure of macroscopic systems such as food webs. Studies of the topology [11, 148] often provide the foundation of future work, informing research into dynamical processes evolving on such networks. These studies again reveal the occurrence of scale-free structure [2]. Clearly, the understanding of properties that arise from whole-cell function require integrated, theoretical descriptions of the relationships between different cellular components.

*Brain.* Perhaps the richest example of a biological network is the brain [39]. Network studies have shown that the brain's structural [179] and functional systems have features of non-biological complex networks, such as small-world topology [24], and heavy tailed degree distributions [72], where highly connected hubs play an important role. This is true at the whole-brain scale of human neuroimaging, also known as functional network where one studies macroscopic areas of correlated activity. Here, entire regions of the brain are represented as nodes, and edges are drawn between areas with correlated electrical activity. This is also true at a cellular scale, where neuronal elements of the brain constitute a structural network. Importantly, the brain is a type of temporal network due to its demonstrated *plasticity*, a term used to broadly describe the dynamic reconfiguration during learning [23].

*Dynamical processes.* In biological networks, dynamical processes are often related to the network structure themselves, since links represent interactions during chemical and biological reactions. For example, an external signal in a cellular network can trigger a cascade of interactions that involves both biochemical reactions and transcriptional regulation, a process that orchestrates gene activity.

Regarding the brain, it is common to study the dynamic emergence of coherent physiological activity, such as phase-locked high-frequency electromagnetic oscillations, that can span the multiple spatially distinct brain regions that make up a functional network. Such networks are thought to provide the physiological basis for information processing and mental representations. Synchronisation represents just one example of the many approaches to dynamics on biological networks. Threshold dynamics, the focus of this thesis, have find one of their main applications in models of neuron firing.

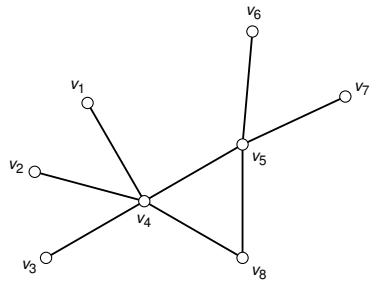


**Figure 1.1 –** An unweighted, undirected, single-layer graph. See Eqn. 1.2 for adjacency matrix and adjacency list representations, and Fig. 1.2 for the stub sequence representation arising from the configuration model.

## 1.3 GRAPH TYPES AND REPRESENTATION

A large number of random graph models have been introduced in order to understand systems observed in the real world. We highlight those aspects of random graph theory that are relevant to dynamics *on* networks. A particularly useful reference for understanding the relation between random graphs and dynamical processes is Durrett's [70]. A classical reference for random graphs in general, without regard to dynamical processes, is provided by Bollobás [32].

The simplest graph theoretical picture of a complex system is obtained by disregarding various edge and node qualities, and considering only the connectivity between nodes. This leads to the undirected, unweighted, single-layer representation in Fig. 1.1. In general, graphs are composed of a set of *nodes* $V$, typically indexed by $u$ or $v$, but also $i$ and $j$. In addition, edge sets are composed of edges $e_{uv} = (u, v)$ where $u, v \in V$ and $E = \{e_{uv} \mid u, v \in V\}$. A graph $G$ is defined as a tuple of these two quantities, such that $G = (V, E)$.

*Weighted networks.* This type of network arises when each edge $e_{uv}$ is attributed a fixed scalar value called a weight, which we denote $w_{uv}$, or $w(u, v)$. We focus on this type of network in Chapter 3. In this work, weights represent the strength of interaction between two nodes, or similarly, the cost of an interaction. Weights can represent arbitrarily different properties, however, depending on the context.

*Mulitiplex networks.* When a set of nodes $V$ is connected by more than one distinct set of edges, we have what is called a multiplex network. We introduce multiple edge sets $E_1, E_2, \ldots, E_M$, where $M$ is the total number of layers, in order to model multiplexity. In these networks, the nature of an edge is meaningful, not just in terms of weight, but

some other context dependent quality that is necessary to preserve. If such differences are not necessary to preserve, we take the single-layer projection, $E = \bigcup_{l=1}^{M} E_l$ and write $G = (V, E)$, recovering a possibly weighted single-layer network.

*Temporal networks.* As its name suggests, this type of network is introduced in order to model systems that evolve over time. This will be the focus of Chapter 5. In studies of dynamical processes on networks, it is important that the time scales of network evolution are commensurate with those of the dynamics. In the so-called quenched regime, network structure evolves so slowly relative to node dynamics that a static approximation is likely sufficient. In contrast, the annealed regime refers to networks that evolve so rapidly that their behaviour must be time-averaged in order to infer its interaction with node dynamics. At time scales between these two extremes, temporal network representations are likely to provide insight.

### 1.3.1 *Adjacency lists and matrices*

There are two standard representations of a graph $G = (V, E)$, either as a collection of adjacency lists or as an adjacency matrix [58], both of which are used extensively in this work. The adjacency-list representation is usually preferred, because it provides a compact way to represent sparse graphs, or those for which $|E|$ is much less than $|V|^2$. This is the case for the real systems considered throughout the remainder of this thesis. The graph algorithms presented in this book assume that an input graph is represented in adjacency-list form. An adjacency-matrix representation may be preferred, however, when the graph is dense, or when $|E|$ is close to $|V|^2$. In the following we give an outline of these data structures, as well as outline the use of edge-lists and event sequences, which are required in the study of temporal networks.

*adjacency matrix*

*Adjacency matrix.* This representation has a storage complexity of $\mathcal{O}(|V|^2)$. We denote the adjacency matrix by $A$. In a single layer unweighted network the term $a_{ij}$ indicates the presence of an edge between two nodes $i$ and $j$, such that

$$a_{ij} = \begin{cases} 1, & (i,j) \in E \\ 0, & \text{otherwise.} \end{cases} \tag{1.1}$$

As such, if two nodes $i$ and $j$ can be said to be connected in a simple graph, one writes $a_{ij} = a_{ji} = 1$, and 0 otherwise. Recording the connectivity of all pairs of nodes in this manner determines the adjacency

matrix $A$. When the network is undirected, a case that we do not generally consider in this thesis, we have $a_{ij} \neq a_{ji}$. As an illustration, the adjacency matrix $A$, and adjacency list $\tilde{A}$ of the undirected graph in Fig. 1.1 are given by

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \qquad \tilde{A} = \begin{cases} 1: & 4 \\ 2: & 4 \\ 3: & 4 \\ 4: & 1,\ 2,\ 3,\ 5,\ 8 \\ 5: & 4,\ 6,\ 7,\ 8 \\ 6: & 5 \\ 7: & 5 \\ 8: & 4,\ 5 \end{cases} \tag{1.2}$$

In the case of a multiplex networks, we introduce the superscript $l$ to index layer, and write $a_{ij}^{l}$. This definition is straightforward to generalise to weighted networks, where $a_{ij} = w_{ij}$ stores the weight of the edge connecting nodes $i$ and $j$. Throughout this work, we can safely ignore the case where nodes are connected to themselves, which we call "self edges," and in so doing set $a_{ii} = 0$.

The size of a network refers to the number of nodes in $V$, and is given by $N = |V|$. Throughout this work we refer to the degree $k_i$ of node $i$, and to the average degree $z = \sum_k k p_k$ across all nodes in the network. We refer also to the notion of *density*. This is the ratio of the number of edges present, $|E|$, to the total number of possible pairwise combinations, $\binom{|V|}{2}$, which gives

$$\frac{2|E|}{|V|(|V|-1)}. \tag{1.3}$$

In practice, $N$ is not only large, but growing, especially when modelling social systems. Since average degree $z$ in these settings remains relatively constant, even as the overall size of the system grows, the notion of density becomes important for graph representation. Consider that the connectivity of individuals in real systems such as social networks is determined by finite resources such as time and cognitive capacity, features that do not depend on the population size $N$. Clearly then, a node's connectivity does not depend upon the size of the overall graph in question, and the type of graph representation must be selected appropriately.

In light of the above considerations, it is clear that the adjacency matrix becomes sparse in the limit of large networks. This is especially

true in the case of the social sciences. Finite energy considerations means the same may be true in biological settings; when the number of entities increases, there is a linear increase in the number of connections, and a quadratic increase in the number of *possible* connections, meaning that the corresponding adjacency matrix is increasingly sparse. As such it's no longer efficient, and we favour alternative representations. In the following, we summarise the main graph representations used variously throughout this work. It is useful to talk about the complexity of various graph representations [58], which we do for the time take to store the same graph in the following ways.

*adjacency list*

*Adjacency list*. By far the most useful and common, solves the sparsity problem of adjacency matrices, and the lookup problem of edge lists. $O(|V| + |E|)$. The adjacency-list representation of a graph $G = (V, E)$ consists of an array $A$ of $|V|$ lists, one for each vertex in $V$. For each $u \in V$, the adjacency list $A[u]$ contains all the vertices $v$ such that there is an edge $(u, v) \in E$. That is, $A[u]$ consists of all the vertices adjacent to $u$ in $G$. The vertices in each adjacency list are typically stored in an arbitrary order. See Table 1.1 for a comparison of the complexity of various operations on adjacency list structures. Finally, note that there exist straightforward transformations between each of the representations, and they store the exact same information. All differences are practical.

*edge list*

*Edge list*. When network structure is not explicitly required, it is available implicitly by storing the set of edges $E$. This would be useful, for example, when all the weights in a weighted network had to be incremented by a certain amount. This would involve cycling through every edge in the network, meaning the network structure itself is irrelevant. Practically speaking, this comes in handy when initialising network simulations. However, this means that searching for a given edge is inefficient. Such an iteration is slightly simpler to code than on an adjacency list or matrix, and its complexity is $\mathcal{O}(|E|)$ rather than $\mathcal{O}(|E| + |V|)$ so can be slightly more efficient. Edge lists are also necessary for building more complex structures such as event sequences, which we discuss further below. The edge list for Fig. 1.1 is given by

$$E = \{(1,4), (2,4), (3,4), (4,5), (4,8), (5,6), (5,7), (5,8)\} \tag{1.4}$$

*half-edge, or stub*

A related structure is the list of so-called "stubs," or half edges. This is an edge where only one adjacent node index is specified, leaving the other available for pairing with other half edges. We use a half-edge or stub list in our implementation of the configuration model, shown in Fig. 1.2.

**Table 1.1 –** Time complexity of performing elementary operations on alternative graph data structures. Assumes graphs $G(V, E)$ to be either directed or undirected, and holds also for weighted graphs.

| | adjacency list | adjacency matrix |
|---|---|---|
| store graph | $\mathcal{O}(\|V\| + \|E\|)$ | $\mathcal{O}(\|V\|^2)$ |
| add vertex | $\mathcal{O}(1)$ | $^{\dagger}\mathcal{O}(\|V\|^2)$ |
| remove vertex | $\mathcal{O}(\|E\|)$ | $^{\dagger}\mathcal{O}(\|V\|^2)$ |
| add edge | $\mathcal{O}(1)$ | $\mathcal{O}(1)$ |
| remove edge | $\mathcal{O}(\|V\|)$ | $\mathcal{O}(1)$ |
| $^{\ddagger}$check adjacency | $\mathcal{O}(\|V\|)$ | $\mathcal{O}(1)$ |

† assumes matrix dimension fixed, and requires copying, otherwise $\mathcal{O}(|V|)$
‡ determine whether two nodes $u$ and $v$ are adjacent, assuming locations are known

## 1.4 RANDOM GRAPH MODELS

In this section, we outline some important generative models of random graphs. Each model produces ensembles that, in one way or another, help to understand the properties of the real systems outlined in Section 1.2. Much of the material in this section can be found in references [3, 70]. In each of the following models, we briefly describe emergent properties such as connected component sizes, average path length and clustering.

### 1.4.1 *Erdős-Rényi model*

*Model description.* In the late 1950's Erdős and Rényi introduced two random graph models [74]. In each there are $N$ vertices. In the first and less commonly used version, one picks $M$ of the $N(N-1)/2$ possible edges between these vertices at random. Investigation of the properties of this model tells us what a "typical" graph with $N$ vertices and $|E|$ edges looks like. However, there is a subtle dependence caused by selecting a fixed number of edges, namely that each edge can only be selected once, and must therefore be removed from the list of possible edges after selection. In any case, we will follow here the more common approach of studying the version in which each of the $N(N-1)/2$ possible edges between these vertices are independently present with probability[3] $p$. When $p = 2|E|/N(N-1)$, the second model is closely

3. Not to be confused with the rate of spontaneous adoption in a spreading process

related to the first. It is to this variant that we refer when we discuss the Erdös-Rényi (ER) random-graph model.

*Component sizes.* Erdős and Rényi discovered that there was a sharp threshold for the appearance of many properties. One of the first properties that was studied, and that will be the focus of much or our attention here, is the emergence of a giant component.

*giant component*

If $p = \frac{c}{N}$ and $c < 1$ then, when $N$ is large, most of the connected components of the graph are small, with the largest having only $\mathcal{O}(\log N)$ vertices.[4] In contrast if $c > 1$ there is a constant $\theta(c) > 0$ so that for large $N$ the largest component has on the order of $\theta(c)N$ vertices and the second largest component $\mathcal{O}(\log N)$. As such there is a critical value at $p = \frac{1}{N}$, where a second order phase transition takes place [180].

*phase transition*

As mentioned above, there is an interesting connection between the Erdős-Rényi model and epidemic processes. In a dynamical interpretation of the Erdős-Rényi model, there are two extremes. In the first all individuals are susceptible and there is a probability $p$ that an infected individual will transmit the infection to a neighbour, in the second only a fraction $p$ of individuals are susceptible, but the disease is so contagious that if an individual gets infected all of their susceptible neighbours will become infected.

In percolation terms, the first model is bond percolation, where edges are retained with probability $p$ and deleted with probability $1 - p$. The second is site percolation, where the randomness is applied to the sites instead of the edges. Percolation is easy to study on a random graph, since the result of retaining a fraction $p$ of the edges or sites is another random graph. Using a branching process heuristic, percolation occurs if and only if the mean of the associated branching process is greater than one, meaning a giant component is likely to occur.

*percolation*

In the $G_{N,p}$ variant of the model, the degree distribution is binomial,[5]

$$p_k = \binom{N-1}{k} p^k (1-p)^{N-k-1}. \tag{1.5}$$

4. If we haven't yet defined asymptotic notation, the $\mathcal{O}$ symbol means that there is a constant $C < \infty$ so that the probability that largest component is $\leq C \log N$ tends to 1 as $N \to \infty$.

5. This is straightforward to show using the identity

$$e^x = \lim_{n\to\infty} (1 + \tfrac{x}{n})^n$$

and that for constant $k$, we have

$$\lim_{n\to\infty} \tfrac{n!}{(n-k)!} = n^k.$$

It is easy to see that in this model the average degree is $\langle k \rangle = (N-1)p$. To see why this is the case, choose a node, iterate through the $N-1$ nodes in the network as potential neighbours, wire with probability $p$. In the limit $N \to \infty$ and $p = \frac{c}{N} > 0$ for some constant $c$, we get

$$p_k = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \tag{1.6}$$

*Poisson distribution*

Of course, this is nothing other than the Poisson distribution. As such, when referring to random graphs, we shall frequently refer to their degree distribution as Poissonian, even if in reality they are finite and generated according to Erdős-Rényi random graph models, which have binomially distributed degrees.

*path length and diameter*

*Average path length.* Path lengths and diameters of random graphs have been studied in [50]. In random graphs, authors use the convention that the diameter of a graph $G$ is the maximum diameter of its connected components. For most values of $p$, almost all graphs with the same $N$ and $p$ have precisely the same diameter. This means that when we consider all graphs with $N$ nodes and connection probability $p$, the range of values in which the diameters $d$ of these graphs can vary is very small, usually concentrated around

*small-world property of random graphs*

$$d = \frac{\ln N}{\ln pN}. \tag{1.7}$$

If $k = pN < 1$, a typical graph is composed of isolated trees and its diameter equals that of the tree's diameter. If $pN > 1$ a giant cluster appears. The diameter of the graph equals the diameter of the giant cluster if $k \geq 3.5$, and is proportional to $\ln N / \ln Np$. If $pN \geq \ln N$, almost every graph is totally connected. The diameters of the graphs having the same $N$ and $Np$ are concentrated on a few values around $\ln N / \ln Np$.

*clustering coefficient*

*Clustering coefficient.* This is straightforward to calculate in the case of a random graph generated according to the Erdős-Rényi model. [6] As such, the clustering coefficient $C$ is simply

$$C = p = \frac{\langle k \rangle}{N}. \tag{1.8}$$

6. Since we haven't introduced it before, the clustering coefficient can be calculated as

$$C = \frac{3 \times \text{number of triangles}}{\text{number of all triplets}},$$

and is a measure of transitivity. That is, if nodes $a$ and $b$ are connected to node $c$, then the clustering coefficient gives the probability with which $a$ and $b$ are also connected to each other.

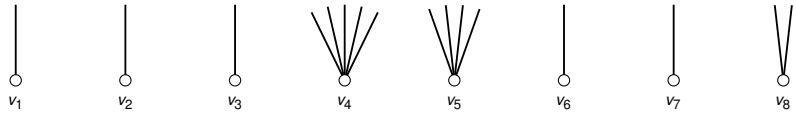


**Figure 1.2 –** A set of nodes with a desired degree sequence, enforced by the use of half-edges or "stubs." Such a configuration of nodes and stubs may lead to the graph in Fig. 1.1 upon randomly wiring pairs of stubs, providing a sample from a much larger ensemble of random graphs with the above degree sequence.

This result presents a problem if one intends to use Erdős-Rényi random graphs as a model of real systems. That is, real world networks have relatively large clustering coefficients. A large part of the motivation for the models of the following sections is to overcome these shortfalls.

### 1.4.2 *Configuration model*

In this section we turn our attention to a generalisation of the much-studied random graphs of Erdős and Rényi, such that the resultant networks are maximally random up to their degree distribution [145]. In contrast to the Erdős-Rényi model, where degree distributions were asymptotically Poisson, this model produces graphs with fixed, arbitrary degree distribution. Such graphs have also been termed "semi-random," to highlight their connection to random graphs as studied by Erdős, Rényi and others [3]. Indeed, in this section we aim to understand the same properties outlined above, namely the thresholds at which such graphs exhibit a giant component, and become fully connected. Similarly, we are interested in determining the clustering coefficient. We base this brief overview on references [3, 70, 145, 150]. The techniques therein exploit a generating function formalism, developed initially in [150]. We shall see that in general, the $G_{N,p}$ model of Erdős and Rényi is easier to analyse than configuration model graphs; each pair of nodes is independent in Erdős-Rényi models, which has only two parameters, $N$ and $p$. In contrast, the input into the configuration model is an entire *distribution*.

*k-regular random graph*

Note that Bollobás [31] proved results for the interesting special case of a random $k$-regular graph,[7] whereas Molloy and Reed [142, 143] were the first to construct graphs with a general distribution of degrees.

7. This is a graph where each node has degree $k$, but is maximally random otherwise

In an Erdős-Rényi random graph, vertices have degrees that have asymptotically a Poisson distribution. However, as discussed in Section 1.2, in social and communication networks, the distribution of degrees differs significantly from Poisson, and in many cases has a power law form, i.e., the fraction of vertices of degree $k$, frequently obeys $p_k \sim Ck^{-\beta}$ as $k \to \infty$.

*Model description.* In order to define the configuration model, it is useful to introduce the degree sequence,

*degree sequence*

$$(d_1, d_2, \ldots, d_N)\,, \tag{1.9}$$

where $d_i \in \mathbb{N}_0$ is the degree of the $i$-th node. The graph in Fig. 1.1, and the stub configuration in Fig. 1.2 have degree sequence $(1, 1, 1, 5, 4, 1, 1, 2)$. Further, we assume that $d_i$ are independent and have $P(d_i = k) = p_k$. Since we want $d_i$ to be the degree of vertex $i$, we condition on the sum of degrees, $\sum_i d_i$, being even. To construct the graph now we imagine $d_i$ half-edges or "stubs" attached to node $i$, and then pair the stubs at random. The resulting graph may have self-edges [8] and multi-edges [9] between points. The number of these self and multi-edges is is $\mathcal{O}(1)$, and as such can be safely ignored for our purposes, as we deal with large, sparse graphs. Nevertheless, one can condition on the event that there are no loops or multiple edges, if this is desired. Again, interest focuses first on the giant component, whose existence can be determined by considering the appropriate branching process. Importantly, the condition for its emergence is not simply that the mean degree satisfies $\sum_k kp_k > 1$, as was the case for random graphs.

*Component sizes.* The derivations of the distribution of component sizes in the configuration model can be found in [150]. When $p < 1$ in the Erdős-Rényi model, the largest component is $\mathcal{O}(\log N)$. This result does not hold for graphs with arbitrary degree distributions. There exist asymptotics of the cluster size distribution of configuration model graphs, because something is known about the coefficients of generating functions. Properties are functions of the input degree distributions. General. The authors focus on power law degree distributions with exponential tails. Theory matches numerical simulation perfectly.

As is common in the analysis of Erdős-Rényi random graph properties, the growth of clusters can be approximated by a branching process for low average degrees. It is possible to show that a giant component, or equivalently an infinite cluster, emerges almost surely using the above mentioned generating function formalism. We reproduce this here since it provides an important verification of results derived in

*percolation in configuration model*

8. Also known as loops or 1-cycle
9. Also known as 2-cycles

Chapter 2. The generating function $G_0(x)$ of the degree distribution [10] is given by

$$G_0(x) = \sum_{k=0}^{\infty} p_k x^k, \tag{1.10}$$

which contains all the information of the distribution $p_k$, since

$$p_k = \frac{1}{k!} \frac{d^k G_0}{dx^k}\bigg|_{x=0}. \tag{1.11}$$

The distribution of outgoing edges is given by

$$G_1(x) = \frac{\sum_k k p_k x^{k-1}}{\sum_k k p_k} = \frac{1}{\langle k \rangle} G_0'(x). \tag{1.12}$$

Note that the breadth-first search (BFS) algorithm is implicitly incorporated into the generating function method. Let $H_1(x)$ be the generating function for the distribution of the sizes of components that are reached by choosing a random edge and following it to one of its ends. This means that

$$H_1(x) = x q_0 + x q_1 H_1(x) + x q_2 \left[H_1(x)\right]^2 + \ldots \tag{1.13}$$

so that using Eqn. 1.12 we can write

$$H_1(x) = x G_1(H_1(x)). \tag{1.14}$$

The generating function for the size of the whole component is

$$H_0(x) = x G_0(H_0(x)). \tag{1.15}$$

In principle, therefore, given the functions $G_0(x)$ and $G_1(x)$, we can solve Eqn. 1.14 for $H_1(x)$ and substitute into Eqn. 1.15 to get $H_0(x)$. Then we can find the probability that a randomly chosen vertex belongs to a component of size $s$ by taking the $s$-th derivative of $H_0$. The mean component size is given by

$$\langle s \rangle = H_0'(1) = 1 + \frac{G_0'(1)}{1 - G_1'(1)} = 1 + \frac{z_1^2}{z_1 - z_2} \tag{1.16}$$

where $z_1 = \langle k \rangle = G_0'^{(1)}$ and $z_2 = \langle k^2 \rangle - \langle k \rangle = G_0'(1) G_1'(1)$. This expression diverges when $G_1'(1) = 1$, or $z_1 = z_2$, indicating the appearance of a giant cluster. It can also be written as

$$\sum_k k(k-2) p_k = 0, \tag{1.17}$$

which is the percolation transition for configuration model networks with degree distribution $p_k$, analogous to when $p = 1$ for Erdős-Rényi model graphs. This expression was also derived in [142].

*percolation in the configuration model*

*Average path length.* Extending the method of calculating the average number of nearest neighbours, we find the average number of $m$-th neighbours,

$$z_m = G_0'(1)\left[G_1'(1)\right]^{m-1} = z_1\left[\frac{z_2}{z_1}\right]^{m-1}, \tag{1.18}$$

where $z_1$ and $z_2$ are the numbers of nearest and next-nearest neighbours. Using this expression, we can derive an approximative relation for the average path length of the graph. Let us start from a given node and find the number of its nearest, next-nearest, ..., $m$-th neighbours. Assuming that all nodes in the graph can be reached within $l$ steps, we have

$$\sum_{m=1}^{l} n(m) = N - 1, \tag{1.19}$$

where $l$ is the number of steps from a randomly selected node in order to reach the remaining $N - 1$ nodes in the network. The term $n(m)$ the number of $m$-th order neighbours. To estimate the average path length, one replaces $n(m)$ with $z_m$. Further, we assume $N \gg z_1$ and $z_2 \gg z_1$, meaning the average path length is

$$\langle l \rangle = \frac{\ln(N/z_1)}{z_2/z_1} + 1, \tag{1.20}$$

meaning it grows logarithmically with $N$.

*Clustering coefficient.* The probability that two randomly selected nodes $i$ and $j$ are connected is $k_i k_j / 2|E|$. The probability that two nodes share a common neighbour and thus form a triangle is the above probability, multiplied by the probability that $i$ has excess degree $k_i$ and $j$ has excess degree $k_j$, for any value of $k_i$ and $k_j$. The clustering coefficient can be shown to be

$$C = \sum_{k_i=0}^{\infty}\sum_{k_j=0}^{\infty} q_{k_i} q_{k_j} \frac{k_i k_j}{2m} = \ldots = \frac{1}{N}\frac{\left[\langle k^2 \rangle - \langle k \rangle\right]^2}{\langle k \rangle^3}. \tag{1.21}$$

Since this expression does not depend on network size $N$, it is a constant of the degree distribution, or more precisely, a function of the first and second moments of the degree distribution. As such, it is a

10. As is convention, we use $x$ as the variable of the generating function, since $k$, which is typical notation outside the network literature, is reserved for degree.

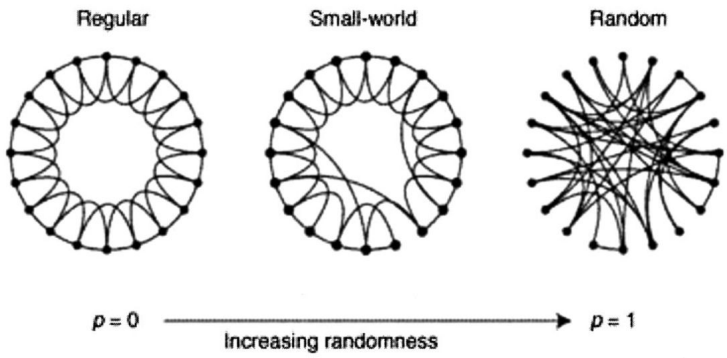


**Figure 1.3 –** Rewiring in Watts-Strogatz model, taken from [199].

### 1.4.3 *Small World model*

The next model, termed the Watts-Strogatz (WS) model, was inspired by the popular concept of "six degrees of separation," which is based on the notion that every one in the world is connected to everyone else through a chain of at most six mutual acquaintances. An Erdős-Rényi random graph for $N = 7 \times 10^9$ people in which each individual has an average of $\langle k \rangle = 43.74$ friends would have average pairwise distance $(\log N)/(\log\langle k \rangle) = 6$, but would lack the property of having substantial clustering coefficients [3]. It was noted that clustering coefficient varies very little as a function of network size $N$. One type of graph that shares this property is a regular lattice. In their seminal work [199], Watts and Strogatz used this idea by introducing a ring lattice [11] as a structure that necessarily contains a positive density of triangles, but also higher order cycles. To construct a network with small diameter and a positive density of triangles, the authors started from a ring lattice with $N$ vertices and $k$ edges per vertex, then rewired each edge with probability $p$, connecting one end to a vertex chosen at random. This construction interpolates between a regular ring lattice where $p = 0$, and complete disorder when $p = 1$.

*Average path length.* As discussed in [199], the origin of the rapid drop in $\langle l \rangle$ is the appearance of shortcuts between nodes. Every rewiring, created at random, is likely to connect widely separated parts of the graph, and thus has a significant impact on the characteristic path length of the entire graph. Even a relatively low fraction of random

11. A regular ring lattice is a graph with $N$ nodes each connected to $k$ neighbours, $k/2$ on each side.

rewirings is sufficient to drastically decrease the average path length, even if locally the network remains highly ordered.

*Clustering coefficient.* An alternative definition[12] of the clustering coefficient was provided by Barrat and Weigt in [17]. Typically, one defines $C$ as the fraction of edges among the $\binom{k}{2}$ neighbours of a node. In other words, this is the local density of edges in the subgraph formed by a node's neighbours. If instead we define $\tilde{C}$ as the ratio of the mean number of links between the neighbours of a vertex and the mean number of possible links between the neighbours of a vertex, we obtain

$$\tilde{C}(p) = \frac{3(k-1)}{2(2k-1)}(1-p)^3, \tag{1.22}$$

as $N \to \infty$. Here, the prefactor is $C(0) = |\mathcal{E}|/2k$, where $|\mathcal{E}| = 3k(k-2)/2$ and $\mathcal{E}$ is the set of edges between a nodes neighbours.

The success of the model is due to that face that for certain values of the rewiring probability $p$, the resulting network captures two crucial features of many real world networks. The first is of course the small-world property, which emerges for relatively small values of $p$, where we observe a large drop in average path length by the time $p = 10^{-2}$. The second is the large clustering coefficient, which remains large for a large interval of $p$ values for which the average path length is small. The fact that this simple, one parameter model is able to simultaneously capture these two effects was a milestone in the study of complex systems, in particular network science.

### 1.4.4 *Barabási-Albert model*

Although the Watts-Strogatz model accounted for the small world property, as well as the high clustering property of real networks, it fails to explain the broad degree distribution of real networks [6, 75, 113]. To this end, Barabási and Albert introduced a simple model [15] that produces such graphs. It belongs to a family of random graphs that work by *growing*. We refer to this model as the Barabási-Albert (BA) model.

*Model description.* The authors start with a graph with a small number of vertices $m_0$, that will act as seed to the resultant network. At every time step, one adds a new vertex with $m$ edges that link the new vertex to $m$ different vertices already present in the graph. The key idea is that nodes attach via preferential attachment, meaning that nodes are biased in their choice of neighbour when attaching to existing nodes

12. Equivalent to "triangle" definition above.

in the network. To this end, Barabási and Albert assume that the probability $\Pi$ that a new vertex will be connected to an existing vertex $i$ depends on the connectivity of that vertex. That is,

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j}. \tag{1.23}$$

Clearly, after $t$ steps, the model leads to a random network with $t + m_0$ vertices and $mt$ edges.

*Clustering coefficient.* State without proving. Derived in [124]

$$C = \frac{m}{8}\frac{(\ln N)^2}{N}, \tag{1.24}$$

where $N$ is the network size as always, and $m$ is the parameter of the Barabási-Albert model. Disregarding the logarithmic correction, the clustering coefficient is the same as that of the Erdős-Rényi model (see Eqn. 1.8), and goes to zero in the limit $N \to \infty$.

*Average path length.* State without proving. Derived in [53]

$$\langle l \rangle = \frac{\ln N}{\ln \ln N} \tag{1.25}$$

when $m \geq 2$, which results in a power law exponent of $\lambda = 3$.

### 1.4.5 *Other random graph models*

Another generative model of random graphs will be used in Chapter 4, called the forest-fire process (FFP). Provided a "seed" graph, the forest-fire process models network densification, or the tendency of a graph's average degree to grow over time.

## 1.5 DYNAMICAL PROCESSES ON NETWORKS

We have already seen that graphs themselves exhibit dynamic behaviour, via random growth and other processes. We have even seen that pure network dynamics can be construed as dynamical process on top of these networks. We generalise this latter concept by explicitly consider dynamical processes as separate from the network. Effectively, this adds another layer to the models of random graphs. See [163] by Porter and Gleeson For an excellent tutorial on dynamical process on networks, as well as [20] by Barrat *et al*. Unsurprisingly, across the many disciplines concerned, a large number of dynamics can be defined. Typical examples are the spread of disease, as well as information, behaviour and trends. The former we term biological or simple

contagion, and the latter social or complex contagion, and. These are examples of so called *binary-state* dynamics, [87], where nodes are restricted to one of two mutually exclusive states. Further examples include voter models [137, 178], language models [1, 193], majority vote [63, 137, 162].

Such processes will be the focus of this section, and indeed the remainder of this work. We outline other dynamical processes on networks that are nevertheless of great importance to science in general. Some examples include random walkers, oscillators, and general spin systems. More broadly still, one can study signals on networks, as well as control theory, however we do not even summarise these fields, despite their depth.

### 1.5.1 *Binary-state dynamics*

While the state of a node in real systems may be complex to fully characterise, it is often useful to abstract to a small, discrete number of states. The state of a nodes is determined by the state of its neighbour, and in our formalism, the state of the edge connecting the two nodes. The local neighbourhood of a node is entirely described by its class $(\mathbf{k}, \mathbf{m})$, and all such nodes become infected at rate $F_{\mathbf{k},\mathbf{m}}$, and recover at rate $R_{\mathbf{k},\mathbf{m}}$.

In the following sections, we detail the formulation of binary-state models that appear throughout various physical, biological and social settings, with a focus on complex contagion and threshold models. In Tables 1.2 and 1.3 we summarise the infection and recovery in each model, amounting to spin flip probabilities.

### 1.5.2 *Biological contagion*

One of the most mature areas of study of dynamical processes on networks relates to epidemiology, or the study of the distribution and determinants of states of health in population. For much of its history, the approach was to assume that the population was fully mixed; this is known as the homogeneous mixing approximation, which assumes that the individuals in the population interact completely at random [8], and that a chosen node is able to make contact with the remaining $N-1$ nodes in the network. A great review is [160]. Epidemic models generally assume that the population can be divided into different classes or "compartments" depending on the stage of the disease [8, 64, 65, 121]. The most important of these compartements are known as the

**Table 1.2 –** Transition rates in various definitions of binary-state dynamics. Nodes with local configuration given by $(\mathbf{k}, \mathbf{m})$ become infected at rate $F_{\mathbf{k},\mathbf{m}}$, and recover with rate $R_{\mathbf{k},\mathbf{m}}$. This table is a generalisation of Table I in [87], allowing for edge state heterogeneity. Generically, the vector $\mathbf{w}$ of continuous variables $w_{ij} \geq 0$ determines the interaction strength, or weight. When setting $w_j = 1\ \forall j$, we recover the edge-homogenous dynamics outlined in [87]

| model | infection rate $F_{\mathbf{k},\mathbf{m}}$ | recovery rate $R_{\mathbf{k},\mathbf{m}}$ |
|---|---|---|
| SI | $\boldsymbol{\lambda} \cdot \mathbf{m}$ | $0$ |
| SIS | $\boldsymbol{\lambda} \cdot \mathbf{m}$ | $\mu$ |
| Bass | $c + \mathbf{d} \cdot \mathbf{m}$ | $0$ |
| Kirman | $c_1 + \mathbf{d} \cdot \mathbf{m}$ | $c_2 + \mathbf{d} \cdot (\mathbf{k} - \mathbf{m})$ |
| Voter | $\dfrac{\mathbf{m} \cdot \mathbf{w}}{\mathbf{k} \cdot \mathbf{w}}$ | $\dfrac{(\mathbf{k} - \mathbf{m}) \cdot \mathbf{w}}{\mathbf{k} \cdot \mathbf{w}}$ |
| Link-update voter | $\dfrac{\mathbf{m} \cdot \mathbf{w}}{\mathbf{z} \cdot \mathbf{w}}$ | $\dfrac{(\mathbf{k} - \mathbf{m}) \cdot \mathbf{w}}{\mathbf{z} \cdot \mathbf{w}}$ |
| Language | $s \left( \dfrac{\mathbf{m} \cdot \mathbf{w}}{\mathbf{k} \cdot \mathbf{w}} \right)^{\alpha}$ | $(1-s) \left( \dfrac{\mathbf{k} \cdot \mathbf{w} - \mathbf{m} \cdot \mathbf{w}}{\mathbf{k} \cdot \mathbf{w}} \right)^{\alpha}$ |
| Majority vote | $\begin{cases} Q, & \mathbf{m} \cdot \mathbf{w} < \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \\ \frac{1}{2}, & \mathbf{m} \cdot \mathbf{w} < \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \\ 1 - Q, & \mathbf{m} \cdot \mathbf{w} > \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \end{cases}$ | $\begin{cases} 1 - Q, & \mathbf{m} \cdot \mathbf{w} < \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \\ \frac{1}{2}, & \mathbf{m} \cdot \mathbf{w} < \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \\ Q, & \mathbf{m} \cdot \mathbf{w} > \frac{1}{2} \mathbf{k} \cdot \mathbf{w} \end{cases}$ |

**Table 1.3 –** See caption in Table 1.2. Also, note that continuous edge weight distributions are accounted for in Chapter 4, and that the vector notation here (**k** and **m** vectors) indicate that edge types are discrete.

| model | infection rate $F_{\mathbf{k},\mathbf{m}}$ | recovery rate $R_{\mathbf{k},\mathbf{m}}$ |
|---|---|---|
| Non-linear voter | $\begin{cases} F_{4,0} = 0 \\ F_{4,1} = \alpha_1 \\ F_{4,2} = \alpha_2 \\ F_{4,3} = 1 - \alpha_2 \\ F_{4,4} = 1 - \alpha_1 \end{cases}$ | $\begin{cases} F_{4,0} = 1 - \alpha_1 \\ F_{4,1} = 1 - \alpha_2 \\ F_{4,2} = \alpha_2 \\ F_{4,3} = \alpha_1 \\ F_{4,4} = 0 \end{cases}$ |
| Ising-Glauber | $\dfrac{1}{1+\exp\left[\frac{2J}{T}(\mathbf{k}\cdot\mathbf{w} - 2\mathbf{m}\cdot\mathbf{w})\right]}$ | $\dfrac{\exp\left[\frac{2J}{T}(\mathbf{k}\cdot\mathbf{w} - 2\mathbf{m}\cdot\mathbf{w})\right]}{1+\exp\left[\frac{2J}{T}(\mathbf{k}\cdot\mathbf{w} - 2\mathbf{m}\cdot\mathbf{w})\right]}$ |
| Ising-Metropolis | $\begin{cases} e^{\frac{2J}{T}(\mathbf{k}\cdot\mathbf{w} - 2\mathbf{m}\cdot\mathbf{w})}, & \mathbf{m}\cdot\mathbf{w} < \frac{1}{2}\mathbf{k}\cdot\mathbf{w} \\ 1, & \mathbf{m}\cdot\mathbf{w} \geq \frac{1}{2}\mathbf{k}\cdot\mathbf{w} \end{cases}$ | $\begin{cases} 1, & \mathbf{m}\cdot\mathbf{w} \leq \frac{1}{2}\mathbf{k}\cdot\mathbf{w} \\ e^{\frac{2J}{T}(\mathbf{k}\cdot\mathbf{w} - 2\mathbf{m}\cdot\mathbf{w})}, & \mathbf{m}\cdot\mathbf{w} > \frac{1}{2}\mathbf{k}\cdot\mathbf{w} \end{cases}$ |

**Table 1.4 –** See caption in Table 1.2. Also, note that continuous edge weight distributions are accounted for in Chapter 4, and that the vector notation here (**k** and **m** vectors) indicate that edge types are discrete.

| model | infection rate $F_{\mathbf{k},\mathbf{m}}$ | recovery rate $R_{\mathbf{k},\mathbf{m}}$ |
|---|---|---|
| Watts threshold | $\begin{cases} 1, & \text{if } \mathbf{m}\cdot\mathbf{w} \geq \phi\mathbf{k}\cdot\mathbf{w} \text{ and } \mathbf{k}\cdot\mathbf{w} > 0 \\ 0, & \text{otherwise} \end{cases}$ | 0 |
| †Watts, *and* | $\begin{cases} 1, & \text{if } m_i w_i \geq \phi_i k_i w_i \text{ and } k_i > 0, \ \forall i \\ 0, & \text{otherwise} \end{cases}$ | 0 |
| †Watts, *or* | $\begin{cases} 1, & \text{if } \exists i \text{ s.t. } m_i w_i \geq \phi_i k_i w_i \text{ and } k_i > 0 \\ 0 & \text{otherwise} \end{cases}$ | 0 |
| ‡Absolute threshold | $\begin{cases} 1, & \text{if } \mathbf{m}\cdot\mathbf{w} \geq M_\phi \text{ and } \mathbf{k}\cdot\mathbf{w} > 0 \\ 0, & \text{otherwise} \end{cases}$ | 0 |

† Note that $i$ is layer index, and $j$ is edge type index. As outlined in Chapter 4, the layer index runs $1 \leq i \leq M$, whereas $1 \leq j \leq 2^M - 1$
‡ This is also referred to as the Centola-Macy threshold rule

the susceptible (denoted by S, those who can contract the infection), infectious (I, those who contracted the infection and are contagious) classes. This allows one to define the susceptible-infected (SI) model, where infected nodes remain infected for the remainder of the process, and the susceptible-infected-susceptible (SIS) model, where nodes are able to recover and rejoin the susceptible class. Clearly, the goal in modelling is to understand and make sense of data, and eventually to predict and control cascades of infection. To this end, Colizza *et al* have developed the global epidemic and mobility model [56].

We allow for a discrete number of infection rates, determining the dimension of the degree and partial degree vectors $\mathbf{k}$ and $\mathbf{m}$, where $\lambda_j dt$ is the probability of infection over an interval $[t, t+dt]$. We store these rates in the vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^T$. Since $m_j$ gives the number of adjacent infected nodes via edges of type $j$, the infection rate is given by

$$\begin{aligned} F_{\mathbf{k},\mathbf{m}} dt &= \lim_{dt \to 0} \sum_j \left[ 1 - \left(1 - \lambda_j dt\right)^{m_j} \right] \\ &= \sum_j \lambda_j m_j dt \\ &= \boldsymbol{\lambda} \cdot \mathbf{m} dt. \end{aligned} \tag{1.26}$$

In the case of the SI process, where dynamics are monotonic, we have $R_{\mathbf{k},\mathbf{m}} = 0$, and in the case of SIS, $R_{\mathbf{k},\mathbf{m}} = \mu$. Having two compartments of course corresponds to what we call binary-state dynamics. Although it is clearly interesting to make more general classes, which many authors have done. Common additional compartments are recovered (R, those who recovered from the disease). Additional compartments can be used to signal other possible states of individuals with respect to the disease, for instance immune individuals. susceptible-infected-recovered (SIR) model. Only the first two are binary state. Another is the SEIR model, a variation of the SIR model including the effects of exposed (E) individuals, which have been infected by the disease but cannot yet transmit it.

### 1.5.3 *Social contagion*

In contrast to biological contagion, where contact with a single infected neighbour is sufficient for a cascade of infection to develop, social contagion refers to the spread of behaviours, ideas and trends where reinforcement plays a role. Examples of social dynamics where repeated exposure to a behaviours is important are fads [28, 82], riots

[97], crime [83], competing technologies [9], and the spread of innovations [10, 191], conventions [171], and cooperation [105]. Common to all of these behaviours is that regardless of the nature of the social epidemic, an individual's binary decision is determined entirely by the state of the local neighbourhood, in particular, the relative number of neighbours who are observed to be in the adopting state, compared to the total number of neighbours [171].

As pointed out in [198], the economic literature refers to this entire class of problems generically as binary decisions with externalities. The formulation of such problems in economic [28, 138, 175] and technological networks [126, 167] can be traced back to the seminal work by Schelling [171]. The idea of network being "robust yet fragile" is discussed in [42], where the authors suggest that power laws in real systems are due to tradeoffs between yield, cost of resources, and tolerance to risks. These tradeoffs lead to highly optimized designs that allow for occasional large events. These events are what we refer to as global cascades or avalanches.

To take an extreme example, the creation of a political coalition or an international treaty is unquestionably a complex, multifaceted process with many potential outcomes. But once the coalition exists or the treaty has been drafted, the decision of whether or not to join is essentially a binary one. Similar reasoning applies to a firm's choice between two technologies, or an individual's choice between two neighbourhood restaurants - the factors involved in the decision may be many, but the decision itself can be considered as binary.

### 1.5.4 *Opinion and voter models*

Another binary-state model is the voter model. In the basic voter model nodes can have two opinions, analogous to the susceptible or infected state in epidemic processes. At each time step (to see what this means see Section 1.7 on computational methods), a node is selected uniformly at random, and chooses a randomly selected neighbour whose opinion the node adopts. For infinitesimal $dt$, the transition rates then becomes

$$F_{\mathbf{k},\mathbf{m}} = \frac{\mathbf{m}\cdot\mathbf{w}}{\mathbf{k}\cdot\mathbf{w}} \qquad \text{and} \qquad R_{\mathbf{k},\mathbf{m}} = \frac{(\mathbf{k}-\mathbf{m})\cdot\mathbf{w}}{\mathbf{k}\cdot\mathbf{w}}. \tag{1.27}$$

In typical applications, edges are of homogeneous interaction strength, or $w_j = 1\ \forall j$, meaning $F_{\mathbf{k},\mathbf{m}} = \frac{m}{k}$ and $R_{\mathbf{k},\mathbf{m}} = 1 - \frac{m}{k}$. The basic voter model can be mapped to coalescing random walkers [52, 102]. An

excellent review of social physics is provided by Ref. [43], where voter models are discussed in detail.

## 1.6 THRESHOLD MODELS

In threshold driven processes, the state of an entity changes when the concentration of incoming stimuli or cumulating force reaches a certain threshold. Some typical examples are neural systems [125, 181], earthquakes [100], and solar flares [30, 47], commonly identified as self-organised critical systems driven by integrate-and-fire mechanisms. Thresholds play a role in some epidemic diseases, such as tuberculosis and dysentery [115], where infection requires the concentration of pathogens in an individual to overcome a threshold. Moreover, thresholds are associated with social contagion phenomena, where social influence from acquaintances may change the behaviour of an individual after reaching a cognitive limit. Studies of so-called *complex contagion* date back to Schelling, Axelrod, and Granovetter, but have recently attracted interest due to the seminal cascade model formulated by Watts [198], and also to the enormous amount of digital data on human behaviour collected to observe, analyse and model social contagion.

In this section we discuss the principle models of threshold dynamics, with a strong focus on those modelling the transmission of information and behaviour. We discuss the original work of Granovetter and Schelling to provide historical context to the Watts model. It is this model that we build upon in the following chapters, and as such, we carefully outline its relevant properties here.

### 1.6.1 *Granovetter*

A pioneering model of collective behaviour using threshold models was introduced by Granovetter in [97]. Binary decisions are defined are those where an actor has two distinct and mutually exclusive behavioural alternatives. Granovetter adapted the idea of behavioural thresholds from Schelling's models of residential segregation, where thresholds determines an actor's decision to leave their neighbourhood, as a function of how many of the actor's own colour also do so. In particular, see the second footnote of [97] for a historical overview of the threshold models of behavioural contagion. Namely, Granovetter builds upon the work of [169, 170, 172]. Some of the systems modelled by threshold processes are; the diffusion of innovations; the propaga-

tion of rumours and diseases; the decision of workers to strike; voting behaviour; educational attainment; the decision to leave social events; and the decision to migration. Granovetter discussed the notion of equilibrium state, or number of people participating in a behaviour, based on the initial distribution of thresholds. He assumed the population to be small and fully mixed; now social networks are allowed to be much larger.

### 1.6.2 *Watts*

Importantly, the model proposed by Granovetter is fully mixed, meaning that any individual can be assumed to interact with every other individual in the system. Such interactions can be represented as a graph, even if its network structure is trivial - it is simply a complete graph. In contrast, real networks are sparse. The Watts model [198] generalised the work of Granovetter, removing the fully mixed assumption. Essentially, this amounts to a networked version of the Granovetter model.

In the Watts model, a network of size $N$ is generated according to the configuration model, with degree distribution $p_k$. Typical experiments for this model vary a control parameter $z$, the average degree of the network. A node's threshold $\phi$ is drawn randomly from a distribution $p_\phi$. At time $t = 0$, a randomly selected node is set to the active state. The seed is therefore infinitesimal; regardless of the size of the network, a single node is chosen as the seed. A remarkable result is that even for a single seed, a positive fraction of an infinite network can be reached by subsequent avalanches of infection. The dynamics of the system are implemented in [198] via an asynchronous update rule, approximating continuous time in the limit of large networks (see Section 1.7 for more information regarding implementation). In our notation, the threshold rule is given by

$$F_{\mathbf{k},\mathbf{m}} = \begin{cases} 1, & m \geq \phi k \text{ and } k > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{1.28}$$

where $k = \sum_j k_j$ and $m = \sum_j m_j$. This is equivalent to saying that edge weights are uniform in Table 1.4, or $k = \mathbf{k} \cdot \mathbf{w}$ and $m = \mathbf{m} \cdot \mathbf{w}$, with $\mathbf{w} = (1, \ldots, 1)^T$. We also assume $k > 0$ for this rule, and that degree $k = 0$ nodes cannot become infected as they receive no influence. In other models (see Section 1.6 for further generalisations), isolated nodes may become infected spontaneously via random noise. Finally, note

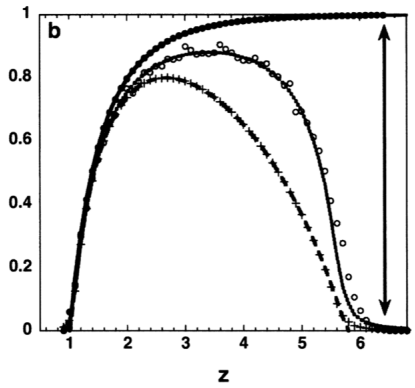


**Figure 1.4 –** Dynamics of the global cascades in the Watts model, as a function of average degree $z$, and uniform threshold $\phi = 0.18$. Filled black circles give the size of global cascades resulting from a single initial seed in an $N = 10^4$ node network. Until the discontinuous transition at $z \simeq 7$, this coincides almost exactly with the underlying black line, the size of the largest connected component. The frequency of global cascades after $10^3$ perturbations is given be unfilled circles. Although these are relatively noisy, they are well approximated by the underlying densely dotted line, which is the size of the extended vulnerable cluster. Finally, the innermost curve gives the size of the vulnerable cluster itself, which is well predicted by theory. Taken from [198].

that in the case of an undirected, unweighted network, we can write $F_{\mathbf{k},\mathbf{m}} = F_{k,m}$.

Watts' famous experiment was to explore a $(\phi, z)$ phase space, varying the average threshold and the average degree. As has remained common in such studies, thresholds are set to a uniform value for most of the study, and the effect of normally distributed thresholds was verified as a generalisation. The effect of a increased variance in the threshold distribution was to broaden the region in $(\phi, z)$ space that allowed global cascades. In contrast, the effect of a power-law degree distribution was to diminish the range of parameters leading to global cascades, relative to a binomially distributed degree $k$.

We briefly outline a number of concepts that are crucial to the results of this thesis. The first is the idea of a node being *vulnerable*, meaning that a single active neighbour is sufficient to induce activation in the node in question. Inactive nodes that are not vulnerable are either stable, or isolated, meaning they have no neighbours and $k = 0$. If the initial seed in the Watts model is not adjacent to a vulnerable node, then no transmission takes place. A set of connected vulnerable nodes is referred to as a vulnerable cluster. By definition, if a single node

*vulnerable cluster*

within or adjacent to a set of vulnerable nodes is active, then it will eventually topple all the nodes in that set. If such a cluster percolates throughout the entire network, it is said to satisfy the condition for the emergence of global cascades. Necessarily, the vulnerable cluster must reside within the giant component of the network. Then, any randomly selected seed within the cluster, or adjacent to it, termed the extended vulnerable cluster, will trigger a global cascade. The frequency with which a single randomly selected node triggers a global cascade is given by the relative size of the extended vulnerable cluster.

*global cascade*

*extended vulnerable cluster*

These concepts are explained in Fig. 1.4, which is extracted from [198]. As expected, the frequency with which global cascades emerge as a result of a single initial seed is given by the relative size of the extended vulnerable cluster. Once such an avalanche is triggered, it generally extends beyond the vulnerable cluster, to reach every node in the giant component. This is reflected in the fact that global cascade size is in precise agreement with the size of the giant component in Fig. 1.4. As expected, the continuous phase transition in cascade size occurs around $z = 1$ since the degree distribution is binomial. The discontinuous transition is marked by the vertical arrow and occurs around $z = 7$. This transition is due to the increased stability of nodes at higher $z$ values, meaning the vulnerable cluster becomes finite.

### 1.6.3 *Variations*

In the original formulation by Watts [198], all neighbours we qualitatively identical, distinguishable only in their state. We generalise this by allowing a discrete number of edge types, so that a node's local neighbourhood is characterised in general by $(\mathbf{k}, \mathbf{m})$, in contrast to simply $(k, m)$. A node's total influence is given by $\mathbf{k} \cdot \mathbf{w}$, and influence due to active neighbours $\mathbf{m} \cdot \mathbf{w}$. This is shown as the Watts threshold rule in Table 1.4, and recovers the homogeneous Eqn. 1.28 when setting $n = 1$.

Further, it is possible to extend all the threshold rules in Table 1.4 by allowing a rate of spontaneous activation $p$, which amounts to random noise. The activation rule then becomes

*spontaneous activation*

$$F_{\mathbf{k},\mathbf{m}} = \begin{cases} 1, & \mathbf{m} \cdot \mathbf{w} \geq \phi \mathbf{k} \cdot \mathbf{w} \text{ and } \mathbf{k} \cdot \mathbf{w} > 0 \\ p, & \text{otherwise.} \end{cases} \tag{1.29}$$

*blocked nodes*

We are also able to introduce blocked nodes to model individuals who refuse to adopt, or equivalently, nodes whose threshold exceeds $\phi > 1$, as was studied in [166]. Further generalisations will be explored espe-

cially in Chapter 4, where the boolean *and* and *or* rules are introduced in a multiplex setting.

An important alternative to the Watts threshold rule is that of Centola and Macy. Whereas the former rule measures relative influence, the latter measures absolute influence. The absolute influence required for activation is given by $M_\phi$, and the rule becomes

*absolute threshold*

$$F_{\mathbf{k},\mathbf{m}} = \begin{cases} 1, & \mathbf{m}\cdot\mathbf{w} \geq M_\phi \text{ and } \mathbf{k}\cdot\mathbf{w} > 0 \\ p, & \text{otherwise.} \end{cases} \tag{1.30}$$

As a result it may be better suited to fitting data, as one does not require a complete mapping of a node's ego-centric network. Incomplete data, at a certain point, doesn't diminish the model. If a node's threshold is $M_\phi = 3$, and it has $m = 3$ infected neighbours, it doesn't matter if the node's total degree is $k = 3$, 10, 100 or 1000, the threshold is still met and the relevant node becomes active.

## 1.7 COMPUTATIONAL METHODS

In this section we provide a brief overview of the technical considerations that are recurrent throughout this thesis. The methodology that we follow relies heavily on numerical simulation of the physical processes in question, and at the same time, numerical integration of master equation systems derived to model such processes. For completeness, we reproduce the fourth-order Runge-Kutta (RK) system that we use throughout this thesis, as well as the Monte Carlo (MC) simulation techniques that are most of use. Since simulations are expansive, both in terms of the size of the networks studied, the statistics that we seek to obtain requiring huge numbers of realisations, and the fact that we study fine-grained parameter spaces, we give a brief sketch of our approach to the distribution of such calculations.

### 1.7.1 *Runge-Kutta solution*

The analytical solution of the systems under examination takes the form of a master equation. This refers to a set of first-order differential equations describing the time evolution of a discrete number of states, which in our context refers to a density with respect to network-node configurations. If the column vector $\mathbf{s}(t)$ contains the time-dependent densities $s_i(t)$, then the matrix $A$ stores the transition rates $A_{ij}$ between states $i$ and $j$. In general, these transition rates may depend on time,

both explicitly [13], and implicitly via a dependence on $\mathbf{s}(t)$. As such, we write either $A = A(t, \mathbf{s})$, or in the case of no explicit time dependence, $A = A(\mathbf{s})$. For a complete outline of the master-equation solution and its derivation, see Chapter 2.

In light of these considerations, the numerical implementation of our analytic solution amounts to solving the first-order ordinary differential equation (ODE),

$$\frac{d}{dt}\mathbf{s} = A(\mathbf{s})\mathbf{s}, \tag{1.31}$$

*configuration space*

where the dimension of $\mathbf{s}$ and $A$ is the dimension of the so-called configuration space. Throughout this work we use a fourth-order Runge-Kutta solution, which for completeness we outline here. We denote by $t_n$ the $n$-th discrete time step, where step size is $\Delta t$, such that $t_n = n\Delta t$, always setting $t_0 = 0$. Then, $\mathbf{s}_n$ is the value of the state vector at time $t_n$. After determining the initial conditions $\mathbf{s}_0$, the fourth-order Runge-Kutta system involves the sum

$$\mathbf{s}_{n+1} = \mathbf{s}_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4) \tag{1.32}$$

$$t_{n+1} = t_n + \Delta t, \tag{1.33}$$

at every time step, or $n = 0, 1, 2, \ldots$, until the system reaches some steady state. If we define $f(t, \mathbf{s}) = A(t, \mathbf{s})\mathbf{s}$, then the coefficients $l$ in Eqn. 1.32 are given by

$$l_1 = \Delta t f\left(t_n, \mathbf{s}_n\right) \tag{1.34a}$$

$$l_2 = \Delta t f\left(t_n + \frac{\Delta t}{2}, \mathbf{s}_n + \frac{l_1}{2}\right) \tag{1.34b}$$

$$l_3 = \Delta t f\left(t_n + \frac{\Delta t}{2}, \mathbf{s}_n + \frac{l_2}{2}\right) \tag{1.34c}$$

$$l_4 = \Delta t f\left(t_n + \Delta t, \mathbf{s}_n + l_3\right). \tag{1.34d}$$

We use this solution for the numerical integration for all the systems in this thesis. In Chapter 2 we compare the configuration space for the full and reduced AMEs, that is, solving for $\mathbf{s}$ and $\boldsymbol{\nu}$. The reduced AMEs has a complete graph, because the $\dot{\nu}_i$ depends on all $\nu_j$ such that $i \neq j$.

### 1.7.2 *Asynchronous vs synchronous update*

The key to the Monte Carlo implementation of dynamical processes on networks, in this setting, is the manner of the execution of the

13. We shall see that this depends on whether the system is *full* or *reduced*. The reduced AMEs allowing spontaneous adoption has an explicit time dependence

update rules that we've seen in Tables 1.2, 1.3 and 1.4. Briefly, it is useful to distinguish between two closely related operations when discussing Monte Carlo simulation. First, we refer to *node selection*, which is the process of choosing a node $v$, either randomly or via some pre-determined sequence. Secondly, we refer to *node update*. This involves determining the state of the selected node $v$ at the next time step. One does this by evaluating the state of $v$'s immediate neighbourhood [14], and applying the rules of Tables 1.2, 1.3 and 1.4. Crucially, a node update does not necessarily lead to a change in a node's state - the result of an update could very well be to remain in its current state, depending on the outcome of the adoption or recovery rules.

Having distinguished between node selection and node update, we now outline two variations of the implementation of these operations. The first involves a node $v$ being selected uniformly at random, and its updated state being immediately visible to its neighbours. This is called the *asynchronous* update rule, since the updating of different nodes is independent. In contrast, one may select nodes sequentially, and the state of a node at time $t_{n+1}$ depends on the state of its neighbours at time $t_n$. In other words, when determining the state of a node $v$ at time $t_{n+1}$, it considers the state of a neighbour $u$ at time $t_n$, even if $u$ was selected and updated before $v$ at the $t_{n+1}$-th time step. This is termed the *synchronous* update rule, and implies a system that advances discretely in time [15]. In contrast, the asynchronous update rule corresponds naturally to a differential equation solution. This is because if time is incremented at every node update, then the time step is normalised to system size. That is, for a network of size $N$, we have

$$\Delta t = \frac{1}{N}, \tag{1.35}$$

meaning that time is quasi-continuous in the large network limit, with $\Delta t \to 0$ as $N \to \infty$. Since we wish to model the systems of Section 1.2, the asynchronous update rule is clearly better adapted.

Since nodes are selected for update uniformly at random following the asynchronous update rule, it is clear that the time spent between consecutive node updates is exponentially distributed. This corresponds to a simple survival process. Since the probability of a node

14. In practice we do not actually do this, since it's wasteful. We prefer to update the state-change probability of all of a node's neighbours after it changes state. That means all that we do after selecting a node is flip a coin and compare to $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$.

15. Although both synchronous and asynchronous systems involve a discrete time step, a synchronous rule implies that time is course grained, typically with $\Delta t = 1$, whereas an asynchronous rule has a time step diminishing with system size, $\Delta t = 1/N$.

being randomly selected at time $t$ is $\Delta t = 1/N$, the probability of becoming infected or recovering is $F\Delta t$ or $R\Delta t$, respectively. Assume a node is in the susceptible state, and that it has just undergone a node update, where no change in state has been recorded. We are interested in the probability that the node has not updated after a time $\tau$. This is clearly $(1 - F\Delta t)^{\tau/\Delta t}$, since there are $\tau/\Delta t$ increments of duration $\Delta t$ over an interval $\tau$. Taking the limit $\Delta t \to 0$ gives

$$\lim_{\Delta t \to 0} (1 - F\Delta t)^{\tau/\Delta t} = e^{-F\tau}, \tag{1.36}$$

so that the "residual" adoption time is exponentially distributed. If we follow the natural interpretation of this quantity as a residual distribution, or the CDF of the inter-update time PDF. The CDF of the inter-update time distribution is $1 - e^{-F\tau}$, meaning the inter-update time PDF is

$$\frac{d}{d\tau}(1 - e^{-F\tau}) = Fe^{-F\tau}. \tag{1.37}$$

Crucially, this corresponds to a Poisson process, where $F$ (or the recovery rate $R$ in the mirroring process) can be construed as constant rates of infection and adoption, as discussed in [163]. This is important in the case of dynamical processes where infection and recovery are defined as constant hazard rates. In fact, since we've defined $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$ in terms of *probabilities* per $dt$, we have implicitly suggested an asynchronous update regime, as well as constant rate infection and recovery processes.

Throughout the remainder of this thesis, we use the asynchronous update rule in our Monte Carlo simulation of dynamical processes.

### 1.7.3 *Event sequences for temporal networks*

In addition to the simulation node dynamics, we will be confronted with the problem of simulating a *temporal network*, which is a network evolving in time, at a scale commensurate with node state dynamics. We shall introduce these networks in more detail in Chapter 5. In the above algorithms, we have not specified whether the underlying network is static, or dynamic. When the state of a link is able to evolve over time, we have what is called a *temporal network*. Whereas nodes can take on one of two possible states, we can do the same with edges, and attribute them with any number of possible states. If these states are fixed, edges are static, and if allowed to evolve over time, can be useful as a model of temporal networks. Following the discussion above regarding asynchronous node update, edge state evolution at

constant-rate could be implemented via the random selection of edges combined with Monte Carlo state flip probabilities. However, we are interested in temporal network models that don't necessarily evolve at a constant rate.

When simulating general classes of temporal networks, we cannot assume constant rates of evolution, and therefore turn to an *event-based* simulation models [16]. This borders on a class of event-based algorithms, going back to Gillespie algorithms [80], Instead of advancing in time by uniform discrete steps, an event based simulation advances according to events. This amounts to a time-ordered queue of events, typically represented as a tuple containing at least the following attributes; the time $t$ at which the event takes place, the adjacent nodes $(u, v)$, as well as the operation that is taking place on the edge, $\xi$. We can thus define an $e$,

*event-based simulation*

$$e = \{t, (u, v), \xi\}, \tag{1.38}$$

which we implement as an ordered set. Such a set can be made to act as a sequence of events, if the set is equipped with a comparison operator, which we choose to be time. Also need to be able to insert and delete elements. Algorithms are straightforward, as the next event we effectuate is the first in the ordering, fast to access. Inserting is logarithmic in the size of the event sequence, $\mathcal{O}(\log |S|)$. Accessing the next event is constant or $\mathcal{O}(1)$.

We have to pay attention to degeneracy, since even with high precision of the event time $t$, multiple events might share a single time $t$. For very large networks, so called "collisions" are possible, and therefore require a more sophisticated comparison operator (instead of $t < t'$, node comparison as well $(u, v) \sim (w, z)$). This is because certain data structures don't allow multiple entries; trying to insert two events $e_1$ and $e_2$. Alternatively, one can exploit pre-defined templates and data structures. In C++, we have available the multiset. However when this is not possible, class can be built manually using the described comparison. This is essential if large networks or event sequences are required, and double precision not available.

### 1.7.4 *System size considerations*

In general we will be interested in large networks with constant average degree. Such networks are sparse, and since we mostly use the

16. As pointed out in [163], constant rate processes can also be simulated using an event-based algorithm, by leveraging our knowledge of the inter-update distribution. That is, one can generate a queue of events, which here means spin flips, based on the known adoption rates $F$ and $R$ of nodes in the network

**Table 1.5 –** Time complexity of performing elementary operations on an edge-activity stream. The set is equiped with a comparison operator applied to the time $t$ key. This may be extended to include the tuple of adjacent nodes, $(u, v)$ to avoid collisions. Assumes a C++ set of size $|S|$. Time degeneracy can be avoided by using a multiset.

| | †access head | ‡erase head | insert element |
|---|---|---|---|
| edge stream | $\mathcal{O}(1)$ | $\mathcal{O}(1)$ | $\mathcal{O}(\log \lvert S \rvert)$ |

† returns a pointer to the next event
‡ assumes input is a pointer to location in set, if input is element, $\mathcal{O}(\log |S|)$

configuration model to generate networks, self-edges and multi-edges can be neglected in principle. In fact, in order to favour speed of implementation, we do not even check for them [17]. Throughout this work we shall sometimes refer to *thermodynamic limit*, which is the same as saying $N \to \infty$, and implies that $\mathcal{O}(1)$ effects can be comfortably neglected.

*thermodynamic limit*

One instance in which the large $N$ assumption is essential, is in the derivation of our master equation solution. We will see that such a derivation takes place independent of network size. [18] As a consequence, the predictions of our master equation solution correspond most closely to large networks. In fact, the larger the simulated network, the better the fit with analytics.

#### 1.7.4.1 *A caution against blindly performing averages*

Fluctuations play a significant role in dynamical processes. As such, it is always good practice to explore these systems using large numbers of realisations, and collect a distribution of outcomes. In simulation, one should never measure the average alone, which can be misleading. As a simple illustration, consider that a global cascade in the Watts model as per Section 1.6.2 is either local, or global. When the network is very large, and the network is connected, cascade sizes can be well approximated by 0, local cascades, or 1, global cascades. Clearly, taking the mean of experimental data gives $0 < \rho < 1$, which is misleading.

17. If we store the networks using an adjacency list (see Section 1.3.1), implementing the adjacency list using a data structure not allowing the repetition of keys, such as sets in C++, multi-edges will be cancelled out anyway.

18. This is in contrast to other popular modelling approaches such as the well-known master stability equation [22, 161, 163], where the dimension of the system is precisely $N$, the number of nodes in the network.

However, our master equation solution is deterministic, and doesn't inform matters of stochasticity encountered in experiment. Since this is the case, we have to be careful, especially when we are averaging spreading curves. In general we store distributions of results from experiments; it is almost never a good idea to perform averages on-the-go. It is preferable to save all experimental data, and consider averages during post-processing.

#### 1.7.4.2 *Asynchronous update vs breadth-first search*

We shall encounter a particular variety of experiment where both asynchronous and synchronous Monte Carlo simulation are hopelessly inefficient. This is the case when studying threshold models with small initial seeds, with no noise (this is the focus of Chapter 4). If initial seed is small, no point updating as before; only the $k$ neighbours of the initial seed have any chance of updating. As such, $N(1 - z\rho_0)$ selections will have been wasted. If $N$ is large, very inefficient. If $N$ small, we don't care. In these experiments, much better to branch outwards from the seed using an adapted breadth-first search algorithm. The gains are magnified in topologies where the cascade dies out quickly. Say you have $N = 10^8$ and your initial seed contains only isolated nodes. If you diligently wait 100 time steps, equivalent to $10^{10}$ random node selections under an asynchronous update rule, you get the same result as for the modified BFS which would require $N\rho_0$ operations. Note $\rho_0 \ll 1$, so amounts to checking the degree of something like 1, 10 or 100 nodes, whatever the seed size is, to verify that no global cascade is possible. The message is that we use a mixture of simulation regimes depending on the context. A Monte Carlo component is still important in such an experiment, namely in the random sampling of seeds from which to start the breadth-first search.

*breadth-first search*

### 1.7.5 *Distribution of experiments*

Throughout this thesis, we shall frequently face the problem of simulating both large networks, as well as smaller networks over a large number of realisations. We briefly sketch here the bottlenecks in computation that present themselves at various points in this work, and the solutions that we employ to overcome them. When we talk about network size, we don't just mean the number of nodes $N$, because storing node information is inexpensive compared to edge information

if the network is dense [19]. Note that simulating a dynamical process on a large network is memory intensive, because of graph storage, but also CPU intensive because the complexity of the simulation is $\mathcal{O}(N + |E|)$. As such, when we say CPU intensive we imply medium sized networks.

#### 1.7.5.1 *CPU intensive experiments*

At various points we shall wish to explore large parameter spaces for various experiments. These parameters are typically network topology, and parameters of node dynamics such as node threshold $\phi$. Further, for each configuration of parameters, we normally wish for a substantial number of realisations in order to collect valuable statistics regarding experimental outcomes. Because of the computational demands of such tasks, we simulate networks that are smaller than the maximum available, favouring instead a large number of network instances.

When this is the case, we are able to exploit the resources of high-performance computing centre [20], many of which have physical resources for computationally expensive, but memory-cheap calculations. Balance between real time, and available resources.

#### 1.7.5.2 *Memory intensive experiments*

Memory intensive networks arise when networks are very large, very dense, or both. Memory bottlenecks are in terms *density* rather than number of nodes, that is, $\mathcal{O}(N + |E|)$. For example, we have a tradeoff in simulating large sparse networks, a typical example involves single runs on a $N = 10^8$ node network (simple static network spreading process) with average degree $z \sim 5$. In contrast, we can afford $N = 10^7$ nodes if $z \sim 50$. If memory resources become sparse, for example if we require the simulation of very large networks, whilst still demanding a large number of realisations and parameter space, we turn to parallelisation (see Fig. 1.5). We save memory by only storing one copy of the graph, which dominates the memory, and running many dynamical processes on top of it in parallel. Note that node dynamics can be stored in just $\mathcal{O}(N)$ space, in the case of discrete state dynamics

19. See Table 1.1, the spatial complexity of an adjacency list is $\mathcal{O}(N + |E|)$, meaning edge information dominates node information for dense networks. For example, in Chapter 4 we probe networks up to average degree $z = 10^3$, meaning edge storage $|E| = \frac{zN}{2} \gg N$ dominates node information.

20. We have benefited immensely from the assistance of the Pôle Scientifique de Modelisation Numérique (PSMN) in Lyon, France, throughout this thesis. Here, the distribution of millions of processes is made straightforward by a batch queueing system (also known as a job scheduler, or grid engine).

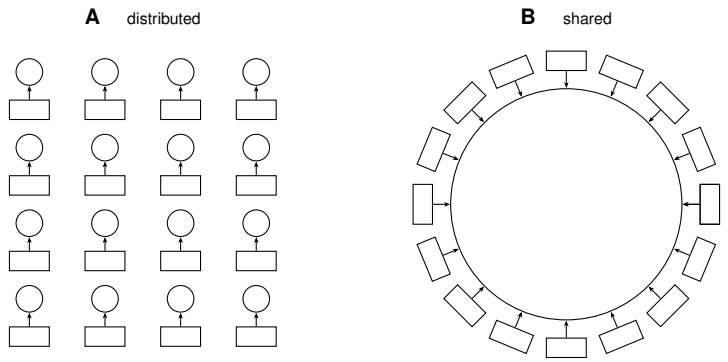


**Figure 1.5 –** Approaches to the distribution of experiments on a computing cluster. (a) Distribution of graph realisations along with dynamical process, where on average each CPU (rectangles) is allocated one graph (circles). This is important when sampling from a random graph ensemble is important. (b) When large graphs (in terms of $|V| + |E|$) are favoured, multiple CPUs simulate independent spreading processes on each core, on an identical graph. The dynamical process is inexpensive, on the order of $|V|$ in size.

(effectively a time-dependent vector of integers detailing the state of each node).

## 1.8 THESIS OUTLINE AND CONTRIBUTIONS

The remainder of this thesis is structured as follows. Throughout the rest of this part, we further develop the background information, including original analytic results, required for the studies in Part II. In Chapter 2 we introduce the analytical framework used throughout this work. Here we briefly discuss existing analytic results, before focussing on results that are novel to this thesis. A major original contribution of this thesis is the development of an extension to analytic tools that incorporate a broad range of edge heterogeneities. This refers to heterogeneity in edge weight, layer, direction, sign, or some otherwise specified state.

Then, in Part II, we apply our master equation framework to a number of systems of interest. In Chapter 3 we outline the results of [190], where we apply the static edge-heterogeneous master equation solution to threshold driven contagion on weighted networks. Analytic tools, combinatorial arguments, simulation on synthetic networks, as well as a data driven study on a number of empirical networks. In addition to prompting the initial development of the analytic tools

introduced in Chapter 2, we provide a microscopic explanation of the impact of weight heterogeneity on spreading dynamics. This manifests as a non-monotonic acceleration and deceleration of the spreading process relative to that on an unweighted network.

In Chapter 4 we study the effect of multiplexity on threshold driven contagion. We shall see that rich dynamics occur when multiplex topology is defined by so-called intimacy circles, an elementary network structure that can be shown to result from entropy maximisation in the presence of cost heterogeneity of forming an edge. We show that the effect of these structures is to fundamentally destabilise resulting multiplexes. This is particularly true at large values of average degree $z$.

In Chapter 5 we examine contagion on network structures that are not fixed in time. We use a simple model of temporal networks, known as stochastic temporal networks. Our main concern is to understand the effect of temporal heterogeneities on spreading processes. In other words, what is the effect of burstiness in interaction patterns on the spread of information over a network.

In Chapter 6 we discuss open questions, the limitations of our approach, as well as potential further work.

# 2

# MATHEMATICAL METHODS

In this chapter we give a detailed outline of a master equation formulation of spreading processes on edge-heterogeneous networks. This formalism is applied in subsequent chapters, and represents the principal analytic contribution of this thesis. This framework builds upon the work of Gleeson [86, 87], who introduced a master equation solution of binary state dynamics on unweighted, single-layer static networks. In this work, we generalise this framework in order to introduce various edge heterogeneities. In particular, this includes edge weight, layer and direction, as well as general framework to incorporate temporality.

After deriving a general master equation solution, we show how the resulting system of ordinary differential equations can be reduced in dimension, allowing qualitative analyses via the tools of dynamical systems theory. This includes linear stability analysis (LSA), as well as vector field analysis (VFA). Additionally, we compare this approach to other analyses of binary state dynamics on networks. This includes the use of tree-based methods, such as branching processes and message passing, but also node based methods such as master stability equations.

## 2.1 MASTER EQUATIONS

In this section we provide a general derivation of the approximate master equation (AME) formalism. A master equation is a phenomenological set of first-order differential equations describing the time evolution of a dynamical system, in terms of the densities of a discrete set of states composing the system, over a continuous time variable $t$.

Our approach is to assume that heterogeneities in edge qualities, such as weight, layer and temporal state, can be modelled with discrete, finite distributions. It then follows that each edge can be assigned an index representing such a state, providing the value of its edge

weight, the layer that it belongs to in a multiplex network, or some time varying quantity in a model of temporal networks.

In order that the resulting system of differential equations be finite, there must exist upper and lower cutoffs in the discrete distribution of possible edge states. We shall see that even coarse discretisations can yield useful insights. Continuous quantities are things like weight distributions which are often smooth [1] can be discretised, with continuous systems readily explorable via simulation, even if the corresponding analytical solution is infinite. On the other hand, many edge-heterogeneous systems are inherently discrete and can be readily approached with this formalism. In Chapter 4, we note that many real systems are multiplexed, where the number of layers provides the level of edge-state heterogeneity. We shall refer to an edge's *type*, a term which is interchangeable with edge state.

*edge type*

Although the following master equation formalism is able to solve for a range of binary-state dynamics, it is particularly appropriate for the use of threshold models. That is, dynamics such as SI and SIS processes can be accurately solved on random networks using the *fully mixed* assumption, requiring only the average degree $\langle k \rangle$. This leads to a single equation describing the spreading process. When the degree distribution is heterogeneous, the *degree-block* approximations is introduced [157], where nodes are partitioned into classes by their degree $k$, where they are assumed to be statistically equivalent. This increases the number of equations from 1 to $k$, reflecting the information cost of degree heterogeneity. While the degree-block approximation is sufficient in the case of epidemic models, where the effect of an additional infected neighbour is independent of a node's configuration, it is insufficient in the case of threshold models. We shall see that *degree-active degree* blocks are necessary here, where the network is further partitioned according to the number of active neighbours $m$, in addition to the total number of neighbours $k$. This allows us to model more complex phenomena such as threshold dynamics, where as we saw in Section 1.5.1, nodes consider neighbours collectively rather than independently.

### 2.1.1 *Node configurations*

Since we are interested in systems where edge properties can be discretised into a finite number of states, we let $n$ be the number of possible edge states, and index these edge states by $j$, such that

1. This depends on the network in question. A mobile call network where weights are defined by the number of calls in a day, gives a discrete distribution.

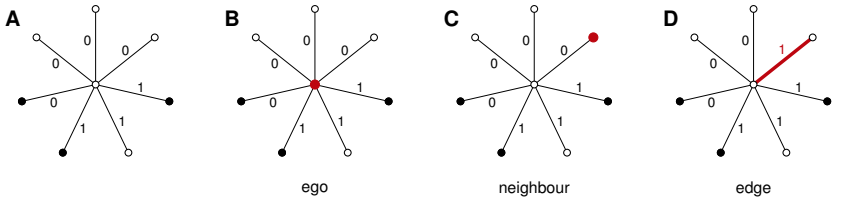


**Figure 2.1 –** A $k = 7$ degree node in a system allowing $n = 2$ edge types, with its allowed transitions. The node configuration in (a) is given by $\mathbf{k} = (4, 3)^T$ and $\mathbf{m} = (1, 2)^T$. Activation of the ego in (b) does not change the local neighbourhood description $(\mathbf{k}, \mathbf{m})$. A neighbour becoming infected in (c) changes the active degree vector to $\mathbf{m} = (2, 2)^T$, while $\mathbf{k}$ remains unchanged. The edge-state increment in (d) gives $\mathbf{k} = (3, 4)^T$, leaving $\mathbf{m}$ unchanged since the adjacent neighbour is inactive.

$1 \leq j \leq n$. We introduce the term $k_j$ to quantify the number of neighbours adjacent to a given node via edges of type $j$. We call this the $j$-th degree, or degree of $j$-type edges. Similarly, $m_j$ is the number of those neighbours that are *active*.[2] The total number of neighbours, or degree, of a given node is defined as $k$, the total degree, such that $k = \sum_j k_j$. While we have $0 \leq k_j \leq k$, we can also state $0 \leq m_j \leq k_j$, and we call $m_j$ the $j$-th partial degree, or partial degree of $j$-type edges. The total number of active neighbours is given by $m$, where $m = \sum_j m_j$ and $0 \leq m \leq k$, and we call $m$ the partial degree. We store the $n$ values $k_j$ and $m_j$ in the degree vector $\mathbf{k}$, and partial degree vector $\mathbf{m}$. We sometimes refer to $\mathbf{k}$ as the structural degree vector, since it depends only on network connectivity, and not node state which is a function of a dynamical process. Thus, a node's instantaneous *local configuration* is given by the tuple of degree vectors $(\mathbf{k}, \mathbf{m})$, which we also refer to as a node *class*. This incorporates both neighbour state, by distinguishing between inactive and active neighbours via the partial degree vector $\mathbf{m}$, as well as neighbour edge type. Note however that ego state is not conveyed by $(\mathbf{k}, \mathbf{m})$, since both inactive and active nodes can have identical neighbour configurations in this notation. Since every node belongs to exactly one class, the set of classes partitions a network.

*degree*

*degree and partial degree vectors*

*configuration*

We now turn to the notion of *configuration space*. While a node's local configuration is given by $(\mathbf{k}, \mathbf{m})$, the set of all configurations allowed by a given topology of an infinite network, provide the points in this space. This space accommodates all configurations $(\mathbf{k}, \mathbf{m})$ that may occur at any time $t$ over the course of a dynamical process, with $t \in [0, \infty)$.

*configuration space*

2. We use interchangeably the terms active, infected and adopting to refer to nodes in this state. Opposing states are termed inactive or susceptible.

A node's configuration is not fixed in time, and may change via two mechanisms. The first is via a change in the partial degree vector $\mathbf{m}$, which occurs when an inactive neighbour becomes active, or an active neighbour recovers to the inactive state. The ego in question transitions between classes $(\mathbf{k}, \mathbf{m})$ as a result of the change in $\mathbf{m}$ (see Section 2.1.9). The second transition is via a change in degree vector $\mathbf{k}$, which occurs in models of temporal networks (see Sections 2.1.10 and 2.1.11). If all nodes in the network are inactive at the beginning of a dynamical process, at $t = 0$, then every node has $\mathbf{m} = \mathbf{0}$. If at the end of this process, all nodes in the network are active, then $\mathbf{m} = \mathbf{k}$ for each node. Since time is continuous, and node dynamics are Markovian, nodes follow a smooth path through configuration space, such that $m_j$ (and $k_j$ in the case of Chapter 5) are only incremented by $\pm 1$ over any $dt$.

*configuration path*

### 2.1.2 *Degree and edge-quality distributions*

In general our approach will be to first define the distribution $p_k$ of total degree $k$, then to specify how those various edge qualities such as weight and layer membership are distributed among those edges. The total degree $k$ is comprised of edge-type degrees $k_i$ summing to $k$. If $p_{k_i}$ is the probability of a node having degree $k_i$ of edge type $i$, then

$$p_{\mathbf{k}} = p_k p(k_1, \ldots, k_n; k, p_{k_1}, \ldots, p_{k_n}), \tag{2.1}$$

gives the probability that a randomly selected node has degree vector $\mathbf{k}$, where the second term is the multinomial distribution. As such, the sum over all degree vectors is normalised, and we write $\sum p_{\mathbf{k}} = 1$.

### 2.1.3 *Configuration space*

As discussed above, a network is partitioned by the configurations $(\mathbf{k}, \mathbf{m})$, since each node has exactly one configuration. We introduce $C_{\mathbf{k},\mathbf{m}}$, the set of all nodes in the network with local configuration $(\mathbf{k}, \mathbf{m})$. We define a number of quantities based on this that will prove useful in following discussions. First, we define $C_{\mathbf{k}}$ to be the set of all sets $C_{\mathbf{k},\mathbf{m}}$ having degree vector $\mathbf{k}$, that is $C_{\mathbf{k}} = \{C_{\mathbf{k}',\mathbf{m}'} \mid \mathbf{k}' = \mathbf{k}\}$. Whereas $C_{\mathbf{k},\mathbf{m}}$ is a set of *nodes*, $C_{\mathbf{k}}$ is a set of *sets*. Similarly, we define $C_k$ to be the set of all sets $C_{\mathbf{k},\mathbf{m}}$ with total degree $k$, that is $C_k = \{C_{\mathbf{k},\mathbf{m}} \mid \sum_j k_j = k\}$. Finally, $C$ is the set of all *possible* sets $C_{\mathbf{k},\mathbf{m}}$, provided a distribution of total degrees $p_k$, and edge dimension $n$. This includes sets for which $C_{\mathbf{k},\mathbf{m}} = \emptyset$ at a given time. The cardinality of this universal set, $|C|$, is determined by the support of $p_k$, in addition to $n$. We discuss the

structure of $C$, in particular the dependence of $|C|$ on $k$ and $n$, in Section 2.1.5.

Since $(\mathbf{k}, \mathbf{m})$ does not convey ego state, just edge and neighbour configuration, we partition $C_{\mathbf{k},\mathbf{m}}$ into inactive and active nodes,[3] such that $C_{\mathbf{k},\mathbf{m}} = S_{\mathbf{k},\mathbf{m}} \cup I_{\mathbf{k},\mathbf{m}}$. Similar definitions allow us to introduce $S_{\mathbf{k}}$ and $I_{\mathbf{k}}$, $S_k$ and $I_k$, as well as $S$ and $I$. Although in general $|S_{\mathbf{k},\mathbf{m}}| \neq |I_{\mathbf{k},\mathbf{m}}|$, the structure of the inactive and active configuration spaces is identical, such that $|C| = |S| = |I|$, $|C_k| = |S_k| = |I_k|$ and $|C_{\mathbf{k}}| = |S_{\mathbf{k}}| = |I_{\mathbf{k}}|$.

### 2.1.4 *Density of states*

The evolution of a dynamical process over a network amounts to a flow of nodes through the sets $S_{\mathbf{k},\mathbf{m}}$ and $I_{\mathbf{k},\mathbf{m}}$. Since the number of nodes $N$ in the network is conserved, it is just their distribution over the sets $S_{\mathbf{k},\mathbf{m}}$ and $I_{\mathbf{k},\mathbf{m}}$ that evolves in time. These distribution provide the *state* of the network at time $t$. Since our formalism is independent of network size, we treat the distributions of nodes rather than the absolute sizes of these sets. To simplify the following discussion, we define

$$\|C_{\mathbf{k}}\| \equiv \sum_{C_{\mathbf{k},\mathbf{m}} \in C_{\mathbf{k}}} |C_{\mathbf{k},\mathbf{m}}| \quad \text{and} \quad \|C_k\| \equiv \sum_{C_{\mathbf{k},\mathbf{m}} \in C_k} |C_{\mathbf{k},\mathbf{m}}|, \tag{2.2}$$

in order to give the number of *nodes* with degree vector $\mathbf{k}$, and total degree $k$, respectively. This is in contrast to $|C_{\mathbf{k}}|$ and $|C_k|$ which give the number of *configurations* with degrees $\mathbf{k}$ and $k$. To convert from absolute node count to densities of nodes, we need to normalise $S_{\mathbf{k},\mathbf{m}}$ and $I_{\mathbf{k},\mathbf{m}}$ by some non-zero quantity that is conserved over the course of a dynamical process. In static networks (see Chapters 3 and 4), such a quantity is the number of nodes with degree vector $\mathbf{k}$ or $\|C_{\mathbf{k}}\|$. This allows us to define

$$s_{\mathbf{k},\mathbf{m}} = \frac{|S_{\mathbf{k},\mathbf{m}}|}{\|C_{\mathbf{k}}\|}. \tag{2.3}$$

We define $i_{\mathbf{k},\mathbf{m}}$ in a similar way, replacing $S_{\mathbf{k},\mathbf{m}}$ with $I_{\mathbf{k},\mathbf{m}}$ in the expression above. The normalised quantities $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$ give the probability that a randomly selected node with degree vector $\mathbf{k}$ is inactive or active, respectively, with partial degree vector $\mathbf{m}$, at time $t$. In contrast, we shall see that $\|C_{\mathbf{k}}\|$ is not fixed in the case of temporal networks (see Chapter 5), therefore we need to normalise by another quantity. See

3. The use of $S$ and $I$ here is a vestige of epidemiology, where inactive and active nodes are termed susceptible and infected, respectively.

Fig. B.1 to see that it is the connected components of configuration space that determine the conserved quantities to be used for normalisation. For the temporal network models in question, the desired quantity is $\|C_k\|$, defined above. The density of inactive nodes in class $(\mathbf{k}, \mathbf{m})$ in this case is given by

$$s_{\mathbf{k},\mathbf{m}} = \frac{|S_{\mathbf{k},\mathbf{m}}|}{\|C_k\|}, \tag{2.4}$$

with $i_{\mathbf{k},\mathbf{m}}$ defined analogously. We have the normalisation condition $\sum_{\mathbf{m}|\mathbf{k}}(s_{\mathbf{k},\mathbf{m}} + i_{\mathbf{k},\mathbf{m}}) = 1$ for a given $\mathbf{k}$ class in the case of static networks. This allows us to write

$$\rho_{\mathbf{k}} = 1 - \sum_{\mathbf{m}|\mathbf{k}} s_{\mathbf{k},\mathbf{m}} \quad \text{and} \quad \rho = \sum_{\mathbf{k}} p_{\mathbf{k}} \rho_{\mathbf{k}} \tag{2.5}$$

where the sum in the first expression is over all partial degree vectors $\mathbf{m}$ that satisfy $0 \leq m_j \leq k_j$. This gives the probability $\rho_{\mathbf{k}}$ that a randomly selected node with degree vector $\mathbf{k}$ will be active at time $t$. A weighted sum over all $\rho_{\mathbf{k}}$ gives the probability $\rho$ that a randomly selected node is active. Similarly, in the case of temporal networks, the node conservation principle leads to the normalisation condition $\sum_{\mathbf{k},\mathbf{m}|k}(s_{\mathbf{k},\mathbf{m}} + i_{\mathbf{k},\mathbf{m}}) = 1$, for a given $k$ class. We then have

$$\rho_k = 1 - \sum_{\mathbf{k},\mathbf{m}|k} s_{\mathbf{k},\mathbf{m}} \quad \text{and} \quad \rho = \sum_{k} p_k \rho_k, \tag{2.6}$$

where the sum in the first expression is over all configurations $(\mathbf{k}, \mathbf{m})$ that satisfy $\sum_j k_j = k$. The term $\rho_k$ gives the probability that a randomly selected node with total degree $k$ will be active, and $\rho$ the probability that any randomly selected node will be active. We turn to temporal networks in detail in Chapter 5.

### 2.1.5 *Size of configuration space*

By far the most important quantity in determining the feasibility of a master equation solution to a given problem, is the number of points in its configuration space, or $|C|$. Crucially, this determines the order of the resultant system of differential equations, with each point associated to a rate equation for $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$. As such, it is the size of $C$ that ultimately determines the efficiency of a given implementation. Even in special cases where one can derive equivalent low dimensional systems, as in Section 2.2, their terms are still evaluated by iterating

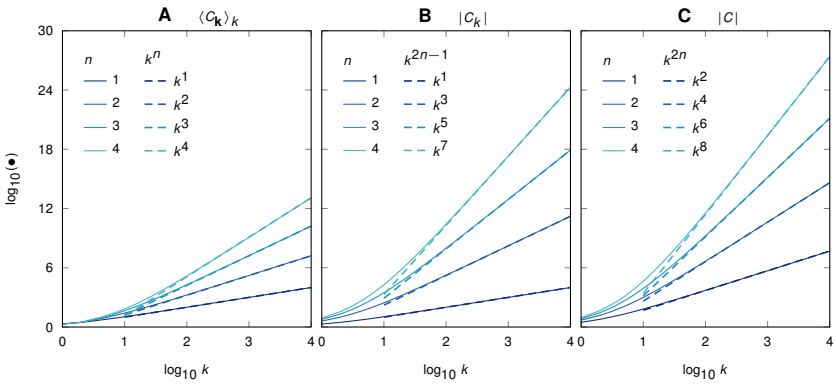


**Figure 2.2 –** Size of configuration space for fixed $n$ and increasing $k$. In (a) we plot $\langle C_{\mathbf{k}} \rangle_k$, the expected number of $\mathbf{m}$ per $\mathbf{k}$, for a given $k$. In (b) we plot $|C_k|$, the number of classes $(\mathbf{k}, \mathbf{m})$ with total degree $k$, or equivalently, the entire space $C$ of a $k$-regular random network. To calculate $|C|$ for a general degree distribution in (c), we assume that the support of the total degree distribution is $\text{supp}(p_k) = \{0, \ldots, k\}$, and as such can be construed as the cumulative sum of the terms in (b). Dashed lines correspond to $k^n$, $k^{2n-1}$ and $k^{2n}$, in plots (a), (b) and (c) respectively, as per Table 2.1. Since scale is log-log, dashed lines are straight with slopes $n$, $2n-1$ and $2n$, with vertical intercepts in Table 2.1. See Fig. 2.3 for side-by-size comparison of corresponding $n$ values.

over $C$. As such, we devote this section to providing approximate bounds [4] on the size of this space.

Importantly, the configuration space is finite if and only the degree distribution has some cutoff and the number of allowed edge states $n$ is finite. In the following we refer to two types of dimension, the first being the dimension of vectors $\mathbf{k}$ and $\mathbf{m}$ which is $n$, as well as the dimension of configuration space $|C|$, which is expressed in terms of $n$ and maximum degree $k$, and whose growth rate as a function of $k$ are shown in Table 2.1.

Our approach will be to first calculate the sizes of $C_{\mathbf{k}}$ and $C_k$, or the number of configurations with degree vector $\mathbf{k}$ and $k$, respectively. These results then determine the total size of a system whose degree distribution has support $\{0, \ldots, k\}$, such that $|C| = \sum_{k'} |C_{k'}|$, where $k$

4. Many of the results are asymptotic, and we use the standard asymptotic notation: for two functions $f$ and $g$, we write $f = \mathcal{O}(g)$ if $f \leq g$ for all sufficiently large values of the variables of the two functions, where $c$ is an absolute positive constant. We write $f = \Omega(g)$ if $g = \mathcal{O}(f)$ and $f = \Theta(g)$ if $f = \mathcal{O}(g)$ and $f = \Omega(g)$. If the limit of the ratio $f/g$ tends to zero as the variables of the functions tend to infinity we write $f = o(g)$. Finally, $f \sim g$ denotes that $f = (1 + o(1))g$, that is $f/g$ tends to 1 when the variables tend to infinity.

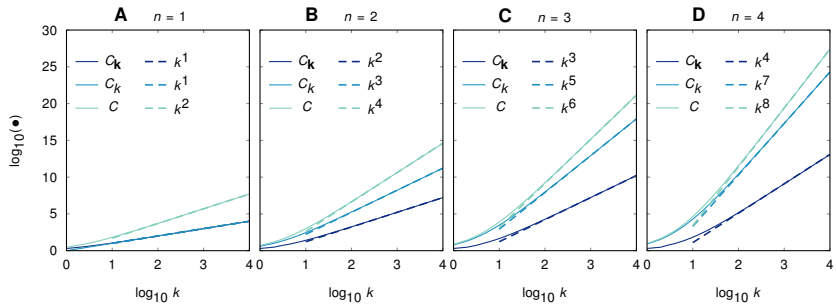


**Figure 2.3 –** Same data as for Fig. 2.2, but comparing $\langle C_{\mathbf{k}} \rangle_k$, $|C_k|$ and $|C|$ for equal $n$. Colours do not necessarily correspond to those in Fig. 2.2.

equals the maximum degree in the remainder of this section. Explicitly, $|C_{\mathbf{k}}|$ and $|C_k|$ are related by

$$|C_k| = \sum_{\mathbf{k}} |C_{\mathbf{k}}| \tag{2.7a}$$

$$= \sum_{\mathbf{k}} \prod_{j=1}^{n} (k_j + 1) \tag{2.7b}$$

$$\leq \sum_{\mathbf{k}} (k+1)^n, \tag{2.7c}$$

where the sum runs over all degree vectors $\mathbf{k}$ with total degree $k$, or $\sum_j k_j = k$. The product in Eqn. 2.7b gives the number of partial degree vectors $\mathbf{m}$ corresponding to a given $\mathbf{k}$. This is because every $j$-type degree has $k_j + 1$ possible inactive, active configurations, given by $m_j = 0, \ldots, k_j$. Enumerating the total number of $\mathbf{m}$ then entails the product of the $j$ edge types. Since it is not clear how to write Eqn. 2.7 in closed form, we attempt now to find a more useful expression by considering its bounds.

We estimate the growth rates of two quantities, namely the number of terms in the sum of Eqn. 2.7, and the average value of its summand $|C_{\mathbf{k}}|$ for a given $k$, which we denote $\langle C_{\mathbf{k}} \rangle_k$. The sum is constrained to all degree vectors $\mathbf{k}$ whose elements add to $k$. There are $\binom{k+n-1}{k}$ such vectors, a result that can be obtained by enumerating the number of ways of placing $k$ objects into $n$ containers. This binomial obeys the inequalities

$$\frac{(k+1)^{n-1}}{(n-1)!} < \frac{(k+n-1)!}{k!(n-1)!} < \frac{(k+n-1)^{n-1}}{(n-1)!}, \tag{2.8}$$

with the three terms converging when $n$ is constant and $k \to \infty$, each term growing as $\Theta(k^{n-1})$. In order to determine the expected value of

**Table 2.1 –** Growth rate of the size of configuration space $C$, as well as the subspaces $C_{\mathbf{k}}$ and $C_k$. These rates are accurate up to a multiplicative constant, which are shown here for the cases $n = 1, \ldots, 4$, and correspond to Figs. 2.2 and 2.3. In these figures we plot these limits against actual set sizes.

| | $\langle C_{\mathbf{k}} \rangle_k$ | $\lvert C_k \rvert$ | $\lvert C \rvert$ |
|---|---|---|---|
| $n$ | $\Theta(k^n)$ | $\Theta(k^{2n-1})$ | $\Theta(k^{2n})$ |
| 1 | 1.00000 | 1.00000 | 0.50000 |
| 2 | 0.16500 | 0.16500 | 0.04200 |
| 3 | 0.01660 | 0.00830 | 0.00140 |
| 4 | 0.00125 | 0.00020 | 0.00002 |

$|C_{\mathbf{k}}|$ for a given $k$, it is clear that for some degree vector $\mathbf{k}$, the number of corresponding $\mathbf{m}$ is bounded above as per Eqn. 2.7c. This inequality holds because $k_j \leq k$ for all $j$, with equality holding only when $n = 1$. Although this is sufficient to write $\mathcal{O}(k^n)$, we propose that the bound is actually stricter,[5] growing as $\Theta(k^n)$.

$$\langle C_{\mathbf{k}} \rangle_k = \Theta(k^n). \tag{2.9}$$

This appears to be confirmed in Fig. 2.2(a), where we plot the average number of active degree vectors $\mathbf{m}$ per degree vector $\mathbf{k}$, for a given $k$. Actual values of $C_{\mathbf{k}}$ are solid, dashed lines are the expected value.

Given that the number of terms in Eqn. 2.7 grows like $k^{n-1}$, and that the mean value of its summand $|C_{\mathbf{k}}|$ grows like $k^n$, we conclude that the total number of configurations with total degree $k$, or $|C_k|$, grows like $k^{2n-1}$. This is confirmed in Fig. 2.2(b), where we plot the actual value of $|C_k|$ against its expected value, for some multiplicative constant that we find manually by tuning.

$$|C_k| = \Theta(k^{n-1}) \times \Theta(k^n) = \Theta(k^{2n-1}). \tag{2.10}$$

Since Fig. 2.2 is on a log-log scale, polynomials appear as straight lines whose slope is given by their leading order. The multiplicative constants of $\Theta$ determine the vertical intercept, and are provided in Table 2.1. Finally, we can generate an upper bound on the *total* number of points in $|C|$, not just for a particular subset of degree $k$. To do this, we observe that the number of $\mathbf{k}$ for each $k$ is monotonically increasing,

5. See footnote 4 for summary of asymptotic notation.

and does so gradually [6] for large $k$. As such, if the cutoff in total degree is $k$, there will be on the order of

$$|C| = k \times \Theta(k^{2n-1}) = \Theta(k^{2n}). \tag{2.11}$$

total configurations in $C$. This result is once again confirmed in Fig. 2.2(c), with the multiplicative constants listed in Table 2.1. Note that our result of $|C| = \Theta(k^{2n})$ is consistent with Gleeson's result of $\Theta(k^2)$, reported in [86], since the setting there was networks of homogeneous edge-type, a case recovered here by setting $n = 1$.

A consequence of these results is that a precise master equation implementation of real systems is quite impractical. Real systems often include hubs, or nodes whose degrees reach into the tens of thousands. In the presence of edge heterogeneity, we have seen that such systems are prohibitively large. For example, a four-level multiplex with a maximum total degree of $k = 10^4$ is out of reach analytically, as it contains $10^{27}$ terms. Nevertheless, a great deal of intuition can be build be studying the behaviour of more tractable systems, by imposing cutoffs on the maximum degree, and formulating the system in terms of low edge heterogeneity $n$.

### 2.1.6 *Time evolution of density of states*

*initial seed*

The initial condition refers to the distribution of inactive and active nodes over the classes $S_{\mathbf{k},\mathbf{m}}$ and $I_{\mathbf{k},\mathbf{m}}$ (or the distribution of densities $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$) at time $t = 0$. A common situation is that the network is entirely inactive at $t = 0$, except for an infinitesimal perturbation which refers to a finite number of nodes that are active in an infinite network. Once the configuration space is defined (by the topology, degree distribution and edge heterogeneity), the time evolution of these distributions is determined entirely by the node dynamics, or the functions $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$. We assume that experiments are run until a steady state is reached.

6. The relative number of $\mathbf{k}$ with total degree $k$ and $k+1$ increases smoothly for large $k$ and constant $n$, with

$$\lim_{k\to\infty} \frac{B_{k+1+n-1,k+1}}{B_{k+n-1,k}} = \lim_{k\to\infty} \frac{k+n}{k+1} = 1.$$

We index the elements of the space $C$ using lexicographic ordering[7] of the configurations $(\mathbf{k}, \mathbf{m})$. This allows us to define one dimensional vectors $\mathbf{s}$ and $\mathbf{i}$ of length $|C|$ to contain the distribution of densities $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$. As such, $s_i$ and $i_i$ give the $i$-th elements of $\mathbf{s}$ and $\mathbf{i}$ according to the lexicographic ordering of $(\mathbf{k}, \mathbf{m})$. We introduce the matrix $A^s$ of transition rates $a_{ij}$ between states $s_i$ and $s_j$. The matrix $A^i$ is defined analogously. The full equations (see following sections) amount to solving $s_i = \sum_j a_{ij} s_j$, effectively a rate equation for the variable $\mathbf{s}$. We shall see that transitions can be time dependent, such that $a_{ij} = a_{ij}(t)$. The time evolution of the system can thus takes the form of the autonomous system When solving for non-recovery dynamics, we simply write $A^s = A$, as there is no need to solve for $\mathbf{i}$. Since dynamics are Markovian, the state of the system $\mathbf{s}$ in this case, evolves according to the rate equation

$$\frac{d}{dt}\mathbf{s} = A\mathbf{s}. \tag{2.12}$$

The meaning of this will become clear over the remainder of the chapter. In the following sections, we show how to determine the transition rates between classes, by explicitly writing the flux equations for $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$ for general binary state dynamics.

It is useful to note that the diagonals are necessarily less than or equal to zero, $a_{ii} \leq 0$, as they represent the loss terms from $s_i$. In contrast, off-diagonal terms are greater than or equal to zero, $a_{ij}$ for $i \neq j$, as they represent gain terms, or flows into $s_i$ from neighbouring classes. Keep in mind that configuration space is amounts to a directed graph, with time-dependent weights, as illustrated in Fig. 2.4, as well as Appendices A and B. Since they are generally sparse, we store $A$ as an adjacency list. The above statements still hold, keeping in mind that there is a one-to-one mapping between adjacency matrices and lists.

### 2.1.7 *Configuration transitions*

In Section 2.1.1 we outlined how the local neighbourhood of a node is fully described by the quantity $(\mathbf{k}, \mathbf{m})$. In this section we derive the transition rates between these configurations. There are slightly different formulations of the AMEs depending on the system in question,

7. Technically, any one-to-one mapping of $i \in \{1, \ldots, |C|\}$ is sufficient in a numerical implementation of the following system. For simplicity, we store the set of configurations $(\mathbf{k}, \mathbf{m})$ in lexicographic order. This defines a map $(\mathbf{k}, \mathbf{m}) \mapsto i$ for some $i \in \{1, \ldots, |C|\}$. In other words, whenever we write a class using a subscript with a single index, such as $s_i$, one can assume there exists a precise $s_{\mathbf{k},\mathbf{m}}$ that corresponds to it, and vice versa.

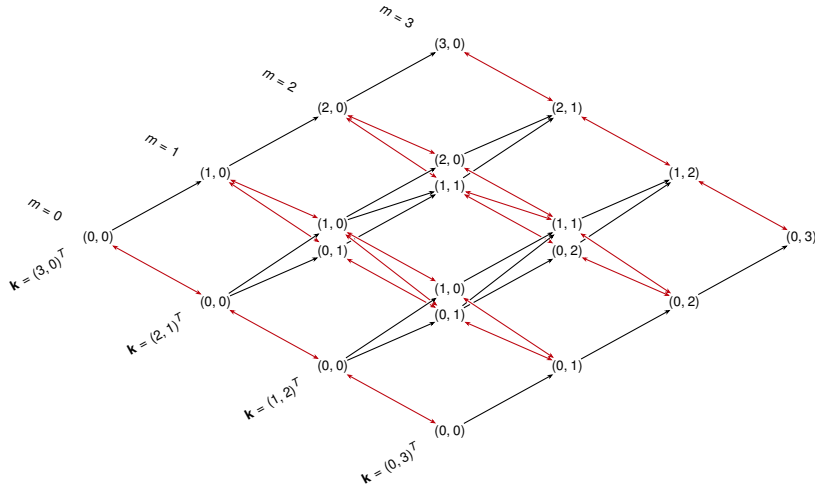


**Figure 2.4 –** Configuration space for nodes with total degree $k = 3$ and $n = 2$ allowed edge states. Points represent node configurations $(\mathbf{k}, \mathbf{m})$, and links the transitions between these configurations due to a dynamical process on nodes, or network temporality. See Appendices A and B for a detailed explanation, as well as more examples.

be it static weighted and multiplex networks, or temporal networks. If the network is static, nodes are confined to two types of transitions, which we call *ego state* transitions, as well as *neighbour state* transitions. The first is where an ego changes between the inactive and active state. The second type of transition is derivative of the first, where an ego observes a change in its local configuration $(\mathbf{k}, \mathbf{m})$ due to the change in state of one of its neighbours. Whereas ego-state transitions are binary, the number of neighbour-state transitions is specified by the state $j$ of the corresponding edge. In Chapters 3 and 4, we study systems where these are the only allowed transitions.

In addition, temporal networks exhibit a third type of transition, where the state of an edge changes according to the dynamics of the temporal network model in question. We term these transitions *edge-state* transitions. As outlined above, we confine ourselves to systems where edges have discrete state, and in Chapter 5, we consider a temporal network model that satisfies this constraint.

The following sections provide the exact form of these transitions.

### 2.1.8 *Ego transitions*

Ego transitions from the inactive state $S$ to the active state $I$, and vice versa, drive the dynamical process. The transitions between these two classes are defined by infection and recovery rates $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$, as discussed extensively in Chapter 1, and summarised in Tables 1.2, 1.3 and 1.4. We assume that node transitions occur homogeneously in time regardless of the underlying dynamics. In other words, a node in the inactive state becomes active over an interval $[t, t+dt]$ with probability $F_{\mathbf{k},\mathbf{m}}dt$, and recovers with probability $R_{\mathbf{k},\mathbf{m}}dt$. If $W(C_{\mathbf{k}_1,\mathbf{m}_1} \to C_{\mathbf{k}_2,\mathbf{m}_2})$ is the rate of transition from class $C_{\mathbf{k}_1,\mathbf{m}_1}$ to $C_{\mathbf{k}_2,\mathbf{m}_2}$, the two ego transitions are given by

$$W(S_{\mathbf{k},\mathbf{m}} \to I_{\mathbf{k},\mathbf{m}})dt = F_{\mathbf{k},\mathbf{m}}dt \tag{2.13}$$

and

$$W(I_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k},\mathbf{m}})dt = R_{\mathbf{k},\mathbf{m}}dt. \tag{2.14}$$

Combining these terms gives the total flux through the classes $S_{\mathbf{k},\mathbf{m}}$ and $I_{\mathbf{k},\mathbf{m}}$, due to ego transitions, which in terms of the corresponding densities gives

$$ds^{ego}_{\mathbf{k},\mathbf{m}} = -F_{\mathbf{k},\mathbf{m}}s_{\mathbf{k},\mathbf{m}}dt + R_{\mathbf{k},\mathbf{m}}i_{\mathbf{k},\mathbf{m}}dt \tag{2.15}$$

and

$$di^{ego}_{\mathbf{k},\mathbf{m}} = -R_{\mathbf{k},\mathbf{m}}i_{\mathbf{k},\mathbf{m}}dt + F_{\mathbf{k},\mathbf{m}}s_{\mathbf{k},\mathbf{m}}dt. \tag{2.16}$$

Ego transitions can be viewed as the transition driving the dynamical process, from which the following transitions are derived.

### 2.1.9 *Neighbour transitions*

Neighbour transitions refer to the change in nodes class due to the change in state of one of its neighbours. Whereas ego transitions do not entail a change in $(\mathbf{k}, \mathbf{m})$, neighbour transitions involve a change in partial degree vector $\mathbf{m}$. We distinguish neighbour transitions by the type $j$ of the corresponding edge, as well as the direction of the transition, i.e., whether it is inactive to active, or inactive to active. The rates at which nodes leave the class $(\mathbf{k}, \mathbf{m})$ due to neighbour adoption are given by

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k},\mathbf{m}+\mathbf{e}_j})dt = \beta^s_j(k_j - m_j)dt \tag{2.17}$$

and

$$W(I_{\mathbf{k},\mathbf{m}} \to I_{\mathbf{k},\mathbf{m}+\mathbf{e}_j})dt = \beta^i_j(k_j - m_j)dt \tag{2.18}$$

The coefficients $\beta_j^s$ and $\beta_j^i$ give the rates at which inactive neighbours of inactive and active nodes, respectively, become active. These quantities are derived below. Influxes to the class $(\mathbf{k}, \mathbf{m})$ due to the same mechanism are

$$W(S_{\mathbf{k},\mathbf{m}-\mathbf{e}_j} \to S_{\mathbf{k},\mathbf{m}})dt = \beta_j^s(k_j - m_j + 1)dt. \tag{2.19}$$

and

$$W(I_{\mathbf{k},\mathbf{m}-\mathbf{e}_j} \to I_{\mathbf{k},\mathbf{m}})dt = \beta_j^i(k_j - m_j + 1)dt. \tag{2.20}$$

Conversely, the outflux of $(\mathbf{k}, \mathbf{m})$ due to the recovery of active $j$-type neighbours is denoted by

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k},\mathbf{m}-\mathbf{e}_j})dt = \gamma_j^s m_j dt \tag{2.21}$$

and

$$W(I_{\mathbf{k},\mathbf{m}} \to I_{\mathbf{k},\mathbf{m}-\mathbf{e}_j})dt = \gamma_j^i m_j dt. \tag{2.22}$$

The coefficients $\gamma_j^s$ and $\gamma_j^i$ give the rates at which active neighbours of inactive and active nodes, respectively, become inactive. These quantities are also derived below. Influxes to the class $(\mathbf{k}, \mathbf{m})$ due to the same mechanism are

$$W(S_{\mathbf{k},\mathbf{m}+\mathbf{e}_j} \to S_{\mathbf{k},\mathbf{m}})dt = \gamma_j^s(m_j + 1)dt \tag{2.23}$$

and

$$W(I_{\mathbf{k},\mathbf{m}+\mathbf{e}_j} \to I_{\mathbf{k},\mathbf{m}})dt = \gamma_j^i(m_j + 1)dt. \tag{2.24}$$

To calculate $\beta$ and $\gamma$ we use straightforward ensemble arguments. That is, we don't calculate the transition rate for a particular class, but an expected rate over the entire network, and assume that such a rate applies to every node. To obtain the expected fraction of neighbours undergoing transitions, we observe the number of egos undergoing transitions at time $t$, and count the number of neighbour transitions thus produced. That is, when an inactive node in class $(\mathbf{k}, \mathbf{m})$ becomes active, which occurs with probability $F_{\mathbf{k},\mathbf{m}}dt$, it produces $k_j - m_j$ susceptible nodes that observe neighbour transitions. The number of such edges across the entire network is given by $\sum_S p_{\mathbf{k}}(k_j - m_j)F_{\mathbf{k},\mathbf{m}}s_{\mathbf{k},\mathbf{m}}$, where the sum is over all susceptible classes. We compare this to the total number of inactive-inactive edges, $\sum_S p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}$, giving the neighbour transition rate $\beta_j^s$. The remaining rates are found in a similar way.

The rates $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are implicitly time-dependent, due to their dependence on the distributions $\mathbf{s}(t)$ and $\mathbf{i}(t)$. They are explicit functions of the rates $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$, and can be written

$$\beta_j^s = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)F_{\mathbf{k},\mathbf{m}}s_{\mathbf{k},\mathbf{m}}}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}}, \tag{2.25}$$

and

$$\beta_j^i = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j F_{\mathbf{k},\mathbf{m}} s_{\mathbf{k},\mathbf{m}}}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j s_{\mathbf{k},\mathbf{m}}}, \tag{2.26}$$

which gives the expected rate at which an inactive neighbour of an inactive or active node, respectively, becomes active. Similar expressions provide the expected rate at which an active neighbour of an inactive or active node, respectively, recovers. That is,

$$\gamma_j^s = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} (k_j - m_j) R_{\mathbf{k},\mathbf{m}} i_{\mathbf{k},\mathbf{m}}}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} (k_j - m_j) i_{\mathbf{k},\mathbf{m}}}, \tag{2.27}$$

and

$$\gamma_j^i = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j R_{\mathbf{k},\mathbf{m}} i_{\mathbf{k},\mathbf{m}}}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j i_{\mathbf{k},\mathbf{m}}}. \tag{2.28}$$

As indicated, these rates may differ by edge type $j$, so that neighbour transition rates are given by the $n$-dimensional vectors $\boldsymbol{\beta}^s$ and $\boldsymbol{\beta}^i$ as well as $\boldsymbol{\gamma}^s$ and $\boldsymbol{\gamma}^i$. Note that when $R_{\mathbf{k},\mathbf{m}} = 0$, we have $\gamma_j^s = \gamma_j^i = 0$ for all $j$. Configuration flows due to these transitions are combined in the term

$$ds_{\mathbf{k},\mathbf{m}}^{neigh} = \sum_{j=0}^{n} \Big[ -\beta_j^s (k_j - m_j) s_{\mathbf{k},\mathbf{m}} + \beta_j^s (k_j - m_j + 1) s_{\mathbf{k},\mathbf{m}-\mathbf{e}_j} - \gamma_j^s m_j s_{\mathbf{k},\mathbf{m}} + \gamma_j^s (m_j + 1) s_{\mathbf{k},\mathbf{m}+\mathbf{e}_j} \Big] dt \tag{2.29}$$

for inactive classes, and

$$di_{\mathbf{k},\mathbf{m}}^{neigh} = \sum_{j=0}^{n} \Big[ -\beta_j^i (k_j - m_j) i_{\mathbf{k},\mathbf{m}} + \beta_j^i (k_j - m_j + 1) i_{\mathbf{k},\mathbf{m}-\mathbf{e}_j} - \gamma_j^i m_j i_{\mathbf{k},\mathbf{m}} + \gamma_j^i (m_j + 1) i_{\mathbf{k},\mathbf{m}+\mathbf{e}_j} \Big] dt \tag{2.30}$$

for active classes. These terms, as well as those corresponding to ego transitions above, are due to *node* dynamics. The remainder of this section deals with flows due to *edge* dynamics, such as those defined by the temporal network model in Chapter 5.

### 2.1.10 *Positive edge transitions*

A positive edge transition refers to the change in a node's configuration due to an increment in the state of one of its edges over an interval

$dt$. Regardless of the interpretation of edge state and the mechanism driving the transition, the probability of this occurring on an edge of type $j$ is $\mu_j dt$. In fact, we delay until Chapter 5 all discussions of models of temporal networks - for now it suffices to assume that they can be represented by networks with dynamic edge state. For a configuration $(\mathbf{k}, \mathbf{m})$, a positive edge transition on a $j$-type edge means losing an edge of that type, and gaining an edge of type $j+1$. For brevity, we introduce the term $\Delta_j = -\mathbf{e}_j + \mathbf{e}_{j+1}$, corresponding to the change in the degree vector $\mathbf{k}$ imposed by such a transition. That is, an adjacent node loses a $j$-type edge, and gains a $j+1$-type edge, all while conserving the overall degree $k$.

The configuration that a $(\mathbf{k}, \mathbf{m})$ node enters when undergoing a positive transition on a $j$-type edge is $(\mathbf{k}+\Delta_j, \mathbf{m})$, or $(\mathbf{k}+\Delta_j, \mathbf{m}+\Delta_j)$, depending on whether the neighbouring node was inactive or active, respectively. If the edge in question is connected to a inactive neighbour, we have

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k}+\Delta_j,\mathbf{m}})dt = \mu_j(k_j - m_j)dt, \tag{2.31}$$

and if the $j$-neighbour is active, we have

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k}+\Delta_j,\mathbf{m}+\Delta_j})dt = \mu_j m_j dt. \tag{2.32}$$

Similarly, nodes may enter the configuration $(\mathbf{k}, \mathbf{m})$ through a positive transition on a $j$-type edge, via the classes $(\mathbf{k}-\Delta_j, \mathbf{m})$ and $(\mathbf{k}-\Delta_j, \mathbf{m}-\Delta_j)$ as follows. If the event occurs between a node and an inactive neighbour, we have

$$W(S_{\mathbf{k}-\Delta_j,\mathbf{m}} \to S_{\mathbf{k},\mathbf{m}})dt = \mu_j(k_j - m_j + 1)dt, \tag{2.33}$$

and if the event occurs between a node and an active neighbour, we have

$$W(S_{\mathbf{k}-\Delta_j,\mathbf{m}-\Delta_j} \to S_{\mathbf{k},\mathbf{m}})dt = \mu_j(m_j + 1)dt. \tag{2.34}$$

Combining these terms gives the flow through the configuration $(\mathbf{k}, \mathbf{m})$ due to positive transitions on $j$-type edges. Note that we typically impose boundary conditions, if not because they arise naturally in many temporal network models, because we require the configurations space to remain finite (see Table 2.1). If $n$ remains the number of allowed edge states, we impose the condition that one cannot observe a positive edge transition on an $n$-type edge. As such, a node cannot

lose an $n$-type edge through a positive edge transition, and we write $\mu_n = 0$. Positive edge transitions combine to give

$$ds^{pos}_{\mathbf{k},\mathbf{m}} = \sum_j \Big[ -\mu_j(k_j - m_j)s_{\mathbf{k},\mathbf{m}} - \mu_j m_j s_{\mathbf{k},\mathbf{m}} + \mu_j(k_j - m_j + 1)s_{\mathbf{k}-\Delta_j,\mathbf{m}} + \mu_j(m_j + 1)s_{\mathbf{k}-\Delta_j,\mathbf{m}-\Delta_j} \Big] dt \tag{2.35}$$

and

$$di^{pos}_{\mathbf{k},\mathbf{m}} = \sum_j \Big[ -\mu_j(k_j - m_j)i_{\mathbf{k},\mathbf{m}} - \mu_j m_j i_{\mathbf{k},\mathbf{m}} + \mu_j(k_j - m_j + 1)i_{\mathbf{k}-\Delta_j,\mathbf{m}} + \mu_j(m_j + 1)i_{\mathbf{k}-\Delta_j,\mathbf{m}-\Delta_j} \Big] dt \tag{2.36}$$

which give the positive edge transition terms for $\mathbf{s}$ and $\mathbf{i}$, respectively.

### 2.1.11 *Negative edge transitions*

A negative edge transition refers to the change in a node's configuration when an adjacent event is forgotten over an interval $dt$, causing a decrease in the number of memorable events on that edge. As defined above, this occurs with probability $\nu_j dt$. When an event terminates on an edge of type $j$, it gains an edge of type $j-1$, and loses and edge of type $j$, conserving the total degree $k$.

The relations between node classes due to negative edge transitions are the reverse of their positive counterparts. We write them below for completeness.

A node in configuration $(\mathbf{k}, \mathbf{m})$ moves to the class $(\mathbf{k} - \Delta_{j-1}, \mathbf{m})$, if an event on a $j$-type edge terminates while connected to a inactive neighbour, such that

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k}-\Delta_{j-1},\mathbf{m}})dt = \nu_j(k_j - m_j)dt. \tag{2.37}$$

If the $j$-type neighbour is active, we have

$$W(S_{\mathbf{k},\mathbf{m}} \to S_{\mathbf{k}-\Delta_{j-1},\mathbf{m}-\Delta_{j-1}})dt = \nu_j m_j dt. \tag{2.38}$$

Similarly, nodes may enter the configuration $(\mathbf{k}, \mathbf{m})$ with a negative edge transition via the classes $(\mathbf{k} + \Delta_j, \mathbf{m})$ and $(\mathbf{k} + \Delta_j, \mathbf{m} + \Delta_j)$ as follows. If the event terminates between a node and a inactive neighbour, we have

$$W(S_{\mathbf{k}+\Delta_{j-1},\mathbf{m}} \to S_{\mathbf{k},\mathbf{m}})dt = \nu_j(k_j - m_j + 1)dt, \tag{2.39}$$

and if the event terminates between a node and an active neighbour, we have

$$W(S_{\mathbf{k}+\Delta_{j-1},\mathbf{m}+\Delta_{j-1}} \to S_{\mathbf{k},\mathbf{m}})dt = \nu_j(m_j+1)dt. \tag{2.40}$$

Combining these terms gives the flow through the configuration $(\mathbf{k},\mathbf{m})$ due to negative edge transitions. Note the boundary condition, namely that a $j=0$ edge is the case where no events have taken place in the last $\eta$ interval. As such, a node cannot lose a 0-type edge through a negative edge transition, and we reflect this by writing $\nu_0 = 0$. Negative edge transitions combine to give the term

$$\begin{aligned} ds^{neg}_{\mathbf{k},\mathbf{m}} = \sum_j \Big[ & -\nu_j(k_j - m_j)s_{\mathbf{k},\mathbf{m}} - \nu_j m_j s_{\mathbf{k},\mathbf{m}} \\ & +\nu_j(k_j - m_j + 1)s_{\mathbf{k}+\Delta_{j-1},\mathbf{m}} + \nu_j(m_j+1)s_{\mathbf{k}-\Delta_{j-1},\mathbf{m}-\Delta_{j-1}} \Big] dt \end{aligned} \tag{2.41}$$

and

$$\begin{aligned} di^{neg}_{\mathbf{k},\mathbf{m}} = \sum_j \Big[ & -\nu_j(k_j - m_j)i_{\mathbf{k},\mathbf{m}} - \nu_j m_j i_{\mathbf{k},\mathbf{m}} \\ & +\nu_j(k_j - m_j + 1)i_{\mathbf{k}+\Delta_{j-1},\mathbf{m}} + \nu_j(m_j+1)i_{\mathbf{k}-\Delta_{j-1},\mathbf{m}-\Delta_{j-1}} \Big] dt \end{aligned} \tag{2.42}$$

which we combine with the other terms below.

#### 2.1.11.1 *Combining the configuration flows*

Combining these terms for positive and negative edge transitions, the total flux through the class $(\mathbf{k},\mathbf{m})$ due to changes in edge state can be written

$$dS^{edge}_{\mathbf{k},\mathbf{m}} = dS^{pos}_{\mathbf{k},\mathbf{m}} + dS^{neg}_{\mathbf{k},\mathbf{m}}. \tag{2.43}$$

As such, the rate equation governing the configuration $(\mathbf{k},\mathbf{m})$ can be written

$$\frac{d}{dt}\mathbf{s} = (A^s_{ego} + A^s_{neigh} + A^s_{edge})\mathbf{s} = A_s\mathbf{s} \tag{2.44}$$

$$\frac{d}{dt}\mathbf{i} = (A^i_{ego} + A^i_{neigh} + A^i_{edge})\mathbf{i} = A_i\mathbf{i} \tag{2.45}$$

where $S^{ego}$ transitions are those due to infection and recovery of nodes, $S^{neigh}$ are transitions due to neighbour infection and recovery, and $S^{edge}$ are transitions due to edge activity. Clearly, if edge state is fixed, as is the case in static networks, then $\mu_j = \nu_j = 0$, then we

have $S^{edge}_{\mathbf{k},\mathbf{m}} = I^{edge}_{\mathbf{k},\mathbf{m}} = 0$. This is the case for our studies of weighted and multiplex networks in Chapter 3 and 4, where we allow edge heterogeneity, $n > 1$, in order to study different weight and layer configurations. Setting $n = 1$ recovers the static edge-homogeneous master equations [86, 87].

The explicit equations above can be viewed as the result of the matrix, vector product $s_i = \sum_j a_{ij} s_j$. Also, diagonal terms are necessarily negative and correspond to outflow from $s_i$, Off-diagonal terms are positive, and correspond to inflow from other classes. In other words, $a_{ij} = (-1)^{\delta_{ij}} |a_{ij}|$, where $\delta_{ij}$ is the Kronecker delta. Explain that basically, the matrix representation is necessary for the numerical implementation.

## 2.2 REDUCED DIMENSION MASTER EQUATIONS

As seen in Section 2.1.5, the system of ODEs resulting from a $(\mathbf{k}, \mathbf{m})$ partition of configuration space is exponentially large in $n$, the dimension of the vectors $\mathbf{k}$ and $\mathbf{m}$. Such as system may still be practical, i.e., much faster than the corresponding Monte Carlo simulation, especially since low dimensional systems can be very useful (we study a non-recovery system, with $k = 3$ and $n = 2$, yielding just 20 equations, which is extremely fast to solve). However, for larger $k$ and $n$, the number of equations $|C|$ quickly becomes overwhelming, in terms of memory, and the Runge-Kutta solution in Section 1.7.1 is linear (in space and time) in system size. In this section we extend a technique introduced in [87] to decrease the system size by transforming the ODE variables.

Note that this is applicable only to threshold models,[8] so nodes cannot recover and we do not need to solve for $i_{\mathbf{k},\mathbf{m}}$. We can write such a system explicitly, with the $|C|$ ordinary differential equations

$$\dot{s}_{\mathbf{k},\mathbf{m}} = -F_{\mathbf{k},\mathbf{m}} s_{\mathbf{k},\mathbf{m}} - \sum_{j=1}^{n} \beta_j^s (k_j - m_j) s_{\mathbf{k},\mathbf{m}} + \sum_{j=1}^{n} \beta_j^s (k_j - m_j + 1) s_{\mathbf{k},\mathbf{m}-\mathbf{e}_j}. \tag{2.46}$$

To reduce the dimension of Eqn. 2.46, we need to consider system-wide quantities that are more aggregated than $s_{\mathbf{k},\mathbf{m}}$. One is the probability that a randomly chosen node is active, i.e., the fraction of active nodes in the network, Eqn. 2.5. The other is the probability $\nu_j(t)$ that a

8. More precisely, we mean any model that can be expressed in terms of a step-function adoption rate. This includes threshold models that allow spontaneous adoption, that are built of conditionals, such as the *and* and *or* multiplex rules, as well as relative and absolute threshold rules

randomly chosen neighbour (across a $j$-type edge) of an inactive node is active,

$$\nu_j(t) = \sum_{\mathbf{k}} p_{\mathbf{k}} \frac{\sum_{\mathbf{m}} m_j s_{\mathbf{k},\mathbf{m}}(t)}{\sum_{\mathbf{m}} k_j s_{\mathbf{k},\mathbf{m}}(t)}. \tag{2.47}$$

We start by proposing an exact solution for the AME system in form of the ansatz

$$s_{\mathbf{k},\mathbf{m}}(t) = e^{-pt} \prod_{j=1}^{n} B_{k_j,m_j}[\nu_j(t)] \quad \text{for} \quad S_{\mathbf{k},\mathbf{m}} \in S|_{F=p}, \tag{2.48}$$

where $B_{k_j,m_j}$ is the binomial distribution[9] and $\nu_j$ is given by Eqn. 2.47. The meaning of the ansatz in Eqn. 2.48 is quite intuitive and considers two processes. First, an inactive agent with $k_j$ edges of type $j$ and $m_j$ active neighbours across these edges is connected to the $m_j$ active nodes with the binomially distributed probability $B_{k_j,m_j}(\nu_j)$. Second, for $q_m < \phi q_k$ an inactive node does not fulfil the threshold rule and can only get active spontaneously with probability $e^{-pt}$, since the system is progressively filled via spontaneous infection. Considering these processes as independent we end up with the product in Eqn. 2.48.

The next step is to insert the ansatz (2.48) into the AME system, written explicitly in Eqn. 2.46, and derive a set of ODEs for the aggregated quantities $\rho$ and $\nu_j$. Taking the time derivative of Eqn. 2.48 (i.e. the left-hand side of Eqn. 2.46) we get

$$\dot{s}_{\mathbf{k},\mathbf{m}} = \left( \sum_{j=1}^{n} \left[ \frac{m_j}{\nu_j} - \frac{k_j - m_j}{1 - \nu_j} \right] \dot{\nu}_j - p \right) s_{\mathbf{k},\mathbf{m}}. \tag{2.49}$$

Then, we use the threshold rule given by Eqn. 1.30, those classes $(\mathbf{k},\mathbf{m})$ for which $q_m < \phi q_k$, the ansatz (2.48) and the binomial identity

$$B_{k_j,m_j-1}(\nu_j) = \frac{1-\nu_j}{\nu_j} \frac{m_j}{k_j - m_j + 1} B_{k_j,m_j}(\nu_j), \tag{2.50}$$

in the right-hand side of Eqn. 2.46 to obtain

$$\left[ -p + \sum_{j=1}^{n} \beta_j^s \left( m_j - k_j + \frac{1-\nu_j}{\nu_j} m_j \right) \right] s_{\mathbf{k},\mathbf{m}} = \\ -F_{\mathbf{k},\mathbf{m}} s_{\mathbf{k},\mathbf{m}} - \sum_{j=1}^{n} \beta_j^s (k_j - m_j) s_{\mathbf{k},\mathbf{m}} + \sum_{j=1}^{n} \beta_j^s (k_j - m_j + 1) s_{\mathbf{k},\mathbf{m}-\mathbf{e}_j}. \tag{2.51}$$

9. The binomial term is given by

$$B_{k_j,m_j}(q) = \binom{k_j}{m_j} q^{m_j} (1-q)^{k_j - m_j}.$$

Equating Eqns. 2.49 and 2.51 as in the AME system (2.46), and separating terms for a given value of $j$ from the rest ($i \neq j$) leads to

$$\frac{(1-\nu_j)m_j + \nu_j(m_j - k_j)}{\nu_j}\left(\frac{\dot{\nu}_j}{1-\nu_j} - \beta_j^s\right) = \sum_{i\neq j}^{n} \frac{(1-\nu_i)m_i + \nu_i(m_i - k_i)}{\nu_i}\left(\beta_i^s - \frac{\dot{\nu}_i}{1-\nu_i}\right). \tag{2.52}$$

Since the left-hand side of Eqn. 2.52 depends on the function $\nu_j$ and its derivative only, while the right-hand side depends on the rest of the functions $\nu_i$, both sides must be equal to some constant $c_j$. For the left-hand side, this means that

$$\frac{\dot{\nu}_j}{1-\nu_j} - \beta_j^s = c_j \frac{\nu_j}{m_j - \nu_j k_j} \quad \forall m_j, k_j. \tag{2.53}$$

For the ODE (2.53) to hold regardless of the values of $m_j$ and $k_j$, we need $c_j = 0$. Then, the condition on $\nu_j$ such that the ansatz (2.48) is a solution of Eqn. 2.46 is

$$\frac{\dot{\nu}_j}{1-\nu_j} = \beta_j^s. \tag{2.54}$$

This ODE has the initial condition $\nu_j(0) = 0$, obtained by evaluating Eqn. 2.48 at $t = 0$ and comparing with the expression $B_{k_j,m_j}(0)$, which corresponds to an infinitesimally small initial infection randomly distributed in the network.

Now, we wish to extend a general result derived by Gleeson in [87] [Eqns. (F6)–(F10) therein] to the case of edge heterogeneous networks. We start by multiplying Eqn. 2.46 by $p_{\mathbf{k}}(k_j - m_j)$ and summing over all classes $(\mathbf{k}, \mathbf{m})$,

$$\frac{d}{dt}\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}} = -\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)F_{\mathbf{k},\mathbf{m}}s_{\mathbf{k},\mathbf{m}} - \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} \sum_{i=1}^{n} \beta_i^s (k_j - m_j)\left[(k_i - m_i)s_{\mathbf{k},\mathbf{m}} - (k_i - m_i + 1)s_{\mathbf{k},\mathbf{m}-\mathbf{e}_i}\right]. \tag{2.55}$$

From the definition of $\beta_j^s$ in Eqn. 2.25, the first term on the right hand side of Eqn. 2.55 may be written as

$$-\beta_j^s \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}. \tag{2.56}$$

As for the second term on the right hand side, when $i = j$ the term telescopes to Eqn. 2.56, and for $i \neq j$ it telescopes to 0. Overall, we can rearrange Eqn. 2.55 and obtain

$$\beta_j^s = -\frac{1}{2}\frac{d}{dt}\ln \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}. \tag{2.57}$$

Since $\beta_j^s = -\frac{d}{dt}\ln(1-\nu_j)$ from Eqn. 2.54, equating Eqns. 2.54 and 2.57 implies that

$$d_j(1-\nu_j)^2 = \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}, \tag{2.58}$$

with $d_j$ a constant that can be determined from initial conditions. By using Eqn. 2.48 we have $s_{\mathbf{k},\mathbf{m}}(0) = \prod B_{k_j,m_j}(0)$, and since $\nu_j(0) = 0$ and $B_{k_i,m_i}(0) = \delta_{m_i,0}$ with $\delta_{ij}$ the Kronecker delta, we have

$$d_j = \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)\prod_{i=1}^{n} B_{k_i,m_i}(0) = \sum_{\mathbf{k}} p_{\mathbf{k}}k_j = z_j, \tag{2.59}$$

where $z_j$ is the average number of $j$-type edges a node has in the network, or average $j$-degree. Thus,

$$\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}} = z_j(1-\nu_j)^2. \tag{2.60}$$

The next step is to use Eqn. 2.60 to find a new expression for $\beta_j^s$ and thus write the ODE (2.54) explicitly in terms of $\nu_j$. Noting that the left-hand side of Eqn. 2.60 is the denominator in the definition (2.25) of $\beta_j^s$, we get

$$\begin{aligned}
\beta_j^s =& \frac{1}{z_j(1-\nu_j)^2}\left[p\sum_{S|_{F=p}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}} + \sum_{S|_{F=1}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}\right] \\
=& \frac{1}{z_j(1-\nu_j)^2}\left[z_j(1-\nu_j)^2 - (1-p)\sum_{S|_{F=1}} p_{\mathbf{k}}(k_j - m_j)s_{\mathbf{k},\mathbf{m}}\right] \\
=& \frac{1}{1-\nu_j}\left[1-\nu_j - (1-p)e^{-pt}\sum_{S|_{F=p}} p_{\mathbf{k}}\frac{k_j}{z_j}B_{k_j-1,m_j}(\nu_j)\prod_{i\neq j}^{n} B_{k_i,m_i}(\nu_i)\right],
\end{aligned}$$

where the sums $\sum_{S|_{F=p}}$ and $\sum_{S|_{F=1}}$ run over all sets $S_{\mathbf{k},\mathbf{m}}$ that lead to adoption rates $F_{\mathbf{k},\mathbf{m}} = p$ and $F_{\mathbf{k},\mathbf{m}} = 1$, respectively. Essentially, we're exploiting the fact that the threshold condition partitions the network into two classes - those that satisfy the threshold rule, the set

of configurations $S|_{F=1}$ and those that don't, the set of configurations $S|_{F=p}$. We have also inserted the ansatz (2.48) and the binomial identity

$$B_{k_j,m_j}(\nu_j) = \frac{k_j(1-\nu_j)}{k_j - m_j} B_{k_j-1,m_j}(\nu_j) \tag{2.61}$$

to simplify the expression of $\beta_j^s$. Moreover, it is useful to introduce the *response function* of the monotone, threshold-driven dynamics of our model,

$$f(\mathbf{k},\mathbf{m}) = \begin{cases} 0 & F_{\mathbf{k},\mathbf{m}} = p \\ 1 & F_{\mathbf{k},\mathbf{m}} = 1, \end{cases} \tag{2.62}$$

a function that activates when a $(\mathbf{k},\mathbf{m})$-class node fulfils the threshold condition and gets active, in order to invert the restricted sum of Eqn. 2.61,

$$\sum_{S|_{F=p}} B_{k_j-1,m_j}(\nu_j) \prod_{i\neq j}^{n} B_{k_i,m_i}(\nu_i) = 1 - \sum_{S|_{F=1}} B_{k_j-1,m_j}(\nu_j) \prod_{i\neq j}^{n} B_{k_i,m_i}(\nu_i). \tag{2.63}$$

Overall, comparing Eqns. 2.54 and 2.61, we can write an explicit ODE for $\nu_j$,

$$\frac{d}{dt}\nu_j = g_j(\boldsymbol{\nu}, t) - \nu_j, \tag{2.64}$$

with $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)^{\mathrm{T}}$, $j = 1, \ldots, n$, and the function $g_j(\boldsymbol{\nu}, t)$ given by

$$g_j(\boldsymbol{\nu}, t) = f_t + (1 - f_t) \sum_{S|_{F=1}} p_{\mathbf{k}} \frac{k_j}{z_j} B_{k_j-1,m_j}(\nu_j) \prod_{i\neq j}^{n} B_{k_i,m_i}(\nu_i), \tag{2.65}$$

where we have defined $f_t = 1 - (1-p)e^{-pt}$.

Even though the system (2.64) is closed and in this sense equivalent to Eqn. 2.46, we may also derive a corresponding ODE for $\rho$, since we are mainly interested in the temporal evolution of the fraction of adopters in the network. Using Eqns. 2.5 and 2.46 we have

$$\begin{aligned} \dot{\rho} = -\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} \dot{s}_{\mathbf{k},\mathbf{m}} &= \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} F_{\mathbf{k},\mathbf{m}} s_{\mathbf{k},\mathbf{m}} \\ &+ \sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} \sum_{j=1}^{n} \beta_j^s \left[ (k_j - m_j) s_{\mathbf{k},\mathbf{m}} - (k_j - m_j + 1) s_{\mathbf{k},\mathbf{m}-\mathbf{e}_j} \right], \end{aligned} \tag{2.66}$$

where the second term in the right-hand side telescopes to zero. Then, we use an algebraic manipulation similar to that of Eqn. 2.61 to obtain

$$\begin{aligned}\dot{\rho} &= p \sum_{S|_{F=p}} p_{\mathbf{k}} s_{\mathbf{k},\mathbf{m}} + \sum_{S|_{F=1}} p_{\mathbf{k}} s_{\mathbf{k},\mathbf{m}} \\ &= 1 - \rho - (1-p)e^{-pt} \sum_{S|_{F=p}} p_{\mathbf{k}} \prod_{j=1}^{n} B_{k_j,m_j}(\nu_j). \end{aligned} \tag{2.67}$$

Thus, the ODE for $\rho$ is

$$\frac{d}{dt}\rho = h(\boldsymbol{\nu}, t) - \rho, \tag{2.68}$$

where the function $h(\boldsymbol{\nu}, t)$ is given by

$$h(\boldsymbol{\nu}, t) = f_t + (1 - f_t) \sum_{S|_{F=1}} p_{\mathbf{k}} \prod_{j=1}^{n} B_{k_j,m_j}(\nu_j). \tag{2.69}$$

Combining all of these results, the AME system (2.46) is reduced to the closed system of $n$ coupled, non-linear ODEs by solving for $\nu_j$, or a system of $n+1$ ODEs by solving for both $\nu_j$ and $\rho$,

$$\dot{\nu}_j = g_j(\boldsymbol{\nu}, t) - \nu_j, \tag{2.70a}$$

$$\dot{\rho} = h(\boldsymbol{\nu}, t) - \rho, \tag{2.70b}$$

with the quantities $g_j(\boldsymbol{\nu}, t)$ and $h(\boldsymbol{\nu}, t)$ given explicitly by Eqns. 2.65 and 2.69. Note that in contrast to the full AMEs, the system is no longer autonomous due the $e^{-pt}$ factors in $\mathbf{g}$ and $h$.

### 2.2.1 *Complexity of reduced system*

The reduced dimension master equation system given by Eqn. 2.70 is of dimension $n$ if one is solving only for $\boldsymbol{\nu}$, or of size $n+1$ when solving simultaneously for both $\boldsymbol{\nu}$ and $\rho$. While the system size is relatively small (with $n = 2$ or 3 in the systems studied in Chapters 3 and 4), each equation involves summing over the entire configuration space. [10] This is very large compared to $n$, growing exponentially as outlined in Table 2.1. In contrast, the corresponding full system, Eqn. 2.46, entails $|C|$ equations, and at most $2n+1$ terms per equation. Thus, the number of operations is of the same order when integrating the two systems is

10. Technically only sums over the subset $S|_{F=1}$, but this is on the order of $|S|$. In practice, this may involve iterating over all $S$ anyway, and determining for each class $(\mathbf{k}, \mathbf{m})$ whether $F = 1$.

**Table 2.2 –** Size comparison of the full and reduced master equation systems. Since the reduced solution, Eqn. 2.70a, assumes a threshold dynamic which is monotonic, the full solution we compare to is Eqn. 2.46.

| | system size | †terms per equation |
|---|---|---|
| full solution | $\lvert C\rvert$ | $\mathcal{O}(2n+1)$ |
| reduced solution | $n$ | $\mathcal{O}(\lvert C\rvert)$ |

† the multiplicative constant for the big $\mathcal{O}$ is 1, $2n+1$ and $|C|$ being strict upper bounds.

approximately the same, with the reduced system shifting the burden of the number of equations $|C|$ to the sum of terms on the right hand side of the equation. Quite simply, instead of having a large system with a small number of terms to sum in each equation, the reduced system has a small number of equations, each involving a sum of a very large number of terms.

Despite this seeming equivalence, note that there are potential savings in space, because it is not necessary to store the entire set of configurations $C$, which can be immensely costly in heterogeneous systems. If one knows the support of the distribution $p_{\mathbf{k}}$ as well as $n$, one can construct the set of configurations on-the-fly, and thus calculate the terms $p_{\mathbf{k}}$, $k_j/z_j$, $B_{k_j-1,m_j}(\nu_j)$ and $B_{k_i,m_i}(\nu_i)$ in Eqn. 2.65. That is, we store only the $n$ values of $\boldsymbol{\nu}$, and generate the full set of $(\mathbf{k},\mathbf{m})$ at each Runge-Kutta step. However, if savings in space are intended to be made in this way, one must sacrifice the large amount of time required to repeatedly generate the space $C$.

A much more concrete saving afforded by the reduced system is in regards to its stability analysis. That is, if the dynamical system in the full solution is described by the density of states $\mathbf{s}$, and its trajectory $\mathbf{s}(t)$, then the reduced solution involves the study of $\boldsymbol{\nu}(t)$. Studying these trajectories under small perturbations of the initial conditions sheds light on the stability of the dynamical system. This qualitative analysis amounts to studying a well-known problem involving eigenvalues of matrices. In the case of the full solution, such matrices are of size $|C|\times|C|$, but only $n\times n$ for the reduced solution. Both reveal system stability in theory, but in practice, finding the eigenvalues of matrix of rank $|C|$ is prohibitive, since $|C|$ can grow to the order of $10^{27}$, for example, in a network allowing $n=4$ edge types, and degrees up to $k=10^4$ (see Figs. 2.2 and 2.3). In contrast, stability analysis is well within reach of the reduced system, which has size $n$, and we turn our attention to this in the next section.

## 2.3 QUALITATIVE METHODS

If we do not desire the entire spreading curve, for example if the size of the space $|C|$ is too large to perform Runge-Kutta integration of the full or reduced master equations, then we can turn to qualitative methods that exploit the low dimension of the reduced system.

### 2.3.1 *Linear stability analysis*

As discussed in the previous section, it is possible to study the master equation system qualitatively via stability theory. In so doing, we derive a cascade condition, as has been done previously for the Watts model [163] and for complex contagion in homogeneous networks [166]. First, note that it is sufficient to consider Eqn. 2.70a and determine the stability of the trajectories $\boldsymbol{\nu}(t)$. The stability or otherwise of $\boldsymbol{\nu}$ informs us of the stability of $\rho$. Then, we observe that by assuming $p = 0$, $f_t = 0$, removing the explicit time dependence in the function $g_j$, such that $g_j = g_j(\boldsymbol{\nu})$. Then, the system simplifies to

$$\dot{\nu}_j = g_j(\boldsymbol{\nu}) - \nu_j, \tag{2.71}$$

where

$$g_j(\boldsymbol{\nu}) = \sum_{S|_{F=1}} p_{\mathbf{k}} \frac{k_j}{z_j} B_{k_j-1,m_j}(\nu_j) \prod_{i\neq j}^{n} B_{k_i,m_i}(\nu_i). \tag{2.72}$$

We perform a linear stability analysis of the reduced AME system in Eqn. 2.70a around the fixed point $\boldsymbol{\nu}^* = \mathbf{0}$, corresponding to a total lack of infection. If $\boldsymbol{\nu}^*$ is unstable, then a small perturbation (such as a single active node at $t = 0$) can drive the system out of equilibrium and create a global cascade of infection where $\rho > 0$, that is, a system where a non-vanishing fraction of nodes is active in the limit $N \to \infty$. Since the system (2.71) is closed, the stability of Eqn. 2.70 is determined by the stability of this equation at $\boldsymbol{\nu}^* = \mathbf{0}$. According to linear stability theory, a local instability exists if the Jacobian matrix of the system evaluated at the fixed point,

$$J^*_{ji} = -\delta_{ji} + \left.\frac{\partial g_j(\boldsymbol{\nu})}{\partial \nu_i}\right|_{\boldsymbol{\nu}=\boldsymbol{\nu}^*}, \tag{2.73}$$

has at least one eigenvalue with a real part larger than zero. We can write the partial derivative in Eqn. 2.73 explicitly by expanding the $B_{k_i,m_i}(\nu_i)$ factors in Eqn. 2.72,

$$\frac{\partial g_j}{\partial \nu_i} = \begin{cases} \sum_{\mathbf{k},\mathbf{m}} \frac{k_j}{c_j} p_{\mathbf{k}} f_{\mathbf{k},\mathbf{m}} \dot{B}_{k_j-1,m_j}(\nu_j) \prod_{i\neq j} B_{k_i,m_i}(\nu_i) & j = i \\ \sum_{\mathbf{k},\mathbf{m}} \frac{k_j}{c_j} p_{\mathbf{k}} f_{\mathbf{k},\mathbf{m}} B_{k_j-1,m_j}(\nu_j) \dot{B}_{k_i,m_i}(\nu_i) \prod_{l\neq j,i} B_{k_l,m_l}(\nu_l) & j \neq i \end{cases}, \tag{2.74}$$

where

$$\dot{B}_{k_j-1,m_j}(\nu_j) = \binom{k_j-1}{m_j} \left[ m_j \nu_j^{m_j-1} (1-\nu_j)^{k_j-1-m_j} - (k_j-1-m_j) \nu_j^{m_j} (1-\nu_j)^{k_j-2-m_j} \right] \tag{2.75}$$

and $\dot{B}_{k_i,m_i}(\nu_i)$ is written similarly (by making the changes $j \to i$ and $k_j - 1 \to k_i$). Then, for $j = i$ we analyse terms in the sum over $\mathbf{m}$ at the fixed point $\boldsymbol{\nu}^* = \mathbf{0}$. That is, for $m_j = 0$ we have $\dot{B}_{k_j-1,0}(0) = 1 - k_j$, but since $B_{k_i,m_i}(0) = \delta_{m_i,0}$ and $f(\mathbf{k}, \mathbf{0}) = \mathbf{0}$ for $\phi > 0$ (for all threshold rules), the associated term in Eqn. 2.74 is zero. For $m_j = 1$ we have $\dot{B}_{k_j-1,1}(0) = k_j - 1$. Finally, for $m_j > 1$ we get $\dot{B}_{k_j-1,m_j}(0) = 0$, so the only non-zero term corresponds to $m_j = 1$. By a similar argument, for $j \neq i$ the only surviving term in Eqn. 2.74 is $\dot{B}_{k_i,1}(0) = k_i$ (for $m_i = 1$).

Combining these results, we can write 2.73 explicitly as

$$J^*_{ji} = -\delta_{ji} + \sum_{\mathbf{k}} \frac{k_j}{c_j} (k_i - \delta_{ji}) p_{\mathbf{k}} f(\mathbf{k}, \mathbf{e}_i), \tag{2.76}$$

where $\mathbf{e}_i$ is the $i$-th basis vector of dimension $n$. The Jacobian matrix $\mathbf{J}^*$ of Eqn. 2.76 encodes the structure of the network, namely the distribution of degrees and heterogeneous quantities such as weights and layer via $p_{\mathbf{k}}$, as well as node dynamics via the response function $f(\mathbf{k}, \mathbf{e}_i)$, which provides the response of a node with degree vector $\mathbf{k}$ to a single active neighbour across an $i$-type edge. The eigenvalues $\lambda_j$ of $\mathbf{J}^*$ are obtained by solving the characteristic equation $\det(\mathbf{J}^* - \lambda \mathbf{1}) = 0$. Then, the cascade condition for complex contagion over multiplex networks (in the case $p = 0$) is

$$\mathrm{Re}(\lambda_j) > 0 \tag{2.77}$$

for at least one $j \in \{1, \ldots, n\}$. Even though we cannot write an algebraic formula for $\lambda_j$ when $n > 4$, we can compute the eigenvalues numerically. We may also find an explicit expression for the cascade condition in simple cases such as when $n = 1, 2$ or 3. This is done in Chapters 3, 4 and 5.

**Table 2.3 –** We provide some approximate bounds on linear stability and phase space analyses of our system of ODEs. These provide upper bounds, asymptotic in the maximum degree of $k$, with constant edge heterogeneity $n$.

| | upper bound |
|---|---|
| linear stability | $\mathcal{O}(k^n)$ |
| vector field | $\mathcal{O}(k^{2n})$ |

### 2.3.2 *Phase space analysis*

Another related qualitative method that we can use to gain insight to the system (2.70) is phase space analysis. Since our system is entirely described by the probabilities $\boldsymbol{\nu}$, the time evolution of the dynamical process is given by paths through the $n$-dimensional space spanned by $\nu_j$. In other words, given some initial condition $\boldsymbol{\nu}^*$, the dynamical process is represented by a trajectory in $\boldsymbol{\nu}$-space. Taking the set of possible initial conditions gives the *flows* through the phase space of the system.

*phase space*

Instead of finding the trajectory itself [this would amount to integrating Eqn. 2.70a over time to obtain the curves $\boldsymbol{\nu}(t)$], it can be easier to study the vector field associated with the velocities of the trajectories, in other words, the tangents $\dot{\boldsymbol{\nu}}$, around a fixed point. When that point corresponds to the initial conditions, we can gain a picture of the stability of the system. Graphically, the stability of the initial conditions can be easily identified; limit cycles as opposed to a divergent velocity field.

Concretely, to find $\dot{\boldsymbol{\nu}}$ we evaluate $g_j(\boldsymbol{\nu}, t)$, given by Eqn. 2.65 at $(\boldsymbol{\nu}^*, 0)$, for a region about $\boldsymbol{\nu}^* = \mathbf{0}$. Linear stability analysis, see Section 2.3.1, via the eigenvalues of the linearised system, implicitly contains the information that we obtain by visualising the velocity field. In particular, combinations of signs of the eigenvalues of the Jacobian describe the velocity field.

### 2.3.3 *Complexity of qualitative methods*

To determine the computational complexity of linear stability analysis, we recall that we require the $n^2$ elements of the Jacobian $J^*$, as per Eqn. 2.76. The computation of each element $J^*_{ij}$ assumes that all sets $S_{\mathbf{k},\mathbf{m}}$ are unpopulated, except for those with $\mathbf{m} = \mathbf{0}$. This is simply the initial condition of $\boldsymbol{\nu} = 0$ at time $t = 0$, and is why there is no sum over

$\mathbf{m}$ in Eqn. 2.76. As a result, it is sufficient to sum over the classes $(\mathbf{k}, \mathbf{0})$. This was precisely our assumption in Section 2.1.5, when estimating a lower bound, asymptotically in $k$, on the number of classes with total degree $k$. This was shown to be $k^{n-1}$, and assuming that each $k' < k$ has this maximum amount provides an upper bound on the number of $\mathbf{k}$ across the network. Thus, $k \times k^{n-1} = k^n$ provides the order of terms required for evaluating each element in $J^*$. We then perform an eigenvalue decomposition of $J^*$, which assuming a QR algorithm [11] requires $\mathcal{O}(n^3)$ floating point operations.

In contrast to linear stability analysis, vector field analysis of $\boldsymbol{\nu}$-phase space involves iterating over the full set of classes $(\mathbf{k}, \mathbf{m})$, if not to include them in the sum, to determine whether they satisfy the condition $F_{\mathbf{k},\mathbf{m}} = 1$ (see Eqn. 2.72). As we saw in Section 2.1.5, such an iteration entails $\mathcal{O}(k^{2n})$ classes. Note that despite the vector field usually being about the fixed point $\boldsymbol{\nu} = \mathbf{0}$, the entire configuration space will be in fact be populated. This is why we cannot apply the $(\mathbf{k}, \mathbf{0})$ reduction above, and is precisely why this analysis works.

Finally, it is useful to underline that one only has to cycle through the relevant configurations once, meaning we do not have to store their values in memory. This is in contrast to Runge-Kutta methods for finding the trajectories $\boldsymbol{\nu}(t)$ over time, which involves $\frac{T}{\Delta t}$ iterations through $C$, if $T$ is the time taken to reach steady state, and $\Delta t$ is the time step. These results are summarised in Table 2.3.

## 2.4 COMPARISON WITH OTHER METHODS

Here we briefly review alternative mathematical methods frequently employed in the literature to model spreading processes on networks. Each differs in some way from our master equation approach, and we highlight the different settings and applications in which such methods are useful. We also briefly mention how a pair-approximation method, as well as a mean field approximation can be derived from our master equation method.

### 2.4.1 *Branching processes*

A central assumption in our master equation derivation is that the dynamics occur on infinite networks, where spreading outcomes that

11. The QR algorithm computes the so-called Schur decomposition of a matrix. It is the workhorse of eigenvalue decomposition, especially for dense matrices. In our case, $J^*$ is clearly dense in general. This can be seen by considering the summand in Eqn. 2.76, which will clearly be non-zero in general. See [197] for more details.

are small and finite in experiment correspond may correspond to $\rho = 0$ analytically (see Section 4.6 for further discussion). As such, we simulate networks that are as large as possible, observing increasing agreement between Monte Carlo simulation and analytic solution as network size grows. In particular, we see reduced fluctuations about the average behaviour. On smaller networks, stochastic effects are important in determining the final outcome. Different approaches are thus required to describe the distribution of results among realisations, and effect that a system of ordinary differential equations are unable to capture. A typical solution here is to use branching processes [99].

Branching processes is a broad term that encompasses a number of analytical techniques. However, all such techniques use some kind of tree structure approximation. Initial works, such as [198] and [149] don't even use the term "branching process", but use its key features of tree-like assumptions and generating functions. Since these works, a large number of studies have used similar techniques to extend to different settings [89, 91–93].

Following the treatment in [89], we outline some of the key quantities that arise in branching process analyses. The offspring distribution, denoted by $q_k$, where $k \in \{0, 1, 2, ...\}$. Roughly speaking, this distribution gives the likelihood that if a cascade of activation reaches a node (i.e., if one of the node's neighbours becomes active), that the node will activate and expose $k$ other neighbouring nodes to potential activation. For an undirected network,

$$\hat{q}_k = \frac{k+1}{z} p_{k+1} v_{k+1}. \tag{2.78}$$

This uses the probability $\frac{k+1}{z} p_{k+1}$ that a randomly selected neighbour of a node in a configuration model network has degree $k+1$. The quantity $v_k$ is the probability that a node $i$ of degree $k$ is vulnerable, meaning that the activation of a single neighbouring node (at time $t_1$) will lead to the activation of node $i$ at some time $t > t_1$, assuming no other neighbour of $i$ becomes active by time $t$. Then, $r = \sum_k \hat{q}_k$ is the probability that a node reached by travelling along a random edge is vulnerable. This is then related to the offspring distribution as follows. An exposed vulnerable node has $k'$ inactive neighbours with probability $\hat{q}/r$, and since each inactive neighbour is vulnerable with probability $r$, the offspring distribution is given by

$$q_k = \frac{1}{r} \sum_{k'=k}^{\infty} \hat{q}_{k'} \binom{k'}{k} r^k (1-r)^{k'-k} \tag{2.79}$$

Finally, the branching number defines whether the dynamical process on a given network is subcritical, critical, or supercritical. The branching number is the mean of the offspring distribution, in other words, the expected number of "children" per "parent," and it can be expressed as:

$$\xi = \sum_k k q_k = \sum_k k \hat{q}_k. \tag{2.80}$$

Then we have three cases, the subcritical case $\xi < 1$, the critical case $\xi = 1$, and the supercritcal case $\xi > 1$. In the critical case, power-law distributions of avalanche sizes are observed.

This is sufficient to highlight a weakness relative to the AMEs; it's not clear how this handles large thresholds, where more than one active node is necessary, because $v_k$ is by definition for vulnerable nodes. Further, network size is infinite, but branching process is finite. AMEs are technically infinite, but deals in probabilities.

### 2.4.2 *Message-passing methods*

Similar to the use of branching processes is message passing. Message passing solutions for spreading on networks were first introduced in [85, 88], and recently summarised in [90]. Once again, this approach assumes configuration model random graphs with degree distribution given by $p_k$, such that cascades propagate in tree-like dynamics. They appear to have been introduced to solve for seed size, as the original threshold model on networks assumed only an infinitesimal seed $\rho_0$. Unlike branching processes, but like master equations, message passing gives the fraction $\rho$ of active nodes in an infinite network.

The original formulation is in discrete time, with time-step indexed by $n$. The average final fraction of nodes $\rho$ is

$$\rho = \rho_0 + (1-\rho_0) \sum_k p_k \sum_m \binom{k}{m} q_\infty^m (1-q_\infty)^{k-m} f(\tfrac{m}{k}) \tag{2.81}$$

where $q_\infty$ is the fixed point of the recursion relation

$$q_{n+1} = \rho_0 + (1-\rho_0) G(q_n), \quad \text{for} \quad n = 0, 1, 2, \ldots \tag{2.82}$$

with $q_0 = \rho_0$. The function $G$ is given by

$$G(q) = \sum_k \frac{k}{z} p_k \sum_m \binom{k-1}{m} q^m (1-q)^{k-1-m} f(\tfrac{m}{k}). \tag{2.83}$$

Like master equations, message passing solutions can be extended to a large number of different settings, but with the cost of increased

computational complexity. Similar to branching process techniques, message passing can be solved for both discrete and continuous time [85]. Conveniently, message passing recovers the cascade condition of [198], that we also see in Chapter 1.

Finally, it is not clear how this handles large $\phi$. Master equations can also solve for arbitrary seed size, by incorporating it into its initial condition, but with the cost of system size. For small threshold values, message passing is likely to be accurate since it only considers $\phi < \frac{1}{k}$, i.e. $m = 1$. In contrast, a master equation naturally stores all $\mathbf{m}$.

### 2.4.3 *Master stability functions*

A master stability equation is a tool used to determine the stability of a systems of coupled oscillators. This approach was originally in [161], and has since been extended in many other works. A concise outline of its use is given in [163], which in turn follows parts of [146]. It's application to a broad class of dynamics of interest is carried out in [22].

An important difference between our master equation formalism and the master stability functions is that the ODE system size here simply the number of nodes $N$ in the network. We associate each node $i$ with a dynamical variable $x_i$ that describes node state, such that the system evolves according to

$$\frac{dx_i}{dt} = W(x_i(t)) + \sum_{j=1}^{N} A_{ij} Q(x_i(t), x_j(t)), \tag{2.84}$$

where $W(x_i(t))$ describes the "self" dynamics, or dynamics that are independent of the node's neighbours. Then, $A$ is the adjacency matrix and $Q(x_i, x_j)$ describes the dynamical mechanism governing the pairwise interactions. Instead of solving the system 2.84 explicitly, one usually performs linear stability analysis, in a manner similar to Section 2.3.1. Such an approach is attractive, since we reduce the numerical integration of the system to an eigenvalue problem. Here, we directly obtain information regarding stability of equilibrium points, and gain

Another difference between this and our master equation approach is that the form of $x_i$ is unspecified, and can correspond to oscillator phase as well as binary state or compartment. In fact it is straightforward to generalise to $\mathbf{x}_i$, where the state of each node is defined by a vector instead of a scalar.

# Part II

# OVERVIEW OF PUBLICATIONS

In this part, we apply the mathematical methods of Part I to a number of systems. In Chapter 3, we examine threshold driven contagion on weighted networks. In Chapter 4, we extend our study of threshold driven contagion to multiplex networks, where we observe reentrant phase transitions in cascade size due to a fundamental and widespread model of multiplexity. In Chapter 5, we generalise previous applications of master equations to network dynamics by allowing for temporal evolution of the network structure.

# 3

# THRESHOLD-DRIVEN CONTAGION ON WEIGHTED NETWORKS

In this chapter we apply the master equation formalism outlined in Chapter 2 to a static, weighted network system. We examine the dynamics of threshold driven contagion on these networks, following closely the work carried out by Unicomb *et al* in [190]. Both the main text as well as its supporting information are followed closely.

Weighted networks capture the structure of complex systems where interaction strength is meaningful. This information is essential to a large number of processes, such as threshold dynamics, where link weights reflect the amount of influence that neighbours have in determining a node's behaviour. Despite describing numerous cascading phenomena, such as neural firing or social contagion, the modelling of threshold dynamics on weighted networks has been somewhat overlooked. We fill this gap by studying a dynamical threshold model over synthetic and real weighted networks with numerical and analytical tools. We show that the time of cascade emergence depends non-monotonously on weight heterogeneities, which accelerate or decelerate the dynamics, and lead to non-trivial parameter spaces for various degree and weight distributions. As discussed in Chapters 1 and 2, our master equation approach applies to arbitrary binary state processes and link properties, and may prove useful in understanding the role of edge heterogeneities in various natural and social phenomena.

## 3.1 INTRODUCTION TO WEIGHTED NETWORKS

Weighted networks provide meaningful representations of the architecture of a large number of complex systems where interacting entities, represented as nodes in a graph, are connected with links weighted by the strength of their interactions. Weighted networks are ubiquitous across a range of systems, from biological [104] and ecological [139],

to infrastructure [18, 158, 203], social [46, 96, 153, 196], information and economic [7, 101] systems. A number of these were highlighted in Section 1.2. Their analysis has been in focus from the early stages of complex networks research [144, 195], with several measures [154, 155, 174] and models [19, 204] introduced. These studies show that link weights in real networks are usually heterogeneous, may be correlated with the network structure [96, 152], and can even capture signed relationships [183]. More importantly, weights help to differentiate links of varying importance, influence, and role. On a microscopic level, weights identify the most relevant neighbours of a node [168]; on a network level, they indicate links with special roles or positions in the system [96, 152]. Such information is crucial for dynamical processes evolving on weighted networks. Examples can be found in epidemiology, where important ties maintained by frequent interactions may enhance the spread of disease locally, while ties with infrequent interactions but located between densely connected parts of the network may suppress spreading globally [152, 208]. Link weights are also relevant in phenomena like random walks, spin models, synchronization, evolutionary games, as well as cascading failures. Despite this, weighted networks have been less studied than their unweighted counterparts, especially for threshold driven processes, which play an essential role in systems of self-organised criticality [12, 59, 112], epidemiology [115], firing neurons [79, 125, 181], or social contagion [96, 198].

In threshold models on networks links are usually considered unweighted, such that the stimuli or influence arriving from each neighbour contributes uniformly to reaching the behavioural threshold. Although this assumption simplifies their modelling, it does not lead to an accurate representation of real world dynamics. For example, in neural systems synaptic connections have weights that quantify the strength of incoming stimuli, and contribute unequally in bringing neurons to an excited state, as recognised recently in models of neural population dynamics [108]. In social systems, link weights are associated with tie strengths that quantify the influence that individuals have on their peers. Measurement of tie strength is a long-standing challenge, but it is generally accepted that social ties are not equal, as some are more influential than others on one's decision making. Surprisingly, apart from some recent studies [60, 106, 122], weights have been commonly overlooked in models of threshold driven phenomena.

Our aim is to close this gap by exploring the effect of weight heterogeneities on threshold driven contagion processes. We first study a dynamical variant of the Watts cascade model on a simple system, a random regular network with a bimodal weight distribution. We

then provide an analytical solution of the dynamics, for arbitrary degrees and weights, together with numerical simulations and combinatorial arguments to show that the speed of spreading depends non-monotonously on the strength of weight heterogeneities and may radically accelerate or decelerate with respect to the unweighted case, even for fixed thresholds. We also observe this effect under more realistic synthetic scenarios, such as scale-free networks and lognormal weight distributions, as well as in data-driven simulations over large-scale empirical weighted networks. Our contribution is a meaningful step forward in the largely unexplored modelling of dynamical processes with heterogeneous interactions, typical in neural systems and social contagion. Moreover, our results may have broader implications as our methodology is not specific to threshold dynamics and may be easily extended to any binary state process, while our study and conclusions may be useful in accurately modelling other dynamical phenomena over weighted networks.

## 3.2 THRESHOLD MODEL AND APPROXIMATE SOLUTIONS

To study threshold driven dynamical processes over weighted networks we build on the seminal model proposed by Watts [198], which we outlined in Section 1.6.2. Following its standard formulation [119, 166, 176, 198], we define a monotone binary-state dynamics over a weighted, undirected network of size $N$. Degrees take discrete values $k = 0, \dots, N-1$ according to the distribution [1] $p_k$, and edge weights $w > 0$ are continuous variables with distribution $p_w$. The edge weight $w_{uv}$ represents the capacity of connected nodes $u$ and $v$ to influence each other. Denoting the set of neighbours of node $u$ by $\mathcal{N}(u)$, the total influence upon $u$ is $q_k = \sum_{v \in \mathcal{N}(u)} w(u, v)$. Like in other conventional models of spreading dynamics [163], nodes can be in one of two mutually exclusive states, inactive (also called susceptible in the epidemiology literature, initially accounting for all nodes), or active (also called infected in epidemiology literature). An inactive node can become active either spontaneously with rate $p$ [119, 166], or if the influence of its active neighbours exceeds a given threshold $\phi$ ($0 < \phi < 1$) of its total received influence. Restricted to active neighbours, $\mathcal{N}(u)|_I$, node influence is given by $q_m = \sum_{v \in \mathcal{N}(u)|_I} w(u, v)$. If the condition

1. Alternatively, $p_k$ is the probability that a randomly selected node has degree $k$. A problem in these model descriptions is that the distribution $p$ is routinely overloaded, used for degree $k$, weight $w$ and threshold $\phi$, which are written $p_k$, $p_w$, $p_\phi$. Incidentally, we follow the notation of Newman, favouring subscripts to deal with this overloading, writing $p_k$ rather than $P(k)$, for example.

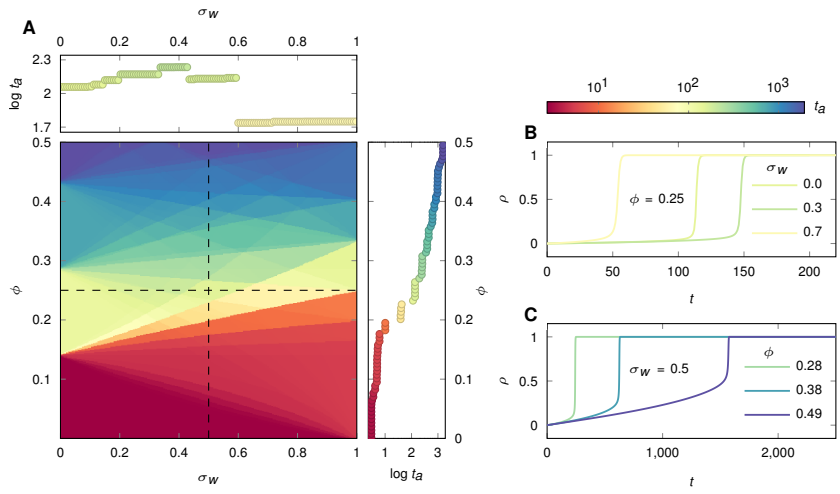


**Figure 3.1 –** Threshold driven contagion and cascade evolution on weighted networks. (a) Parameter dependence of the time $t_a$ of cascade emergence (main panel) on a random regular network with degree $k = 7$, and bimodal weight distribution with mean $\mu_w = 1$ and standard deviation $\sigma_w$ (for further details see text). Cascade speed is measured by the time $t_a$ to reach $\rho = 0.75$ activation. For fixed threshold $\phi$ and varying $\sigma_w$, $t_a$ changes non-monotonously, while for fixed $\sigma_w$ and varying $\phi$, dynamics slows down for increasing $\phi$ (top/right panels, corresponding to horizontal/vertical dashed lines in main panel). (b-c) Spreading time series $\rho(t)$ for selected parameter values in (a). Simulation results in (b-c) are averages of 25 simulations with $p = 2 \times 10^{-4}$ and $N = 10^4$.

$q_m \geq \phi q_k$ is fulfilled, a node becomes active and remains so indefinitely. In the special case of discrete weight distributions, the above expressions can be written $q_k = \mathbf{k} \cdot \mathbf{w}$ and $q_m = \mathbf{m} \cdot \mathbf{w}$. For simplicity we assume that all nodes have the same threshold $\phi$, just as in many other studies [176, 198].

We explore this model analytically using the master equation formalism laid out in Chapter 2. In particular, since we assume discrete edge-weight distributions, each weight value $w_j$ can be directly associated to an edge type $j$. Then, a node's local configuration is entirely described by its class $(\mathbf{k}, \mathbf{m})$. Since the network topology is static, the edge transition rates $\mu_j$ and $\nu_j$ are zero, and the configuration flows are given by the matrices $A_{ego}$ and $A_{neigh}$. In the case of non-recovery dynamics, such as the threshold rule to be explored, the matrix $A$ can be assumed to correspond to the inactive configuration space $S$. In other words, it is sufficient to solve only for the inactive densities $s_{\mathbf{k},\mathbf{m}}$.

In threshold driven contagion, an inactive node can become active in two ways, either spontaneously with rate $p$, or if its weighted threshold $\phi$ is reached. As such, the activation rate of inactive nodes in class $(\mathbf{k}, \mathbf{m})$ is $F_{\mathbf{k},\mathbf{m}} = 1$ when $q_m \geq \phi q_k$, and $p$ otherwise. Isolated nodes, or the $(\mathbf{0}, \mathbf{0})$ class, are unable to activate unless they make up part of the initial seed. These properties are specified by the "Watts threshold" rule in Table 1.4. The stepwise nature of $F_{\mathbf{k},\mathbf{m}}$ is exactly what allowed us to map the rate equations for $s_{\mathbf{k},\mathbf{m}}$ and $i_{\mathbf{k},\mathbf{m}}$ to the reduced-dimension system in Chapter 2, and we solve Eqn. 2.70 provided the initial conditions $\rho = 0$ and $\boldsymbol{\nu} = \mathbf{0}$ at $t = 0$.

## 3.3 REGULAR NETWORKS WITH BIMODAL WEIGHTS

In this section we consider a number of tools in order to understand the role of edge weights in threshold driven contagion. To appreciate the role of both degree and weight heterogeneity, we examine in detail the case of degree-regular networks with bimodal weight distributions. The results obtained in such a setting motivate the use of straightforward "combinatorial" arguments that build a microscopic picture of spreading dynamics. These arguments can be extended by examining asymmetries in active-active, and inactive-active edge types.

### 3.3.1 *Relative versus absolute spreading times*

To study the dynamics of our model we first consider a simple structure, namely a configuration-model $k$-regular network, with $k = 7$. Edge weights are sampled from a bimodal distribution with $n = 2$ values, denoted strong and weak, or $w_1$ and $w_2$, respectively. The weight distribution is characterised by its average $\mu_w$, standard deviation $\sigma_w \geq 0$, and the fraction $\delta_z$ of strong links[2]. Thus, weights take the values

$$w_1 = \mu_w + \sigma_w \sqrt{\frac{1-\delta_z}{\delta_z}} \quad \text{and} \quad w_2 = \mu_w - \sigma_w \sqrt{\frac{\delta_z}{1-\delta_z}}. \tag{3.1}$$

The parameter $\delta_z$ contributes to the skewness of $p_w$, initially fixed to the symmetric case $\delta_z = 0.5$. The parameter $\sigma_w$ interpolates weight heterogeneity between the homogeneous case of an unweighted network ($\sigma_w = 0$), and the most heterogeneous case of a diluted network ($\sigma_w = \mu_w \sqrt{(1-\delta_z)/\delta_z}$, obtained by setting $w_2 = 0$ in Eqn. 3.1), where

2. Whereas $\delta_z$ is the fraction of strong links here, with $0 < \delta_z < 1$, in Chapter 4 it will be the ratio of strong to weak links.

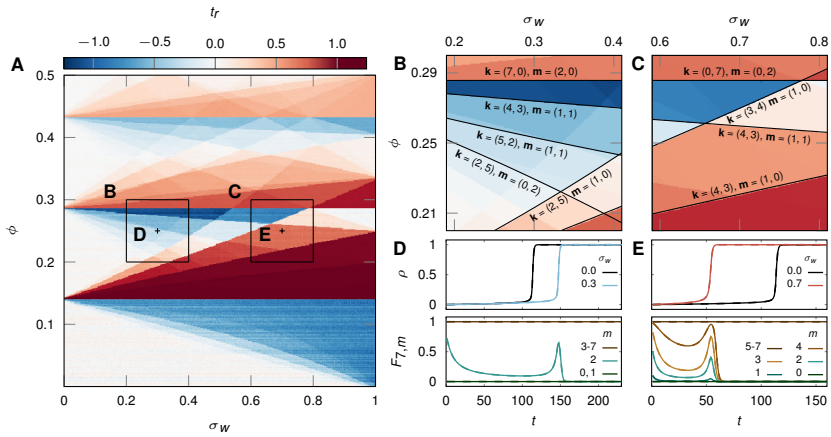


**Figure 3.2 –** Relative speed of threshold driven cascades on weighted networks. (a) Relative time $t_r$ of cascade emergence on $(\sigma_w, \phi)$ space, simulated over $k$-regular regular networks ($k = 7$) with $\mu_w = 1$, $\delta_z = 0.5$, $p = 2 \times 10^{-4}$, $N = 10^4$ and averaged over 25 realisations. Time of cascades for given $\phi$ is either higher or lower than the corresponding case $(0, \phi)$ of an unweighted network. (b-c) Selected regions of parameter space in (a), where $t_r$ is instead calculated from the numerical solution of the AME systems in Eqn. 2.70. Boundaries are obtained from a combinatorial argument (see Section 3.3.2) for various $(\mathbf{k}, \mathbf{m})$ classes. For example, the boundary $\mathbf{k} = (2, 5)$, $\mathbf{m} = (1, 0)$ separates networks where nodes with $k_1 = 2$ strong links and $k_2 = 5$ weak links may (or may not) be activated by $m_1 = 1$ strong active neighbour. (d-e) Quantities characterising the dynamics in simulations (solid lines) and AMEs (dashed lines) for $\phi = 0.25$ and $\sigma_w$ corresponding to the unweighted case, as well as to a slow (d) or fast (e) cascade. Quantities are the activation density $\rho(t)$ (upper panel), and the aggregated activation rate $F_{k,m}(t)$ for various numbers of active neighbours $m$ (lower panel).

only strong links have influence and the weak are functionally absent. After fixing the spontaneous activation rate $p$ and edge skewness $\delta_z$, our model has only two parameters, $\sigma_w$ and $\phi$ (Fig. 3.1a). Similar to other dynamical cascade models [119, 166], contagion initially evolves at a linear rate close to $p$ until the density $\rho(t)$ of active nodes reaches a critical value, triggering a rapid cascade of activation that spreads through the whole network (sample scenarios in Fig. 3.1b-c). Thus, to characterise the speed of dynamics we introduce the quantity $t_a$, the time when activation density reaches a set value ($\rho = 0.75$), called the absolute time of cascade emergence. We measure $t_a$ via numerical simulations of the $(\sigma_w, \phi)$ space (Fig. 3.1a), which shows

unexpected dependencies on both parameters. On one hand, for fixed $\sigma_w$ and increasing $\phi$ the dynamics slows down, since nodes with higher thresholds require more active neighbours in order to become active themselves. This can be seen in the right hand plot of Fig. 3.1a, where the spreading time increases *monotonously* for increasing $\phi$. On the other hand, for fixed $\phi$ the dynamics depends *non-monotonously* on $\sigma_w$, where cascades may evolve either faster or slower as we increase weight heterogeneity, relative to the unweighted case, i.e., $\sigma_w = 0$. This can be seen in the upper plot of Fig. 3.1a.

We concentrate on the $\sigma_w$ dependency by calculating

$$t_r = \frac{t_a(0,\phi) - t_a(\sigma_w,\phi)}{t_a(0,\phi)}, \tag{3.2}$$

the time of cascade emergence relative to the unweighted case with the same $\phi$ value. (Fig. 3.2a). The relative time $t_r$ will be positive if the weighted process evolves faster than the unweighted case, zero if they evolve at the same speed, and negative if slower than the unweighted case, such that $-\infty < t_r < 1$. The $(\sigma_w, \phi)$ space for $t_r$ is highly structured and driven by competing effects of key $(\mathbf{k}, \mathbf{m})$ classes, which either reduce or enhance the speed of the spreading process with respect to the unweighted case. We also explore the corresponding numerical solution of the AME systems in Eqn. 2.70, as well as an independent combinatorial solution for the boundaries between regions of low and high cascade speed (Fig. 3.2b-c) (see Section 3.3.2). Both the AME and combinatorial solutions perfectly recover the parameter space obtained by simulations. To explore *how* weight heterogeneities produce slow or fast cascades, we partition the system according to the total number $m = \sum_j m_j$ of active neighbours required for activation, and measure the aggregated activation rate

$$F_{k,m}(t) = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} F_{\mathbf{k},\mathbf{m}} s_{\mathbf{k},\mathbf{m}}(t)}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} s_{\mathbf{k},\mathbf{m}}(t)}, \tag{3.3}$$

as well as other determinant quantities in several spreading scenarios (Fig. 3.2d-e).

In the neutral scenario, all $(\mathbf{k}, \mathbf{m})$ classes of the weighted network share the same dynamics as the corresponding $(k, m)$ class in an unweighted network, so $F_{k,m} = p$ or 1 and weights have no impact on contagion, meaning $t_r = 0$. In a decelerating scenario like $\phi = 0.25$ and $\sigma_w = 0.3$ (blue curve in Fig. 3.2d), $F_{k,m}$ for any $m$ is equal to its unweighted counterpart, except for the $m = 2$ class, whose activation rate is 1 in the unweighted case but strongly suppressed in the weighted case. Consider that $m = 2$ active neighbours are always sufficient to

trigger activation in the unweighted case, but configurations such as $\mathbf{k} = (5, 2)$, $\mathbf{m} = (0, 2)$ are stable in the weighted case, since two weakly influencing active neighbours are unable to overcome a node's threshold, due to the presence of strongly influencing inactive neighbours. As such, the overall spreading speed is reduced.

In contrast, in an accelerating scenario, like $\phi = 0.25$ and $\sigma_w = 0.7$, competing effects from various $(\mathbf{k}, \mathbf{m})$ classes determine the emergent dynamics (red curve in Fig. 3.2e). The rate $F_{k,m}$ for $m = 2, \ldots, 4$ is lower than its corresponding value of 1 observed in the unweighted case. This is due to the fact that $m = 2, \ldots, 4$ active neighbours are always sufficient to cause activation in the unweighted case, but not so when $\sigma_w = 0.7$. Clearly this amounts to a decelerative effect, as in the previous case of Fig. 3.2d. However, a key difference here is that for the class $m = 1$, $F_{k,1}$ *exceeds* its rate of $p$ in the unweighted case, even if only by a small amount, due to the presence of classes like $\mathbf{k} = (1, 6)$, where a single active neighbour can induce activation, if that neighbour is strong, i.e., when $\mathbf{m} = (1, 0)$. This effect can be seen in the smallest bump in Fig. 3.5e. Since the $m = 1$ class is critical in early stages of contagion, vastly outnumbering classes with $m > 1$, spreading evolves rapidly to an early cascade. In other words, the accelerative effect of the $m = 1$ class outweighs the decelerative effect of the higher $m$ classes, due to their relative abundance at early spreading times.

This simple example is illustrative of more general settings, where the observed dynamics is the result of competition between a large number of classes with accelerative and decelerative effects. We examine this effect more systematically in the following section.

### 3.3.2 *Combinatorial description of parameter space*

The dynamics of threshold driven contagion on weighted networks depends on the stepwise activation rate $F_{\mathbf{k},\mathbf{m}}$, given by the "Watts threshold" rule of Table 1.4. Considering the case of equality, $q_m = \phi q_k$, and writing $q_k$ and $q_m$ explicitly, we obtain $\phi = \mathbf{m} \cdot \mathbf{w} / \mathbf{k} \cdot \mathbf{w}$. Noting that the $\sigma_w$ dependence is contained in the weight vector $\mathbf{w}$, we solve for $\mathbf{k}$ and $\mathbf{m}$, and associate the solution with a unique boundary in $(\sigma_w, \phi)$ space separating regions of differing $t_r$, the relative time of cascade emergence (see Fig. 3.3). In other words, boundaries for $t_r$ in $(\sigma_w, \phi)$ space separate network configurations where the corresponding $(\mathbf{k}, \mathbf{m})$ class does and does not satisfy the threshold rule $q_m \geq \phi q_k$, thus promoting or hindering spreading. In Fig. 3.3, we enumerate all possible boundaries for up to two active neighbours in the case of a $k$-regular random network ($k = 7$) and a bimodal weight distribution

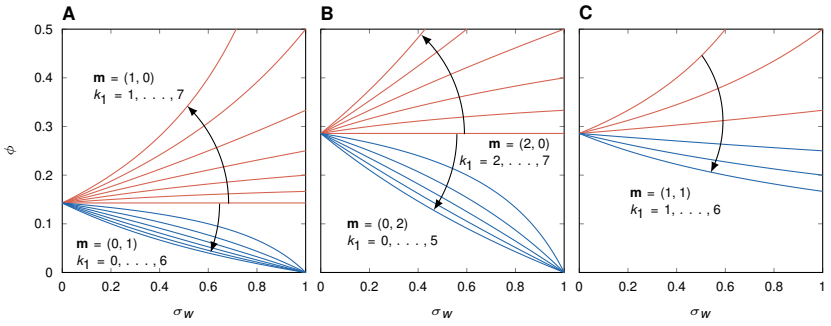


**Figure 3.3 –** Phase boundaries in $(\sigma_w, \phi)$ space for a $k$-regular random network ($k = 7$) with $m = 1, 2$. (a) Boundaries of regions where just one active neighbour of type $j = 1, 2$ is sufficient to induce activation. Curves in red indicate that the associated $(\mathbf{k}, \mathbf{m})$ class produces a speed-up effect on the spreading process relative to the same process on an unweighted network. Conversely, classes associated with the curves in blue produce a slow-down effect on cascades. (b) Similar boundaries for the networks where two active neighbours of the same type are sufficient to cause induced activation, over a range of degree vectors. (c) Boundaries where one active neighbour of each type causes activation.

($n = 2$). This is the setting of the parameter space in Fig. 3.2. Plot 3.3a shows the case where one strong active neighbour, $\mathbf{m} = (1, 0)$, is sufficient to cause activation for nodes with $k_1 = 1, \ldots, k$ strong neighbours. These curves are shown in red, since the corresponding node classes induce a faster cascade of spreading compared to the same process carried out on an unweighted network. Since the weight vector is $\mathbf{w} = (\mu_w + \sigma_w, \mu_w - \sigma_w)^T$ for weight mean $\mu_w$ and skewness $\delta_z = 0.5$, boundaries can be written explicitly as

$$\phi = \frac{m_1 w_1 + m_2 w_2}{k_1 w_1 + k_2 w_2} = \frac{\mu_w(m_1 + m_2) + \sigma_w(m_1 - m_2)}{\mu_w(k_1 + k_2) + \sigma_w(k_1 - k_2)}. \tag{3.4}$$

Curves in blue are the boundaries where one weak active neighbour, $\mathbf{m} = (0, 1)$, is sufficient to induce activation (and a slower cascade than the unweighted case). These curves are enumerated by the number of strong neighbours, $k_1 = 0, \ldots, k - 1$. Curves in Fig. 3.3b are analogous with Fig. 3.3a, except replacing $\mathbf{m} = (1, 0)$ and $\mathbf{m} = (0, 1)$ with $\mathbf{m} = (2, 0)$ and $\mathbf{m} = (0, 2)$. Finally, Fig. 3.3c corresponds to the boundaries due to node classes with $\mathbf{m} = (1, 1)$, where having two active neighbours, one of each type, is sufficient for these classes to undergo induced activation.

As a concrete example, the boundary $\mathbf{k} = (3, 4)$, $\mathbf{m} = (1, 0)$ in Fig. 3.2c separates networks where nodes with $k_1 = 3$ strong links

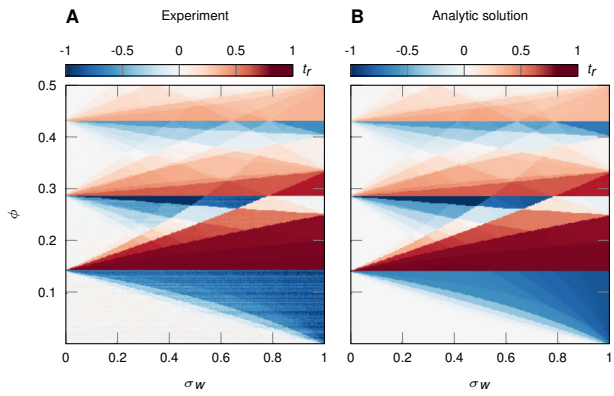


**Figure 3.4 –** Comparison between numerical simulations and AME solutions. $(\sigma_w, \phi)$ space for the relative time $t_r$ of cascade emergence, obtained by Monte Carlo numerical simulations (a) and the numerical solutions of the full and reduced AME systems (b), the last two of which are indistinguishable. Numerical simulations consider $k$-regular random networks ($k = 7$) with $N = 10^4$, $p = 2 \times 10^{-4}$, and averages over 25 realisations.

and $k_2 = 4$ weak links may be activated by just $m_1 = 1$ strong active neighbour. If two networks differ only in the rate of activation of nodes in this $(\mathbf{k}, \mathbf{m})$ class (such that one is vulnerable to activation and the other is not), we observe a difference in spreading time. In the case of Fig. 3.2d, the resultant spreading process is faster than its unweighted counterpart, due precisely to this mechanism.

### 3.3.3 *Accuracy of master equation, an aside*

As discussed in Section 3.3.1, the simulated behaviour of threshold driven contagion over weighted networks, and their corresponding analytic solutions, are in close agreement. We have illustrated this similarity by comparing $t_r$ values over $(\sigma_w, \phi)$ space, as well as by plotting the temporal evolution of the activation density $\rho$ and other quantities characterizing the dynamics. To further compare the $(\sigma_w, \phi)$ space produced by Monte Carlo simulation and AME solution, we show them side-by-side in Fig. 3.4, and quantify their similarity by computing the mean absolute difference $D$, given by

$$D[t_r(\sigma_w, \phi)] = \frac{1}{N_{\sigma_w} N_\phi} \sum_{\sigma_w} \sum_{\phi} \left| \frac{t_r^{\text{sim}}(\sigma_w, \phi) - t_r^{\text{theo}}(\sigma_w, \phi)}{t_r^{\text{sim}}(\sigma_w, \phi)} \right|, \qquad (3.5)$$

where $N_{\sigma_w}$ and $N_\phi$ are the number of points considered in each dimension of the parameter space, and $t_r^{\text{sim}}(\sigma_w, \phi)$ and $t_r^{\text{theo}}(\sigma_w, \phi)$ are the relative times of cascade emergence for a given $(\sigma_w, \phi)$ point, measured by numerical simulations or AME solutions. This quantity is very small, $D = 1.3 \times 10^{-4}$, indicating that despite making a number of assumptions during the derivation of the full and reduced AME systems [Eqns. 2.46 and 2.70], they provide an excellent approximation of the spreading process with differences due only to small statistical fluctuations in finite-size numerical simulations.

It is worthwhile pointing out that in many settings, the numerical solution can be computed orders of magnitude more rapidly than the corresponding Monte Carlo simulation. Recall that simulation requires large networks, or a large number of realisations to remove fluctuations, which are expensive in terms of space and time, respectively (see Chapters 1 and 2). In contrast, if these systems have small cutoff in degree $k$, the cost of computing them analytically is negligible. For instance, the system in Fig. 3.4 requires only 120 equations each with less than 10 terms, so an entire parameter space can be produced within seconds. In contrast the corresponding simulations, if done on large networks, may take hours.

### 3.3.4 *Bulk and interface of the contagion cluster*

As a consequence of the arguments outlined in Section 3.3.2, an asymmetry is observed to emerge in the fractions of weak and strong links connecting active nodes, $E^{II}(t)$, or inactive and active nodes, $E^{SI}(t)$. Since strong ties contribute the most in reaching the threshold of a node, they are critical early in the contagion process, and comprise most ties within the active subgraph. Conversely, weak ties dominate the surface of the cascade by connecting active and inactive nodes. This asymmetry in edge type is an essential feature of weighted contagion that is trivially absent in the unweighted case. This asymmetry evolves both in cases of accelerated and decelerated spreading, with amplitude dependent on the absolute value of the relative speed of contagion.

Explicitly, we compute the fraction of $j$-type links per node, connecting two active nodes. We term this the cascade "bulk", with the corresponding edge density given by

$$E_j^{II}(t) = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j i_{\mathbf{k},\mathbf{m}}(t)}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m i_{\mathbf{k},\mathbf{m}}(t)}. \tag{3.6}$$

Edges between inactive and active nodes, which we term the cascade "surface", are characterised by their densities

$$E_j^{SI}(t) = \frac{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m_j s_{\mathbf{k},\mathbf{m}}(t)}{\sum_{\mathbf{k},\mathbf{m}} p_{\mathbf{k}} m s_{\mathbf{k},\mathbf{m}}(t)}, \tag{3.7}$$

such that $\sum_j E_j^{II} = \sum_j E_j^{SI} = 1$. See Fig. 3.5b and Fig. 3.5c for the case of $n = 2$. These quantities diverge from $\delta_z$ and $1 - \delta_z$, their values when the network is unweighted, with amplitude dependent on the relative importance of strong links in triggering activation. When weights are uniform, $\sigma_w = 0$, we have $E_1^{II} = E_1^{SI} = \delta_z$ and $E_2^{II} = E_2^{SI} = 1 - \delta_z$.

### 3.3.5 *Asymmetry in edge types*

Up until now we have considered the symmetric case $\delta_z = 0.5$ with equal numbers of strong and weak links. However, by skewing the weight distribution we observe an additional effect of weight heterogeneities on the spreading behaviour. When $\delta_z = 0.2$ the size of the cascade decreases for large $\sigma_w$ with respect to the unweighted case (Fig. 3.5a). In this case, despite their sparsity, strong links again drive the contagion, but are soon exhausted causing spreading to slow down and continue via spontaneous or infrequent threshold driven activations over weak ties (Fig. 3.5b). Indeed, strong links dominate the bulk of the active component, but disappear quickly from its surface (Fig. 3.5c). These so-called *partial cascades*, which do not activate the whole system through the cascade, are associated with relative sparsity in the distribution of strong edges, $\delta_z < 0.5$, as well as sufficiently large standard deviation in the weight distribution. This effect is reminiscent of the slow spreading caused by immune nodes, as well as low connectivity networks in unweighted complex contagion, see Refs. [166] and [198], respectively. Overall, we identify non-monotonous spreading behaviour and partial cascades as the main consequences of weight heterogeneities in threshold driven contagion.

## 3.4 HETEROGENEOUS SYNTHETIC AND REAL NETWORKS

Although regular networks and bimodal weights are useful in characterising the qualitative impact of weights on contagion, they are rather unrealistic since real complex networks commonly appear with broad degree and weight distributions [18]. Thus, in the following we explore how threshold driven contagion is influenced by weights using simulations in heterogeneous synthetic and real weighted networks (Fig. 3.6).

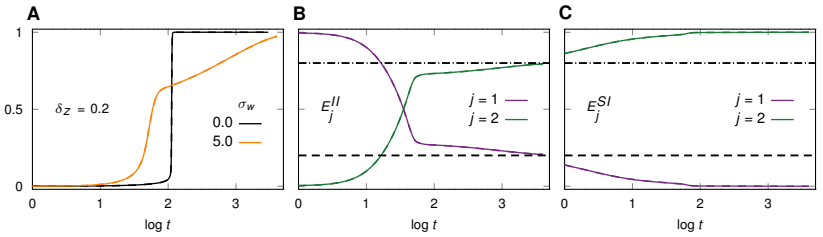


**Figure 3.5 –** Effect of skewed weight distributions on cascade evolution. (a) Infection density $\rho(t)$ on $k$-regular networks ($k = 7$) and a bimodal weight distribution with $\mu_w = 3$ and $\delta_z = 0.2$, both for unweighted ($\sigma_w = 0$) and heterogeneous ($\sigma_w > 0$) cases. (b-c) Fractions of strong ($j = 1$) and weak ($j = 2$) links connecting two active nodes in the bulk of the active component [$E^{II}(t)$, b] and inactive and active nodes on its surface [$E^{SI}(t)$, c] in the heterogeneous spreading scenario of (a). Simulations (solid lines) are averaged over 25 realisations with $p = 2 \times 10^{-4}$ and $N = 10^4$, and compared with the corresponding AME solution of Eqn. 2.70 (dashed lines). Horizontal dashed lines are the expected fractions of weak and strong links as determined by $\delta_z$ and $1 - \delta_z$.

We expect degree heterogeneities to affect threshold driven processes since thresholds are defined relative to the degree (or strength) of nodes. As a first step, we construct synthetic scale-free networks using the configuration model, with degree distribution $p_k \sim k^{-\tau}$ and exponent $\tau = 2.5$, but maintain a bimodal weight distribution with $\mu_w = 1$ and $\delta_z = 0.5$ as in the previous section (Fig. 3.6a). The increased number of $(\mathbf{k}, \mathbf{m})$ classes (due to degree heterogeneity) fragment $(\sigma_w, \phi)$ space with respect to $t_r$, but its structure still exhibits areas of slow and fast cascades and can be explained using the same arguments used in the $k$-regular case. Real world examples of this synthetic structure are signed social networks, like the network of Wikipedia editors [200], where degrees are broadly distributed and edge signs indicate the parity (or binary weight) of a social interaction like trust, intimacy, or influence. We simulate our threshold model over the Wikipedia social network by associating $+$ and $-$ tie signs with strong ($w_1$) and weak ($w_2$) links, thus obtaining a weighted network with $\delta_z = 0.88$ and arbitrary $\sigma_w$ (Fig. 3.6b) (see Appendix C for data description). Despite complex structural correlations potentially present in the real data, the Wikipedia $(\sigma_w, \phi)$ space is qualitatively similar to the case of a synthetic scale-free degree distribution, although these correlations and the high $\delta_z$ value modify the areas of relative acceleration and deceleration. To further validate these observations, we have also analysed configuration-model random networks and another empirical signed

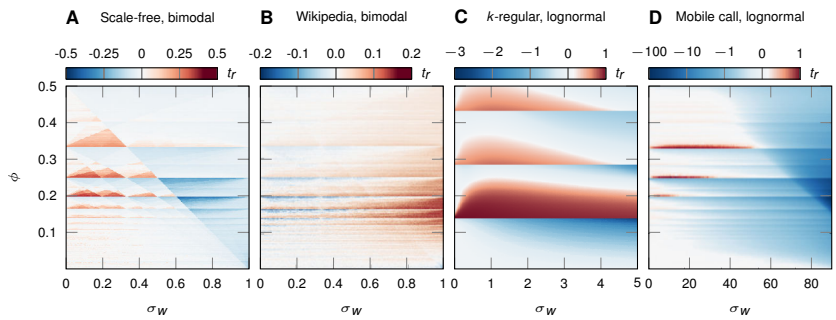


**Figure 3.6 –** Threshold driven contagion on heterogeneous synthetic and real networks, with network type and weight distribution as indicated. (a) Relative time $t_r$ of cascade emergence on $(\sigma_w, \phi)$ space, simulated over synthetic scale-free networks with degree exponent $\tau = 2.5$, average degree $z = 4.54$ and minimum degree $k_{\min} = 2$. Link weights are bimodally distributed with $\mu_w = 1$ and $\delta_z = 0.5$. (b) A signed social network of Wikipedia editors with heterogeneous degrees and skewed bimodal weight distribution, with $\delta_z = 0.88$ (see Appendix C for a description of data). (c) Same as (a) but over a $k$-regular network ($k = 7$) and a lognormal weight distribution with $\mu_w = 1$. (d) Lognormally distributed weights over a mobile phone call network with heterogeneously distributed degrees, and $\mu_w = 37.7$. Synthetic networks in (a) and (c) have $N = 10^4$ and parameter spaces are averaged over 25 realisations. Parameter space in (b) is averaged over $10^3$ realisations, while (d) is the result of a single realisation. All simulations have rates of spontaneous activation $p = 2 \times 10^{-4}$.

network, the Pardus dataset [183] (see Section 3.5 for the corresponding $(\sigma_w, \phi)$ spaces, and Appendix C for data description).

As with degree heterogeneities, weights in empirical networks are broadly distributed and approximated by scale-free or lognormal distributions, which we address next by exploring the threshold model on $k$-regular networks ($k = 7$) and a lognormal weight distribution with average $\mu_w = 1$ (Fig. 3.6c). Even though all nodes have the same degree, a broad distribution of weight values increases the number of $(\mathbf{k}, \mathbf{m})$ classes, smoothing out the $(\sigma_w, \phi)$ space with respect to the bimodal case but qualitatively maintaining its non-monotonous patterns of slow and fast cascades. The standard deviation $\sigma_w$ controls the skewness of the weight distribution and determines the temporal evolution of contagion, promoting partial cascades for large $\sigma_w$. Finally, we consider threshold driven contagion in a large empirical weighted social network, an aggregated mobile phone call (MPC) network, where weights are proportional to the number of calls between individuals

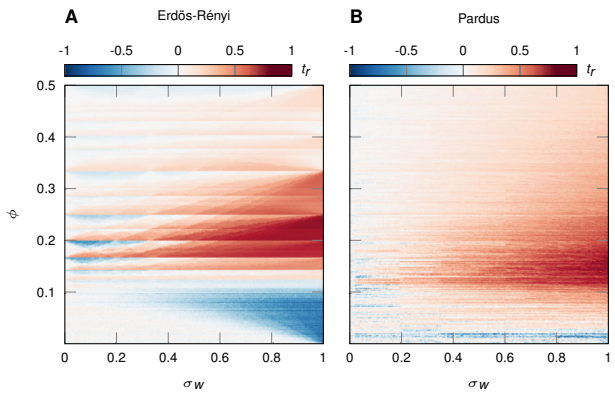


**Figure 3.7 –** Other heterogeneous synthetic and real networks. $(\sigma_w, \phi)$ space for the relative time $t_r$ of cascade emergence, simulated on a configuration-model random network (a) and the Pardus signed social network (b). Numerical simulations on (a) consider networks with $N = 10^4$, average degree $z = 7$ and averages over 25 realisations. All simulations correspond to $p = 2 \times 10^{-4}$.

(Fig. 3.6d) (see Appendix C for a description of data). This network has broad degree and weight distributions [153], communities, degree correlations and Granovetter-type degree-weight correlations [96]. Despite this added complexity, the $(\sigma_w, \phi)$ space of the MPC network is qualitatively similar to previous cases, apart from the magnitude of the decelerative effect when weights are strongly heterogeneous. As before, skewness in the weight distribution temporally inhibits contagion and induces partial cascades. Our data-driven simulations show that, even in empirical networks of vastly different origins, threshold driven contagion strongly depends on link weights via simple mechanisms that can be understood by master equations or combinatorial arguments. This dependence may be responsible for the diverse dynamical scenarios of threshold driven contagion observed in nature, like the diffusion of information in techno-social networks, which typically reaches a limited population, but can occasionally unfold globally through slow or fast cascades of activation.

## 3.5 FURTHER HETEROGENEOUS NETWORKS

In the previous section we have seen that, even for heterogeneous synthetic and real world networks, threshold driven contagion strongly depends on link weights via simple mechanisms that can be understood by master equations or combinatorial arguments, and develops

spreading cascades that are either faster of slower than their counterparts in unweighted contagion, depending on the values of $\sigma_w$ and $\phi$. Here we further support this argument by exploring two additional examples of synthetic and empirical networks. The synthetic structure is a configuration-model random network with average degree $z = 7$, and bimodal weight distribution with average $\mu_w = 1$ and skewness $\delta_z = 0.5$ (Fig. 3.7a). The $(\sigma_w, \phi)$ space for $t_r$ is qualitatively very similar to the ones observed for configuration-model $k$-regular or scale-free networks, with several fast and slow cascade regimes that start from values on the $\phi$ axis determined by the harmonic series of degrees present in the network. The empirical structure is a signed social network, the alliance / enemy network of the Pardus massive multiplayer online game [183] (see Appendix C for details). The $(\sigma_w, \phi)$ space for $t_r$ (Fig. 3.7b) is somewhat less structured than in the other explored networks, but still shows regions of fast and slow cascades with respect to the unweighted case.

## 3.6 DISCUSSION

In complex networks, weights quantify the strength of interactions between nodes and distinguish neighbours by the relevance or influence among them. Threshold driven contagion in empirical settings is particularly sensitive to link weights, since influence between connected nodes may vary enormously, thus changing the temporal pattern of global spreading. Examples of real-world threshold driven processes over weighted networks can be found in epidemiology, where contagion is enabled by direct human interactions that occur at varying frequencies thoughout the population [115]. In the case of social contagion, like the spreading of information, service or product adoption, participation in collective movements, or the adoption of behavioural patterns, heterogeneities in social tie strengths are relevant as they may reflect the strength of mutual influence between individuals [46, 96, 152, 208]. In neural systems, the synaptic weight (a function of several variables) may vary heterogeneously between connected neurons and even in time due to synaptic plasticity [78, 108]. Despite this broad set of real-world examples, threshold driven contagion has mostly only been studied over unweighted networks where links are considered equal.

Our aim in this chapter has been to address this shortfall by systematically studying a threshold model on synthetic and empirical weighted networks. We explore networks with increasing complexity, from configuration-model networks with bimodal or lognormal

weights, to real world networks with broad degree and weight distributions as well as multiple correlations. We show that threshold driven contagion depends non-monotonously on weight heterogeneity, creating slow or fast cascades relative to the equivalent unweighted spreading process. Via numerical simulations, master equations and combinatorial arguments, we find that this effect is the result of competing configurations of degree, weight, and active neighbours that slow down or speed up contagion. We also observe that an imbalance in the amount of strong and weak weights leads to partial cascades, and smoother temporal patterns of spreading than those in unweighted networks. By analysing many degree and weight distributions, we show that these features are systematic and thus may drive a variety of real world contagion phenomena.

Our contribution in [190] opens up directions of research in the largely unexplored area of dynamical processes with heterogeneous interactions. First, the weight-based, master equation formalism described here can be modified to consider any interaction quality like direction and type, thus providing analytical tools to characterise threshold driven contagion in temporal and multiplex networks [103, 209]. Second, our methodology may be used to describe any binary-state dynamics and thus a broad class of empirical processes over weighted networks. We expect our results to find meaningful applications in fields where threshold driven contagion is relevant, like computational epidemiology, neural networks, and social contagion. In these fields our modelling framework, which distinguishes the varied roles and influence of links, may lead to advances in the understanding and prediction of specific temporal features of global pandemics, collective neural firing, or the adoption of innovations and behavioural norms.

In the following chapter, we extend the weighted network model by introducing multiplexity. Here, weighted edges will be associated with separate layers of a multiplex network, which allows individual applications of the threshold rule by edge type. We shall see that the effects of edge weights in a single layer case, in particular the competing accelerative and decelerative effects of a given weight distribution, inform the dynamics when the number of network layers is increased.

# 4

# THRESHOLD-DRIVEN CONTAGION ON MUTIPLEX NETWORKS

Models of threshold driven contagion explain the cascading spread of information, behaviour, systemic risk, and epidemics on social, financial and biological networks. At odds with empirical observation, these models predict that single-layer unweighted networks become resistant to global cascades after reaching sufficient connectivity. We investigate threshold driven contagion on weight heterogeneous multiplex networks and show that they can remain susceptible to global cascades at any level of connectivity, and with increasing edge density pass through alternating phases of stability and instability in the form of reentrant phase transitions of contagion. Our results provide a novel theoretical explanation for the observation of large scale contagion in highly connected but heterogeneous networks.

## 4.1 INTRODUCTION TO MULTIPLEX NETWORKS

Information-communication technology has radically transformed social and economic interaction [34], introducing new means of transmitting ideas, behaviour, and innovation [46, 165], overcoming limitations imposed by time and cognitive constraints [68, 95]. The same technology provides an increasingly accurate picture of human interaction, mapping the underlying network structures that mediate dynamical processes, like epidemics [115, 184]. In complex contagion [46], characteristic of the spreading of innovation, rumors, or systemic risk, transmission is a collective phenomenon in which all social ties of an individual may be involved. Node degree, or number of links, is therefore critical to the dynamical outcome [198]; large relative neighbour influence is easier to achieve the smaller the ego network. This behaviour is well captured by threshold models of social contagion on single-layer unweighted networks, which predict large-scale cascades of activation in relatively sparse networks [97, 117, 166,

198]. In empirical social networks, however, individuals can maintain hundreds of ties [69, 95], with interaction strength varying across social contexts [40, 114, 190], yet still exhibit frequent system-wide cascades of social contagion [13, 66, 89, 119, 189].

We address this issue by incorporating relevant features of empirical social networks into a conventional threshold model. We consider that network ties are heterogeneous, and can be characterized by edge "types". In the case of social networks, these edge types vary in "quality" [96, 120], usually associated with the intimacy or perceived importance of a relationship between individuals [206], and scale with the strength of interpersonal influence [51, 187]. Heterogeneity in tie quality is well modelled by multiplex structures, as has been recognized in both network [29, 123] and social science [35, 194], particularly regarding social contagion [37, 135, 202, 209]. In multiplex models of social networks, individual layers represent the social context of a relationship (e.g. kinship, acquaintance), allowing us to classify ties by social closeness, as recognised by Dunbar's intimacy circle theory [206]. According to this theory, due to cognitive and time resources being finite but necessary to maintaining social ties, individuals actively cultivate a limited number of relationships, organising them into intimacy circles that increase in size as they decrease in importance. Ego networks thus comprise a small but high-intimacy circle of close relationships, like family and long term friends, followed by large but low-intimacy circles of distant friends and acquaintances. Empirical evidence shows the distribution of dyadic social commitments (number of interactions or time devoted to peers) to be strongly heterogeneous [140, 152]. Strikingly, this inverse relation between the cost of maintaining an edge type, and the abundance of that edge type, can be seen as an entropy maximisation process [185] that applies to any system with heterogeneous cost of edge formation and finite node resources. As such, although we use the language of social networks, our results are of relevance to other systems, e.g., financial [7, 26, 73] and biological [115, 184] contagion.

Using analytical and numerical tools, we show that layer hierarchy can lead to global cascades in multiplexes with average degree in the hundreds or thousands, perturbed by a single initial activation. We report the novel observation that in a multiplex network with increasing link density a sequence of phase transitions occur, resulting in alternating phases of stability and instability to global cascades.

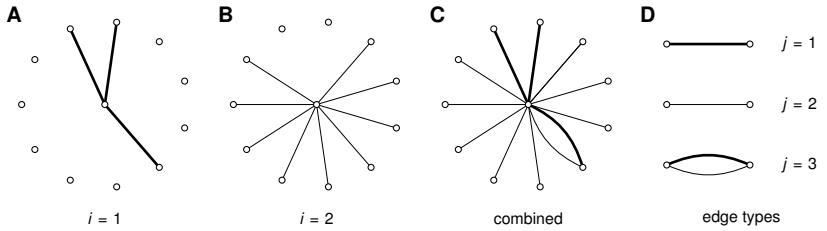


**Figure 4.1 –** (a-b) Egocentric view of multiplex structure with $M = 2$ layers, where edge density increases ($\delta_z > 1$) and edge weight decreases ($\delta_w < 1$) by layer $i$. (c) Egocentric network overlap between layers. (d) Emergent edge types in the overlapping network. In the multiplex, the central node has degree vector $\mathbf{k} = (2, 8, 1)^T$, encoding layer overlap.

## 4.2 MULTIPLEX NETWORK MODEL

### 4.2.1 *Multiplex models of intimacy circles*

Our model builds upon previous studies of threshold driven processes [97, 117, 166, 198] and multiplex networks [29, 123]. We define contagion as a binary-state dynamics over a weighted, undirected multiplex network of $N$ nodes connected throughout $M$ layers (Fig. 4.1). A node represents an individual $u$, and layer $i$ the social context in which individuals interact, $1 \leq i \leq M$. The degree of $u$ in each layer $i$ takes discrete values $k_i = 0, \ldots, N-1$ according to the degree distribution $P_i(k)$. Edge weights $w_i(u, v)$ follow the continuous distribution $P_i(w)$ and capture the total capacity of nodes $u$ and $v$ to influence each other via layer $i$. The network allows for layer overlap [44] as nodes may be connected in multiple layers, modelling individuals who share several social contexts [Fig. 4.1(c)]. For simplicity, we assume that node degree is independent across layers, and that degree and weight distributions $P_i(k)$ and $P_i(w)$ differ by layer only in their means $z_i = \sum_k k P_i(k)$ and $w_i = \int w P_i(w) dw$, otherwise retaining their functional form. In order to reproduce the hierarchical organization of edges suggested by intimacy circle theory [206], we assume that the mean degree $z_i$ and weight $w_i$ scale with layer index $i$ as

$$z_{i+1} = \delta_z z_i \quad \text{and} \quad w_{i+1} = \delta_w w_i, \tag{4.1}$$

with $\delta_z \geq 1$ and $\delta_w \leq 1$. In other words, ego networks comprise a small number of high-intimacy neighbours [Fig. 4.1(a)] and a larger number of low-intimacy neighbours [Fig. 4.1(b)]. We fix the average total degree $z = \sum_i z_i$ as well as $\delta_z$, which determines $z_i$. We also

**Table 4.1 –** Extensions of the Watts threshold rule to multiplex networks. Node state is determined by a single threshold $\phi$ and a weighted sum of influence over layers, or by individual layer thresholds $\phi_i$ and influence within each layer. In the former the multiplex can be projected to a single weighted layer without loss of information relevant to the dynamics.

| weighted sum | multiplex *or* | multiplex *and* |
|---|---|---|
| $q_m \geq \phi q_k$ | $\exists i$ s.t. $q_{m_i} \geq \phi_i q_{k_i}$ | $q_{m_i} \geq \phi_i q_{k_i} \ \ \forall i$ |

impose the arbitrary constraint $\langle w \rangle = 1$ and fix $\delta_w$, which determines $w_i$ (see Section 4.2.3 for more details).

### 4.2.2 *Multiplex threshold rules*

In a binary-state model of contagion, nodes are in one of two mutually exclusive states, inactive or active. Since nodes must be either connected or disconnected via each of the $M$ network layers, their interaction is characterized by one of $2^M - 1$ resultant edge types [Fig. 4.1(d)], disregarding nodes disconnected in all layers, and indexing by $j$ such that $1 \leq j \leq 2^M - 1$. Node configuration is thus described by the number of neighbours $k_j$ and active neighbours $m_j$ across edges of type $j$, with $0 \leq m_j \leq k_j$. We store $k_j$ and $m_j$ in the degree vector $\mathbf{k}$ and partial degree vector $\mathbf{m}$, respectively (of dimension $2^M - 1$). Note that we consistently index layer by $i$ and resultant edge type by $j$.

The threshold rule proposed by Watts [97, 117, 166, 198] defines the fraction $\phi$ of neighbours that must be active for an inactive ego to activate. This rule can be extended to multiplex networks in several ways (Table 4.2). Denoting the set of neighbours of node $u$ in layer $i$ by $\mathcal{N}_i(u)$, the total influence upon $u$ in layer $i$ is $q_{k_i} = \sum_{v \in \mathcal{N}_i(u)} w_i(u, v)$. Restricted to active neighbours, $\mathcal{N}_i(u)|_I$, this gives $q_{m_i} = \sum_{v \in \mathcal{N}_i(u)|_I} w_i(u, v)$. In one variant of the threshold rule, nodes perceive influence in aggregate, summed over layers (reminiscent of neural networks [79, 108]) and activate with respect to a single threshold if $q_m \geq \phi q_k$, where $q_k = \sum_i q_{k_i}$ and $q_m = \sum_i q_{m_i}$ (weighted sum rule). In another variant, node state is determined by $M$ layer thresholds $\phi_i$, along with influence $q_{k_i}$ and $q_{m_i}$ within layers. A node activates when $q_{m_i} \geq \phi_i q_{k_i}$ in every layer (multiplex *and* rule by Lee [135]), or in at least one layer (multiplex *or* rule [135]). Our aim is to show that multiplex networks following the structure of intimacy circle theory exhibit reentrant phase transitions for both the weighted sum and the multiplex *or* threshold rules. Note that if weights are uniform within each layer and node

state is determined by decisions within layers (*and* and *or* rules), then the structure is effectively unweighted. We show that even with this loss of weight information, reentrant phase transitions can still emerge due to contagion within layers.

### 4.2.3 *Degree and weight distributions*

In the previous section we argued that the salient features of real multiplexes, scaling in the mean degree $z_i$ and mean weight $w_i$ from layer to layer, can be modelled as following the scaling relations $z_{i+1} = \delta_z z_i$ and $w_{i+1} = \delta_w w_i$, as per Eqn. 4.1. We term $\delta_z$ the density scaling factor, or density skewness, and $\delta_w$ the weight scaling factor, or weight skewness. Crucially, setting $\delta_z > 1$ and $\delta_w < 1$ recovers the class of structure outlined in intimacy circle theory, indicating a multiplex growing in link density, and decreasing in mean interaction strength or weight, by layer. Both $\delta_z$ and $\delta_w$ are constant in our model, and induce exponentially distributed layer average degrees and layer average weights, since

$$\{z_1,\ z_2,\ z_3,\ \ldots,\ z_M\} = \{z_1,\ \delta_z z_1,\ \delta_z^2 z_1,\ \ldots,\ \delta_z^{M-1} z_1\}, \tag{4.2}$$

with a similar expression holding for the distribution of $w_i$ values. Although $z_1$ appears in the latter expression, this choice is arbitrary and it is not in fact a free parameter in the following experiments. For example, we could equally have written $\{\delta_z^{-1} z_2,\ z_2,\ \delta_z z_2,\ \ldots,\ \delta_z^{M-2} z_2\}$, for some $z_2$, giving an identical set of values. In experiments we set the total average connectivity $z$, defined as $z \equiv z_1 + z_2 + \ldots + z_M$, as well as the density skewness $\delta_z$. Since each expression in Eqn. 4.1 has only two degrees of freedom, choosing $z$ and $\delta_z$ effectively prescribes the individual layer averages $z_1,\ z_2,\ \ldots,\ z_M$. Similarly, the distribution of weight means $w_1, \ldots, w_M$ has only two degrees of freedom. As for the distribution of $z_i$ values, we do not explicitly set $w_i$, rather, we constrain the average weight $\langle w \rangle$, which along with the weight scaling constant $\delta_w$, determines each $w_i$. The total mean weight $\langle w \rangle$ over the multiplex can be found by summing over edge type means, $\langle w \rangle = \sum_j c_j w_j / c$, where $c_j$ is the average degree of the $j$-th edge type, equal to $z_i$ in the case of zero overlap, and $w_j$ is the sum of edge weights constituting the resultant edge of type $j$. In all experiments throughout this work we set the network-wide average weight to $\langle w \rangle = 1$. This allows us to isolate the effect of varying the skewness $\delta_z$ in edge density, and the skewness $\delta_w$ in interaction strength, across layers. Note that we impose the additional constraint that all edge weights be positive.

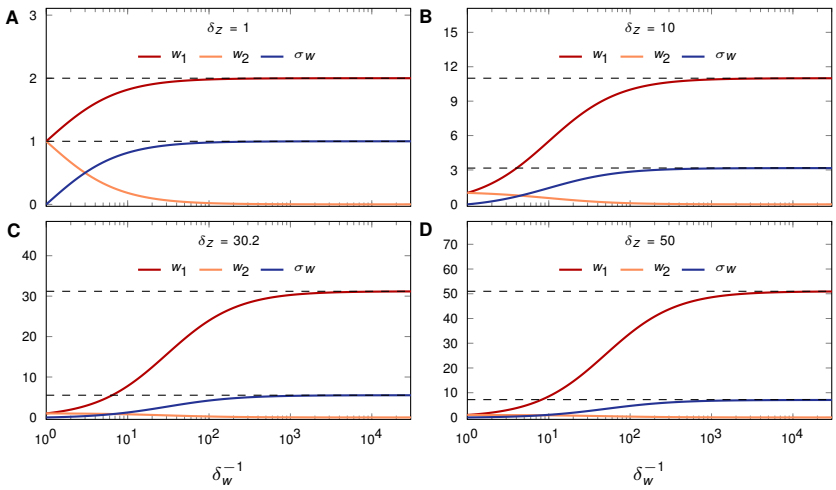


**Figure 4.2 –** The dependence of weight means $w_i$ and standard deviation $\sigma_w$ of the weight means on $\delta_w$. Weight distributions correspond to an $M = 2$ layer multiplex, with values of $\delta_z$ providing the density skewness of various networks used in this text. We refer to the limit $\delta_w \to 0$ as maximal weight heterogeneity, where $w_1 = 1 + \delta_z$, $w_2 = 0$ and $\sigma_w = \sqrt{\delta_z}$.

### 4.2.4 *Maximal weight heterogeneity*

In this section we discuss the consequences of our definition of average degree and weight scaling. In particular, given the constraints $\langle w \rangle = 1$ and $w_i > 0 \; \forall i$, the weight scaling constant $\delta_w$ is defined over the interval $(0, 1]$, with $\delta_w = 1$ giving identical mean weights $w_i$ in each layer, and $\delta_w \to 0$ producing maximal weight heterogeneity. The interpolation between these limits is shown in Fig. 4.2 for weight means $w_i$ and the standard deviation of the means [1] $\sigma_w$, for various values of $\delta_z$ used in this work. In $M = 2$ layers, assuming density skewness $\delta_z$, it is straightforward to show that the weight heterogeneous limit of these quantities is

$$\lim_{\delta_w \to 0} w_1 = 1 + \delta_z, \tag{4.3}$$

$$\lim_{\delta_w \to 0} w_2 = 0, \tag{4.4}$$

$$\lim_{\delta_w \to 0} \sigma_w = \sqrt{\delta_z}, \tag{4.5}$$

1. Not to be confused with $\sigma_{w_j}$, the weight standard deviation of edge type $j$

where the standard deviation of edge weights across the multiplex is defined as $\sigma_w^2 = \sum z_i \left(w_i - \langle w \rangle\right)^2 / z$. Clearly, increasing weight heterogeneity by decreasing $\delta_w$ has a saturating effect on the values $w_i$ and $\sigma_w$. Furthermore, decreasing $\delta_w$ below the limiting value will have diminishing effect on the actual threshold processes evolving over the network. In the experiments described in the following sections, we first set a density skewness $\delta_z$, and then increase weight skewness subject to the constraints that all weight values are positive ($w_i > 0$ for all layers $i$) and that $\langle w \rangle = 1$. In Fig. 4.1, for example, we are interested in the approach towards maximal weight heterogeneity from a uniform distribution given by $\delta_w = 1$. The maximal weight distribution leads to results given by the white contour in Fig. 4.5(c), corresponding to a value of $\delta_w = 10^{-6}$. Again, any positive value in the range $0 < \delta_w < 10^{-6}$ would give identical results due to the limiting effect of $\delta_w \to 0$. In Fig. 4.15, we conduct a similar experiment in an empirical Twitter network. There, we study the approach to maximal weight heterogeneity $\delta_w \to 0$, beginning with a uniform weight distribution $\delta_w = 1$.

### 4.2.5 *Poisson and log-normal degree distributions*

The scaling conditions in Eqn. 4.1 specify only the average $c_j$ of the degree distribution of edge type $j$, so we are free to choose the actual form of the distribution. Specifically, we use Poisson distributions (PO) in Fig. 4.7, as well as Figs. 4.8 and 4.9. Otherwise, we use log-normal distributions (LN). A Poisson degree distribution is prescribed entirely by its mean $c_j$,

$$P_j(k_j \mid c_j) = e^{-c_j} \frac{c_j^{k_j}}{k_j!}, \tag{4.6}$$

which is defined for integer $k_j$, as required, and the average degree $c_j$ of each edge type $j$. In contrast, the log-normal distribution is defined for continuous variables, so it remains to define an appropriate discretisation. Since the structural percolation transition is of critical importance to the Watts model and its extensions, any discretisation must allow for a non-zero mass of $k = 0$ degree nodes.[2] To this end, we round to the nearest integer, although taking the floor function would also be appropriate. In the numerical implementation of the configuration model, it is straightforward to randomly sample continuous $k$, and

*percolation transition*

2. This rules out, for example, discretisations based on the ceiling function, since values in the range $0 < k < 1$ will be rounded to $k = 1$, resulting in $P_i(0) = 0$.

round up or down to the nearest integer. The masses of the $k$ values thus obtained can be found by integrating the log-normal probability density function over the interval $(k - \frac{1}{2}, k + \frac{1}{2})$, or equivalently, evaluating the cumulative distribution at the limits. As such, degrees $k_j$ are distributed as

$$P_j(k_j \mid c_j, \sigma_{k_j}) = F(k_j + \tfrac{1}{2} \mid c_j, \sigma_{k_j}) - F(k_j - \tfrac{1}{2} \mid c_j, \sigma_{k_j}), \tag{4.7}$$

where $F$ is the cumulative distribution of the continuous log-normal distribution, which we assume is defined $F(k_j) = 0$ for $k_j \leq 0$, $c_j$ the mean of the $j$-th type degree, and $\sigma_{k_j}$ its standard deviation. The actual value of $F$ is given by the error function

$$F(k_j \mid c_j, \sigma_{k_j}) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln k_j - c'_j}{\sqrt{2}\sigma'_{k_j}}\right), \tag{4.8}$$

where $c'_j$ and $\sigma'_{k_j}$ are the mean and standard deviation of the underlying normal distribution.[3] This approach leads to precise agreement between the simulated multiplex and analytic solution, and most importantly, allows for a structural percolation transition in the resulting configuration model multiplex, at low $z$ due to the presence of $k_j = 0$ degree nodes.

Finally, since our experiments generate log-normal distributions with means $c_j$ spanning several orders of magnitude, we must select experimental values of standard deviation $\sigma_{k_j}$ appropriately. That is, the standard deviation must scale with the average, and we set $\sigma_{k_j} = 2c_j$. To see why this is necessary, consider that $\sigma_{k_j} = 10$ is a relatively high variance if the average degree $c_j = 1$, with a qualitatively broad spread. However, if $c_j = 1000$, a standard deviation of $\sigma_{k_j} = 10$ leads to a very narrow overall distribution, qualitatively giving a sharp peak rather than a broad tail.

## 4.3 ANALYTIC SOLUTION

### 4.3.1 *Autonomous reduced-dimension solution*

In this chapter we apply the master equation solution outlined in Chapter 2. In multiplex networks comprised of $M$ layers, there are

3. Normal moments, $\mu'$ and $\sigma'$, are related to log-normal moments, $\mu$ and $\sigma$, via the relations $\mu' = \log\left[\mu\left(1 + \frac{\sigma^2}{\mu^2}\right)^{-\frac{1}{2}}\right]$ and $\sigma' = \left(1 + \frac{\sigma^2}{\mu^2}\right)^{\frac{1}{2}}$.

$2^M - 1$ possible edge types, which determines the edge heterogeneity parameter $n$. To ensure a finite space of configurations $(\mathbf{k}, \mathbf{m})$, we require discrete edge types, and therefore a discrete set of weights. For simplicity, we assume a uniform weight distribution with each layer when solving the AMEs, such that each edge type is associated with a single weight. It is straightforward to relax this assumption to general discrete weight distributions, but at the cost of computational complexity.

Since we restrict ourselves to non-recovery dynamics with stepwise activation rate $F_{\mathbf{k},\mathbf{m}}$, the reduced master equations are applicable. Whereas Eqn. 2.70 allows for spontaneous activation, and thus contains an explicit time dependence, we are mostly in the case of $p = 0$. Here, aside from an initial seed of active nodes, activation only occurs via the threshold mechanism, and the corresponding reduced system is autonomous.

$$\dot{\nu}_j = g_j(\boldsymbol{\nu}) - \nu_j, \tag{4.9a}$$

$$\dot{\rho} = h(\boldsymbol{\nu}) - \rho, \tag{4.9b}$$

where the functions $g_j(\boldsymbol{\nu})$ and $h(\boldsymbol{\nu})$ are obtained by setting $t = 0$ in Eqns. 2.65 and 2.69, such that

$$g_j(\boldsymbol{\nu}) = \sum_{\mathbf{k}} \frac{k_j}{c_j} p_{\mathbf{k}} \sum_{\mathbf{m}} f(\mathbf{k}, \mathbf{m}) B_{k_j - 1, m_j}(\nu_j) \prod_{i \neq j}^{2^M - 1} B_{k_i, m_i}(\nu_i) \tag{4.10}$$

and

$$h(\boldsymbol{\nu}) = \sum_{\mathbf{k}} p_{\mathbf{k}} \sum_{\mathbf{m}} f(\mathbf{k}, \mathbf{m}) \prod_{j=1}^{2^M - 1} B_{k_j, m_j}(\nu_j), \tag{4.11}$$

with $B_{k_i, m_i}(\nu_i)$ the binomial distribution. The function $f(\mathbf{k}, \mathbf{m})$, implementing the response of a node with degree vector $\mathbf{k}$ to a set of active neighbours encoded by $\mathbf{m}$, is equal to 1 if one of the conditions in Table 4.2 is satisfied, and 0 otherwise (see Section 4.3.2 for explicit expressions of the response function for $M = 2$ in all multiplex threshold rules explored here). Except in the experiments presented in Figs. 4.3 and 4.4, we set $p = 0$ and apply the time-independent solution provided by Eqn. 4.9. The computational complexity between the two systems is unchanged.

In Eqns. 4.10 and 4.11, $P(\mathbf{k})$ is the probability that a randomly selected node has degree vector $\mathbf{k}$. Given that our multiplex network is maximally random up to the degree distribution $P_j(k_j)$ of each edge

type $j = 1, \ldots, 2^M - 1$, and that the corresponding degrees $k_j$ are uncorrelated, $P(\mathbf{k})$ is the product of all edge-type degree probabilities,

$$P(\mathbf{k}) = \prod_{j=1}^{2^M-1} P_j(k_j), \tag{4.12}$$

with $c_j = \sum_{k_j} k_j P_j(k_j)$ the average degree for edge type $j$. If $G_j(u) = \sum_{k_j} P_j(k_j) u^{k_j}$ is the probability-generating function associated with edge type $j$, then the aggregate degree $k = \sum_j k_j$ has probability-generating function $G(u) = \sum_k P(k) u^k = \prod_j G_j(u)$, from which the aggregate degree distribution $P(k)$ can be obtained.

### 4.3.2 *Eigenvalues for $M = 2$ layers*

Here we analyse the simple case of a multiplex network of $M = 2$ layers, with or without overlap ($\gamma > 0$ or $\gamma = 0$, respectively), and with Poissonian degree distributions for all composite edge types. There are three edge types; two resulting from node pairs connected in exactly one layer ($j = 1, 2$), and one composite edge resulting from node pairs connected in both layers ($j = 3$). The degree, partial degree, and weight vectors are $\mathbf{k} = (k_1, k_2, k_3)$, $\mathbf{m} = (m_1, m_2, m_3)$, and $\mathbf{w} = (w_1, w_2, w_3)$, respectively, subject to the constraints $k = k_1 + k_2 + k_3$, $m = m_1 + m_2 + m_3$, and $w_3 = w_1 + w_2$ (assuming that weights are additive over composite edges).

We consider Poissonian degree distributions for all edge types,

$$P_j(k_j) = \frac{{c_j}^{k_j} e^{-c_j}}{k_j!}, \tag{4.13}$$

such that the only tunable parameter is the average degree for edge type $j$, $c_j = \sum_{k_j} k_j P_j(k_j)$. The probability-generating function of Eqn. 4.13 is $G_j(u) = e^{c_j(u-1)}$, from which the probability-generating function of the aggregate degree $k = \sum_j k_j$ takes the form $G(u) = \prod_j G_j(u) = e^{\sum_j c_j(u-1)}$. Then, $k$ also follows a Poisson distribution $P(k) = c^k e^{-c}/k!$ with average aggregate degree $c = \sum_j c_j$. The total number of edges a node has in layer $i = 1, 2$ is the sum of its composite edges of type $i$ plus the overlap edges of type $j = 3$, i.e. $k_i + k_3$, which is also Poisson distributed. Then, the average degrees $z_i$ in layer $i = 1, 2$ and the total average degree $z = z_1 + z_2$ are given by

$$z_i = c_i + c_3, \tag{4.14}$$

$$z = c_1 + c_2 + 2c_3. \tag{4.15}$$

As stated in Section 4.2.1, we implement intimacy circle theory by considering the scaling $z_2 = \delta_z z_1$ ($\delta_z \geq 1$) and $w_2 = \delta_w w_1$ ($\delta_w \leq 1$). Since layer overlap is defined as $\gamma = |E_1 \cap E_2|/|E_1|$ with $E_i$ the edge set in layer $i = 1, 2$, we may also write $\gamma = (Nc_3/2)/(Nz_1/2) = c_3/z_1$, where $N$ is the size of the network. Assuming that all edges in layer $i = 1, 2$ have the same weight $w_i$ (and $w_3 = w_1 + w_2$), the average weight in the network is $\langle w \rangle = \sum_j c_j w_j / c$. Overall, we can choose a set of four parameters, say $z$, $\delta_z$, $\delta_w$, and $\gamma$, together with the arbitrary constraint $\langle w \rangle = 1$, and use these relations to write the remainder of the network variables as

$$\begin{cases} z_1 = \dfrac{z}{1+\delta_z} \\ z_2 = \delta_z z_1, \end{cases} \qquad \begin{cases} w_1 = \dfrac{1+\delta_z-\gamma}{1+\delta_z\delta_w} \\ w_2 = \delta_w w_1 \end{cases} \qquad \text{and} \qquad \begin{cases} c_1 = z_1(1-\gamma) \\ c_2 = z_1(\delta_z - \gamma) \\ c_3 = z_1\gamma. \end{cases} \tag{4.16}$$

Finally, we write the response function $f(\mathbf{k}, \mathbf{m})$ explicitly in the case of a multiplex network of $M = 2$ layers, as per Table 4.2. For a single threshold $\phi$ for all nodes in the network, the weighted sum threshold rule implies $f(\mathbf{k}, \mathbf{m}) = 1$ for

$$\mathbf{m} \cdot \mathbf{w} \geq \phi \mathbf{k} \cdot \mathbf{w} \quad \text{and} \quad k > 0, \tag{4.17}$$

and $f(\mathbf{k}, \mathbf{m}) = 0$ otherwise. If a threshold $\phi_i$ is defined in layer $i = 1, 2$, the multiplex *and* and *or* rules imply $f(\mathbf{k}, \mathbf{m}) = 1$ for

$$m_i + m_3 \geq \phi_i(k_i + k_3) \quad \text{and} \quad k_i + k_3 > 0, \tag{4.18}$$

for either $i = 1$ or 2, in the case of the *or* rule, and $i = 1$ and 2, in the case of the *and* rule. Otherwise, $f(\mathbf{k}, \mathbf{m}) = 0$. Eqns. 4.13 and 4.18 allow us to write Eqn. 2.76 explicitly and solve its characteristic equation, which we do below for the cases of non-overlapping and overlapping layers.

#### 4.3.2.1 *Non-overlapping layers, $\gamma = 0$*

In the case of no overlap, $\gamma = 0$, $k_3 = m_3 = 0$ for all nodes, effectively reducing the Jacobian matrix $\mathbf{J}^*$ of Eqn. 2.76 to two dimensions. Then, the characteristic equation $\lambda^2 - \lambda \mathrm{Tr}\mathbf{J}^* + J^*_{11}J^*_{22} - J^*_{12}J^*_{21} = 0$ has solutions

$$\lambda_\pm = \frac{1}{2}\left[\mathrm{Tr}\mathbf{J}^* \pm \sqrt{(J^*_{11} - J^*_{22})^2 + 4J^*_{12}J^*_{21}}\right], \tag{4.19}$$

where $\mathrm{Tr}\mathbf{J}^* = J^*_{11} + J^*_{22}$ is the trace of the Jacobian matrix. From Eqn. 2.76 we have $J^*_{12}, J^*_{21} \geq 0$, so the eigenvalues in Eqn. 4.19 are

real numbers (with the largest corresponding to the + sign). We may write the cascade condition of complex contagion in multiplex networks (for $p = \gamma = 0$) as

$$\mathrm{Tr}\mathbf{J}^* + \sqrt{(J_{11}^* - J_{22}^*)^2 + 4J_{12}^* J_{21}^*} > 0, \tag{4.20}$$

an equation determining the region in $(\phi, z)$ space where infinitesimal perturbations can trigger global cascades [163, 166].

#### 4.3.2.2 *Overlapping layers, $\gamma > 0$*

In an overlapping multiplex network, $\gamma > 0$, with $M = 2$ layers, the Jacobian matrix $\mathbf{J}^*$ of Eqn. 2.76 is three-dimensional. [4] The characteristic equation is $j_0 + j_1\lambda + j_2\lambda^2 + j_3\lambda^3 = 0$, where

$$j_0 = \det(\mathbf{J}^*), \quad j_1 = -\frac{1}{2}\left[\mathrm{Tr}^2(\mathbf{J}^*) - \mathrm{Tr}(\mathbf{J}^{*2})\right], \quad j_2 = \mathrm{Tr}(\mathbf{J}^*), \quad j_3 = -1. \tag{4.21}$$

Instead of using the general methods of Cardano or Lagrange, we may find a trigonometric solution by making the affine transformation $\mathbf{J}^* = a\mathbf{A} + b\mathbf{1}$ for arbitrary constants $a$ and $b$. If $\mathbf{v}$ is the eigenvector of $\mathbf{J}^*$ associated with eigenvalue $\lambda$ ($\mathbf{J}^*\mathbf{v} = \lambda\mathbf{v}$), then $\mathbf{A}\mathbf{v} = c\mathbf{v}$ with $\lambda = ac + b$. By solving the eigenvalue problem for the affine transformation $\mathbf{A}$ we can find the eigenvalues of the original Jacobian. We take

$$a = \sqrt{\frac{1}{6}\mathrm{Tr}(\mathbf{J}^* - b\mathbf{1})^2}, \qquad b = \frac{1}{3}\mathrm{Tr}(\mathbf{J}^*), \tag{4.22}$$

for which $\mathrm{Tr}(\mathbf{A}) = 0$ and $\mathrm{Tr}(\mathbf{A}^2) = 6$. The characteristic equation for $\mathbf{A}$ is $\det(\mathbf{A}) + 3c - c^3 = 0$ with discriminant $\Delta = 4 - \det^2(\mathbf{A})$, so we have three (distinct or multiple) real roots for $|\det(\mathbf{A})| < 2$. By making the change of variable $c = 2\cos\alpha$ and using the trigonometric identity $\cos 3\alpha = 4\cos^3\alpha - 3\cos\alpha$, we finally write the eigenvalues $\lambda_l$ of the Jacobian matrix $\mathbf{J}^*$ as

$$\lambda_l = 2a\cos\left[\frac{1}{3}\arccos\left(\frac{1}{2}\det\left[\frac{1}{a}\left(\mathbf{J}^* - b\mathbf{1}\right)\right]\right) + \frac{2\pi}{3}l\right] + b, \tag{4.23}$$

for $l = 0, 1, 2$. The cascade condition for complex contagion in multiplex networks (for $p = 0$ and $\gamma > 0$) is for the leading eigenvalue in Eqn. 4.23 to be positive, $\max\{\lambda_l\} > 0$.

4. A possible exception is the case of maximal overlap where $\gamma = 1$, and edges of type $j = 1$ are absent. This is due to the assumption that $|E_1| < |E_2|$. Here, the Jacobian can again be reduced to two dimensions, as was the case for $\gamma = 0$ where $j = 3$ edges were absent.

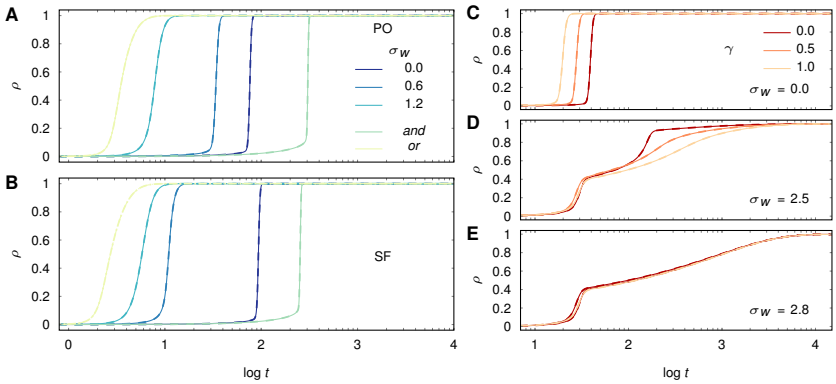


**Figure 4.3 –** Density $\rho(t)$ of active nodes as a function of time $t$. (a-b) Accelerative effect of increasing weight standard deviation $\sigma_w$ in Poisson (PO) and scale-free (SF) configuration-model duplexes, with relative density $\delta_z = 2$ and threshold $\phi = 0.18$. The multiplex *and* and *or* threshold rules can be viewed as upper and lower bounds, respectively, on spreading time. Layer thresholds are $\phi_i = 0.18$ for $i = 1, 2$. (c-e) Diminishing effect of layer overlap $\gamma$ when increasing weight standard deviation $\sigma_w$ in an ER duplex with $\delta_z = 10$ and $\phi = 0.15$. Simulations (solid lines) are averaged over $10^3$ realisations of a network of $N = 10^7$ nodes and average degree $z = 10$. Rate of spontaneous activation is $p = 2 \times 10^{-4}$. AME solutions (dashed lines) are indistinguishable from numerical simulations for all definitions of the threshold rule.

### 4.3.3 *Dynamic solution of approximate master equation*

As a first look at the emergent dynamics of our model, it is useful to consider the case with non-zero rate of spontaneous activation $p$, which is permitted in our model by using an activation function of the form

$$F_{\mathbf{k},\mathbf{m}} = \begin{cases} 1, & \mathbf{m} \cdot \mathbf{w} \geq \mathbf{k} \cdot \mathbf{w} \\ p, & \text{otherwise,} \end{cases} \tag{4.24}$$

which is essentially the weighted sum rule for a uniform weight distribution $P_i(w)$ (see Table 4.2), with noise at rate $p$. Using this update rule, we explore the effects of varying edge weight heterogeneity and layer overlap, Fig. 4.3. The former case can be studied by tuning the weight scaling factor $\delta_w$, or equivalently, the total weight standard deviation $\sigma_w$ defined by $\sigma_w^2 = \sum_{i=1}^{M} c_i (w_i - \langle w \rangle)^2 / c$, where $c_i/c$ is the probability that a randomly selected edge has weight $w_i$, $c_i$ being the average degree in that layer. In configuration model networks with Poisson (PO) or scale-free (SF) degree distributions, varying edge weight het-

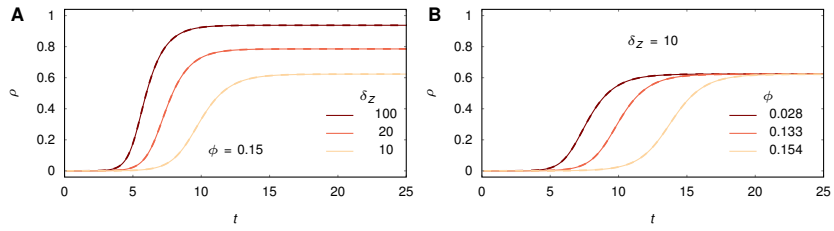


**Figure 4.4 –** Density $\rho(t)$ of active nodes as a function of time $t$. (a-b) Decelerative effect of decreasing density skewness $\delta_z$ and increasing threshold $\phi$ in a Poisson distributed configuration-model multiplex, with $M = 2$. (a) decreasing density skewness $\delta_z$ increases the frequency of "immune" configurations (see text), which serve to slow the spreading process, and decrease the size of global cascades. Similarly, increasing the threshold $\phi$, in (b), slows the time taken to reach maximal saturation, by increasing the reliance on spontaneous activation. Simulations (solid lines) are averaged over $10^3$ realisations of a network of $N = 10^7$ nodes and average degree $z = 10$. Rate of spontaneous activation is $p = 2 \times 10^{-4}$. AME solutions (dashed lines) are indistinguishable from numerical simulations for all definitions of the threshold rule.

erogeneity can accelerate or decelerate spreading [Fig. 4.3(a) and (b)]. Such effects have been studied in detail previously [190], where spreading speed was shown to be determined by competition between node configurations $(\mathbf{k}, \mathbf{m})$ whose activation is either facilitated or hindered by a given weight distribution. To quantify the effects of layer overlap on the dynamics, we consider the simplest case of $M = 2$ layers, and vary the overlap parameter $\gamma$ defined in Section 4.4.1 Via a similar mechanism as edge weight heterogeneity, layer overlap tends to speed up or slow down the process of contagion [see Fig. 4.3(c-e)]. This occurs since a change in $\gamma$ implicitly varies the weight distribution in the monoplex projection of the social network, with weights being additive for overlapping edges. However, when weight heterogeneity is large, a composite edge is always dominated by the layer with the strongest weight, thus diminishing the effect of layer overlap. This is most easily seen in Fig. 4.3(e), where interpolation between zero and maximal layer overlap have negligible effect on the resultant spreading process. Moreover, we observe partial cascades [190] as plateaus in $\rho(t)$ where spreading is temporarily slower.

Results obtained via master equations (dashed) and simulations (solid) are in excellent agreement in Fig. 4.3, becoming indistinguishable for very large networks. Interestingly, experiments allowing noise have reduced fluctuations in the resulting spreading curves compared to

the $p = 0$, or single seed case. This is because noise at rate $p$ can be construed as the addition of randomly chosen seeds sampled uniformly over the entire network. Whereas a single seed experiment may exhibit large variations in the final outcome due to the dependence on the local topology of the seed, spontaneous activation ensures a global cascade. As such, if $p > 0$ there is little sensitivity to the local configurations $\mathbf{k}$ of the initial active nodes and their nearest neighbours. For example, if $p = 10^{-4}$ and $N = 10^7$, $10^3$ nodes are expected to activate in the first "global" time step, uniformly sampled over the whole network.

## 4.4 WATTS PHASE SPACE FOR MULTIPLEX NETWORKS

In Section 4.3.3 we examined spreading dynamics over time, where due to the presence of random noise, the system was guaranteed to reach non-zero levels of activation $\rho$. For the remainder of this chapter, we will be interested the response of the system to an infinitesimal initial seed, over the Watts phase space $(\phi, z)$ of threshold and total average degree. Analytically, this entails linear stability analysis of space to determine the regions allowing global cascades (see for example the dash-dotted lines in Figs. 4.5 and 4.6, or the shaded intervals in Fig. 4.7).

In what follows we study the response of the network to an infinitesimal perturbation, or single active seed, and record the relative frequency $f_g$ of global cascades via Monte Carlo (MC) simulations. Regions in $(\phi, z)$ space with non-zero $f_g$ in the $N \to \infty$ limit are well predicted by the spectrum of Eqn. 2.76. For simplicity we assume uniform edge weights with value $w_i$ within layers, which can be easily generalised (see Section 4.5.3).

### 4.4.1 *Emergence of reentrant phase transitions*

The weighted sum rule leads to a high-$z$ cascading phase, and thus reentrant phase transitions for constant $\phi$, in an $M = 2$ layer multiplex with a log-normal (LN) degree distribution in each layer (Fig. 4.5, distribution details in Section 4.2.3). In two layers, we define layer overlap as $\gamma = |E_1 \cap E_2|/|E_1|$, where $E_i$ is the edge set in layer $i = 1, 2$ ($|E_1| < |E_2|$). We can increase weight heterogeneity by decreasing the weight scaling factor $\delta_w$, resulting in a second cascading regime. As explained in [198], global cascades are due to "vulnerable" nodes with sufficiently low threshold so that a single neighbour can infect them. A cascading phase is formed in $(\phi, z)$ space when vulnerable nodes form a

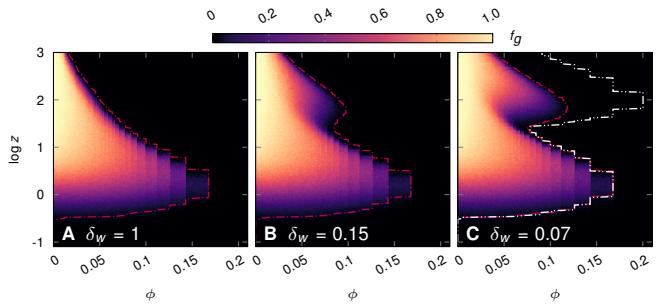


**Figure 4.5 –** Emergence of a high-$z$ cascading phase in $(\phi, z)$ space for the weighted sum rule, for LN degree distribution, fixed $\delta_z = 50$, $\gamma = 0.5$ and decreasing $\delta_w$. MC simulations provide the relative frequency $f_g$ of global cascades, after $10^3$ instances of single node perturbation, in a configuration-model multiplex with $N = 10^6$. In (a) we recover the classic Watts phase diagram ($\delta_w = 1$). The constraint $\langle w \rangle = 1$ means $\mathbf{w} = (1, 1)^T$, $(6, 0.9)^T$ and $(11, 0.8)^T$, from (a) to (c). The outer contour (dash-double dotted white line) in (c) shows the case $\delta_w \to 0$ [$\delta_w = 10^{-3}$; see heat map in Fig. 4.6(a)]. Dash-dotted red lines show agreement with LSA prediction.

percolating cluster. In single-layer unweighted networks, large $z$ results in most nodes being stable against neighbour activation, and cascades becoming exponentially rare. However, under the weighted sum rule, weight heterogeneity allows one high-influence active neighbour to dominate a node's total received influence if remaining neighbours have low influence. Crucially, such configurations are abundant when the conditions $\delta_z > 1$ and $\delta_w < 1$ are satisfied simultaneously, resulting in a percolating vulnerable cluster at high $z$.

In the low-$z$ phase, cascades are mediated by the connectivity of the weak layer, since the strong layer is too sparse to percolate. In the high-$z$ phase, strong edges percolate and determine the stability of adjacent nodes that are otherwise stable to the dense weak layer. Both regions are accurately predicted by LSA [see Fig. 4.5 and the velocity field analysis of Section 4.5.1]. Note that other mechanisms are able to generate additional transitions in $(\phi, z)$ space (e.g., degree assortativity in [209]).

### 4.4.2 *Comparison of multiplex threshold rules*

We compare the behaviour induced by the threshold rules of Table 4.2 for configuration-model multiplexes with LN degree distributions and a real-world multiplex extracted from Twitter (TW) (Fig. 4.6). TW

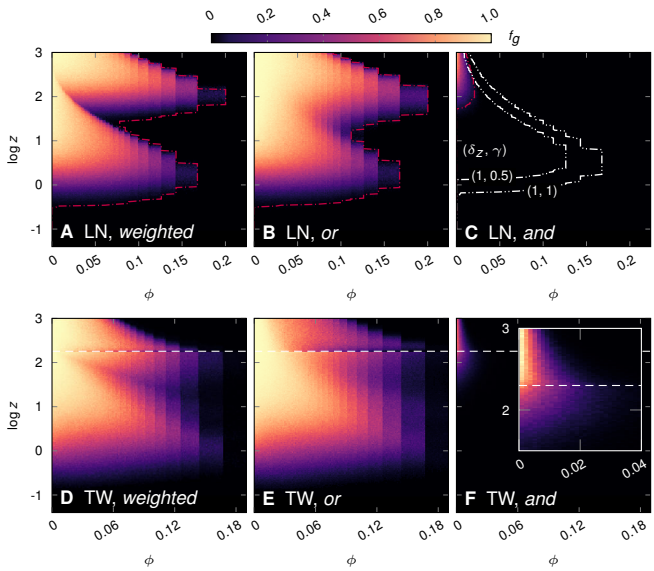


**Figure 4.6 –** Relative frequency $f_g$ of global cascades in LN (top) and TW (bottom) multiplexes with $M = 2$ layers. LN networks in (a-c) are synthetic (standard deviation $\sigma_{k_i} = 2z_i$, overlap $\gamma = 0.5$, and density scaling $\delta_z = 50$). (a) Maximal weight heterogeneity ($\delta_w = 10^{-3}$) leads to reentrant transitions in the weighted sum rule. (b) Reentrant phase transitions also appear for the *or* threshold rule. (c) Under the *and* rule only one global cascading phase emerges, which vanishes when $\gamma = 0$. Decreasing $\delta_z$ and increasing $\gamma$ expands the region of susceptibility to global cascades. See the outer dash-double dotted white contours (the LSA solution for $\delta_z = 1$, with $\gamma = 0.5$ and 1). (d-f) Reentrant phase transitions under the weighted sum and *or* rules in an empirical Twitter network ($\delta_z = 30.2$ and $\gamma = 0.45$). The dashed horizontal line at $z = 166$ is the empirical density, with sparsification providing lower $z$ values, and densification higher $z$ (see Sections 4.7.1 and 4.7.2). (f) A single phase region observed in the *and* multiplex rule. LN and TW networks have size $N = 10^5$ and $N = 3.7 \times 10^5$. We obtain $f_g$ via $10^3$ realisations of single node perturbation. Dash-dotted red lines show the LSA prediction.

comprises a sparse, strongly interacting layer ($z_1 = 5.4$) formed by mutual-mention interactions between $N = 3.7 \times 10^5$ users, and a dense layer of weak links ($z_2 = 163$) formed by the follower network of the same users. The two layers (taken as undirected; data details in Section 4.7) exhibit an overlap $\gamma = 0.45$. In order to explore the effect of single node perturbation over $(\phi, z)$ space, we remove edges uniformly at random from TW, decreasing its average degree $z$ below its observed value of 165.8 [dashed lines in Fig. 4.6(d-f)]. Conversely, we use a

model of network densification known as the Forest-Fire process [136] to extrapolate to higher $z$ values (details in Sections 4.7.1 and 4.7.2).

Assuming the weighted sum threshold rule [Fig. 4.6(a) and (d)], we find reentrant cascading phases under maximal weight heterogeneity ($\delta_w = 10^{-3}$) [for the approach to maximal heterogeneity see Fig. 4.5(a-c) for LN, and Section 4.7.3 for TW]. The multiplex *or* condition also leads to reentrant transitions in both LN and TW networks [Fig. 4.6(b) and (e)]. The onset of the high-$z$ cascading phase, and thus of the reentrant transition, is triggered by the structural percolation of the sparse layer. Since the *or* rule considers influence within layers, and $P_i(w)$ is uniform here, the structure is effectively unweighted, underlining that density skewness is sufficient to trigger a reentrant phase when thresholds are layered. For both LN and TW networks, overlap $\gamma$ and density skewness $\delta_z$ determine the stability under the *and* threshold rule [Fig. 4.6(c) and (f)]. Being the most restrictive condition, the *and* rule suppresses reentrant phase transitions and confines global cascades to a single phase at low $\phi$, with cascades vanishing when $\gamma = 0$. As $\delta_z$ decreases and $\gamma$ increases [Fig. 4.6(c)], overlapping edges, necessary for mediating cascades under the *and* rule, become more abundant and increase the area of the unstable phase [Fig. 4.6(c)]. For simplicity, we set $\phi_i = \phi$ for the *and* and *or* rules. Inspection of the contours of Fig. 4.6(a-c) reveals that the weighted sum rule occupies an area intermediate between the *and* and *or* rules; we perform a comparative eigenvalue analysis in Section 4.5.2 to argue that this is generally the case.

### 4.4.3 *Effect of additional layers*

We illustrate using the weighted sum rule that density skewness $\delta_z$ determines the average degree $z$ at which reentrant phases are triggered [Fig. 4.7(a) and (b)]. This is because the structural percolation transition of individual layers is necessary for the percolation of a subgraph of vulnerable nodes; the value of $z$ at which this occurs depends on $\delta_z$. Increasing the number of layers in the network also creates additional phases of contagion [see Fig. 4.7(c-e) for $M = 3, 4$]. When $\delta_z$ differs between layers, the onset of contagion phases may be delayed or promoted [Fig. 4.7(d)]. In lower phases, strong edges that are too sparse to percolate structurally inhibit cascades driven by edges that are denser but weaker, leading to "partial" cascades that are global but do not fill the network [e.g., lower phase in Fig. 4.7(a)]. This is due to the immunizing effect of strong edges in information diffusion; pairs of inactive nodes connected by a sufficiently strong edge are impossible to infect if all other neighbours are weak, even if all

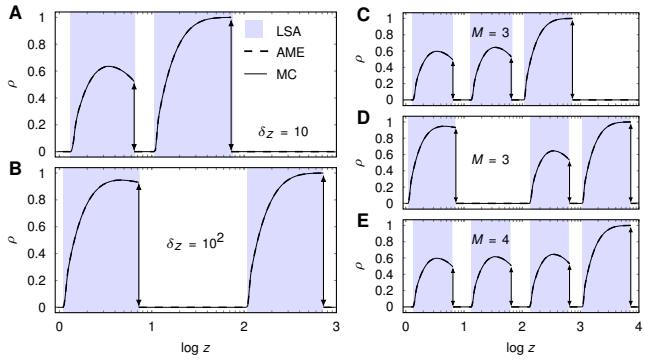


**Figure 4.7 –** Steady state global cascade size as a function of average degree $z$, for constant threshold $\phi = 0.15$ and maximal weight heterogeneity ($\delta_w = 10^{-6}$), using the weighted sum rule. Degree distributions are Poisson and overlap is $\gamma = 0$. Shaded intervals due to LSA indicate systems with a positive leading eigenvalue (see Chapter 2); dashed lines indicate the steady state solution of Eqn. 4.9; and MC solutions are given by the solid curve (error bars narrower than line width). (a, b) Increasing density skewness $\delta_z$ delays the onset of high-$z$ phases of contagion, and allows larger cascades in low-$z$ phases in an $M = 2$ layer multiplex. (c-e) Increasing the number of layers to $M = 3$ and 4 induces 6 and 8 phase transitions in cascade size, respectively. (d) Varying $\delta_z$, such that $z_2/z_1 = 100$, and $z_3/z_2 = 10$. MC results are averaged over $10^3$ realizations of single node perturbation, with $N = 10^7$.

those weak neighbours are active. These configurations are abundant when the strong layer is yet to undergo structural percolation.

## 4.5 UNDERSTANDING REENTRANT PHASE TRANSITIONS

### 4.5.1 *Velocity field analysis*

In this section we illustrate the typical results of the above linear stability analysis [Eqn. 4.20], and compare with the output of Monte Carlo simulation, as well the velocity field of Eqn. 2.70, as shown in Fig. 4.8. We do this for the weighted sum threshold rule, as well as the *or* threshold rule, for identical multiplexes. In Fig. 4.8(a), top and bottom, the region enclosed by dashed lines corresponds to $(\phi, z)$ configurations where the leading eigenvalue $\lambda_+$ is positive, and thus satisfies Eqn. 4.20. In Fig. 4.8(b-f) we show the corresponding velocity field analysis at five points along the $\phi = 0.15$ axis: below the low-$z$ cascade phase at $z = 0.5$, within the low-$z$ phase at $z = 3$, between

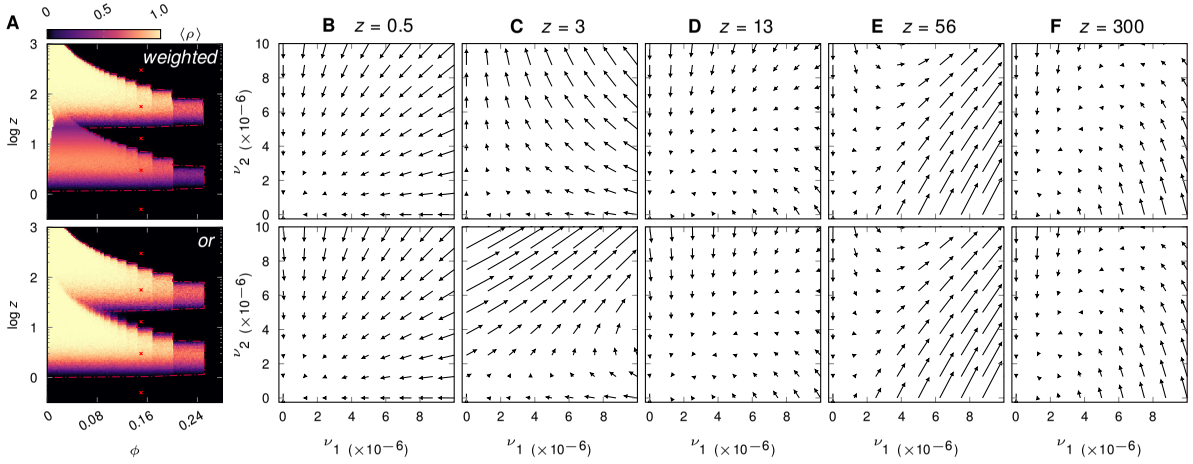


**Figure 4.8 –** Linear stability analysis (a) and velocity field analysis (b–f) around the fixed point $\boldsymbol{\nu} = (\nu_1, \nu_2) = \mathbf{0}$ for two threshold rules, in an $M = 2$ layer non-overlapping multiplex network with Poisson degree distributions in all layers and density skewness $\delta_z = 20$. Top row corresponds to the weighted sum threshold rule, and bottom row to the *or* threshold rule, for otherwise identical configurations. Overlap is $\gamma = 0$, so the third component $\nu_3 = 0$ everywhere, since composite links are absent. The heat map in (a) shows numerical calculations of the fraction of active nodes $\rho$ as a function of threshold $\phi$ and average total degree $z$. The dashed curve in (a) encloses the region where Eqn. 4.20 is satisfied, and the velocity vectors in (b-f) are found by evaluating Eqn. 2.70 for small $\boldsymbol{\nu}$. In (a), numerical calculations fit analytical results perfectly.

cascade phases at $z = 13$, within the high-$z$ phase at $z = 56$, and above the high-$z$ phase at $z = 300$. For $z = 0.5$ in both case, the system is clearly stable, with the initial condition $\boldsymbol{\nu} = \mathbf{0}$ being an attractor. This is due to the lack of connectivity; a giant connected component forms only at $z = 1$ for a Poisson distributed network, meaning the multiplex consists of many small, disconnected components, and a small perturbation cannot develop into a global cascade. In the lower phase, low-weight links ($i = 2$) provide most of the connectivity, being $\delta_z = 10$ times more abundant, and allow for the emergence of a percolating vulnerable cluster. Hence, the system is unstable along the $\nu_2$ axis for both threshold rules. In the case of the weighted sum rule, the sparse but high-weight links of layer one inhibit the size of cascades driven by the sparse layer. This effect is absent for the *or* rule, where layer one links only serve to facilitate cascades, resulting in the increased $\nu_1$ component in Fig. 4.8(c), bottom compared to top.

Between cascade regions at $z = 13$, Fig. 4.8(d), the fixed point $\boldsymbol{\nu} = 0$ is again an attractor, since nodes are stable to low-weight neighbour activation from layer two, and high-weight neighbours from layer one are too sparse to percolate structurally, for both the weighted sum and the *or* threshold rule. At $z = 56$, Fig. 4.8(e), nodes are mostly connected through low-weight neighbours to whom they are stable, but sparse, high-weight neighbours ($i = 1$) now percolate structurally, and dominate the strength of adjacent nodes since weight heterogeneity is maximal in this experiment ($\delta_w = 10^{-3}$). As such, a percolating vulnerable cluster is able to form, and the system becomes unstable along the $\nu_1$ axis. Beyond this phase, at $z = 300$ for example, Fig. 4.8(f), all nodes are stable against active neighbours of all weights, since both layers are excessively dense.

### 4.5.2 *Comparative eigenvalue analysis*

In this section, we develop the claim that the stability of the weighted sum threshold rule is intermediate between that of the *and* and *or* threshold rules. We do so without proof, performing instead a comparative eigenvalue analysis of the Jacobian, Eqn. 2.76, for each of the three definitions of threshold rule, outlined in Table 4.2. By means of elementary limiting arguments over $(\delta_z, \delta_w, \gamma)$, the space of parameters broadly defining our model multiplex, we see that the various rules converge in terms of stability at certain extremities of this space. This allows us to conclude that in general, in response to an initial perturbation our model multiplex is most stable under the *and* rule, least

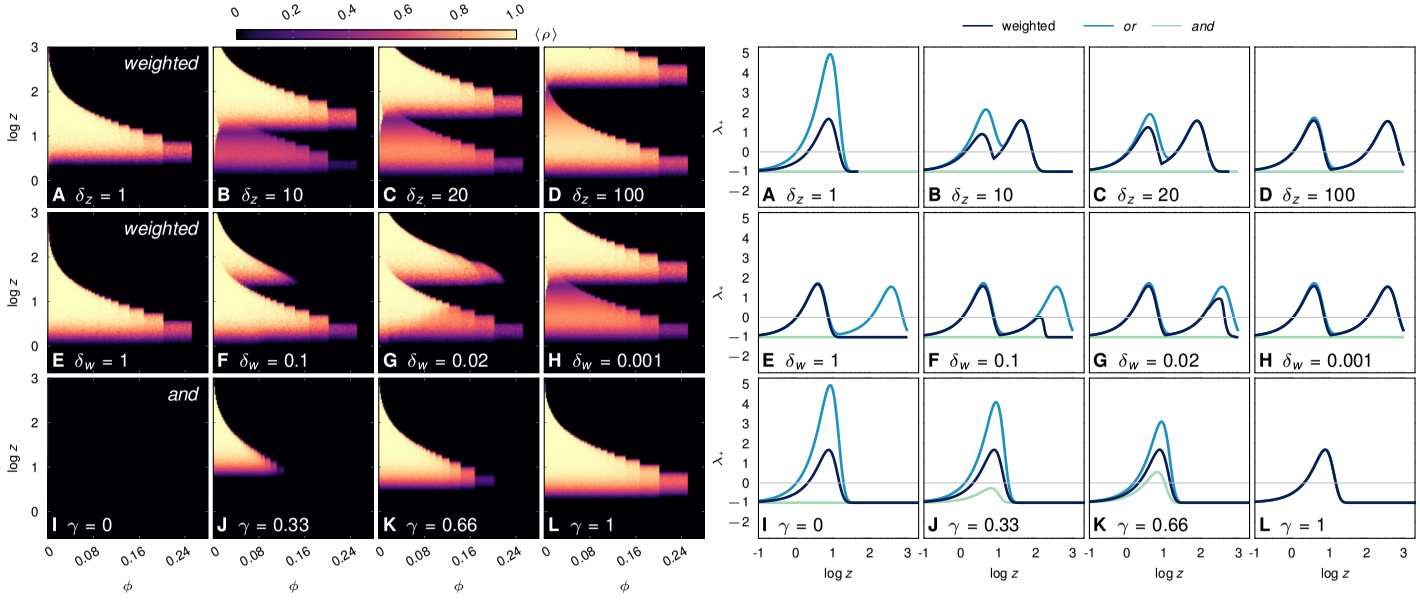


**Figure 4.9 –** Comparison of Monte Carlo simulation with leading eigenvalues $\lambda_+$ of Eqn. 2.76 as a function of increasing $\delta_z$, $\delta_w$ and $\gamma$, for each threshold rule. Heat maps for a selected threshold rule are on the left. As a general rule, $\lambda_+$ for the weighted sum rule is bounded above and below by the *or* and *and* rules, respectively. Heat maps result from $10^3$ single node perturbations of an $N = 10^5$ node multiplex. Eigenvalues are along a constant $\phi = 0.15$ slice of the corresponding heat map.

stable under the *or* rule, with the weighted sum rule providing a level of stability intermediate between the two.

In Fig. 4.9, left, we plot the Watts phase space $(\phi, z)$ along each of the axes of the parameter space $(\delta_z, \delta_w, \gamma)$. On the right of Fig. 4.9, we plot the leading eigenvalues $\lambda_+$ of Eqn. 4.19 for a constant $\phi = 0.15$ slice of each Watts phase space, for each response function defined in Table 4.2. That is, we plot $\lambda_+$ as a function of average total degree $z$ for a given $\phi$, for each threshold rule (in contrast, we do not plot Monte Carlo simulations of each threshold rule, just a representative one). When $\lambda_+ > 0$, the condition for global cascades of contagion is satisfied. Gray horizontal lines on the right of Fig. 4.9 correspond to $\lambda_+ = 0$, the value above which the system becomes unstable, and an infinitesimal perturbation triggers global cascades. When $\lambda_+ < 0$, the system is stable, and no global cascades emerge. It is worthwhile noting that for all $\delta_z$, $\delta_w$ and $\gamma$, in the limit of $z \to 0$ and $z \to \infty$, the eigenvalues of each rule converge at $\lambda_+ = -1$, the minimum value arising from Eqn. 2.76 when the response function is $f = 0$ for all configurations $(\mathbf{k}, \mathbf{m})$. Trivially, this means that cascades are impossible if the multiplex is exceedingly sparse or dense, respectively.

What we observe across all values of $\delta_z$, $\delta_w$ and $\gamma$ in Fig. 4.9, is that the *and* and *or* rules bound the weighted sum rule below and above, respectively, in the magnitude of the leading eigenvalue $\lambda_+$. In other words, the system is always least stable under an *or* response function, and most stable under an *and* response function, with the weighted sum rule intermediate between the two. In particular, these rules converge at the limiting values of $\delta_z$, $\delta_w$, and $\gamma$. Consider first the leading eigenvalue $\lambda_+$ under the *or* rule, which is always larger than or equal to that associated with the weighted projection, as seen in the top two rows of Fig. 4.9. This can be interpreted as being due to the permissiveness of the *or* rule; a node will activate if its threshold $\phi_i$ is satisfied in any layer $i$. To trigger global cascades, an edge type must be of sufficient density such that it percolates structurally, but not so dense that nodes are stable against a single active neighbour of that edge type. This condition roughly determines when the *or* rules leads to global cascades. The same is true of the weighted sum rule, with the additional constraint that edges in this range of density must dominate the local neighbourhood in terms of weight. Clearly, this coupling between layers via the weighted sum rule can only serve to increase the system's stability with respect to the or rule. This results in the *or* rule being at least as unstable everywhere as the weighted sum rule. In all experiments conducted, cascading phases due to the *or* rule begin earlier, and finish later as function of $z$, with respect to corresponding

experiments using the weighted sum rule. This is evident for all $\lambda_+$, from (a) to (l) in Fig. 4.9.

It is relatively straightforward to see why eigenvalues $\lambda_+$ are smaller in the weighted sum rule than in the *or* rule. As explained in Section 4.4.3, this increased stability is due to certain "blocked" configurations that are formed when weight heterogeneity is large. That is, pairs of inactive nodes connected via a high-weight edge, remainder of their neighbourhoods are week. Even if all these weakly interacting neighbours are active, the strong interaction mutually "immunises" each inactive node, ensuring that they remain in this state forever. This leads to partial cascades, evident for example in the lower phase of Fig. 4.9(b). Such cascades spread more slowly due to the presence of these immune configurations. See for example [166], where spreading speed decreases as a result of "blocked" configurations. As such, the effective coupling between layers due to the weighted sum rule increases the stability of these systems with respect to *or* dynamics under the same settings. In other words, the independence of each layer in the *or* rule. This coupling is minimised in the $\delta_z \gg 1$, top row, where the effect of partial cascade diminishes [compare color of bottom phase in heat map of Fig. 4.9(b) and (f), for example], and the dynamics of the two rules converge.

Now consider the *and* rule, which can be viewed as the most restrictive, requiring that a node's threshold is satisfied in every layer before activation takes place. When overlap $\gamma$ is zero, $\lambda_+ = -1$ and the system is stable everywhere, illustrated in the top two rows of Fig. 4.9. This is because overlapping links, which are necessary in order for cascades to develop under the *and* rule, are absent in this case. When overlap is present small perturbations can trigger global cascades, as in the bottom row where we interpolate between no overlap, and maximal overlap. Even when the *and* rule allows a cascading phase, the system is more robust than the corresponding weighted sum response (with cascades emerging later and disappearing sooner in terms $z$; see third column of the bottom row in the rightmost array). An elegant illustration of the relative stability of each rule is when we $\delta_z = 1$, and $\gamma$ increases from 0 to 1. Here we observe the "sandwiching" of the weighted sum rule below and above by the *and* and *or* rules, respectively. This occurs for any value of weight heterogeneity, which is controlled by the skewness parameter $\delta_w$.

The above eigenvalue analysis highlights the contrasting stability of multiplex and aggregated systems, and even suggests that a weighted aggregate has an intermediate stability between the two multiplex behaviours, namely the *and* and *or* dynamics).

### 4.5.3 *Uniform and log-normal weight distributions*

In Section 4.4, for simplicity of presentation as well as ease of analytic solution, we assume that the weight distribution with each layer $i$ is uniform with value $w_i$. In this section we verify the robustness of our principal results when the weight distributions on each layer are no longer uniform. In Fig. 4.10, we progressively increase the standard deviation $\sigma_{w_j}$ of a log-normal weight distribution *within* each layer with mean $w_j$, with row one using the weighted mutliplex rule, row two using the multiplex *or* rule, and row three using the multiplex *and* rule. The mean weight $w_j$ in layer $j$ is determined as before. That is, the means of the weight distributions within each layer are related by $w_{i+1} = \delta_w w_i$. Since weights in layer $i$ have mean $w_i$, and we have applied the constraint $\langle w \rangle = 1$, the system wide average is also 1. The means between layers can be made more heterogeneous by decreasing $\delta_w \in (0, 1]$, and the distributions within layers made more heterogeneous by increasing $\sigma_{w_j} \in [0, \infty)$. As is the case with log-normal degree distributions, we provide a scale to the weight standard deviation by varying it with respect to the mean $w_j$.

When $\sigma_{w_j} = 0$, we recover Fig. 4.6(a-c), where edge weights were uniform with layers. This provides the first column in Fig. 4.10, namely plots (a), (g) and (m). As discussed in Section 4.4.2, when node dynamics are determined by the *and* and *or* rules, i.e., the second and third rows of column one, the multiplex is effectively unweighted. In this case, high-$z$ cascades, and the associated reentrant phase transition, are entirely driven by layer density skewness $\delta_z$ along with individual layer thresholds.

Increasing $\sigma_{w_j} = 0$ across the top row of Fig. 4.10(a-f), using the weighted sum rule, reentrant transitions are observed to emerge even in the presence of very large weight heterogeneity within layers. This is surprising, and highlights that it is the first moment of the layer weight distributions, the mean $w_i$, along with the inter-layer skewness $\delta_w$ that drive high-$z$ cascades under this threshold rule. Similarly, reentrant phase transitions persist under the *or* threshold rule, as seen along row two of Fig. 4.10(g-l), although are not visible under the most strongly heterogeneous distributions ($\sigma_{w_j} = 5w_j$ and $10w_j$). Further, we observe that the cascading phase shrinks to lower and lower $\phi$, for increasing $\sigma_{w_j}$ in row one. In contrast, the effect of weight heterogeneity in the *or* rule in row two is to extend the cascading phase to higher $\phi$. This is surprising, and represents another instance where explicit multiplexity leads to global cascades in settings where a pure weighted structure cannot.

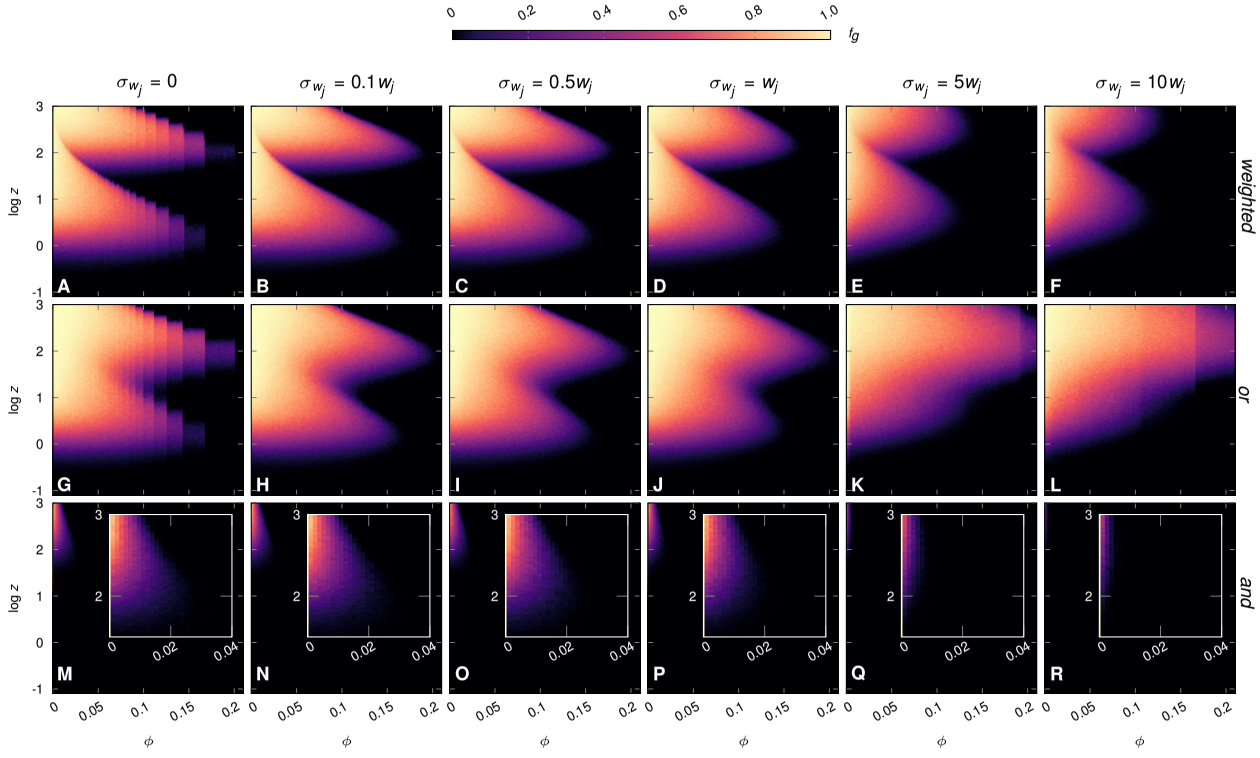


**Figure 4.10 –** Varying the standard deviation of weights in layer $i$, using a log-normal weight distribution with average $w_i$ and standard deviation $\sigma_{w_j}$. The first column corresponds to Fig. 4.6(a-c), with $\sigma_{w_j} = 0$ recovering the uniform weight distribution. We use the weighted sum threshold rule in row one, the *or* threshold rule in row two, and the *and* threshold rule in row three. Networks are of size $N = 10^5$, and we record the frequency $f_g$ of global cascades after $10^3$ realisations of single node perturbation.

Finally, in row three of Fig. 4.10(m-r), where node dynamics are determined by the *and* rule, we see that the effect of intra-layer weight heterogeneity is to diminish the size of the cascading phase. That is, for larger and larger $\sigma_{w_j}$, the cascading phase extends to smaller and smaller $\phi$. This was the case in row one, for the weighted sum rule, and the opposite of the case in row two for the *or* rule.

### 4.5.4 *Weak and strong conditions for reentrant phases*

In this section we summarise the conditions on $\delta_z$ and $\delta_w$ such that a cascading regime at high $z$ is formed, resulting in reentrant phase transitions in cascade size for constant $\phi$ intervals in $(\phi, z)$ space. By way of illustration, we consider an $M = 2$ layer multiplex and vary $\delta_z$ and $\delta_w$, demonstrating three behavioural regimes as follows. Case I represents the null condition in Fig. 4.11(a), where reentrant transitions are not observed in either the weighted sum threshold rule, or the *or* threshold rule. In case II, reentrant transitions are observed only under multiplex-*or* dynamics, called the weak condition in Fig. 4.11(b). In case III, reentrant transitions are observed under both the weighted and multiplex-*or* dynamics, called the strong condition in Fig. 4.11(b). We show the values $w_i$ and their masses $z_i/(z_1 + z_2)$, for a low-$\delta_z$ configuration in (a), and a high $\delta_z$ configuration in (b), each comparing a low and high variance weight distribution (blue and red masses, respectively). For the purposes of this illustration we can assume layer overlap $\gamma = 0$. Further, we use synonymously the expressions "high-$z$ cascades" and reentrant phase transitions.

If density scaling $\delta_z$ is not substantial, as is the case in Fig. 4.11(a), then no reentrant transitions occur regardless of the activation rule. Consider first the dynamics of the weighted sum threshold rule, in a system where all nodes are in the inactive state except for an infinitesimal seed. If $\delta_z$ is close to unity, a node $u$ with a single active neighbour of the strong type is unlikely activate. That is because when $\delta_z \sim 1$, $u$ is likely to have other strong neighbours contributing to its overall influence, meaning one active neighbour is insufficient to overcome the threshold $\phi$, even if that neighbour is of the strong type. This is the case even when weight heterogeneity is maximal, i.e., when $\delta_w \to 0$. Likewise, in the dynamics of the multiplex *or* rule when $\delta_z \sim 1$, a node is likely to have a similar number of neighbours of each type. At high-$z$, this means a typical node configuration has large number of links both layers, meaning that a small perturbation is unlikely to satisfy the threshold in any layer. We thus term $\delta_z \sim 1$ the null condition.

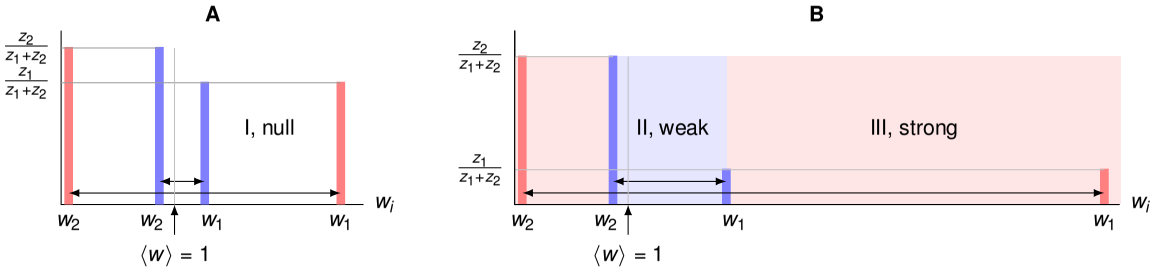


**Figure 4.11 –** Conditions for the observations of a cascading phase at high $z$, resulting in reentrant phase transitions. The null case, when $\delta_z \sim 1$, does not exhibit reentrant transitions, under any dynamic. If the weak condition is met, $\delta_z > 1$ but with low weight variance $\delta_w \sim 1$, reentrant transitions only emerge under the multiplex-*or* case, where there is the increased cost of considering thresholds in individual layers. When the strong condition is met, $\delta_z > 1$ and $\delta_w < 1$, reentrant transitions occur under all dynamics.

Now consider the case where $\delta_z$ is substantially larger than 1, but where weight heterogeneity between layers is low, $\delta_w \sim 1$. This corresponds to the weak condition, case II in Fig. 4.11(b), where reentrant transitions are observed in the *or* rule, but not in the weighted sum rule. Again, let us first consider the weighted sum threshold rule in the presence of an infinitesimal seed. At high-$z$ configurations, a node is likely to have a small number of links from the strong layer, and a large number of links from the weak layer. If a node $u$ has only one neighbour from the strong layer, and that neighbour is active, its influence may still be insufficient to overcome the threshold $\phi$ and infect $u$, as long as the condition $\delta_w \sim 1$ is in place. That is, the sum of the influence from the weak neighbours of $u$, which are numerous when $\delta_z > 1$, is enough for $u$ to remain stable. In contrast, the *or* rule readily leads to activation of $u$ in this setting. By considering its configuration layer by layer and applying the threshold rule, $u$ activates due to all influence in the strong layer being active. However, this comes with a trade-off in complexity; $u$ now has to consciously consider $M$ layer configurations, and make $M$ decisions.

Finally, consider case III in Fig. 4.11(b), identical to that described above, except that weight heterogeneity between layers is now substantial, or $\delta_w < 1$. Under the weighted sum threshold rule, the scenario described above may now lead to the activation of node $u$. This is because the active influence of the one strong neighbour now overwhelms the influence of the inactive weak neighbours, despite them being more numerous. Cascades still occur in the *or* case, since this effect was driven by the density scaling factor $\delta_z$, which has not changed. As such, we recover the strong condition for the observation of reentrant phase transitions. Namely that when $\delta_z > 1$ *simultaneously* with $\delta_w < 1$, the system is vulnerable to cascades at high-$z$ regardless of the functional form of the threshold rule. This leads to reentrant transitions, where the high-$z$ cascading regime is separated from the low-$z$ cascading regime by an intermediate stable phase. As shown in Fig. 4.7(a) and (b), the actual size of this intermediate phase depends on the magnitude of $\delta_z$.

We conclude that a necessary condition for the observation of high-$z$ cascades is that $\delta_z > 1$, regardless of whether dynamics are defined by a weighted sum threshold rule, or the *or* multiplex rule. We refer to this as the weak condition, since it is necessary for high-$z$ cascades under all response functions, but sufficient only in the case of the *or* threshold rule. When the weak condition is not met, we recover the null case of Fig. 4.11(a). In node dynamics following the weighted sum threshold rule, a necessary and sufficient condition for the emergence

of high-$z$ cascades entailing reentrant phase transitions is that $\delta_z > 1$ simultaneously with $\delta_w < 1$. Of course, if these conditions are satisfied, dynamics following the *or* threshold rule also lead to reentrant phases.

## 4.6 MONTE CARLO SIMULATION, AN ASIDE

We implement numerically a multiplex network using the multivariate configuration model, which entails $2^M - 1$ independent applications of the single layer configuration model, one for each composite link type. Since layer overlap is accounted for by the use of composite edges, we do not allow for double edges in the resultant network. Complex contagion is implemented numerically via Monte Carlo simulations of a monotone binary-state dynamics, where nodes are selected uniformly at random for update in asynchronous order, generating a series of time steps. Once a node state changes from inactive to active, it remains so for the rest of the dynamics, thus ensuring a steady state in a finite simulation. Each time step consists of $N$ node updates, where a randomly selected node activates only if the threshold rule is satisfied.

In all experiments in this paper, we are interested in the steady state of the system after single node perturbation. This state is captured by $\rho$, the total fraction of the $N$ nodes in network that are active at $t \to \infty$. In [198], Watts defines a cascade to be "global" if its size is non vanishing in the infinite network limit. In other words, the cascade size is not constant, and occupies a positive fraction of an infinite network. It is straightforward to identify local cascading regimes in Fig. 4.12 via finite size scaling. Here, we plot the outcome $\rho$ of individual realisations, over a range of $z$ and constant $\phi$. Clearly, the noise at $\rho \to 0$ is of constant size, occupying a range $0 < \rho N < 10^2$, regardless of network size. A simple approach to distinguishing local and global cascades is to simulate sufficiently large networks (we choose $N = 10^5$, $10^6$ and $10^7$ throughout this work), and to set a relatively high threshold $\rho_g$ for what constitutes a global cascade. In (b,c) and (e,f), given the magnitude of $N$, cascades larger than $\rho_g = 10^{-2}$ will with high probability be global, and scale with network size. This allows us to define $f_g$, the frequency of occurrence of global cascades, used throughout this work. We also use $\langle \rho \rangle$, the expected final size of cascades that are determined global by the cutoff $\rho_g$.

Finally, it is useful to note that global cascades, once they occur, display very little variance in size. This can be seen in Fig. 4.12(a-c), with variance of global cascade size diminishing for larger and larger $N$. As such, error bars indicating the variance after multiple realisations would be smaller than the point sizes in (b) and (c).

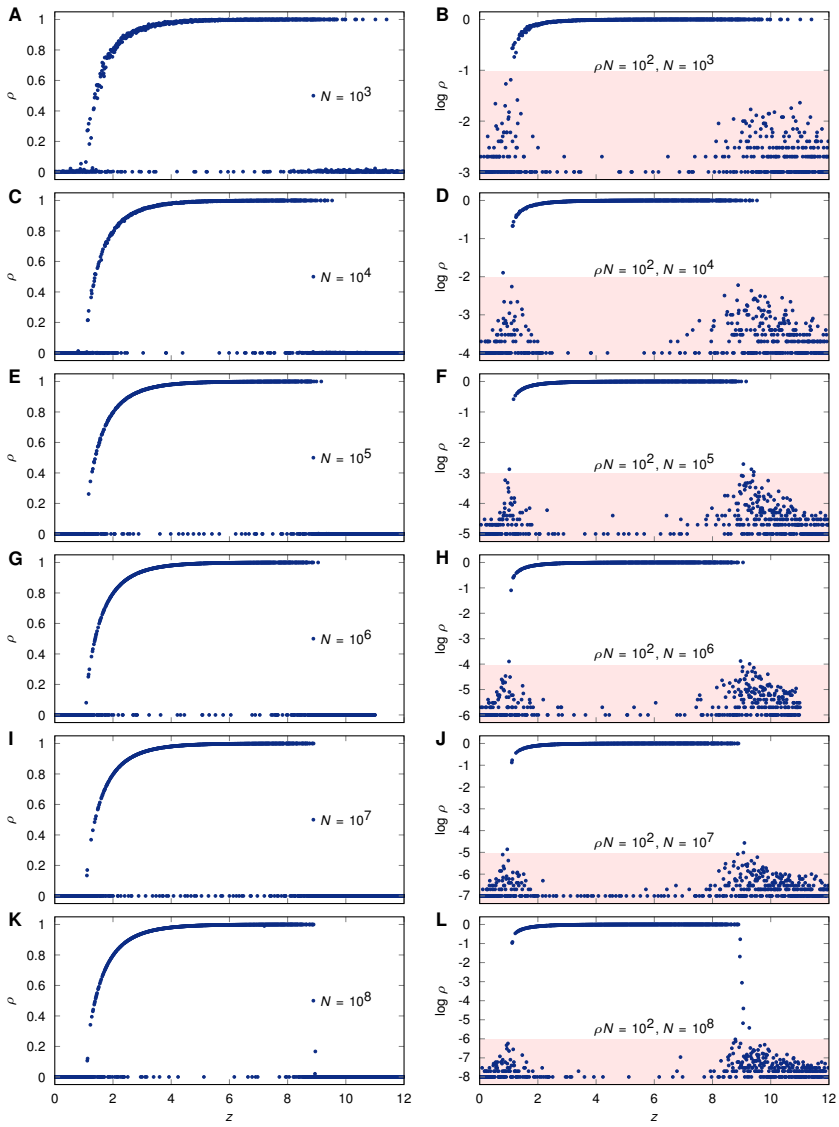


**Figure 4.12 –** Finite-size scaling of individual realisations of single node perturbation on an $M = 2$ layer multiplex with $\delta_z = 1$, $\delta_w = 1$, overlap $\gamma = 0$, with Poisson degree distributions on each layer. Thresholds are uniform, with $\phi = 0.13$, and the "weighted sum" rule applied. As such identical results are expected on and equivalent single layer network. Each point gives the size $\rho N$ of a cascade triggered by a single initial activation. Adjacent plots illustrate identical data, left on a linear scale, right on a log scale. Shaded red interval indicates a low-$\rho$ range of constant size or "local" cascades, where cascade size does not scale with $N$, appearing to exhibit a sharp cutoff of size less than $\rho N = 100$ nodes.

**Table 4.2 –** Number of occurrences $|\bullet|$, and average degree $z$, of each multiplex edge type in an empirical Twitter multiplex. $E_1$ and $E_2$ are the sets of edges in the mutual-mention and follower networks, respectively. Sizes of $E_{10} \equiv E_1 \setminus E_2$, $E_{01} \equiv E_2 \setminus E_1$ and $E_{11} \equiv E_1 \cap E_2$ are shown, and $N = 370,544$.

| | $E$ | $E_1$ | $E_2$ | $E_{10}$ | $E_{01}$ | $E_{11}$ |
|---|---|---|---|---|---|---|
| $\lvert\bullet\rvert$ | 30,717,559 | 999,182 | 30,168,645 | 548,914 | 29,718,377 | 450,268 |
| $z$ | 165.8 | 5.393 | 162.8 | 2.963 | 160.4 | 2.430 |

## 4.7 VERIFICATION ON A TWITTER MULTIPLEX

We validate out model on an empirical Twitter dataset, which was collected over the period of June 2014 and October 2018 through the Twitter Powertrack API provided by Datasift with an access rate of 15%. The data records microblog posts of 140 characters, called tweets, posted in French in the GMT and GMT+1 time zones, together with user profile information. In order to construct a multiplex representation of the proxy social network, first we followed user interactions defined as direct mentions. In Twitter, a mention represents a direct interaction between users in the content of a tweet using the @ symbol (@username). When a user $u$ mentions another user $v$, the tweet containing the mention is visible directly in the feed of user $v$. After creating the network of users who at least once mentioned each other mutually during the observation period, we extracted the second largest connected component of this structure for further investigation. The obtained network contained $N = 370,544$ nodes and 999,182 mutual-mention links. In order to construct a multiplex structure, we considered as a second layer of interaction all follower/followee links between the same set of users, which in turn provided us 30,717,559 links. We argue that while the first mutual-mention layer corresponds to the relatively sparse but strongly interacting layer, the second layer corresponds to the densely connected but weakly interacting layer in our model. [5]

As such, the Twitter multiplex qualitatively supports intimacy circle theory. Consider that mutual-mentions consume both the time and cognitive capacities of each user involved, constraining the number of mentions that can be made by a given individual. In contrast, following another user is inexpensive in terms of time and concentration, and a

5. Following the data handling policy of the company and the GDPR regulations of the EC regarding privacy, the utilized dataset cannot be shared directly. However, similar dataset can be collected via the open API maintained by Twitter or could be constrained using already open datasets.

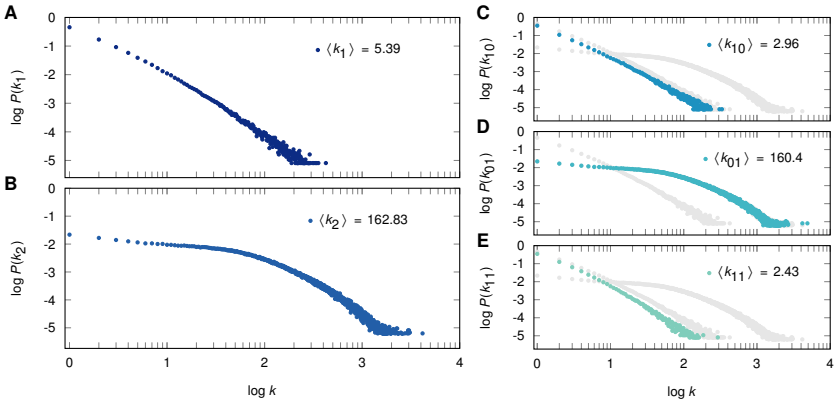


**Figure 4.13 –** Degree distributions for an empirical Twitter dataset. Plots (a) and (b) show layer edge types $i = 1$ and 2, (c) to (e) the resultant edge types $j = 1$, 2 and 3 due to layer overlap.

greater quantity of these relationships can be afforded. Once a user follows another, the cost of maintaining that link is minimal. It can be maintained passively, unlike mutual-mention relationships. As expected, the average mutual-mention degree of $\langle k_1 \rangle = 5.39$ is much smaller than the average follower degree of $\langle k_2 \rangle = 162.83$. In our model, the degree skewness factor therefore equals $\delta_z = 30.2$.

The various degree distributions for the Twitter multiplex are shown in Fig. 4.13. The degree distributions within layers are shown in Fig. 4.13(a) and (b), and are clearly broad tailed. Given that the densities in each layer are relatively skewed, with $\delta_z = 30.2$, and that edge set overlap is substantial, with $\gamma = 0.45$, the degree distributions of the three resultant link types are well approximated by the degree distributions within each layer, Fig. 4.13(c-e). That is, the distributions of $k_{10}$ and $k_{01}$ degrees follow closely those of layer one and layer two degrees, respectively. Finally, the degree $k_{11}$ of the overlapping edge type is well approximated by that of the sparse layer, layer one. This is to be expected given the density skewness $\delta_z = 30.2$, since overlapping links constitute a much larger sample of layer one than layer two links, proportionally. In other words, $|E_{11}|/|E_1| \simeq 0.45$ whereas $|E_{11}|/|E_2| \simeq 0.015$. The former quantity provides the overlap $\gamma = 0.45$.

### 4.7.1 *Sparsification*

Throughout previous sections, we were interested in studying the dependence of the global cascade condition on the average degree $z \equiv z_1 + z_2$ of the underlying multiplex. That is, we wish to explore a phase space $(\phi, z)$, which requires producing networks of a desired average degree $z$. The Twitter network as collected has an average degree of 165.8, and thus requires sparsification to produce samples with a desired average degree lower than this initial value, and densification for average degrees that are higher. Any choice of algorithm can only approximate the network at higher and lower $z$, since we do not have access to historical data showing the multiplex at differing levels of connectivity, nor do we know how the network will evolve beyond its actual state. As such, sparsification and densification algorithms must be used to suggest extrapolations of the empirical network to desired values of connectivity $z$, with the caveat that each algorithm introduces its own biases.

To obtain average degrees of $0.1 \leq z < 165.8$, we sparsify by removing links uniformly at random. Quite simply, we randomly select and remove links sequentially until the original empirical network is reduced to the desired $z$ value. This has the advantage that certain correlations are preserved, such as degree-degree correlations, clustering and community structure. Importantly, this algorithm preserves density skewness $\delta_z$ and overlap $\gamma$, while keeping the overall shape of the degree distribution relatively unaltered. As shown in Fig. 4.6(d-f), sparsified Twitter multiplexes behave in accordance with predictions made on configuration model networks. That is, using the weighted sum and the *or* threshold rules, the sparsifying network undergoes three transitions; the first by exiting the upper cascading phase, the second and third by entering and exiting the lower cascading phase, respectively. Using the *and* multiplex rule the sparsified network passes through a single transition by exiting the cascading phase [Fig. 3(f)], as expected given the Twitter network's layer density skewness $\delta_z = 30.2$ and overlap $\gamma = 0.45$

### 4.7.2 *Densification*

Preliminary experiments indicate that at its original value of $z = 165.8$, the multiplex is susceptible to global cascades under the weighted sum rule, as well as the *or* multiplex rule, over a large interval of thresholds $\phi$. In particular, since both layers have average degree $z_i > 1$, we expect the observed network to be situated within the

**Algorithm 1:** Forest-Fire Process$(G_0 = (V, E_1, E_2), z)$

$G \leftarrow G_0$
**while** $z_G < z$ **do**
  $v \leftarrow$ node chosen u.a.r from $V$
  $V_s \leftarrow \text{Burn}(G, v)$
  $E_1 \leftarrow E_1 \cup \{(v, w) \mid (w, i, j) \in V_s \text{ and } i = 1\}$
  $E_2 \leftarrow E_2 \cup \{(v, w) \mid (w, i, j) \in V_s \text{ and } j = 1\}$
  $G \leftarrow (V, E_1, E_2)$
  $z_G \leftarrow$ average degree of $G$

upper cascading phase. By applying the sparsification algorithm of the previous section, this is indeed found to be true; the network undergoes three phase transitions in susceptibility to global cascades when lowering $z$ from 165.8, to 0.1 [see below the horizontal dashed line in Fig. 4.15(f), and Fig. 4.6(d-f)]. The goal of this section will be to show that by increasing the connectivity of the multiplex, the fourth and final phase transition is traversed. That is, by increasing its average degree, the empirical multiplex can be made stable against global cascades, thus exiting the upper cascading phase.

To extrapolate the empirical Twitter network to average degrees $165.8 < z \leq 1000$, it is not desirable to add links uniformly at random. This is because by the time it grows to $z = 1000$, the network will be almost entirely random, significantly reducing the correlations typical of empirical networks. To incorporate the original structure and preserve as much as possible its empirical correlations, we use a model of densification known as the forest-fire process [116, 136], which is described in Algs. 1 and 2. This process amounts to an extrapolation of the original network to higher $z$, through the probabilistic addition of links biased by the existing structure. The forest-fire model originally proposed by Leskovec *et al.* in [136] has a simple intuitive justification. It is based on having new nodes attach to the network by "burning" through existing edges in an epidemic manner. For example, a new node $u$ attaches to a randomly selected node $v$ in the existing graph, and begins burning through the edges of $v$, attaching to any new node it encounters following a certain probability distribution. In the context of Twitter network growth, this would be interpreted as a new user randomly selecting initial accounts to follow, then browsing the followers of those users in order to find additional accounts to follow, which he does with some probability. The new user continues recursively, using accounts discovered in previous steps to extend their list of contacts,

**Algorithm 2:** Burn$(G = (V, E_1, E_2), v)$

**output:** set $V_s$ of stubs $(u, i, j)$, tuple of $u \in V$ and $i, j \in \{0, 1\}$ indicating $(u, v) \in E_1$, $E_2$

```
V_s ← ∅;
D ← ∅;
enqueue Q with v;
while Q not empty do
    u ← dequeue Q;
    for w ∈ N(u) do
        if w ∉ D then
            p ← real number chosen u.a.r from (0, 1);
            if p < min{1, α/|N(u)|} then
                D ← D ∪ {w};
                enqueue Q with w;
                if (u, w) ∈ E_1 then i ← 1;
                if (u, w) ∈ E_2 then j ← 1;
                V_s ← V_s ∪ {(w, i, j)}
```

until the process dies out. We modify this algorithm in several ways to suit our framework; the authors of [116, 136] consider single layer directed networks, whereas we require a process corresponding to undirected multiplexes. Further, [136] uses a geometric distribution to determine whether to burn through a particular edge. In contrast, we traverse edges adjacent to a node $v$ with probability $\frac{\alpha}{|\mathcal{N}(v)|}$, where $\mathcal{N}(v)$ is the set of neighbours of $v$, meaning $|\mathcal{N}(v)|$ provides the degree of $v$. Here, $\alpha$ is a parameter of the model determining the average number of edges to burn per node. For results in Fig. 4.6(d-f), we set $\alpha = 1$, meaning that when a node $u$ is exploring the neighbours of a node $v$, it selects on average one with whom to connect. If $w \in \mathcal{N}(v)$, the edge type formed between $u$ and $w$ is determined by the type of edge

**Table 4.3 –** Same as Table 4.2, but for the Twitter network extrapolated to $z = 1000$ using forest-fire densification. The layer density skewness $\delta_z = 29.6$, compared to 30.2 in the original. Overlap is $\gamma = 0.45$, as before.

| | $E$ | $E_1$ | $E_2$ | $E_{10}$ | $E_{01}$ | $E_{11}$ |
|---|---|---|---|---|---|---|
| $\lvert\bullet\rvert$ | 185,323,675 | 6,139,522 | 181,954,417 | 3,369,258 | 179,184,153 | 2,770,264 |
| $z$ | 1000 | 33.13 | 982.1 | 18.19 | 967.1 | 14.95 |

$(v, w)$. In other words, if $u$ discovers $w$ via $v$, and $v$ is connected to $w$ via layer two and not layer one, then $u$ also connects to $w$ via layer two and not layer one. In the context of a growing Twitter multiplex, this corresponds to a form of cyclic closure [128, 164]; if a user $u$ discovers a user $w$ via $v$, and $v$ has a follower / followee relationship with $w$ (not going so far as to form a mutual-mention bond), then $u$ is inclined to also form an inexpensive follower relationship with $w$. Although this argument may not apply in every instance, the assumption appears reasonable for a simple model of multiplex densification, and most importantly, preserves the balance of edge types in the multiplex (with density skewness maintained at $\delta_z \simeq 30.2$ even when the original network is extrapolated far beyond its original density, i.e., $z \gg 165.8$).

Finally, to remain consistent with our approach to sparsification which preserves the network size $N$, we implement a variant of the forest-fire process that adds edges without adding nodes. In [136], a new node randomly attaches to a node of the existing graph (termed the "ambassador" node by Leskovec and coauthors), and then starts the forest-fire process at that node. In contrast, we begin each step by randomly selecting an existing node, and performing the forest-fire process from that node. As such, our implementation of the forest-fire process amounts to a randomised version of a graph traversal algorithm, such as a BFS. In fact, in the limiting case of $\alpha \to \infty$, the forest-fire process recovers the BFS. In this case, the node at which we start the forest-fire process eventually discovers every node in the graph. If this occurs at each step of our algorithm, eventually we will have a complete graph. If $\alpha = 0$, no new nodes are discovered and no edges added. As such, $\alpha$ parameterises a Twitter user's tendency to recursively build its network. A consequence of this algorithm is that we're mostly adding short cycles to the network, with edges added after one hop forming 3-cycles, edges added after two hops forming 4-cycles, and so on. [6]

### 4.7.3 *Emergence of unstable phase*

In the final section of this study, we demonstrate the emergence of a high-$z$ cascading phase in the empirical Twitter multiplex, using the weighted sum rule. In so doing, we summarise the arguments of

6. This is desirable since $(i)$ the formation of cycles is well motivated empirically [128, 164], and $(ii)$, our goal in carrying out experiments on the Twitter dataset is to test our results beyond a configuration model setting, where networks are maximally random up to degree distribution, by construction. Clearly, one way this is achieved is through the addition of short cycles.

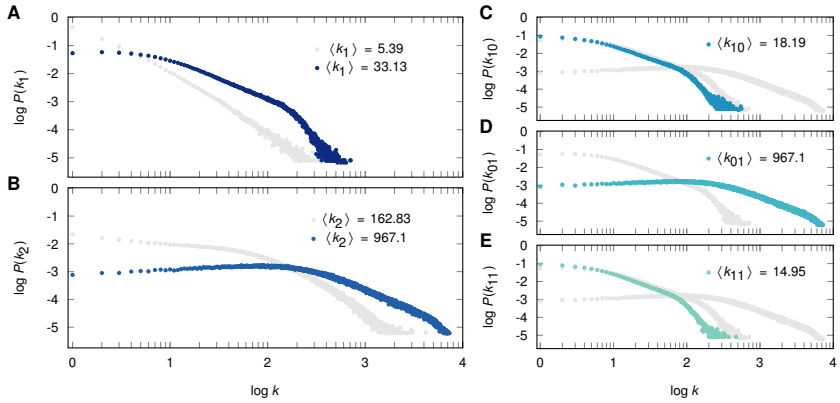


**Figure 4.14 –** Degree distributions of the Twitter network extrapolated to $z = 1000$. Gray curves in (a) and (b) are the original distributions of $k_1$ and $k_2$ degrees [see Fig. 4.13(a) and (b)]. Gray curves in (c-e) are the blue curves in (a) and (b).

previous sections in the language of real networks. Since this experiment involves varying the relative weights $w_i$ in each layer using $\delta_w$, which are assumed to be uniform here, the dynamics of the *and* and *or* threshold rule would be unchanged. In Fig. 4.15(a-f), we vary weight skewness $\delta_w$ from 1, meaning weights are of equal strength in the mutual-mention and follower layers, to $10^{-4}$, such that the strength of interaction in the mutual-mention layer is $10^4$ times stronger than that in the follower layer. Weights are additive in overlapping links, which represent a substantial fraction $\gamma = 0.45$ of the mutual-mention layer. In each panel of Fig. 4.15, we explore a complete $(\phi, z)$ phase space, meaning we perform sparsification and densification to explore below and above the dashed horizontal line, respectively. This experiment is the empirical analogue of that presented in Fig. 4.5, where we vary $\delta_w$ in a configuration model multiplex with log-normal degree distribution. The values of weight skewness $\delta_w$ in Fig. 4.15(a-f) correspond to weights $\mathbf{w} = (1, 1), (7, 0.8), (13, 0.6), (19, 0.4), (25, 0.2)$ and $(30.7, 0.01)$, such that average weight across the entire multiplex gives $\langle w \rangle = 1$, for all plots.

When $\delta_w = 1$, and the mutual-mention network is of equal interaction strength to the follower network, a large overall $z$ ensures that global cascades are exponentially rare [Fig. 4.15(a)]. That is, if the threshold $\phi = 0.1$ for example, it suffices that the total average degree be $z > 30$ to ensure stability against global cascades. In this setting, no cascading phase is observed at high $z$, where the follower network

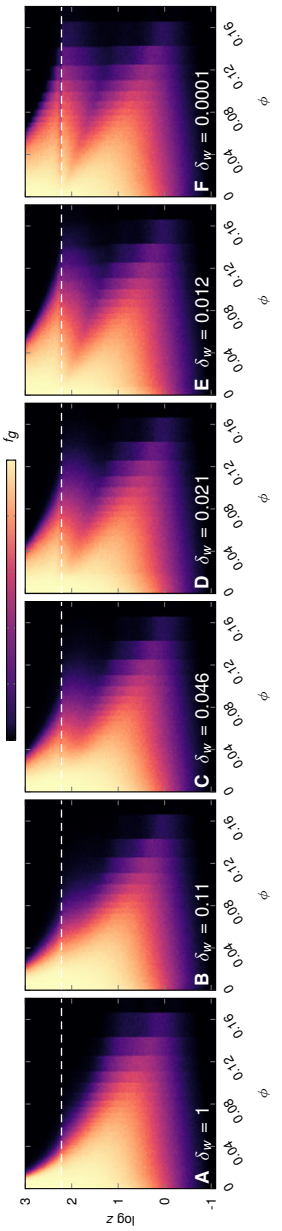


**Figure 4.15 –** Emergence of a high-$z$ cascading phase due to increasing edge weight heterogeneity, or decreasing $\delta_w$, in an empirical Twitter network. Heat maps are the result of $10^3$ realisations of single node perturbation, and give the frequency $f_g$ of the emergence of global cascades. Here, this is an cascade whose steady state is of size $\rho_g = 5 \times 10^{-2}$. Alternatively, this is when a single node perturbation results in cascades of size 18500 or higher. Horizontal dashed line is the original network density of $z = 165.8$.

overwhelms the influence of the mutual-mention network. In contrast, if weight heterogeneity is large, say $\delta_w = 0.012$ as in Fig. 4.15(e), then the multiplex has to reach a much higher average degree $z$ than in the previous case, before becoming stable to perturbations. This is due to the presence of a high-$z$ cascading phase. Now if $\phi = 0.1$, then the overall connectivity must surpass $z > 300$ before global cascades become exponentially rare. This can be understood by noting that the mutual-mention average degree lags behind that of the follower network. In the case of $\delta_w = 1$, by the time the mutual-mention network undergoes structural percolation, follower links are too abundant for influence from mutual-mention links to be perceptible. Taking the mutual-mention layer alone at $\phi = 0.1$ and $z = 5.4$, i.e., at its original value corresponding to the horizontal dashed line in (a), the network would be of ideal density to undergo global cascades. Since, however, the mutual-mention layer is coupled to the follower layer, and here they are of equal influence, $\delta_w = 1$, the nodes in the mutual-mention layer become highly stable. Thus, no cascades are observed at this point. In contrast, when $\delta_w = 0.012$, at that same point $(\phi = 0.1, z_1 = 5.4)$ but in Fig. 4.15(e), mutual-mention links are of sufficient strength to overwhelm the influence from the follower network, despite being vastly outnumbered. Since the strong links are of ideal connectivity to trigger global cascades (percolating structurally and forming a giant component, but not too dense so as to stabilize adjacent nodes), global cascades emerge even from a single initial perturbation.

The above example suggests a straightforward manner to evaluate the susceptibility of an observed network to global cascades at an observed value of $z$. First, one attempts to determine whether the network is comprised of links of heterogeneous interaction strength. If so, and if this heterogeneity is substantial, even a high overall $z$ does not guarantee stability against external shocks. This is not the case when links are of homogeneous strength, where relatively low connectivity $z$ may be sufficient to suppress global cascades.

## 4.8 DISCUSSION

Our results demonstrate that global information cascades emerge in arbitrarily dense networked systems, typically viewed as stable against small perturbations. The types of multiplex structure triggering this behaviour are elementary, and have even been derived from an entropy maximisation process. We have shown that skewness in edge density by layer is necessary for the emergence of reentrant phase transitions under all variants of the threshold rule, but sufficient only

when thresholds are layered and the *or* rule applied. When influence is summed over layers and evaluated with respect to a single threshold, an additional weight skewness condition is necessary. We confirm these phenomena using an analytical formalism that we have extended to multiplex networks, as well as simulation, both on synthetic networks and an empirical Twitter multiplex where all results are recovered. Our results suggest approaches to network design that may promote or suppress system-wide cascades of threshold driven contagion.

In the next chapter we relax one of the major assumptions of this work, namely, that the underlying network is static.

# 5

# THE DYNAMICS OF CASCADES ON BURSTY TEMPORAL NETWORKS

## 5.1 INTRODUCTION TO TEMPORAL NETWORKS

Consideration of real systems make it clear that their underlying networks are far from static, as was assumed throughout the previous chapters in this thesis. Whether it is due to network growth and evolution, or an overlying dynamical process, real networks invariably exhibit temporal behaviour. Whether this is relevant to dynamics such as the spread of information depends on the relative time scale at which the network is evolving. As discussed in Section 1.3 of Chapter 1, the quenched regime refers to network structures that evolves so slowly relative to node dynamics that a static approximation is likely sufficient, whereas the annealed regime refers to networks that evolve so rapidly that their behaviour must be time-averaged in order to infer its interaction with node dynamics. The study of dynamics on networks concerns itself in the intermediate regime, where temporality is critical to the process's evolution.

In this final chapter, we study a general model of binary-state dynamics on temporal networks. In particular, when temporality is defined by a stochastic process, a network may exhibit bursts of event activity [14]. So-called burstiness is a well-studied phenomenon, and refers to the tendency of inter-event times to be heavy tailed. This is in contrast to constant rate or Poisson processes, where very long waiting times between events is exponentially rare. Burstiness is characterised by the presence of these long intervals of inactivity, and corresponding intervals of very high activity.

In the following sections we outline our approach to studying binary-state dynamics on these stochastic temporal networks. We study the effect of our model on threshold driven contagion in the presence of exogenous noise.

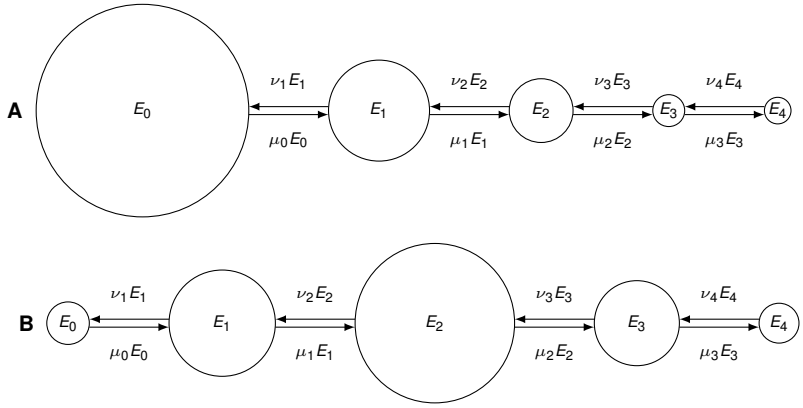


**Figure 5.1 –** Relative sizes of the edge sets $E_j$. In an ensemble of edges following a renewal process $\psi(\tau)$ with memory $\eta$, $E_j$ is the fraction of edges at time $t$ that have observed $j$ events over the interval $[t-\eta, t]$. Alternatively, a single edge following an infinitely long renewal process occupies state $j$ for a fraction $E_j$ of the experiment's duration. Plot (a) illustrates the $E_j$ sets resulting from a heavy-tailed inter-event time distribution $\psi$, where long inter-event times are rare, but are of sufficient duration to cause edges to spend most of their time in state $j = 0$. Plot (b), illustrates the set sizes $E_j$ resulting from an exponentially distributed $\psi(\tau)$. Here, long waits between events are rare, and the expected edge state will recall several events. Edges move between states with rates $\mu_j$ and $\nu_j$.

## 5.2 MODEL DESCRIPTION

### 5.2.1 *Renewal processes*

Our first application of this AME formalism involves a straightforward yet compelling interpretation of temporality. First, we assume a renewal process is taking place on every edge, each following a distribution $\psi(\tau)$, with $\tau$ being the time between consecutive, independent events on a given edge. Edge events represent interactions between the adjacent nodes, such as information exchange or physical contact. Consistent with the framework of the AMEs, the renewal process is continuous in time, and we interpret events as points in time with zero duration. For the time being, the only constraint that we place on the renewal process $\psi$ is that no more than one event may take place over an infinitesimal interval $dt$. Moreover, we attribute to each node a memory of duration $\eta$, allowing it to recall the number of events $j$ that have occurred over the interval $[t-\eta, t]$ for each adjacent edge. As such, at any point in time, we can attribute to every edge a state $j$,

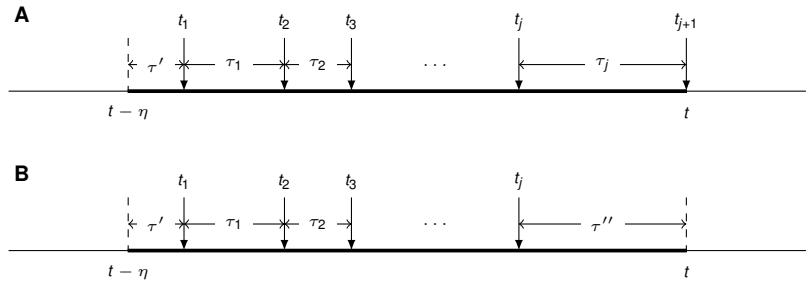


**Figure 5.2 –** Configuration of $j$ events occuring between $[t-\eta, t]$, where $\eta$ is the memory time, used to calculate the rate of positive edge transition $\mu_j dt$. In plot **a** we enumerate all configurations of edges in state $j$ at time $t$, with a $(j+1)$-th event occurring over the interval $t < t_{j+1} < t + dt$. In plot **b** we enumerate all the configurations of edges in state $j$. Inter-event times $\tau_1, \ldots, \tau_j$ are drawn from $\psi$, whereas $\tau'$ and $\tau''$ are drawn from $\Psi$, defined in the text.

corresponding to the number of events that have taken place on that edge over the last $\eta$-window. Further, the state of each edge increments by at most $\pm 1$ over an interval $[t, t + dt]$. In our AME framework, the type $j$ of an edge will correspond exactly to the state of an edge undergoing a renewal process with memory $\eta$. As such, $k_j$ enumerates the number of neighbours connected to a node via edges in state $j$, and $m_j$ represents the corresponding number of infected neighbours.

It is worth noting that the window of memory $\eta$ can in some cases be construed as the duration of an event. That is, one can interpret an instantaneous event happening at time $t$, and the event is remembered by adjacent nodes for a duration $\eta$. This is equivalent, at least with respect to our model, of an event of duration $\eta$ commencing at time $t$, with adjacent nodes forgetting the event immediately after the event terminates, provided that the events in question are not mutually exclusive and can occur simultaneously. The event-duration interpretation may therefore be appropriate for systems whose maximum edge state is constrained to $j = 1$; for example, a distribution $\psi$ whose minimum $\tau$ exceeds $\eta$.

The renewal process that we described above is ergodic, meaning that the properties of an ensemble of such processes at any given time $t$ correspond to the properties of a single process of infinite duration. We introduce the quantity $E_j$ to represent the fraction of time that a single edge spends occupying a state $j$ over the course of an infinite renewal process, and we shall refer to $E$ as the distribution of edge states. We assign $\mu_j dt$ and $\nu_j dt$ to the probabilities that, if the edge is

in state $j$ at a randomly chosen time $t$, it will be in state $j+1$ or $j-1$ at time $t+dt$, respectively. Of course, for a single edge undergoing a renewal process, the interval between an arbitrary time $t$ and the next event is not probabilistic, and is described by the residual time $\tau'$. The edge transition rates $\mu_j dt$ and $\nu_j dt$ apply when nothing is known about a node *a priori*, other than its current state. In the ensemble picture, $E_j$ is the fraction of edges that have state $j$, and the rates $\mu_j dt$ and $\nu_j dt$ are the fraction of edges in class $j$ that increment or decrement state over an interval $dt$, respectively. This idea is illustrated in Figure 5.1 for heavy-tailed and exponential distributions. We shall see that heterogeneity in the distribution $\psi(\tau)$ is captured by heterogeneity in the distribution of edge set sizes $E$, as well as the rates of transitions between them, $\mu_j$ and $\nu_j$. In the following calculations, quantities are most easily interpreted as those of a single edge following an infinitely long renewal process.

*Renewal processes and non-Markovianity*

In this section we clarify we discuss the Markovian and non-Markovian features [1] of our renewal process with an $\eta$ memory window. On the level of a single edge, it is clear that the state $j$ of an edge at time $t+dt$ is determined not by its state at time $t$, but by the time until the next event, or $\tau'$. In other words, once an event is observed at time $t$, the state of the edge is predetermined by the inter-event time $\tau$. It could be argued, on the other hand, that in the so-called "event-space" representation of the renewal process, single-edge dynamics are Markovian. This representation is discrete, and incrementing the system from one point to the next involves randomly sampling from the inter-event time distribution $\psi$. Since consecutive inter-event times are uncorrelated, the system can be described as Markovian in this representation.

Finally, although individual edges are non-Markovian in time $t$, if we define our "system" as being the ensemble of edges in an infinite network, the system as a whole is Markovian.

### 5.2.2 *Density and mixing rates*

A consequence of the "renewal process with memory" is that we induce a distribution of edge states, where $E_j$ gives the distribution of edges in state $j$ and $\mu_j$ and $\nu_j$ the rates of positive and negative edge transition, as outlined in Chapter 2.

1. A process is Markovian if it exhibits the Markov property, i.e., where the state of the system at time $t+dt$, depends only on the state at time $t$.

The rate of positive edge transition $\mu_j dt$ is calculated by finding the probability of a $(j+1)$-th event occurring over an interval $[t, t+dt]$ if there have already been $j$ events in the interval $[t-\eta, t]$. This is illustrated in Figure 5.2, where we shall set $t = \eta$ for convenience. In Figure 5.2, the inter-event times $\tau_1, \ldots, \tau_j$ are drawn from $\psi(\tau)$. The times $\tau'$ and $\tau''$ are drawn from $\Psi(\tau)$, defined as

$$\Psi(\tau) = \int_\tau^\infty \psi(\tau')d\tau', \tag{5.1}$$

which is the probability that the time between events is of duration at least $\tau$. As such, the probability of observing the configuration in Figure 5.2a is $\Psi(\tau')\psi(\tau_1)\ldots\psi(\tau_j)$, which is the same as $\Psi(t_1)\psi(t_2 - t_1)\ldots\psi(\eta - t_j)$ given that we've set the time $t$ to $\eta$. The sum over all such configurations can be expressed as

$$\int_T dt_1 \ldots dt_j \Psi(t_1)\psi(t_2 - t_1)\ldots\psi(\eta - t_j) = \Psi * \psi^{*j}, \tag{5.2}$$

where $\psi^{*j}$ is the $j$-th convolution power of $\psi$, and integration is over the domain $T$ defined by

$$T = (0, \eta] \times (0, \eta - t_1] \times (0, \eta - t_2] \times \ldots \times \left(0, \eta - t_{j-1}\right] \subseteq \mathbf{R}_+^j. \tag{5.3}$$

To obtain the rate $\mu_j dt$, the above quantity must be normalised by the probability $E_j$ that at a randomly selected time $t$, an edge is in state $j$. This corresponds to the illustration in Figure 5.2b, and involves a similar calculation to the one above. The probability of observing the configuration in Figure 5.2b is $\Psi(\tau')\psi(\tau_1)\ldots\psi(\tau_{j-1})\Psi(\tau'')$, which is the same as $\Psi(t_1)\psi(t_2 - t_1)\ldots\psi(t_j - t_{j-1})\Psi(\eta - t_j)$. The normalisation constant is therefore

$$\int_T dt_1 \ldots dt_j \Psi(t_1)\psi(t_2 - t_1)\ldots\psi(t_j - t_{j-1})\Psi(\eta - t_j) = \Psi * \psi^{*(j-1)} * \Psi, \tag{5.4}$$

where the domain $T$ is as defined above. As such $\mu_j dt$, the probability of a positive edge transition from state $j$ to $j+1$ over an interval $dt$, can be written compactly as

$$\mu_j dt = \frac{\Psi * \psi^{*j}}{\Psi * \psi^{*(j-1)} * \Psi} dt, \tag{5.5}$$

which incorporates the details of both the renewal process $\psi$ and the memory function $\eta$. Similar arguments can be used to calculate $\nu_j dt$, the probability of a negative edge transition from $j$ to $j-1$ occurring over an interval $dt$. This allows us to write

$$\nu_j dt = \frac{\Psi * \psi^{*(j-1)}}{\Psi * \psi^{*(j-1)} * \Psi} dt. \tag{5.6}$$

Note that the denominator is the same in the rates for $\mu_j$ and $\nu_j$, as the normalisation constant corresponds to the size of the set $E_j$. Finally, we can argue that by induction from $j = 0$, and using the fact that the distribution $E$ is constant at sationarity of the renewal process, that

$$\mu_j E_j = \nu_{j+1} E_{j+1}, \tag{5.7}$$

which is illustrated in Figure 5.1. The rates $\mu_j$ and $\nu_j$ can be calculated either numerically or analytically, depending on the tractability of the chosen distribution $\psi$. In general, if $\psi(\tau)$ is locally integrable, then the Laplace transform of $\psi$ and $\Psi$ exists and allows us to calculate the convolution as a product in the frequency domain, which will be useful especially if $j$ is large. In particular, if we denote the Laplace transform of $\psi(\tau)$ by

$$\mathcal{L}\{\psi(\tau)\} = \hat{\psi}(s) = \int_0^\infty \psi(\tau) e^{-s\tau} d\tau \tag{5.8}$$

then the transform of Equation 5.1 can be written as

$$\hat{\Psi}(s) = \frac{1 - \hat{\psi}(s)}{s}. \tag{5.9}$$

As illustrated in Figure 5.3, the observed experimental rates $\mu_j$, $\nu_j$ and $E_j$ closely match the predicted values. In this Figure we've simulated a large temporal network, with $N = 5 \times 10^7$ nodes, constituting an ensemble of edges each following a renewal process with inter-event time distribution $\psi$. Node state is irrelevant to these statistics, so we consider only edge state. At the beginning of each interval $\Delta t = 10^{-3}$, we count the number of edges in state $j$, then count those who make a transition to state $j \pm 1$ over the interval $\Delta t$. The number of $j$-state edges becoming $(j+1)$ and $(j-1)$-type edges over this interval, normalised by the total number of edges in state $j$, give the positive and negative edge transition rates $\mu_j$ and $\nu_j$, respectively.

### 5.2.3 *Initial conditions*

If we assume that at the start of a spreading process edge activity is already at equilibrium, then the probability of a node having degree vector $\mathbf{k}$ is calculated directly using the distribution $E_j = \Psi * \psi^{*(j-1)} * \Psi$. Remember that if edge-state $j$ describes the number of events recalled within $\eta$-memory, then $k_j$ gives the total number of edges adjacent to a node in state $j$, while $\mathbf{k}$ enumerates edge states over the local neighbourhood, $\mathbf{k} = (k_1, \ldots, k_n)^T$, for some maximum allowed edge-state $n$. In contrast, one might imagine initialising a network at $t = 0$

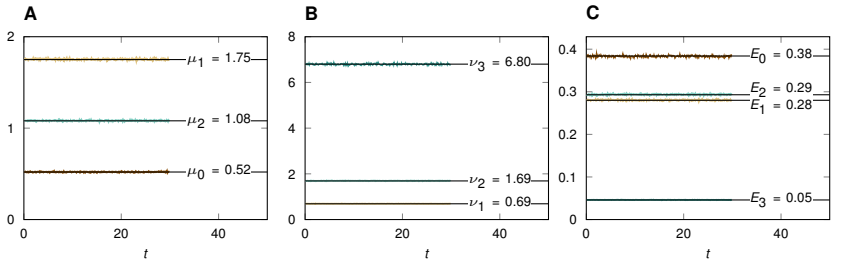


**Figure 5.3 –** Simulation of an ensemble of $10^7$ edges following a power-law inter-event time distribution $\psi(\tau)$, and corresponding theoretical values. Snapshots of $\Delta t = 10^{-3}$ are used to calculate edge state flows. Experiment is coloured, theory results are solid black lines. Power law has exponent $\alpha = 2.5$ and average $\langle\tau\rangle = 1$, and memory is $\eta = 1$.

by assigning all edges to the $E_0$ state. This means that at $t = 0$ no edge activity has yet taken place, and edge states evolve to equilibrium concurrently with the spreading process. To remove the effect of this changing topology, in the following we assume that edge activity has been occurring for a sufficiently long time when we begin the dynamical process at $t = 0$, meaning $E_j$ will be distributed according to their stationary values from the beginning. The Laplace transform of the probability $E_j$ can be written as

$$\begin{aligned} \mathcal{L}\left\{\langle\tau\rangle E_j\right\} &= \mathcal{L}\left\{\Psi * \psi^{*(j-1)} * \Psi\right\} && (5.10)\\ &= \hat{\Psi}\cdot\hat{\psi}^{j-1}\cdot\hat{\Psi} && (5.11)\\ &= \frac{1-\hat{\psi}}{s}\hat{\psi}^{j-1}\frac{1-\hat{\psi}}{s} && (5.12) \end{aligned}$$

in any case where $\psi(\tau)$ is locally integrable.

### 5.2.4 *Model properties*

In this section we list a number of properties that are useful in verifying the correctness of Monte Carlo simulations. First, we have the relation

$$\sum_j \Psi * \psi^{*(j-1)} * \Psi = \langle\tau\rangle, \tag{5.13}$$

the average inter-event time. This is very useful in checking the numerical implementation. Also, consider carefully the case $j = 0$ in the above equations. We get negative convolution powers. Shouldn't sweep this

under the rug. The number of configurations in which an edge can be in state zero, and therefore the average time spent in this state, is

$$E_0 = \int_\eta^\infty \Psi(\tau')d\tau'. \tag{5.14}$$

This is the expression corresponding to probability that there is no event at time $t$, *and* no event until time $t-\eta$, *or* that there is no event at time $t$, *and* no event until time $t-\eta-dt$, *or* that there is no event at time $t$, *and* no event until time $t-\eta-2dt$, and so on. Slightly more compactly, $\Psi(t-t')dt$ is the probability that there has been no event from time $t$ to $t-t'$, including the intervals $[t-dt]$ and $[t-\eta, t-\eta-dt]$. Summing all of these configurations, from $t'=\eta,\ldots,\infty$ gives

$$\begin{aligned} E_0 &= \Psi(t-\eta)+\Psi(t-(\eta+dt))+\Psi(t-(\eta+2dt))+\ldots && (5.15)\\ &= \int_{t-t'}^{-\infty}\Psi(t-t')dt' && (5.16)\\ &= \int_\eta^\infty \Psi(\tau')d\tau', && (5.17)\end{aligned}$$

making the substitution $t-t'=\tau'$. Finally, to calculate $\mu_0$, the fraction of $E_0$ edges that increment over $[t, t+dt]$, we perform a weighted sum of configurations in which an event occurs between $t$ and $t+dt$, and compare this to the total number of configurations. That is,

$$\mu_0 = \frac{\int_\eta^\infty \psi d\tau}{\int_\eta^\infty \Psi d\tau} \quad \text{and} \quad \nu_1 = \frac{\int_\eta^\infty \psi d\tau}{\int_0^\eta \Psi d\tau}, \tag{5.18}$$

as well as

$$E_0 = \int_\eta^\infty \Psi d\tau \quad \text{and} \quad E_1 = \int_0^\eta \Psi d\tau, \tag{5.19}$$

which we use to veryify the emergent transition rates in an $n=2$-level system.

### 5.2.5 *Residual time distribution and CDFs*

With heavy tailed distributions, it is especially important to be able to initialise simulations using the residual time distribution, $\Psi(\tau)$. This is because an ensemble of edges undergoing renewal process takes longer to reach equilibrium when long interevent times, $\tau \gg \langle\tau\rangle$, are significant. Without use of the residual time distribution, the convergence to equilibrium distributions of $E$, $\mu$ and $\nu$ can take an extremely long time.

For simulation purposes it's crucial to relate the CDF of $\psi$ to the residual time distribution. If the probability density function $\psi(\tau)$ is normalised, and $\psi(\tau < 0) = 0$, we have the relation

$$\begin{aligned}\int_{-\infty}^{\infty} \psi(\tau)d\tau &= \int_{0}^{\infty} \psi(\tau)d\tau && (5.20)\\ &= \int_{0}^{\tau} \psi(\tau')d\tau' + \int_{\tau}^{\infty} \psi(\tau')d\tau' && (5.21)\\ &= \Psi(\tau) + \Psi_r(\tau) && (5.22)\\ &= 1, && (5.23)\end{aligned}$$

and we refer to the residual time distribution as $\Psi_r(\tau) = \int_{\tau}^{\infty} \psi(\tau')d\tau'$, and hence make use of the relation $\Psi_r = 1 - \Psi$. Now we concentrate on the lognormal distribution. Note that to initialise our system in a steady state, we apply rejection sampling of the residual distribution $\Psi(\tau)$, rather than inverse sampling.

### 5.2.6 *Networks and dynamics*

Networks are undirected and composed of a single layer, generated using the configuration model. Edge state is determined by our renewal process with memory model, and node state is binary, determined by neighbour and edge state. Here we argue that neighbour influence is proportional to the number of events in $\eta$-memory. As such, the class $(\mathbf{k}, \mathbf{m})$ is determined by the local configuration of event activity and neighbour state, and node activation and recovery occurs at rates $F_{\mathbf{k},\mathbf{m}}$ and $R_{\mathbf{k},\mathbf{m}}$. Thus, fits the framework outlined in Chapter 2. Rather than a single initial seed, we allow random noise, or spontaneous activation at rate $p$.

Up until this point, we have been discussing the edge statistics induced by our temporal network model. The results in the previous section, namely the calculation of rates $\mu_j$ and $\nu_j$, pertain only to the edge dynamics, without saying anything in particular about node dynamics. It remains therefore to specify exactly how a spreading process might result from this time-dependent topology.

We apply our temporal network model to node dynamics defined by the Watts threshold model. It is useful to mention that thanks to the generality of the AME formalism, there are a number of ways of defining the infection rates, $F_{\mathbf{k},\mathbf{m}}$ in the context of renewal processes with memory $\eta$. The most straightforward approach is to say that influence of a node upon its neighbour is directly proportional to the number of events that have occurred on that edge over the preceding

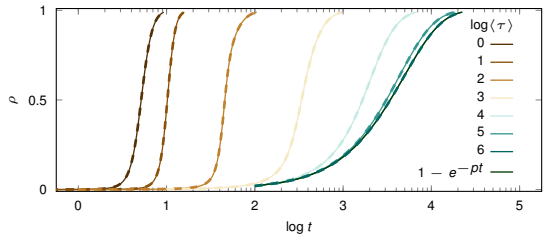


**Figure 5.4 –** Solution of AMEs, dashed, and corresponding Monte Carlo experiment, solid. Exponential inter-event time distribution. Network is a 5-regular random network with $\phi = 0.15$ with an event memory of $\eta = 1$. Increasing the average time between events, $\langle\tau\rangle$ decreases the spreading time to the theoretical slowest value, namely $\rho = 1 - e^{-pt}$, or the solution to $\dot{\rho} = (1-\rho)p$. Network size is $N = 10^6$.

$\eta$-interval. That is, edges in state $j = 0$ have weight $w = 0$, edges in state $j = 1$ have weight $w = 1$, and so on.

We shall set a uniform threshold $\phi$ across the network. Nodes adopt at rate $F_{\mathbf{k},\mathbf{m}} = 1$ if $\mathbf{m}\cdot\mathbf{w} \geq \phi\mathbf{k}\cdot\mathbf{w}$, and $F_{\mathbf{k},\mathbf{m}} = p$ otherwise. We shall see that bursts in activity along inactive-active edges lead to thresholds being overcome.

## 5.3 EXPONENTIAL INTER-EVENT TIME DISTRIBUTION

In this section we explicitly calculate the rates $\mu_j$ and $\nu_j$ for an exponential inter-event time distribution $\psi$. Consider such a distribution with average $\langle\tau\rangle$ defined by

$$\psi(\tau) = \frac{1}{\langle\tau\rangle} e^{-\tau/\langle\tau\rangle} \tag{5.24}$$

and

$$\Psi = e^{-\tau/\langle\tau\rangle}. \tag{5.25}$$

The Laplace transforms of these distributions can be written explicitly as

$$\hat{\psi}(s) = \frac{1}{\langle\tau\rangle s + 1} \tag{5.26}$$

and

$$\hat{\Psi}(s) = \frac{1-\hat{\psi}}{s} = \frac{\langle\tau\rangle}{\langle\tau\rangle s + 1}. \tag{5.27}$$

Substituting these transforms into Equation 5.10, and applying the convolution theorem allows us to calculate the expected size of the set

of edges in state $j$, which is also the normalising constant in the rates $\mu_j$ and $\nu_j$, as the expression

$$\begin{aligned}
\mathcal{L}\left\{\Psi * \psi^{*(j-1)} * \Psi\right\} &= \hat{\Psi} \cdot \hat{\psi}^{j-1} \cdot \hat{\Psi} && (5.28)\\
&= \frac{\langle\tau\rangle}{\langle\tau\rangle s+1}\left(\frac{1}{\langle\tau\rangle s+1}\right)^{j-1}\frac{\langle\tau\rangle}{\langle\tau\rangle s+1} && (5.29)\\
&= \frac{1}{\langle\tau\rangle^{j-1}}\left(\frac{1}{s+\frac{1}{\langle\tau\rangle}}\right)^{j+1}. && (5.30)\\
& && (5.31)
\end{aligned}$$

Using the fact that $\mathcal{L}^{-1}(\frac{1}{s^{j+1}}) = \eta^j/j!$, the scaling property $\mathcal{L}^{-1}(a\hat{\psi}) = a\mathcal{L}^{-1}(\hat{\psi})$, as well as the translation property $\mathcal{L}^{-1}[\hat{\psi}(s+a)] = e^{-a\eta}\psi$, the inverse Laplace transform of the above expression can be written explicitly as

$$\Psi * \psi^{*(j-1)} * \Psi = \langle\tau\rangle \frac{e^{-\eta/\langle\tau\rangle}}{j!}\left(\frac{\eta}{\langle\tau\rangle}\right)^j, \tag{5.32}$$

meaning $E$ is simply the Poisson distribution, as expected. Similarly, the Laplace transform of the numerator in Equations 5.5 and 5.6 can be used to calculate

$$\Psi * \psi^{*j} = \frac{e^{-\eta/\langle\tau\rangle}}{j!}\left(\frac{\eta}{\langle\tau\rangle}\right)^j \tag{5.33}$$

and

$$\Psi * \psi^{*(j-1)} = \frac{e^{-\eta/\langle\tau\rangle}}{(j-1)!}\left(\frac{\eta}{\langle\tau\rangle}\right)^{j-1}. \tag{5.34}$$

Normalising these quantities using Equation 5.32 we obtain

$$\mu_j dt = \frac{1}{\langle\tau\rangle} dt \tag{5.35}$$

and

$$\nu_j dt = \frac{j}{\eta} dt, \tag{5.36}$$

which we are able to derive by much simpler arguments, namely with the definition of the Poisson process. Crucially, $\mu$ has no $j$ dependence in the case of the Poisson process, which we interpret as the "memoryless" property of that distribution. These rates may be verified to be correct in one of two ways. First, by simulating an ensemble of independent renewal processes, one observes the change in the system over an interval $[t, t+dt]$. The flow through the set $E_j$ over that interval,

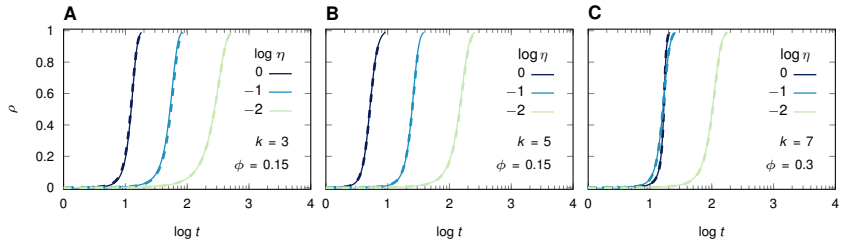


**Figure 5.5 –** Solution of approximate master equation, dashed, and corresponding Monte Carlo experiment, solid. Here we vary the parameter $\eta$, controlling the event duration. Decreased memory slows down threshold driven contagion. This feature is consistent for a number of $k$ and $\phi$ values, even in the case of $k = 7$. Note the curious overlap for $\eta = 1$ and 0.1 here.

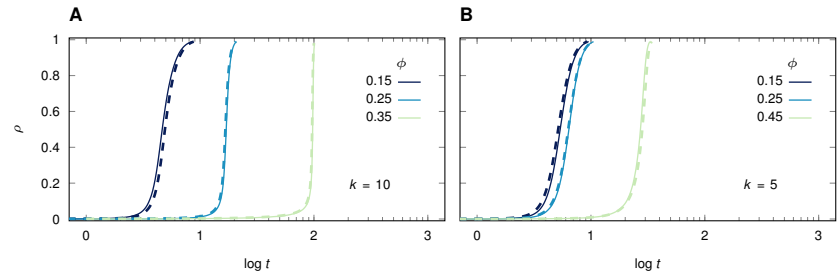


**Figure 5.6 –** Identical to Fig. 5.5, but with $\eta$ constant and increasing threshold $\phi$.

normalised by the size of that set, give the rates $\mu_j$ and $\nu_j$. Alternatively, by simulating a single renewal process for a sufficiently long time, and observing its change in behaviour over each interval $dt$, one obtains the rates $\mu_j$ and $\nu_j$ which are identical to those calculated in the ensemble. See Figure 5.3 for the case of a power law.

## 5.4 POWER-LAW INTER-EVENT TIME DISTRIBUTION

A typical model of heavy tailed distributions is the power law, with $\psi = C\tau^{-\alpha}$ for some positive constant $\alpha$. Given that our study will eventually involve a comparison of the effects of different inter-event time distributions, it is important here to precisely obtain a desired $\langle\tau\rangle$ for any $\alpha$, to ensure a fair comparison against different $\psi$ distributions. Since the power-law as stated is a one parameter function, its average depends explicitly on $\alpha$ and thus cannot be held constant in a study of

the effects of varying $\alpha$. To counter this, we introduce the minimum and maximum cutoffs, $\tau_0$ and $\tau_1$ respectively, and will eventually tune them provided the exponent $\alpha$ and the desired $\langle\tau\rangle$. As such we have

$$\psi(\tau) = \begin{cases} 0 & \tau < \tau_0 \\ C\tau^{-\alpha} & \tau_0 \leq \tau \leq \tau_1 \\ 0 & \tau > \tau_1, \end{cases} \tag{5.37}$$

where $C$ is the normalisation constant

$$C = \frac{1-\alpha}{\tau_1^{1-\alpha} - \tau_0^{1-\alpha}}. \tag{5.38}$$

This function has the related distribution $\Psi(\tau) = \int_\tau^\infty \psi(\tau')d\tau'$ given by

$$\Psi(\tau) = \begin{cases} 1 & \tau < \tau_0 \\ \dfrac{\tau_1^{1-\alpha} - \tau^{1-\alpha}}{\tau_1^{1-\alpha} - \tau_0^{1-\alpha}} & \tau_0 \leq \tau \leq \tau_1 \\ 0 & \tau > \tau_1, \end{cases} \tag{5.39}$$

which is piecewise continuous. Note that the residual distribution $\Psi$ is not normalised. The integral $\int_0^\infty \Psi d\tau$ with $\Psi$ defined as above gives $C/\langle\tau\rangle$, with $C$ defined below. To implement a temporal network whose edges undergo renewal processes following a distribution $\psi$, one is normally interested in initialising the system at steady state. To do this one performs inverse transform sampling of both $\psi$ and $\psi^r$, the residual time distribution. As such, we require the CDF of both these functions. First, we have

$$\int_{-\infty}^{\tau} \psi(\tau')d\tau' = \begin{cases} 0 & \tau < \tau_0 \\ \dfrac{\tau^{1-\alpha} - \tau_0^{1-\alpha}}{\tau_1^{1-\alpha} - \tau_0^{1-\alpha}} & \tau_0 \leq \tau \leq \tau_1 \\ 1 & \tau > \tau_1, \end{cases} \tag{5.40}$$

which can be inverted to allow inverse transform sampling of the $\psi$ distribution, giving

$$\tau = \left[\tau_0^{1-\alpha} + (\tau_1^{1-\alpha} - \tau_0^{1-\alpha})\Psi\right]^{\frac{1}{1-\alpha}}. \tag{5.41}$$

As such, $\Psi$ is set to a value uniformly sampled from the interval $(0, 1)$, and we evaluate Eqn. 5.41 to get the corresponding $\tau$ value.

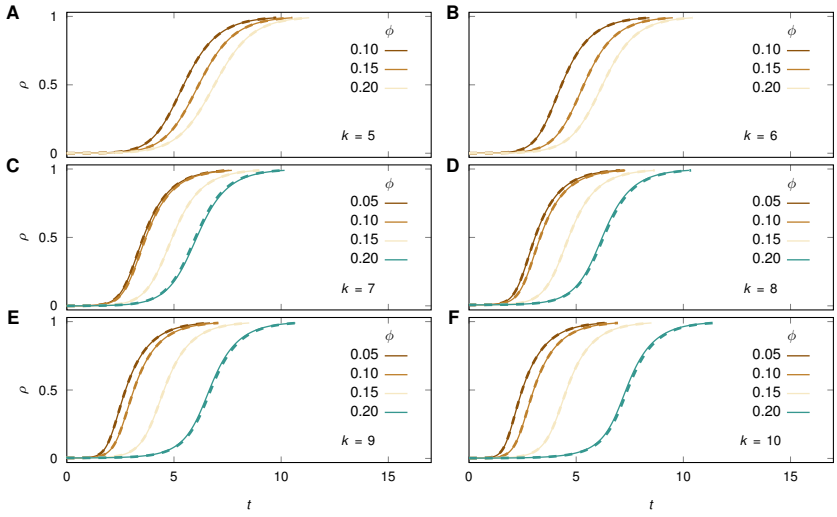


**Figure 5.7 –** Power law inter-event time distribution $\psi(\tau)$ with exponent $\alpha = 2.5$, and average $\langle\tau\rangle = 1$. Networks are random-regular with degree $k$, and of size $N = 10^6$. Output from simulation are solid curves, and result from 50 realisations. results from theory are dashed lines. There is strong agreement between theory and experiment for values of $\phi$ in the range $0 < \phi < 0.25$, but as shown below, theory starts to deviate for larger threshold values.

### 5.4.1 *Laplace transform of power law*

Without introducing the upper and lower bounds, $\tau_1$ and $\tau_0$ respectively, obeying the relation $0 < \tau_0 < \tau_1 < \infty$, the Laplace transform

$$\int_0^\infty \tau^{-\alpha} e^{-s\tau} d\tau \tag{5.42}$$

is clearly divergent for $\alpha > 1$. But writing

$$\psi(\tau) = C\tau^{-\alpha} \left[H(\tau - \tau_0) - H(\tau - \tau_1)\right], \tag{5.43}$$

where $H(\tau)$ is the Heaviside step function, ensures the Laplace transform converges if $\tau_0$ and $\tau_1$ are finite. For now, to avoid a large amount of algebra, we calculate Equations 5.5 and 5.6 numerically using Equations 5.37 and 5.39, by manually computing the convolutions of $\psi$ and $\Psi$. In the following sections we discuss the close agreement between Monte-Carlo simulation and the AMEs for small and intermediate ranges of threshold values $\phi$, and somewhat worse agreement for high thresholds.

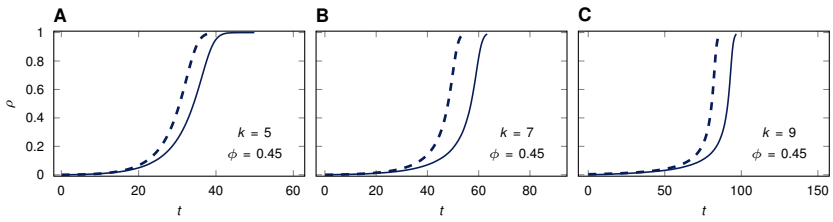


**Figure 5.8 –** Divergence of theory from experiment for high threshold values $\phi$. Power law has exponent $\alpha = 2.5$ and average $\langle \tau \rangle = 1$, while memory has duration $\eta = 1$. Networks are degree regular, with $N = 10^6$ nodes and $10^2$ realisations of the spreading process.

### 5.4.2 *Effectiveness of master equation solution*

For small and intermediate values of the threshold $\phi$, we see strong agreement between the approximate master equations and the corresponding Monte Carlo simulation. As illustrated in Figure 5.7, the AME theory recovers the behaviour of threshold driven contagion for a range of degrees $k$, and for the threshold range $0 < \phi < 0.3$. In particular, the AMEs capture the decelerative effect induced by increasing the threshold $\phi$. This is important, because it means that the rates $\mu_j$ and $\nu_j$ hold even for node configurations $(\mathbf{k}, \mathbf{m})$ with a number of infected neighbours. However, for very large thresholds $\phi$, a disparity begins to emerge between simulation and theory, as depicted in Fig. 5.8. Although solid lines (experiment) and dashed lines (theory) no longer converge, the qualitative decelerative effect is still captured. The error can be traced back to a deviation in the edge transition rates $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$, which are are really system-wide rates. In reality, edge transition rates for each *class* appear to vary slightly over time. This can be remedied, but at the cost of increased system size, and is beyond the scope of the work in this chapter.

### 5.4.3 *Phenomenology*

In this section we change speed and turn our attention to the consequences of our model. Our primary interest is in determining the effect of the interevent time distribution $\psi$ on threshold driven contagion. We wish to compare a homogeneous process, where events arrive at a constant rate, with a heterogenous process. Such a study can be performed by comparing an exponential distribution $\psi$, with a heavy-tailed distribution, such as the power law defined in Eq. 5.37. To ensure a valid

comparison of the two dynamics, we ensure that the average interevent time $\langle\tau\rangle$ is constant across experiments (and maybe even control for $\tau_0$, because this parameter controls the maximum edge-state in the power law, and therefore the maximum weight). This is straightforward in the exponential distribution 5.24, where $\langle\tau\rangle$ is the sole parameter, however in the case of the power law, the three parameters $\tau_0$, $\tau_1$ and $\alpha$ must be adjusted for a desired mean. $\alpha$ controls the heaviness of the tail, $\tau_1$ controls its length, and $\tau_0$ controls where it begins.

It is critical to control for the average inter-event time $\langle\tau\rangle$, which we shall do for a power law by varying the cutoffs $\tau_0$ and $\tau_1$ after selecting a desired exponent $\alpha$. These values are related by the expression

$$\langle\tau\rangle = \frac{1-\alpha}{2-\alpha}\frac{\tau_1^{2-\alpha}-\tau_0^{2-\alpha}}{\tau_1^{1-\alpha}-\tau_0^{1-\alpha}}. \tag{5.44}$$

When $\alpha < 2$, the mean $\langle\tau\rangle$ of the integral $\int_{\tau_0}^{\tau_1}\tau^{-\alpha}d\tau$ diverges when $\tau_1 \to \infty$. As such, if $\tau_0$ is fixed as some small value, there is guaranteed to exist an upper cutoff $\tau_1$ that produces the desired mean. In contrast, when $\alpha > 2$, the mean is well defined. This restricts the $(\tau_0, \tau_1)$ parameter space, since the desired mean must be smaller than its limiting value

$$\lim_{\tau_1\to\infty}\langle\tau\rangle = \frac{1-\alpha}{2-\alpha}\frac{\tau_0^{2-\alpha}}{\tau_0^{1-\alpha}} = \frac{1-\alpha}{2-\alpha}\tau_0, \tag{5.45}$$

for constant $\tau_0$. This provides a lower bound for $\tau_0$ in order that the desired $\langle\tau\rangle$ be achievable by tuning $\tau_1$ when $\alpha > 2$. For example, if we require $\langle\tau\rangle = 1$ and have set $\alpha = 2.5$, then $\tau_0 > 1/3$. In general we have the relation

$$\frac{\tau_0}{\langle\tau\rangle} > \frac{1-\alpha}{2-\alpha}, \tag{5.46}$$

This has a number of implications for the interpretation of the model. For example, if we set the average inter-event time to $\langle\tau\rangle = 1$, the event memory to $\eta = 1$, and select $\alpha = 2.5$ as the exponent, then we limit the range of the lower bound to $1/3 < \tau_0 < 1$, meaning that no more than 3 events can ever be recalled on a given edge. This is not a problem per se, except in the case where one wishes to model a broad weight distribution. The easiest solution here would be to increase the memory length $\eta$.

## 5.5 SAMPLING THE LOG-NORMAL AND GAMMA DISTRIBUTIONS

In this section we study log-normal inter-event time distributions. The log-normal distribution has the advantage of directly inputting

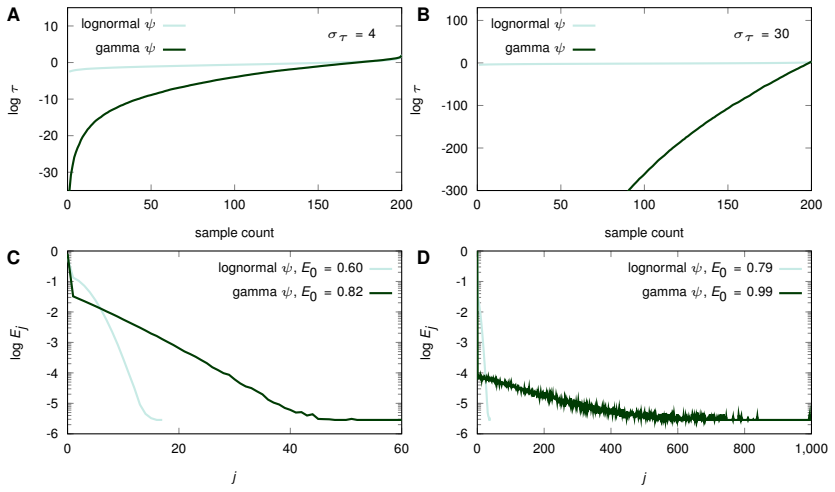


**Figure 5.9 –** Random sample of sorted $\tau$ for log-normal and gamma $\psi$ distributions, all with mean 1, and standard deviation $\sigma_\tau = 4$, (a) and (c), and 30, (b) and (d). The gamma distribution samples with much higher probability from small $\tau$ values. The distributions $E_j$ resulting from the log-normal and gamma $\psi$, with $\sigma_\tau = 4$ and 30. Because of the large number of small $\tau$ values drawn from the gamma distribution, higher $E_j$ states are reached in this distribution compared to the log-normal. Crucially, the network is effectively much sparser under the gamma distribution than the lognormal (see the mass of $j = 0$ edge states in the legends, $E_0$).

its average $\mu_\tau$ and standard deviation $\sigma_\tau$, which was not possible in the case of the power law, above. The Gamma distribution is equally convenient, albeit after some minor transformation of typically used parameterisations. For the log-normal distribution, the probability density and cumulative distribution functions are given by

$$\psi(\tau) = \frac{1}{\tau\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln\tau - \mu)^2}{\sqrt{2\pi\sigma^2}}\right] \tag{5.47}$$

and

$$1 - \Psi(\tau) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\ln\tau - \mu}{\sqrt{2\sigma^2}}\right), \tag{5.48}$$

where $\Psi$ is the residual distribution of $\psi$. It goes without saying that $\mu$ is the mean of the distribution, not the positive edge transition rate $\mu$ defined above.

In addition to a log-normal distribution of inter-event times, we conduct identical experiments using a gamma distribution. This has the

advantage of being able to interpolate between exponential and power law distributions. The gamma distribution is of the two parameter family, with probability density given by

$$\psi(\tau) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} \tau^{\alpha-1} e^{-\frac{\tau}{\beta}}, \tag{5.49}$$

and CDF given by

$$1 - \Psi(\tau) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \tfrac{\tau}{\beta}). \tag{5.50}$$

Here, $\gamma(\alpha, \frac{\tau}{\beta})$ is the lower incomplete gamma function [2]. In contrast to the log-normal, the Gamma distribution does not input the mean and standard deviation of the inter-event times, $\mu_\tau$ and $\sigma_\tau$. They are related by the expressions

$$\mu_\tau = \alpha\beta \quad \text{and} \quad \sigma_\tau^2 = \alpha\beta^2, \tag{5.51}$$

which can be straightforwardly inverted to give

$$\alpha = \frac{\mu_\tau^2}{\sigma_\tau^2} = \frac{1}{\sigma_\tau^2} \quad \text{and} \quad \beta = \frac{\sigma_\tau^2}{\mu_\tau} = \sigma_\tau^2, \tag{5.52}$$

as we set $\mu_\tau = 1$ everywhere. Under this parameterisation, setting $\sigma_\tau = 1$ and fixing as always $\mu_\tau = 1$, we get $\alpha = \beta = 1$. Then the inter-event distribution reduces to the exponential, with $\psi(\tau) = e^{-\tau}$.

We begin our study of these distributions by examining their steady state properties in the context of our renewal process model, where edge state is determined by the number of events $j$ to have been observed in the preceding time window of duration $\eta$. In the top row of Fig. 5.9 we plot a random sample of $\tau$ values drawn from both the log-normal and gamma distribution, with means $\mu_\tau = 1$ and $\sigma_\tau = 4$ and 30 (left and right plots in the top row, respectively). From this first glance at the distributions, it is clear the gamma distribution weights much more highly values of $\tau$ in the limit $\tau \to 0$. Despite having identical mean and standard deviation, almost every sample from the gamma distribution is smaller than that of the log-normal distribution, excepting of course that the gamma distribution compensates by having a proportionally heavier tail. This is true for both values of the standard

2. Whereas the ordinary gamma function is defined as $\Gamma(s) = \int_0^\infty t^{s-1}e^{-t}$, the upper and lower incomplete gamma functions are given by $\Gamma(s,x) = \int_x^\infty t^{s-1}e^{-t}$ and $\gamma(s,x) = \int_0^x t^{s-1}e^{-t}$, respectively. They obey the relations $\Gamma(s) = \gamma(s,x) + \Gamma(s,x)$, and $\Gamma(s) = \Gamma(s,0) = \lim_{x\to\infty} \gamma(s,x)$. Sampling from the lower incomplete gamma function will be crucial when initialising our temporal network system at steady state.

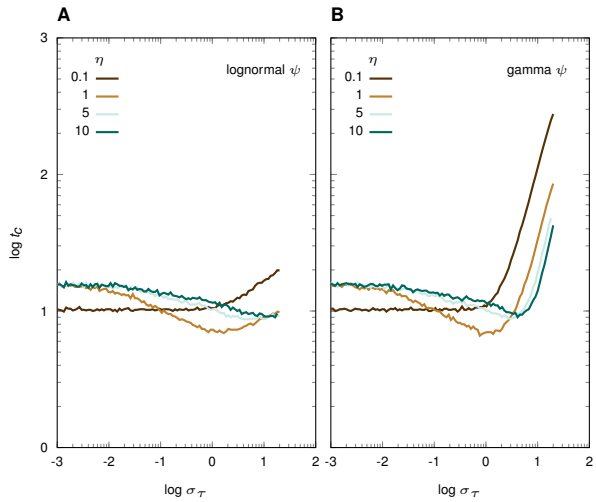


**Figure 5.10 –** Spreading time as a function of inter-event time standard deviation $\sigma_\tau$, for different values of memory $\eta$ (see legend). Left is gamma distributed $\tau$, right is log-normally distributed $\tau$. Network has log-normal degree distribution with $N = 5 \times 10^5$. See below for comparisons of gamma and log-normal results for each $\eta$. Thresholds are uniformly distributed with $\phi = 0.25$.

deviation of $\tau$, but the effect is clearly exaggerated for larger values, that is, when $\sigma_\tau = 30$.

We are interested, of course, in the impact of the inter-event time distribution $\psi$ on the resultant distribution of edge states $E_j$ (bottom row of Fig. 5.9). Clearly, the effect of the greater skewness in $\psi$ demonstrated by the gamma distribution, will be to increase the frequency of higher valued edge states. This is seen in the heavy tail of the $E$ distribution for gamma-distributed $\tau$ (green curves), and the relatively sharp cutoff in the edge state distribution for log-normally distributed $\tau$ (purple curves). Importantly, we observe that the $E$ distributions here are bimodal; there is a spike in the value of the edge state $E_0$ for each distribution, corresponding the frequency of edges that have observed $j = 0$ events in the last $\eta$ window. The heavier the tail of the $E$ distribution, the sharper the spike appears to be. This is key, as it captures the effects of burstiness; a relative abundance of edges in large states (those that observed bursts of activity, $\tau \ll 1$) is compensated by a large number of edges caught in a long waiting time between events (those that have drawn from the heavy tail of $\psi$, where $\tau \gg 1$).

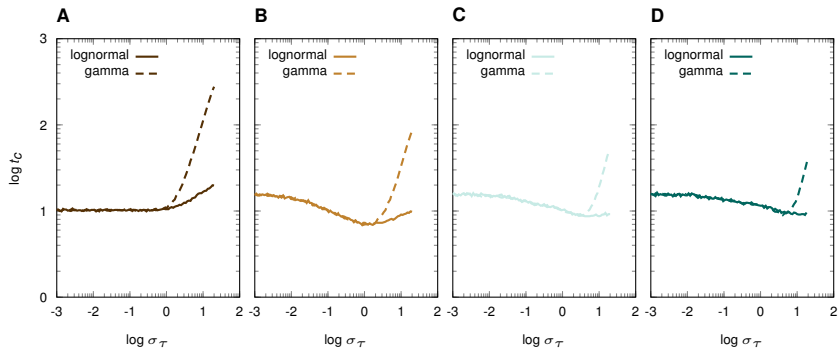


**Figure 5.11 –** Identical to Fig. 5.10, but with log-normal and gamma distributed $\psi(\tau)$ plotted side-by-side for corresponding parameter values. Solid is log-normal $\psi$, dashed is gamma.

### 5.5.1 *Spreading speed*

We are interested in the time $t_c$ required to reach some arbitrary density of active nodes in the network, which we set to $\rho_c = 0.5$. In Fig. 5.10, we compare the value of $t_c$ obtained as a function of inter-event time standard deviation $\sigma_\tau$, for a select number of memory values $\eta$, for both log-normally distributed $\tau$, in (a), and gamma distributed $\tau$, in (b). For $\eta = 1$, 5 and 10, we recover a static network when $\sigma_\tau \to 0$, a consequence of setting $\mu_\tau = 1$. This is because when events are periodic, and memory is an integer multiple of the uniform inter-event time $\tau = 1$, edges are always in state $j = 1$, 5 and 10, respectively. This leads to a relatively rapid spreading speed, $t_c \simeq 16$, a time that is accelerated by decreasing the memory duration to $\eta = 0.1$. Here, a static network is no longer recovered when taking $\sigma_\tau \to 0$. Rather, edges occupy state $j = 1$ for 0.1 units of time after an event has occurred, before waiting in state $j = 0$ for the remaining 0.9 units until the next event. Despite the network being sparsified by 90%, this decrease in memory length is able to increase the spreading speed by introducing *mixing*. By this we mean that a node's local configuration $(\mathbf{k}, \mathbf{m})$ is no longer frozen, and edges are allowed to fluctuate between $j = 0$ and 1, meaning that nodes may occupy configurations that are favourable to activation at one point in time, and not in another.

*"mixing" effect of temporality*

It is precisely the effect outlined above that accelerates the process for increasing $\sigma_\tau$ when $\eta = 1$, 5 and 10, in both Fig. 5.10(a) and (b). That is, nodes that are stable with a given $\mathbf{k}$ have the chance of passing through configurations that leave them vulnerable, once $\sigma_\tau$ increases

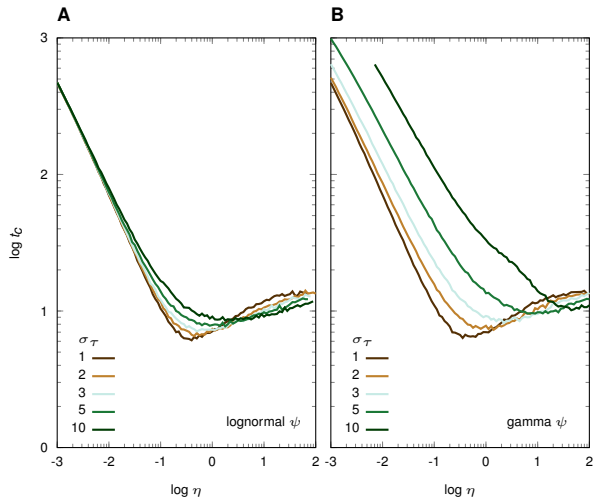


**Figure 5.12 –** Spreading time as a function of memory $\eta$, for different values of memory $\sigma_\tau$ (see legend). Left is gamma distributed $\tau$, right is log-normally distributed $\tau$. Network has log-normal degree distribution with $N = 5 \times 10^5$. See below for comparisons of gamma and log-normal results for each $\eta$.

to 0 or 1 and the spreading time reaches a minimum. Here of course, events are no longer periodic. When $\eta = 0.1$, the increased variance of the inter-event time distribution has no such accelerative effect. For all values of $\eta$ depicted in Fig. 5.10, larger values of $\sigma_\tau$ exaggerate the sparsifying effect outlined above, so much so that the accelerative effect of mixing is overwhelmed.

This deceleration occurs at a faster rate in Fig. 5.10(b), for a given value of $\sigma_\tau$, due to the effects of the gamma distribution discussed in Section 5.5. That is, for equal values of $\sigma_\tau$, the gamma distribution samples more heavily from very small $\tau$, meaning larger bursts, and correspondingly, more edges in state $j = 0$. To emphasise this effect, we plot the curves in Fig. 5.10 side-by-side in Fig. 5.11, where a gamma-distributed $\psi(\tau)$ is always equal to or slower than the corresponding process with log-normally distributed $\psi(\tau)$.

In Fig. 5.12 we expand the above experiment in order to better understand the role of memory. In so doing we perform the inverse experiment, and we fix $\sigma_\tau$ at selected values and smoothly vary the length of the memory duration $\eta$. When taking $\eta \to 0$, for any value of $\sigma_\tau$, and regardless of the inter-event time distribution $\psi$, we effectively sparsify the network. Clearly, when node memory goes to zero, an increasing number of edges are in the $j = 0$ state, and activation can

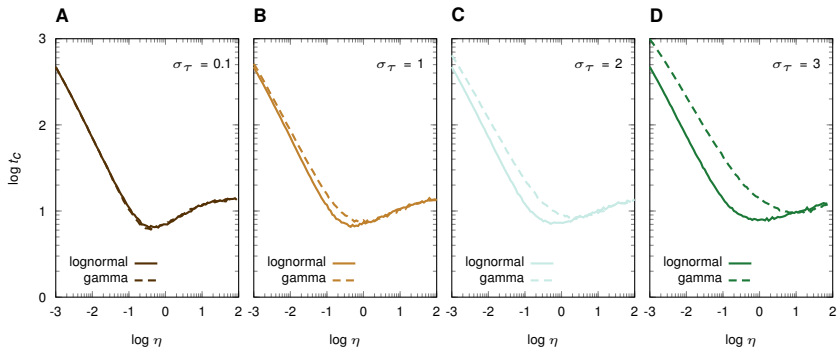


**Figure 5.13 –** Identical to Fig. 5.12, but with log-normal and gamma distributed $\psi(\tau)$ plotted side-by-side for corresponding parameter values. Solid is log-normal $\psi$, dashed is gamma.

only occur spontaneously at rate $p$. Spreading process relying purely on noise will always be slower than processes allowing activation through their interactions. It is for this reason that spreading accelerates (or equivalently, $t_c$ decreases), as $\eta$ grows. For intermediate $\eta$, processes allow mixing, which is an accelerative effect as discussed above. When $\eta$ is very large, large memory duration allows edges to increasingly "sample" from the inter-event time distribution, and the edge state distribution becomes uniform. Even if the average edge state is very large,[3], accelerative effects due to mixing start to diminish. Since we are using a relative threshold $\phi$ here, our temporal-network extension of the Watts model, we begin to recover a static network.

These changes occur at different rates as a function of $\eta$, depending on whether the distribution $\psi$ is log-normal or gamma, as we show in Fig. 5.13. As before, the small-$\tau$ sampling effect of the gamma distribution means that the network will always be as or more sparse than the corresponding log-normal distribution. This is especially evident for short memory $\eta$.

### 5.5.2 *Limiting cases of $\eta$ and $\sigma_\tau$*

In summary, we outline the four limiting scenarios that we recover regardless of whether the system follows a log-normal or gamma inter-event time distribution $\psi$.

3. The average edge state $\langle E \rangle$ increases linearly for large enough $\eta$.

*Constant $\eta$, with $\sigma_\tau \to 0$*

In this case, we have fixed memory, with interevent times tending towards a periodic signal. In the special case of $\eta = \mu_\tau$, interaction strength is constant across the network, with all edges in state $j = 1$. This is equivalent to a static network, and in the case of the Watts model where influence is normalised over a node's local neighbourhood, an unweighted network. Here, the spreading speed is rapid compared to the case of $\sigma_\tau \to \infty$. In this case, the network is maximally filled, with $E_0 = 0$ and $E_1 = 1$.

*Constant $\eta$, with $\sigma_\tau \to \infty$*

In this case, we have fixed memory, with interevent times tending towards extreme bursts. In the infinite network limit, the average edge state is still $\frac{\mu_\tau}{\eta}$, but interaction strength is concentrated among those edges undergoing bursts in activity. We refer to this as the sparsifying effect of increasing burstiness, as $\sigma_\tau \to \infty$. The effect of this is $E_0 \to 1$, with $E_0$ increasing monotonically as a function of increasing $\sigma_\tau$.

*Constant $\sigma_\tau$, with $\eta \to 0$*

Clearly the effect of this is sparsifying, or $E_0 \to 1$. Regardless of the dynamics, no influence occurs when memory $\eta$ goes to zero. Since we allow for an exogenous input noise at rate $p$ in these experiments, spreading curve given by $\rho = 1 - e^{-pt}$, for all dynamics.

*Constant $\sigma_\tau$, with $\eta \to \infty$*

Clearly the effect of this is densifying, or $E_0 \to 0$. Nodes recall a larger and larger number of events for fixed $\sigma_\tau$ and increasing $\eta$. As $\eta \to \infty$, the distribution of edge states $E_j$ narrows, and the network approaches a uniform weight distribution, whose value increases linearly with $\eta$.

In the case of the Watts model, which is normalised, this will be identical to a static, unweighted network. This is a consequence of the threshold $\phi$ being relative. For absolute thresholds, as described by the Centola-Macy model, spreading speeds will be saturated; since the threshold is fixed, increasing $\eta$ has a limited effect on node's whose thresholds have already been overcome. In contrast, if effective infection rate in the SI model is proportional to edge state, we can expect the spreading speed to increase uniformly for increasing $\eta$.

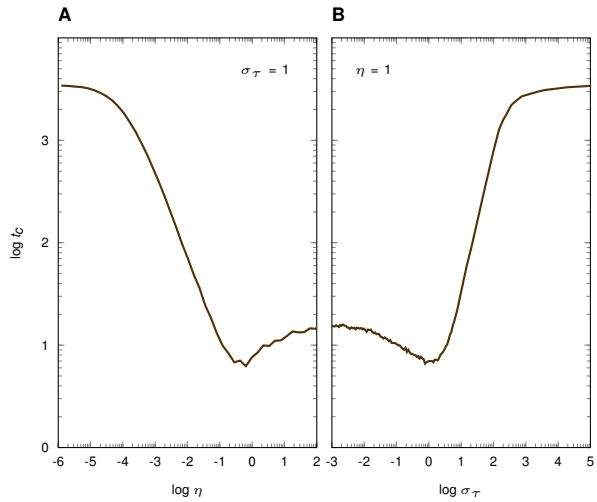


**Figure 5.14 –** Memory $\eta$ and inter-event time standard deviation $\sigma_\tau$ behaving like conjugate variables. The upper plateau represents the theoretical slowest spreading process, driven entirely by spontaneous adoption. This $t_c$ is given by the solution to $\dot{\rho} = p(1-\rho)$. The lower plateau gives the corresponding static network.

### 5.5.3 *Conjugate behaviour of memory and burstiness*

Considering the scenarios outlined above, it appears that memory $\eta$ and inter-event time standard deviation $\sigma_\tau$ have mirroring effects on the spreading process. That is, taking $\eta \to 0$ for constant $\sigma_\tau$, and $\sigma_\tau \to \infty$ for constant $\eta$, tend towards an empty network where $E_0 = 1$, where almost all edges are in the $j = 0$ state. Further, taking $\eta \to \infty$ for constant $\sigma_\tau$, and $\sigma_\tau \to 0$ for constant $\eta$, tend towards a static network,[4] where all edges are in state $E_l = 1$, for some $l > 0$. This is summarised in Fig. 5.14.

## 5.6 DISCUSSION

We have seen that in a model of stochastic temporal networks, burstiness has a substantial effect on the dynamics of threshold-driven contagion. There emerges a non-monotonic dependence of spreading speed on both $\sigma_\tau$, the inter-event time standard deviation, and event memory

4. We require that $\eta$ is an integer multiple of $\mu_\tau$ in order to exactly recover a static network using this mechanism.

$\eta$. This is due to two competing effects. First, is the accelerative effect of mixing, where edge activity allows the local configuration of a given node to evolve over time. This leads to nodes that are otherwise stable to active neighbours, becoming temporarily vulnerable. At the same time, as inter-event times become more heterogeneous, the network becomes sparsified, which is clearly a decelerative effect. These two features of our model combine to produce a rich array of dynamics, that may help to inform our understanding of information spread in the presence of heterogeneous dyadic activity.

This work contributes a general and effective theoretical framework with which to study binary-state dynamics on temporal networks.[5] In addition, we have shown that in a simple model of the spread of contagion on temporal networks, dynamical processes are highly sensitive to temporal heterogeneities such as burstiness.

5. Although the present version of this work is unpublished, it is expected to be finished and submitted by the time of my PhD defence.

# Part III

# OUTLOOK

In Chapter 6 we discuss the limitations of this work, and suggest a number of promising future research directions.

# 6

# DISCUSSION AND FUTURE WORK

In this thesis we have examined threshold-driven contagion on complex networks. We have extended a master-equation formalism to account for edge-based properties that are of importance in real world systems, such as edge weights, edge multiplexity, and network temporality, which we have successfully interpreted as a type of time-dependent edge property. We have applied this formalism alongside an extensive experimental study of these systems, by means of Monte Carlo simulations on synthetic and real networks.

## 6.1 LIMITATIONS OF THE PRESENT WORK

Generically, the same limitations that apply to the complex systems methodology as a whole also apply to this work. The abstraction of a complex system as a network comes with a number of assumptions, in particular, that models lose information exogenous to the network itself.

More particular to this work is the assumption of binary-state dynamics. With regards to the spread of information, it is doubtless that the interaction among various content plays an important role. While the binary-assumption is exact in many respects, for example that exposure in the form of receiving and sharing information is plainly a binary fact, the competition between the different *types* of content is likely to be critical. Further, it must be pointed out that gathering empirical evidence for threshold models is notoriously difficult. Even determining whether genuine social influence is behind instances of spreading over a network, as opposed to homophily, is extremely challenging,

On the analytical side, a number of limitations are evident. Chief among them is the sheer computational cost of solving high-dimensional master equations. We saw that system sizes grow like $\mathcal{O}(k^{2n})$, which becomes prohibitively expensive when either $k$ or $n$ is large. Since degrees and edge heterogeneities are frequently large in real world

scenarios, the master equation approach is limited to academic use. In other words, although applying a master-equation solution can deepen one's understanding of spreading dynamics, it is unlikely to be of practical use when attempting to draw conclusions from industrial data, be it telecommunicative, financial or economic. At least in the way that we have employed it, the master equations are most effective when anatomising a theoretical system - by explicitly considering what constitutes a local configuration, and what is driving its conditions, one begins to build up a low-level intuition for the dynamics.

Finally, although the infinite network assumption massively simplifies analytic calculations, reducing computational complexity at the same time, it places finite dynamics out of reach. In other words, while the master-equation solution, including all of its variations, are able to provide global cascade conditions, they say nothing about finite-sized cascades. The spread of information through finite populations is an important avenue to explore, and our configuration-based master equations are inapplicable here.

Note also that in a master equation framework, we lose all notion of cascade depth, path length, and other such topological features.

## 6.2 FUTURE WORK

A great strength of the master-equation approach used throughout this work is its modularity. In other words, the simplicity of edge-heterogenous assumption invites the study of a broad range of systems that we have left untouched. For instance, it is well known that directedness can be critical to dynamical processes which we have ignored entirely. This is particularly the case in Chapter 5, where the direction of temporal contacts is often important. Further, since our framework applies to arbitrary binary-state dynamics, we are encouraged to look beyond the threshold models that have been the focus of this thesis.

Our interpretation of temporality was extremely confined, and renewal process are just one of a large variety of stochastic processes that may define activity on edges. Extensions to this are already underway in view of deepening the results of Chapter 5.

# Part IV

# APPENDIX

In Appendices A and B we provide illustrations of a number of configuration space structures in order to build intuition for the master equation solution. Finally, in Appendix C we describe some of the data sets used in this work.

# A

## ILLUSTRATIONS OF CONFIGURATION SPACE

In this appendix we provide concrete illustrations of so-called configuration space, introduced in Chapter 2. Whereas the networks that we simulate are produced by random graph models, configuration graphs provide the highly regular relations between node classes $(\mathbf{k}, \mathbf{m})$, defined in Chapter 2. In Fig. A.1, we show the set of classes satisfying the constraints $k = 0, 1, 2$ and 3, with $n = 2$ allowed edge states. The degree vectors $\mathbf{k}$ are indicated on the bottom-left of each set, and total number of active neighbours $m$ on the bottom-right. The active degree vector $\mathbf{m}^T$ is the label of the points in the space. Right diagonal transitions, in black, are neighbour transitions and are shown to be unidirectional here, indicating a non-recovery dynamics. These transitions preserve $\mathbf{k}$, and incrememnt $m$ by one. Left diagonal transitions in red indicate edge-state transitions, which are shown to be bidirectional, indicating a temporal network model where edge states can alternate repeatedly between two states, $j = 0$ and 1.

The number of classes $|C_k|$ in each component of total degree $k$ are $|C_0| = 1$, $|C_1| = 4$, $|C_2| = 10$ and $|C_3| = 20$, for plots Fig. A.1(a), (b), (c) and (d), respectively. An illustration of $k = 4$, where $|C_4| = 35$, is shown in Appendix B. Note that although the degree vector $\mathbf{k}$ of a node may change due to edge dynamics, the total degree $k = \sum_j k_j$ is fixed over the course of the dynamical processes. As such, nodes cannot jump between the various connected components of Fig. A.1. The $k = 0$ class in (a) may appear trivial, but the density of nodes in this class is critical in determining the existance of a percolating vulnerable cluster under the Watts model, or equivalently, the possibility of global cascades.

The solution of the master equation used throughout this work amounts to tracking the density of nodes belonging to each class $(\mathbf{k}, \mathbf{m})$. Clearly the matrices $A_{neigh}$ and $A_{edge}$ are block-diagonal, since nodes can only transition between $k$-components.

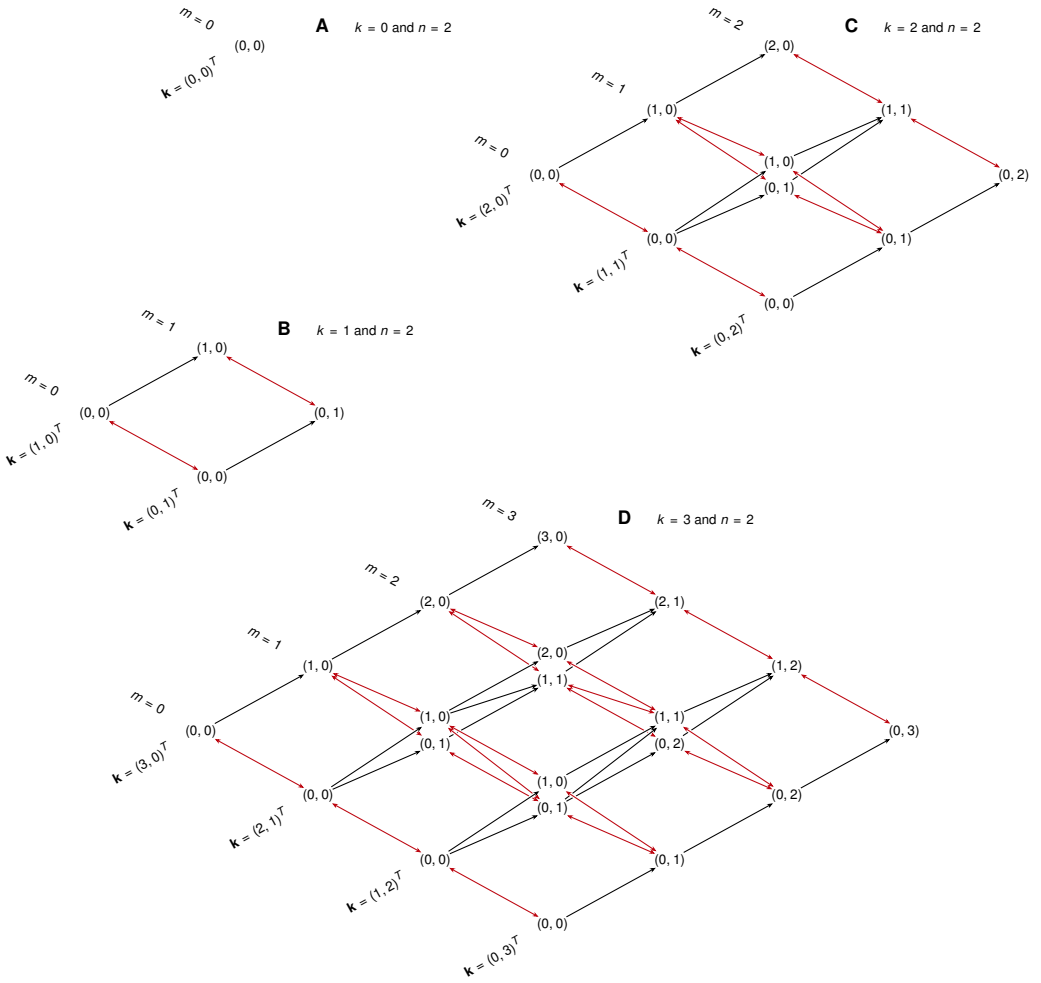


**Figure A.1 –** Temporal network transitions for nodes of total degree $k = 0$, 1, 2 and 3, as labelled, with $n = 2$ allowed edge states. Right-diagonal transitions indicate neighbour adoption, and left-diagonals the increments and decrements between edge-states $j = 0$ and 1 that may occur in models of temporal network. Labelled points in the graph give the active degree vector $\mathbf{m}^T$. Right-diagonal transitions preserve the degree vector $\mathbf{k}$, indicated at the bottom of these diagonals. Left-diagonal transitions preserve total active neighbour count $m$, shown at the top of these diagonals.

# B

## STATIC AND TEMPORAL CONFIGURATION SPACE

In this appendix, we briefly discuss the extension of the static, edge-heterogeneous master equation solution to temporal networks. As outlined in Chapter 2, a node's local *configuration* is described by the tuple $(\mathbf{k}, \mathbf{m})$. The dimension of the degree vectors $\mathbf{k}$ and $\mathbf{m}$ is $n$, giving the number of allowed edge types in the system.

In the case of a static network undergoing non-recovery dynamics, such as the SI process or the Watts threshold model, a node's degree vector $\mathbf{k}$ is fixed, and its dynamics over the course of a spreading process is given by its path from the $m = 0$ position in the inactive configuration space, up to the point where the node becomes active. Such a space is indicated in Fig. B.1(a), for a node with total degree $k = 4$ and $n = 2$ possible edge states, containing 35 total configurations $(\mathbf{k}, \mathbf{m})$. Crucially, connected components are isolated according to the 5 degree vectors $\mathbf{k}^T = (4,0), (3,1), \ldots, (0,4)$. That is, once $\mathbf{k}$ is determined for a node (via a bimodal weight distribution, or an $M = 2$-layer multiplex, for example), it is fixed throughout the entire dynamical process. Its path through the configuration space in Fig. B.1(a) is given by its neighbour configurations, $\mathbf{m}$, which are labelled for each point.

In contrast, a temporal-network configuration space makes the crucial step of allowing a node's edge-configuration, given by $\mathbf{k}$, to vary. A system allowing $n = 2$ edge states might be defined by renewal processes, such as that described in Chapter 5. A node with a total of $k = 4$ neighbours in such a system is given by Fig. B.1(b). This space is comprised of an identical set of 35 configurations $(\mathbf{k}, \mathbf{m})$, as the static network in Fig. B.1(a), but allowing additional transitions. Note that temporal transitions are "symmetric" to neighbour transitions. The number of non-zero entries in $A_{neigh}$ and $A_{edge}$ are equal, depending on directionality. There is a $(4n + 1)$-th allowed transition, the $A_{ego}$ transition, from a configuration $(\mathbf{k}, \mathbf{m})$ to the corresponding transition in the opposing space.

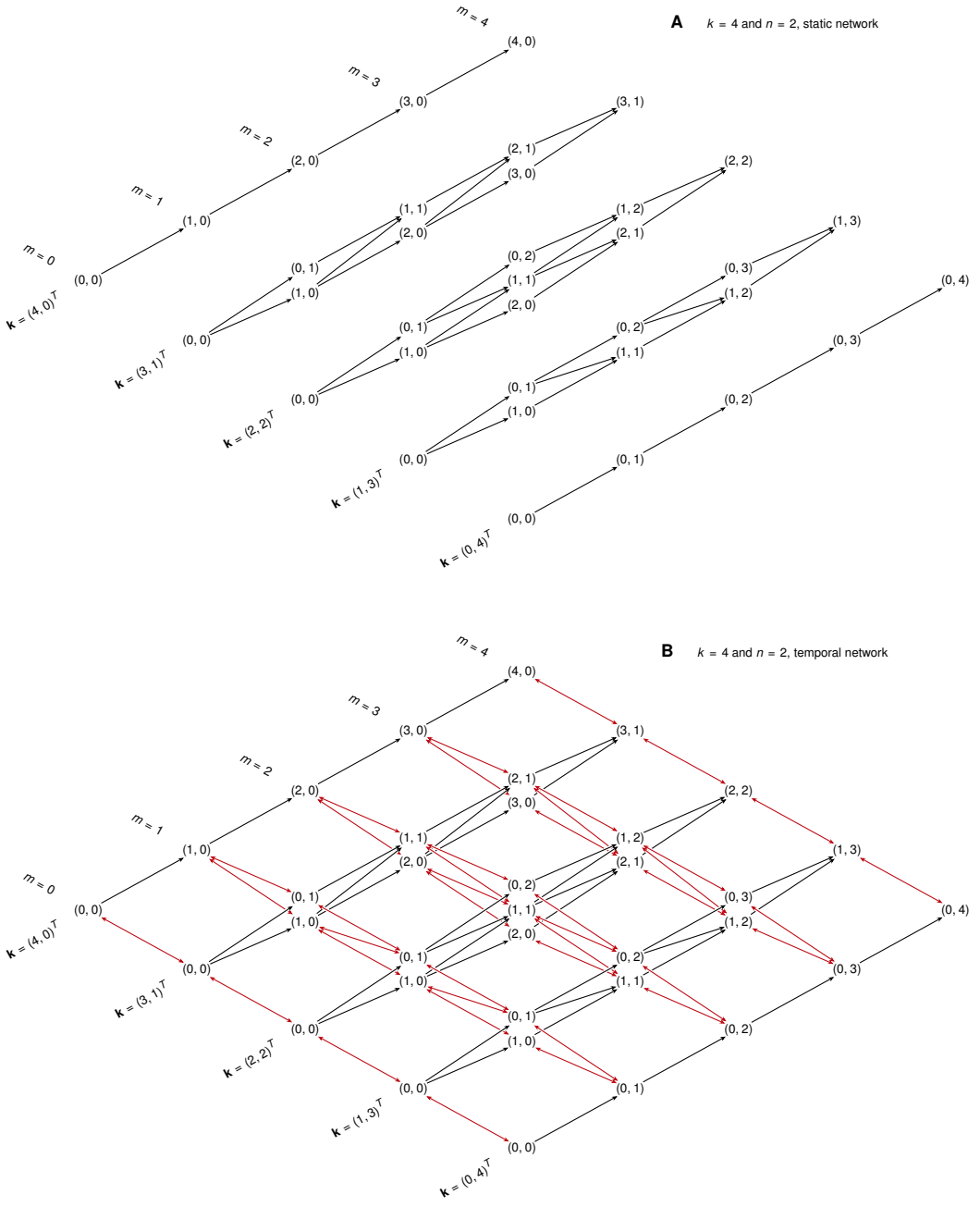


**Figure B.1 –** Allowed transitions in static networks, (a), and temporal networks, (b), for a $k = 4$ degree node with an $n = 2$ level edge-state set. See caption in Appendix A for interpretation.

# C

# DATA DESCRIPTION

In this appendix, we give an overview of the various empirical datasets used throughout this thesis.

## C.1 WIKIPEDIA

In Chapter 3 we perform data-driven simulations of our threshold model in a network of $N = 138,592$ English Wikipedia editors contributing to articles about politics. Each of the $740,397$ directed links (defining an edit, revert, restore, or vote action in an article) has a sign ($\pm$), interpreted as the parity of trust between connected editors (for free access online and details see [200]). In our study we remove self-loops and assume bidirectional links appear as undirected links with their original sign (if they shared the same sign), while choosing a sign randomly in the case where they appear with different signs (such edges only form 0.96% of the network, so their effect is not significant). Unidirectional links are also regarded as undirected with their original sign. Finally, we associate $+$ and $-$ tie signs to strong ($w_1$) and weak ($w_2$) links. The network has a broad degree distribution, a fraction $\delta = 0.88$ of strong links and average weight $\mu_w = 2.7$.

## C.2 MOBILE PHONE CALL

In Chapter 3 we perform data-driven simulations of our threshold model in a network of $N = 6,243,322$ individuals connected by $16,783,865$ undirected links with weights defined as the number of phone calls between people in an observation period of 6 months (a link exists if people have mutually called each other at least once). All individuals are customers of a single phone provider with 20% market share in an undisclosed European country. Degree and weight distributions are broad and can be approximated by power-law and log-normal distributions, respectively (for details see [153]). Since for the

MPC network $p_w$ is fixed, we introduce a method to scale $\sigma_w$ without changing the shape of the distribution, described as follows. We first assume that the MPC network has a weight set $W = \{w_1, \ldots, w_{|E|}\}$, where $w_i$ is the weight of the $i$-th edge, and $|E|$ is the number of edges in the network. This set has mean and variance

$$\mu_w = \frac{1}{|E|}\sum_{i=1}^{|E|} w_i \quad \text{and} \quad \sigma_w^2 = \frac{1}{|E|}\sum_{i=1}^{|E|}(w_i - \mu_w)^2. \tag{C.1}$$

Now we consider a new weight set $W' = \{\mu_w + \alpha(w_1 - \mu_w), \ldots, \mu_w + \alpha(w_{|E|} - \mu_w)\}$, where we have applied the transformation $w_i' = \mu_w + \alpha(w_i - \mu_w)$, $i = 1, \ldots, |E|$, and $0 \leq \alpha \leq 1$ is a tuning parameter. The limits of this transformation give a Dirac delta distribution ($\alpha = 0$) or $p_w$ ($\alpha = 1$). Substituting $w_i'$ into the expression for $\sigma_w$, we see that the mean and standard deviation of the transformed weight set are $\mu_w' = \mu_w$ and $\sigma_w' = \alpha\sigma_w$. Then, we may obtain a new weight distribution retaining the shape of $p_w$ by applying the transformation $w_i \mapsto w_i'$. If $\sigma_w'$ is the desired standard deviation, the required tuning parameter is $\alpha = \sigma_w'/\sigma_w$.

## C.3 PARDUS

The empirical structure is a signed social network, the alliance / enemy network of the Pardus massive multiplayer online game [183]. Pardus is set in a futuristic universe where players interact and compete in space. It is an open-ended game with a player-driven economy. Players travel through hundreds of solar systems while trading, building or battling with non-player characters and other players. The network consists of $N = 4650$ nodes connected by $66,580$ links, of which a fraction $\delta_z = 0.64$ have a positive alliance sign (and are considered as strong ties by us), while the rest of the links have a negative enemy sign and are interpreted as weak.

# BIBLIOGRAPHY


1. Abrams, D. M. & Strogatz, S. H. Modelling the dynamics of language death. *Nature* **424** (2003).
2. Albert, R. Scale-free networks in cell biology. *Journal of cell science* **118,** 4947–4957 (2005).
3. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics* **74,** 47–97 (2002).
4. Albert, R., Jeong, H. & Barabási, A.-L. Internet: Diameter of the world-wide web. *nature* **401,** 130 (1999).
5. Alon, U. Biological networks: the tinkerer as an engineer. *Science* **301,** 1866–1867 (2003).
6. Amaral, L. A. N., Scala, A., Barthelemy, M. & Stanley, H. E. Classes of small-world networks. *Proceedings of the national academy of sciences* **97,** 11149–11152 (2000).
7. Amini, H., Cont, R. & Minca, A. Resilience to Contagion in Financial Networks. *Mathematical Finance* **26,** 329–365 (2016).
8. Anderson, R. M. & May, R. M. *Infectious Diseases in Humans* (Oxford University Press, 1992).
9. Arthur, W. B. Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *Economic Journal* **99,** 116–31 (1989).
10. Arthur, W. B. & Lane, D. A. Information contagion. *Structural Change and Economic Dynamics* **4,** 81 –104. ISSN: 0954-349X (1993).
11. Assenov, Y., Ramírez, F., Schelhorn, S.-E., Lengauer, T. & Albrecht, M. Computing topological parameters of biological networks. *Bioinformatics* **24,** 282–284 (2007).
12. Bak, P. & Sneppen, K. Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* **71,** 4083–4086. https://link.aps.org/doi/10.1103/PhysRevLett.71.4083 (24 1993).
13. Bakshy, E., Hofman, J. M., Mason, W. A. & Watts, D. J. *Everyone's an influencer: quantifying influence on twitter* in *Proceedings of the fourth ACM international conference on Web search and data mining* (2011), 65–74.

14. Barabasi, A.-L. The origin of bursts and heavy tails in human dynamics. *Nature* **435,** 207 (2005).

15. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286,** 509–512 (1999).

16. Barabási, A.-L. & Oltvai, Z. N. Network biology: understanding the cell's functional organization. *Nature reviews genetics* **5,** 101 (2004).

17. Barrat, A. & Weigt, M. On the properties of small-world network models. *The European Physical Journal B-Condensed Matter and Complex Systems* **13,** 547–560 (2000).

18. Barrat, A., Barthelemy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America* **101,** 3747–3752 (2004).

19. Barrat, A., Barthélemy, M. & Vespignani, A. Weighted Evolving Networks: Coupling Topology and Weight Dynamics. *Phys. Rev. Lett.* **92,** 228701. https://link.aps.org/doi/10.1103/PhysRevLett.92.228701 (22 2004).

20. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical processes on complex networks* (Cambridge university press, 2008).

21. Barthélemy, M. Spatial networks. *Physics Reports,* 2011 (2010).

22. Barzel, B. & Barabási, A.-L. Universality in network dynamics. *Nature physics* **9,** 673 (2013).

23. Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M. & Grafton, S. T. Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences* **108,** 7641–7646 (2011).

24. Bassett, D. S. & Bullmore, E. Small-world brain networks. *The neuroscientist* **12,** 512–523 (2006).

25. Battiston, S., Puliga, M., Kaushik, R., Tasca, P. & Caldarelli, G. Debtrank: Too central to fail? financial networks, the fed and systemic risk. *Scientific reports* **2,** 541 (2012).

26. Battiston, S., Delli Gatti, D., Gallegati, M., Greenwald, B. C. & Stiglitz, J. E. Liaisons Dangereuses: Increasing Connectivity, Risk Sharing, and Systemic Risk. *Journal of Economic Dynamics and Control* **36,** 1121–1141 (2012).

27. Bhagat, S., Burke, M., Diuk, C., Onur Filiz, I. & Edunov, S. *Three and a half degrees of separation* https://research.fb.com/blog/2016/02/three-and-a-half-degrees-of-separation/.

28. Bikhchandani, S., Hirshleifer, D. & Welch, I. A Theory of Fads, Fashion, Custom, and Cultural Change in Informational Cascades. *Journal of Political Economy* **100,** 992–1026. https://EconPapers.repec.org/RePEc:ucp:jpolec:v:100:y:1992:i:5:p:992-1026 (1992).

29. Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardenes, J., Romance, M., Sendina-Nadal, I., Wang, Z. & Zanin, M. The structure and dynamics of multilayer networks. *Physics Reports* **544,** 1–122 (2014).

30. Boffetta, G., Carbone, V., Giuliani, P., Veltri, P. & Vulpiani, A. Power laws in solar flares: self-organized criticality or turbulence? *Physical review letters* **83,** 4662 (1999).

31. Bollobás, B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics* **1,** 311–316 (1980).

32. Bollobás, B. Random graphs (1985).

33. Borgatti, S. P., Mehra, A., Brass, D. J. & Labianca, G. Network analysis in the social sciences. *Science* **323,** 892–895 (2009).

34. Borge-Holthoefer, J., Baños, R. A., González-Bailón, S. & Moreno, Y. Cascading behaviour in complex socio-technical networks. *Journal of Complex Networks* **1,** 3–24 (2013).

35. Bruce, K. Norms and the Manipulation of Relationships in a Work Context. *Social Networks in Urban Situations* (1969).

36. Bruch, E. E. & Newman, M. Aspirational pursuit of mates in online dating markets. *Science Advances* **4,** eaap9815 (2018).

37. Brummitt, C. D., Lee, K.-M. & Goh, K.-I. Multiplexity-facilitated cascades in networks. *Physical Review E* **85,** 045102 (2012).

38. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464,** 1025 (2010).

39. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience* **10,** 186 (2009).

40. Burkholz, R., Garas, A. & Schweitzer, F. How damage diversification can reduce systemic risk. *Physical Review E* **93,** 042313 (2016).

41. Capocci, A., Servedio, V. D., Colaiori, F., Buriol, L. S., Donato, D., Leonardi, S. & Caldarelli, G. Preferential attachment in the growth of social networks: The internet encyclopedia Wikipedia. *Physical review E* **74,** 036116 (2006).

42. Carlson, J. M. & Doyle, J. Highly optimized tolerance: A mechanism for power laws in designed systems. *Phys. Rev. E* **60,** 1412–1427. https://link.aps.org/doi/10.1103/PhysRevE.60.1412 (2 1999).

43. Castellano, C., Fortunato, S. & Loreto, V. Statistical physics of social dynamics. en. *Reviews of Modern Physics* **81,** 591–646. ISSN: 0034-6861, 1539-0756. https://link.aps.org/doi/10.1103/RevModPhys.81.591 (2019) (May 2009).

44. Cellai, D., López, E., Zhou, J., Gleeson, J. P. & Bianconi, G. Percolation in multiplex networks with overlap. *Physical Review E* **88,** 052811 (2013).

45. Centola, D. The spread of behavior in an online social network experiment. *science* **329,** 1194–1197 (2010).

46. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *American journal of Sociology* **113,** 702–734 (2007).

47. Charbonneau, P., McIntosh, S. W., Liu, H.-L. & Bogdan, T. J. Avalanche models for solar flares (Invited Review). *Solar Physics* **203,** 321–353 (2001).

48. Christakis, N. A. & Fowler, J. H. The spread of obesity in a large social network over 32 years. *New England journal of medicine* **357,** 370–379 (2007).

49. Christakis, N. A. & Fowler, J. H. The collective dynamics of smoking in a large social network. *New England journal of medicine* **358,** 2249–2258 (2008).

50. Chung, F. & Lu, L. The diameter of sparse random graphs. *Advances in Applied Mathematics* **26,** 257–279 (2001).

51. Cialdini, R. B. & Goldstein, N. J. Social influence: Compliance and conformity. *Annu. Rev. Psychol.* **55,** 591–621 (2004).

52. Clifford, P. & Sudbury, A. A Model for Spatial Conflict. *Biometrika* **60,** 581–588. ISSN: 00063444. http://www.jstor.org/stable/2335008 (1973).

53. Cohen, R. & Havlin, S. Scale-free networks are ultrasmall. *Physical review letters* **90,** 058701 (2003).

54. Cohen, R., Erez, K., Ben-Avraham, D. & Havlin, S. Resilience of the internet to random breakdowns. *Physical review letters* **85,** 4626 (2000).

55. Cohen, S. Social relationships and health. *American psychologist* **59,** 676 (2004).

56. Colizza, V., Barrat, A., Barthelemy, M. & Vespignani, A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci USA* **103,** 2015–2020 (Jan. 2005).

57. Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J.-P., Sanchez, A., *et al.* Manifesto of computational social science. *The European Physical Journal Special Topics* **214,** 325–346 (2012).

58. *Introduction to algorithms* 2nd ed. en (eds Cormen, T. H. & Cormen, T. H.) ISBN: 978-0-262-03293-3 (MIT Press, Cambridge, Mass, 2001).

59. Corral, Á., Pérez, C. J., Díaz-Guilera, A. & Arenas, A. Self-organized criticality and synchronization in a lattice model of integrate-and-fire oscillators. *Physical review letters* **74,** 118 (1995).

60. Cox, S., Horadam, K. & Rao, A. *The spread of ideas in a weighted threshold network* in *Complex Networks & Their Applications V: Proceedings of the 5th International Workshop on Complex Networks and their Applications (COMPLEX NETWORKS 2016)* **693** (2016), 437.

61. Croft, D. P., James, R. & Krause, J. *Exploring animal social networks* (Princeton University Press, 2008).

62. David, E. & Jon, K. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World* ISBN: 0521195330, 9780521195331 (Cambridge University Press, New York, NY, USA, 2010).

63. De Oliveira, M. J. Isotropic majority-vote model on a square lattice. *Journal of Statistical Physics* **66,** 273–281 (1992).

64. Diekmann, O. & Heesterbeek, J. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation* (John Wiley & Sons, New York, 2000).

65. Diekmann, O., Heesterbeek, J. & Britton, T. *Mathematical Tools for Understanding Infectious Disease Dynamics* (Princeton University Press, NJ, 2012).

66. Dow, P. A., Adamic, L. A. & Friggeri, A. The Anatomy of Large Facebook Cascades. *ICWSM* **1,** 12 (2013).

67. Doyle, J. C., Alderson, D. L., Li, L., Low, S., Roughan, M., Shalunov, S., Tanaka, R. & Willinger, W. The "robust yet fragile" nature of the Internet. *Proceedings of the National Academy of Sciences* **102,** 14497–14502 (2005).

68. Dunbar, R. I. Neocortex size as a constraint on group size in primates. *Journal of human evolution* **22,** 469–493 (1992).

69. Dunbar, R. I. Social cognition on the Internet: testing constraints on social network size. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **367,** 2192–2201 (2012).

70. Durrett, R. *Random graph dynamics* (Cambridge University Press, Cambridge, 2006).

71. Dutta, B. & Jackson, M. O. *Networks and groups: Models of strategic formation* (Springer Science & Business Media, 2013).

72. Eguiluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M. & Apkarian, A. V. Scale-free brain functional networks. *Physical review letters* **94,** 018102 (2005).

73. Elliott, M., Golub, B. & Jackson, M. O. Financial Networks and Contagion. *American Economic Review* **104,** 3115 (2014).

74. Erdős, P. & Rényi, A. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* **5,** 17–60 (1960).

75. Faloutsos, M., Faloutsos, P. & Faloutsos, C. *On power-law relationships of the internet topology* in *ACM SIGCOMM computer communication review* **29** (1999), 251–262.

76. Fewell, J. H. Social insect networks. *Science* **301,** 1867–1870 (2003).

77. Gallotti, R., Porter, M. A. & Barthelemy, M. Lost in transportation: Information measures and cognitive limits in multilayer navigation. *Science advances* **2,** e1500445 (2016).

78. Gerrow, K. & Triller, A. Synaptic stability and plasticity in a floating world. *Curr. Opin. Neurobiol.* **20,** 631–639 (2010).

79. Gerstner, W., Kistler, W. M., Naud, R. & Paninski, L. *Neuronal dynamics: From single neurons to networks and models of cognition* (Cambridge University Press, 2014).

80. Gillespie, D. T. Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry* **81,** 2340–2361 (1977).

81. Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proceedings of the national academy of sciences* **99,** 7821–7826 (2002).

82. Gladwell, M. *The Tipping Point: How Little Things Can Make a Big Difference* (Little, Brown and Company, 2000).

83. Glaeser, E. L., Sacerdote, B. & Scheinkman, J. A. Crime and Social Interactions*. *The Quarterly Journal of Economics* **111,** 507–548. ISSN: 0033-5533. https://doi.org/10.2307/2946686 (May 1996).

84. Glasserman, P. & Young, H. P. How likely is contagion in financial networks? *Journal of Banking & Finance* **50,** 383–399 (2015).

85. Gleeson, J. P. Cascades on correlated and modular random networks. *Physical Review E* **77,** 046117 (2008).

86. Gleeson, J. P. High-accuracy approximation of binary-state dynamics on networks. *Physical Review Letters* **107,** 068701 (2011).

87. Gleeson, J. P. Binary-state dynamics on complex networks: Pair approximation and beyond. *Physical Review X* **3,** 021004 (2013).

88. Gleeson, J. P. & Cahalane, D. J. Seed size strongly affects cascades on random networks. *Physical Review E* **75,** 056103 (2007).

89. Gleeson, J. P. & Durrett, R. Temporal profiles of avalanches on networks. *Nature Communications* **8,** 1227 (2017).

90. Gleeson, J. P. & Porter, M. A. in *Complex Spreading Phenomena in Social Systems* 81–95 (Springer, 2018).

91. Gleeson, J. P., Ward, J. A., O'Sullivan, K. P. & Lee, W. T. Competition-induced criticality in a model of meme popularity. *arXiv preprint arXiv:1305.4328* (2013).

92. Gleeson, J. P., Cellai, D., Onnela, J.-P., Porter, M. A. & Reed-Tsochas, F. A simple generative model of collective online behavior. *Proceedings of the National Academy of Sciences* **111,** 10411–10415 (2014).

93. Gleeson, J. P., O'Sullivan, K. P., Baños, R. A. & Moreno, Y. Effects of network structure, competition and memory time on social spreading phenomena. *Physical Review X* **6,** 021019 (2016).

94. Goh, K.-I., Kahng, B. & Kim, D. Fluctuation-driven dynamics of the Internet topology. *Physical review letters* **88,** 108701 (2002).

95. Gonçalves, B., Perra, N. & Vespignani, A. Modeling users' activity on twitter networks: Validation of Dunbar's number. *PloS ONE* **6,** e22656 (2011).

96. Granovetter, M. S. The Strength of Weak Ties. *American Journal of Sociology* **78,** 1360–1380. ISSN: 00029602 (1973).

97. Granovetter, M. Threshold models of collective behavior. *American journal of sociology* **83,** 1420–1443 (1978).

98. Handcock, M. S., Raftery, A. E. & Tantrum, J. M. Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **170,** 301–354 (2007).

99. Harris, T. E. The theory of branching process (1964).

100. Herz, A. V. M. & Hopfield, J. J. Earthquake Cycles and Neural Reverberations: Collective Oscillations in Systems with Pulse-Coupled Threshold Elements. *Phys. Rev. Lett.* **75,** 1222–1225. https://link.aps.org/doi/10.1103/PhysRevLett.75.1222 (6 1995).

101. Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317,** 482–487 (2007).

102. Holley, R. A. & Liggett, T. M. Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model. *The Annals of Probability* **3,** 643–663. ISSN: 00911798 (1975).

103. Holme, P. & Saramäki, J. Temporal Networks. *Phys. Rep.* **519,** 97–125 (2012).

104. Horvath, S. *Weighted network analysis: applications in genomics and systems biology* (Springer Science & Business Media, 2011).

105. Huberman, B. A. & Glance, N. S. Evolutionary games and computer simulations. *Proceedings of the National Academy of Sciences of the United States of America* **90,** 7716–7718. ISSN: 0027-8424. (2019) (Aug. 1993).

106. Hurd, T. R. & Gleeson, J. P. On Watts' cascade model with random link weights. *Journal of Complex Networks* **1,** 25–43 (2013).

107. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nature communications* **10,** 2485 (2019).

108. Iyer, R., Menon, V., Buice, M., Koch, C. & Mihalas, S. The influence of synaptic weight distribution on neuronal population dynamics. *PLoS Comput Biol* **9,** e1003248 (2013).

109. Jackson, M. O. *Social and economic networks* (Princeton university press, 2010).

110. Jackson, M. O. & Watts, A. The evolution of social and economic networks. *Journal of Economic Theory* **106,** 265–295 (2002).

111. Jackson, M. O. & Wolinsky, A. A strategic model of social and economic networks. *Journal of economic theory* **71,** 44–74 (1996).

112. Jensen, H. J. *Self-organized criticality: emergent complex behavior in physical and biological systems* (Cambridge university press, 1998).

113. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. & Barabási, A.-L. The large-scale organization of metabolic networks. *Nature* **407,** 651 (2000).

114. Jo, H.-H., Murase, Y., Török, J., Kertész, J. & Kaski, K. Stylized facts in social networks: Community-based static modeling. *Physica A: Statistical Mechanics and its Applications* **500,** 23–39 (2018).

115. Joh, R. I., Wang, H., Weiss, H. & Weitz, J. S. Dynamics of indirectly transmitted infectious diseases with immunological threshold. *Bulletin of mathematical biology* **71,** 845–862 (2009).

116. Kanade, V., Levi, R., Lotker, Z., Mallmann-Trenn, F. & Mathieu, C. *Distance in the Forest Fire Model How far are you from Eve?* in *Proceedings of the twenty-seventh annual ACM-SIAM symposium on Discrete algorithms* (2016), 1602–1620.

117. Karampourniotis, P. D., Sreenivasan, S., Szymanski, B. K. & Korniss, G. The Impact of heterogeneous thresholds on social contagion with multiple initiators. *PloS one* **10,** e0143020 (2015).

118. Karsai, M., Kivelä, M., Pan, R. K., Kaski, K., Kertész, J., Barabási, A.-L. & Saramäki, J. Small but slow world: How network topology and burstiness slow down spreading. *Physical Review E* **83,** 025102 (2011).

119. Karsai, M., Iñiguez, G., Kikas, R., Kaski, K. & Kertész, J. Local cascades induced global contagion: How heterogeneous thresholds, exogenous effects, and unconcerned behaviour govern online adoption spreading. *Scientific Reports* **6,** 27178 (2016).

120. Kawachi, I. & Berkman, L. F. Social ties and mental health. *Journal of Urban health* **78,** 458–467 (2001).

121. Keeling, M. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, New Jersey, 2007).

122. Kempe, D., Kleinberg, J. & Tardos, É. *Maximizing the spread of influence through a social network* in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining* (2003), 137–146.

123. Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y. & Porter, M. A. Multilayer networks. *Journal of complex networks* **2,** 203–271 (2014).

124. Klemm, K. & Eguiluz, V. M. Growing scale-free networks with small-world behavior. *Physical Review E* **65,** 057102 (2002).

125. Koch, C. & Segev, I. *Methods in neuronal modeling: from ions to networks* (MIT press, 1998).

126. Kosterev, D. N., Taylor, C. W. & Mittelstadt, W. A. Model validation for the August 10, 1996 WSCC system outage. *IEEE Transactions on Power Systems* **14,** 967–979 (1999).

127. Kulshrestha, J., Kooti, F., Nikravesh, A. & Gummadi, K. P. *Geographic dissection of the twitter network* in *Sixth International AAAI Conference on Weblogs and Social Media* (2012).

128. Kumpula, J. M., Onnela, J.-P., Saramäki, J., Kaski, K. & Kertész, J. Emergence of communities in weighted networks. *Physical review letters* **99,** 228701 (2007).

129. Kwak, H., Lee, C., Park, H. & Moon, S. *What is Twitter, a social network or a news media?* in *Proceedings of the 19th international conference on World wide web* (2010), 591–600.

130. Lambiotte, R., Blondel, V. D., De Kerchove, C., Huens, E., Prieur, C., Smoreda, Z. & Van Dooren, P. Geographical dispersal of mobile communication networks. *Physica A: Statistical Mechanics and its Applications* **387,** 5317–5325 (2008).

131. Latora, V. & Marchiori, M. Efficient behavior of small-world networks. *Physical review letters* **87,** 198701 (2001).

132. Latora, V. & Marchiori, M. Vulnerability and protection of infrastructure networks. *Physical Review E* **71,** 015103 (2005).

133. Latour, B. *Reassembling the Social:An Introduction to Actor-Network-Theory: An Introduction to Actor-Network-Theory* ISBN: 9780199256044. https://books.google.fr/books?id=DlgNiBaYo-YC (OUP Oxford, 2005).

134. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., *et al.* Computational social science. *Science* **323,** 721–723 (2009).

135. Lee, K.-M., Brummitt, C. D. & Goh, K.-I. Threshold cascades with response heterogeneity in multiplex networks. *Physical Review E* **90,** 062816 (2014).

136. Leskovec, J., Kleinberg, J. & Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)* **1,** 2 (2007).

137. Liggett, T. M. *Interacting particle systems* en. ISBN: 978-3-540-22617-8 (Springer, Berlin ; New York, 2005).

138. Lohmann, S. The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91. *World Politics* **47,** 42–101. http://www.jstor.org/stable/2950679 (1994).

139. Luczkovich, J. J., Borgatti, S. P., Johnson, J. C. & Everett, M. G. Defining and measuring trophic role similarity in food webs using regular equivalence. *Journal of Theoretical Biology* **220,** 303–321 (2003).

140. Marsden, P. V. & Campbell, K. E. Measuring tie strength. *Social forces* **63,** 482–501 (1984).

141. Milgram, S. The small world problem. *Psychology today* **2,** 60–67 (1967).

142. Molloy, M. & Reed, B. A critical point for random graphs with a given degree sequence. *Random structures & algorithms* **6,** 161–180 (1995).

143. Molloy, M. & Reed, B. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, probability and computing* **7,** 295–305 (1998).

144. Newman, M. E. J. Analysis of weighted networks. *Phys. Rev. E* **70,** 056131. https://link.aps.org/doi/10.1103/PhysRevE.70.056131 (5 2004).

145. Newman, M. E. J., Watts, D. J. & Strogatz, S. H. Random graph models of social networks. en. *Proceedings of the National Academy of Sciences* **99,** 2566–2572. ISSN: 0027-8424, 1091-6490. http://www.pnas.org/cgi/doi/10.1073/pnas.012582999 (2019) (Feb. 2002).

146. Newman, M. E. J. *Networks* (Oxford University Press, New York, 2018).

147. Newman, M. E. Clustering and preferential attachment in growing networks. *Physical review E* **64,** 025102 (2001).

148. Newman, M. E. Assortative mixing in networks. *Physical review letters* **89,** 208701 (2002).

149. Newman, M. E. Spread of epidemic disease on networks. *Physical review E* **66,** 016128 (2002).

150. Newman, M. E., Strogatz, S. H. & Watts, D. J. Random graphs with arbitrary degree distributions and their applications. *Physical review E* **64,** 026118 (2001).

151. Onnela, J.-P., Kaski, K. & Kertész, J. Clustering and information in correlation based financial networks. *The European Physical Journal B* **38,** 353–362 (2004).

152. Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., Lazer, D., Kaski, K., Kertész, J. & Barabási, A.-L. Structure and tie strengths in mobile communication networks. *Proceedings of the national academy of sciences* **104,** 7332–7336 (2007).

153. Onnela, J.-P., Saramäki, J., Hyvönen, J., Szabó, G., De Menezes, M. A., Kaski, K., Barabási, A.-L. & Kertész, J. Analysis of a large-scale weighted network of one-to-one human communication. *New Journal of Physics* **9,** 179 (2007).

154. Opsahl, T. & Panzarasa, P. Clustering in weighted networks. *Social Networks* **31,** 155 –163. ISSN: 0378-8733. http://www.sciencedirect.com/science/article/pii/S0378873309000070 (2009).

155. Opsahl, T., Agneessens, F. & Skvoretz, J. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks* **32,** 245 –251. ISSN: 0378-8733. http://www.sciencedirect.com/science/article/pii/S0378873310000183 (2010).

156. Palchykov, V., Kaski, K., Kertész, J., Barabási, A.-L. & Dunbar, R. I. Sex differences in intimate relationships. *Scientific reports* **2,** 370 (2012).

157. Pastor-Satorras, R. & Vespignani, A. Epidemic dynamics and endemic states in complex networks. *Physical Review E* **63,** 066117 (2001).

158. Pastor-Satorras, R. & Vespignani, A. *Evolution and structure of the Internet: A statistical physics approach* (Cambridge University Press, 2007).

159. Pastor-Satorras, R., Vázquez, A. & Vespignani, A. Dynamical and correlation properties of the Internet. *Physical review letters* **87,** 258701 (2001).

160. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87,** 925–979 (3 2015).

161. Pecora, L. M. & Carroll, T. L. Master stability functions for synchronized coupled systems. *Physical review letters* **80,** 2109 (1998).

162. Pereira, L. F. C. & Moreira, F. G. B. Majority-vote model on random graphs. *Physical Review E* **71,** 016123 (Jan. 2005).

163. Porter, M. A. & Gleeson, J. P. Dynamical systems on networks. *Frontiers in Applied Dynamical Systems: Reviews and Tutorials* **4** (2016).

164. Rapoport, A. in *Social Networks* 389–409 (Elsevier, 1977).

165. Rogers, E. M. *Diffusion of innovations* (Simon and Schuster, 2010).

166. Ruan, Z., Iñiguez, G., Karsai, M. & Kertész, J. Kinetics of social contagion. *Physical Review Letters* **115,** 218702 (2015).

167. Sachtjen, M. L., Carreras, B. A. & Lynch, V. E. Disturbances in a power transmission system. *Phys. Rev. E* **61,** 4877–4882. https://link.aps.org/doi/10.1103/PhysRevE.61.4877 (5 2000).

168. Saramäki, J., Leicht, E. A., López, E., Roberts, S. G. B., Reed-Tsochas, F. & Dunbar, R. I. M. Persistence of social signatures in human communication. *Proceedings of the National Academy of Sciences* **111,** 942–947. ISSN: 1091-6490 (Jan. 2014).

169. Schelling, T. C. *Neighborhood tipping* (Harvard Institute of Economic Research, Harvard University, 1969).

170. Schelling, T. C. Dynamic models of segregation. *The Journal of Mathematical Sociology* **1,** 143–186 (1971).

171. Schelling, T. C. Hockey Helmets, Concealed Weapons, and Daylight Saving: A Study of Binary Choices with Externalities. *The Journal of Conflict Resolution* **17,** 381–428 (1973).

172. Schelling, T. in *The Corporate Society* (ed Marris, R.) 19–64 (Macmillan Education UK, London, 1974). ISBN: 978-1-349-01977-9. https://doi.org/10.1007/978-1-349-01977-9_2.

173. Schweitzer, F., Fagiolo, G., Sornette, D., Vega-Redondo, F., Vespignani, A. & White, D. R. Economic networks: The new challenges. *science* **325,** 422–425 (2009).

174. Serrano, M. A., Boguñá, M. & Vespignani, A. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences* **106,** 6483–6488. eprint: http://www.pnas.org/content/106/16/6483.full.pdf. http://www.pnas.org/content/106/16/6483.abstract (2009).

175. Shiller, R. Conversation, Information, and Herd Behavior. *American Economic Review* **85,** 181–85. https://EconPapers.repec.org/RePEc:aea:aecrev:v:85:y:1995:i:2:p:181-85 (1995).

176. Singh, P, Sreenivasan, S, Szymanski, B. K. & Korniss, G. Threshold-limited spreading in social networks with multiple initiators. *Scientific Reports* **3,** 2330 (2013).

177. Song, C., Havlin, S. & Makse, H. A. Origins of fractality in the growth of complex networks. *Nature physics* **2,** 275 (2006).

178. Sood, V. & Redner, S. Voter model on heterogeneous graphs. en. *Physical Review Letters* **94,** 178701. (2019) (May 2005).

179. Sporns, O., Chialvo, D. R., Kaiser, M. & Hilgetag, C. C. Organization, development and function of complex brain networks. *Trends in cognitive sciences* **8,** 418–425 (2004).

180. Stanley, H. E. *Introduction to phase transitions and critical phenomena* (Oxford University Press, 1971).

181. Stein, R. B. Some models of neuronal variability. *Biophysical journal* **7,** 37–68 (1967).

182. Suh, B., Hong, L., Pirolli, P. & Chi, E. H. *Want to be retweeted? large scale analytics on factors impacting retweet in twitter network* in *2010 IEEE Second International Conference on Social Computing* (2010), 177–184.

183. Szell, M., Lambiotte, R. & Thurner, S. Multirelational organization of large-scale social networks in an online world. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 13636–13641 (2010).

184. Takaguchi, T., Masuda, N. & Holme, P. Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics. *PloS one* **8,** e68629 (2013).

185. Tamarit, I., Cuesta, J. A., Dunbar, R. I. & Sánchez, A. Cognitive resource allocation determines the organization of personal networks. *Proceedings of the National Academy of Sciences* **115,** 8316–8321 (2018).

186. Tu, Y. How robust is the Internet? *Nature* **406,** 353 (2000).

187. Turner, J. C. *Social influence.* (Thomson Brooks/Cole Publishing Co, 1991).

188. Ugander, J., Karrer, B., Backstrom, L. & Marlow, C. The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503* (2011).

189. Ugander, J., Backstrom, L., Marlow, C. & Kleinberg, J. Structural diversity in social contagion. *Proceedings of the National Academy of Sciences* **109,** 5962–5966 (2012).

190. Unicomb, S., Iñiguez, G. & Karsai, M. Threshold driven contagion on weighted networks. *Scientific reports* **8,** 3094 (2018).

191. Valente, T. W. Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory* **2,** 163–164. ISSN: 1572-9346. https://doi.org/10.1007/BF00240425 (1996).

192. Vázquez, A., Pastor-Satorras, R. & Vespignani, A. Large-scale topological and dynamical properties of the Internet. *Physical Review E* **65,** 066130 (2002).

193. Vazquez, F, Castelló, X & Miguel, M. S. Agent based models of language competition: macroscopic descriptions and order–disorder transitions. *Journal of Statistical Mechanics: Theory and Experiment* **2010,** P04007 (2010).

194. Verbrugge, L. M. Multiplexity in adult friendships. *Social Forces* **57,** 1286–1309 (1979).

195. Wang, W.-X., Wang, B.-H., Hu, B., Yan, G. & Ou, Q. General Dynamics of Topology and Traffic on Weighted Technological Networks. *Phys. Rev. Lett.* **94,** 188702. https://link.aps.org/doi/10.1103/PhysRevLett.94.188702 (18 2005).

196. Wasserman, S. & Faust, K. *Social network analysis: Methods and applications* (Cambridge university press, 1994).

197. Watkins, D. S. Understanding the QR algorithm. *SIAM review* **24,** 427–440 (1982).

198. Watts, D. J. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences* **99,** 5766–5771 (2002).

199. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *nature* **393,** 440 (1998).

200. *WikiSigned network* http://konect.uni-koblenz.de/networks/wikisigned-k2. Accessed: 2017-06-07.

201. Wimmer, A. & Lewis, K. Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *American Journal of Sociology* **116,** 583–642 (2010).

202. Yağan, O. & Gligor, V. Analysis of complex contagions in random multiplex networks. *Physical Review E* **86,** 036103 (2012).

203. Yang, R., Wang, W.-X., Lai, Y.-C. & Chen, G. Optimal weighting scheme for suppressing cascades and traffic congestion in complex networks. *Phys. Rev. E* **79,** 026112 (2009).

204. Yook, S. H., Jeong, H., Barabási, A.-L. & Tu, Y. Weighted Evolving Networks. *Phys. Rev. Lett.* **86,** 5835–5838. https://link.aps.org/doi/10.1103/PhysRevLett.86.5835 (25 2001).

205. Zhou, S. & Mondragón, R. J. The rich-club phenomenon in the Internet topology. *IEEE Communications Letters* **8,** 180–182 (2004).

206. Zhou, W.-X., Sornette, D., Hill, R. A. & Dunbar, R. I. Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society of London B: Biological Sciences* **272,** 439–444 (2005).

207. Zhu, X., Gerstein, M. & Snyder, M. Getting connected: analysis and principles of biological networks. *Genes & development* **21,** 1010–1024 (2007).

208. Zhu, Y.-X., Wang, W., Tang, M. & Ahn, Y.-Y. Social contagions on weighted networks. *Phys. Rev. E* **96,** 012306. https://link.aps.org/doi/10.1103/PhysRevE.96.012306 (1 2017).

209. Zhuang, Y., Arenas, A. & Yağan, O. Clustering determines the dynamics of complex contagions in multiplex networks. *Physical Review E* **95,** 012312 (2017).

## DECLARATION

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

*Lyon, November 2019*

Samuel Unicomb

## COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede and Ivo Pletikosić. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/

Thank you very much for your feedback and contribution.

*Final Version* as of November 19, 2019 (classicthesis v4.6).