2-1-2016

Statistics of Epidemics in Networks by Passing Messages

Munik Shrestha

Munik Shrestha

Candidate

Physics and Astronomy

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Carl Caves , Chairperson

Cris Moore

David Dunlap

Mousomi Roy

Statistics of Epidemics in Networks by Passing Messages

by

Munik Kumar Shrestha

Bachelor of Science, Haverford College, 2009

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY
PHYSICS

The University of New Mexico

Albuquerque, New Mexico

December, 2015

Dedication

To the land of enchantment

Across the Rio Grande-eo, across the lazy river

- Grateful Dead

Acknowledgments

I thank New Mexico for showing me what it is like to spend some few years in the american southwest. Known for its rugged and colorful landscape, my time here across the Rio Grande river has been full of enchantment both intellectually and personally.

I am grateful and thankful to my adviser Cris Moore for his honest feedbacks and guidance through out my graduate school experience. I thank Cris for giving me the freedom to explore my research interests and to ask simple and naive questions.

I thank my dissertation committee for their time in reviewing this manuscript and giving me the opportunity to present and discuss the work in this manuscript. I thank Carl Caves for guiding me as an official adviser at the university and making it possible to work with Cris at the Santa Fe Institute.

I thank the University of New Mexico for its inclusive, laid-back and rebel-full atmosphere. I thank the Santa Fe Institute, where I also spent much time as a graduate fellow, for a vibrant multidisciplinary environment.

I thank all my friends, lovers, colleagues, acquaintances, and strangers for sharing their time with me. Thanks to my cat—friend Leroy; he is indeed Leroy the brown, the baddest cat in town.

I am always grateful and thankful to my mother Muna Shrestha, my father Mohan Shrestha, and my brothers Monish Shrestha and Monir Shrestha for their love and support. I thank my mother for teaching me how to add and multiply numbers.

Statistics of Epidemics in Networks by Passing Messages

by

Munik Kumar Shrestha

Bachelor of Science, Haverford College, 2009

Doctor of Philosophy, Physics, University of New Mexico, 2015

Abstract

Epidemic processes are common out-of-equilibrium phenomena of broad interdisciplinary interest. In this thesis, we show how message-passing approach can be a helpful tool for simulating epidemic models in disordered medium like networks, and in particular for estimating the probability that a given node will become infectious at a particular time. The sort of dynamics we consider are stochastic, where randomness can arise from the stochastic events or from the randomness of network structures.

As in belief propagation, variables or messages in message-passing approach are defined on the directed edges of a network. However, unlike belief propagation, where the posterior distributions are updated according to Bayes' rule, in message-passing approach we write differential equations for the messages over time. It takes correlations between neighboring nodes into account while preventing causal signals from backtracking to their immediate source, and thus avoids "echo chamber effects" where a pair of adjacent nodes each amplify the probability that the other is infectious.

In our first results, we develop a message-passing approach to threshold models of behavior popular in sociology. These are models, first proposed by Granovetter, where individuals have to hear about a trend or behavior from some number of neighbors before adopting it themselves. In thermodynamic limit of large random networks, we provide an exact analytic scheme while calculating the time dependence of the probabilities and thus learning about the whole dynamics of bootstrap percolation, which is a simple model known in statistical physics for exhibiting discontinuous phase transition.

As an application, we apply a similar model to financial networks, studying when bankruptcies spread due to the sudden devaluation of shared assets in overlapping portfolios. We predict that although diversification may be good for individual institutions, it can create dangerous systemic effects, and as a result financial contagion gets worse with too much diversification. We also predict that financial system exhibits "robust yet fragile" behavior, with regions of the parameter space where contagion is rare but catastrophic whenever it occurs.

In further results, we develop a message-passing approach to recurrent state epidemics like susceptible-infectious-susceptible and susceptible-infectious-recovered-susceptible where nodes can return to previously inhabited states and multiple waves of infection can pass through the population. Given that message-passing has been applied exclusively to models with one-way state changes like susceptible-infectious and susceptible-infectious-recovered, we develop message-passing for recurrent epidemics based on a new class of differential equations and demonstrate that our approach is simple and efficiently approximates results obtained from Monte Carlo simulation, and that the accuracy of message-passing is often superior to the pair approximation (which also takes second-order correlations into account).

Contents

List of Figures

Chapter 1

Introduction

But what does it mean, the plague?

- **Albert Camus**

In this manuscript, we study mathematical approaches for epidemics in networks. Epidemics, also known as contact processes, are common non-equilibrium phenomena that have gained vigorous attention in recent years because of important interdisciplinary applications. Examples include contagion of diseases [1, 2] and social behaviors [3, 4, 5, 6], cascades of forest fires [7, 8, 9] and banking failures [10, 11], propagation of marginal probabilities in constraint satisfaction problems [12, 13], the dynamics of magnetic and glassy systems [14], and so on.

Mathematical models of epidemics are intrinsically non-linear and multiplicative. The classical approach to modeling epidemics, such as the SIR (susceptible-infectious-recovered) model where each node is Susceptible, Infectious, or Recovered, assumes that at any given time each individual exists in a single state or "compartment" [1, 2]. To make these models analytically tractable, it is often assumed that the population is well mixed, so that interaction between any two individuals is

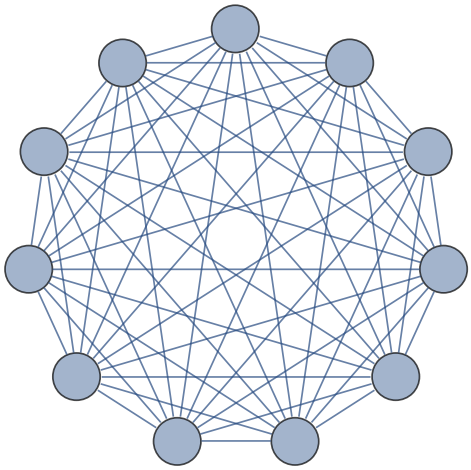

Figure 1.1: An example of a complete network with eleven individuals or nodes, where every node is connected via an edge to every other node.

equally likely as if the epidemics are happening in a complete network as shown in Fig. 1.1. In physical terms, we assume the model is mean-field (also known as mass-action mixing). In spite of this simple but unrealistic assumption, mean-field models capture some essential features of epidemics, such as a threshold above which we have an endemic phase with a non-zero fraction of infected individuals, and below which we have outbreaks of size $o(n)$ so that the equilibrium fraction of infected individuals is zero.

In reality, the mixing of individuals is often sparse and highly structured, with some pairs of individuals much more likely to interact than others due to location or demographics [15, 6]. We mean sparse in the sense that the number of neighbors with which individuals interact is an intensive quantity, i.e. the number of neighbors is independent of the size of the population. Analogous to low dimensional lattices in statistical physics, correlations in sparse networks play an important role, making problems in sparse networks elude solutions derived through a mean-field approximation that washes out spatial correlations. The specific realization of sparse networks thus becomes essential to consider [16, 17, 18]. However, the network structure sub-

stantially increases a model's complexity.

One reasonable goal is to compute the one-point marginals, e.g., for each node i, the probability $I_i(t)$ that i is infectious at time t. In addition to being of direct interest, these marginals help us perform tasks such as inferring the originator of an epidemic, determining an optimal set of nodes to immunize in order to minimize the final size of an outbreak, or calculating the probability that an entire group of nodes will remain uninfected after a fixed time [19, 20, 21, 22, 23].

We can always compute these marginals by performing Monte Carlo experiments. However, since we need to perform many independent trials in order to collect good statistics, this is computationally expensive on large networks. This problem is compounded if we need to scan through parameter space, or if we want to explore many different initial conditions, vaccination strategies, etc. Therefore, it would be desirable to compute these marginals using, say, a system of differential equations, with variables that directly model the probabilities of various events.

The most naive way to do this, as we review below, uses the one-point marginals themselves as variables. However, this approach completely ignores correlations between nodes. At the other extreme, to model the system exactly, we would need to keep track of the entire joint distribution: but if there are n individuals, each of which can be in one of k states, the result is a coupled system with k^n variables. This exponential scaling quickly renders most models computationally intractable, even on moderately sized networks.

In between these two extremes, we can approximate the joint distribution by "moment closure," assuming that higher-order marginals can be written in terms of lower-order ones. This gives a hierarchy of increasingly accurate (and computationally expensive) approximations, familiar in physics as cluster expansions. At the first level of this hierarchy we assume that the nodes are uncorrelated, and approximate

two-point marginals such as $[I_i(t)I_j(t)]$ (the probability that i and j are both infectious at time t) as $I_j(t)I_j(t)$. At the second level, commonly referred to as the pair approximation, we close the hierarchy at the level of pairs $[I_i(t)I_j(t)]$ by assuming that three-point correlations can be factored in terms of two-point correlations. We review these methods in Chapter 2.

In this thesis, we study an alternative method, namely *Dynamic Message-Passing* (DMP). As in belief propagation [24, 25], here variables or "messages" are defined on a network's *directed* edges: for instance, $I_{j\to i}$ denotes the probability that j was infected by one of its neighbors other than i, so that the epidemic might spread from j to i. However, unlike belief propagation, where the posterior distributions are updated according to Bayes' rule, here we write differential equations for the messages over time.

We will see that the directional nature of the messages prevents causal signals from backtracking to their immediate source, and thus avoids "echo chamber effects" where a pair of adjacent nodes each amplify the probability that the other is infectious. DMP was first applied to disease propagation on networks by Karrer and Newman (2010) [26], who investigated non-recurrent state epidemiological models such as the SI (susceptible-infectious), the SIR (susceptible- infectious-recovered) and the SEIR (susceptible-exposed-infectious-recovered) models.

In this thesis, we make the following contributions:

- To prepare the reader, in Chapter 2, we provide background materials for three substantive chapters that are to follow.

- In Chapter 3, in collaboration with my adviser Cristopher Moore, we extend Karrer and Newman (2010) [26] by generalizing DMP to threshold models where healthy individuals get infected only when a certain number of their

neighbors pass infection, a process thought to be important for the propagation of memes and first proposed by Granovetter, where each individual has to hear about a trend or behavior from some number of neighbors before adopting it themselves [27]. In thermodynamic limit of large random networks, we provide an exact analytic scheme while calculating the time dependence of the probabilities and thus learning about the whole dynamics of bootstrap percolation. [Note that bootstrap percolation model is known and studied in statistical physics for exhibiting discontinuous phase transition.] Our approach is general enough to incorporate non-Markovian processes and to include heterogeneous thresholds, and thus can be applied to explore rich sets of complex heterogeneous agent-based models.

The work described in this chapter was published in Phys. Rev. E 89, 022805 (2014). [27].

- In Chapter 4, in collaboration with Sam Scarpino and Cristopher Moore, we extend DMP to a general class of epidemics models including "recurrent" models such as SIS (susceptible-infectious-susceptible) and SIRS (susceptible-infectious-recovered-susceptible), where multiple waves of infection can pass through the population. To date, DMP has been exclusively applied to non-recurrent models, where individuals cannot return to previously occupied state [26, 28, 27, 29, 30]. We have developed a new class of differential equations for these models. Our methods are much faster than direct simulation and are also far more efficient and conceptually simpler than the pair approximations (which take second-order correlations into account) currently used in epidemiology.

  The work described in this chapter has been submitted for publication in Phys. Rev. E.

- In Chapter 5, in collaboration with Fabio Caccioli, Cristopher Moore and J.

Doyne Farmer, we apply an approach similar to threshold models (Chapter 3) to financial networks, studying when bankruptcies spread due to the sudden devaluation of shared assets in overlapping portfolios. Our model estimates the circumstances under which systemic instabilities are likely to occur as a function of parameters such as leverage, market crowding, diversification, and market impact. We predict that although diversification may be good for individual institutions, it can create dangerous systemic effects, and as a result financial contagion gets worse with too much diversification; additionally, financial system exhibits "robust yet fragile" behavior, with regions of the parameter space where contagion is rare but catastrophic whenever it occurs.

The work described in this chapter was published in J. of Banking & Finance Volume 46, (2014). [11].

- We conclude in Chapter 6 by summarizing our contributions and discussing some open questions and potential directions towards future work.

Chapter 2

Background

If a man's wit be wandering, let him study the mathematics.
- **Francis Bacon**

2.1 Epidemics in Networks

For the purpose of illustration, let us consider a simple recurrent state epidemic model, the susceptible-infectious-susceptible (SIS) model. In this model, each individual or node is either Infectious (I) or Susceptible (S). Infectious nodes infect their Susceptible neighbors independently at rate λ and recover back independently to a Susceptible state at rate ρ. Note that this model becomes a non-recurrent state model when infectious node do not recover from infection, i.e. the susceptible-infectious (SI) model if the recovery rate is 0. We denote the probability that that i is Infectious and Susceptible by I_i and S_i respectively. Our goal then is to efficiently and accurately compute these probabilities as a function of continuous time t.

The exact system of ordinary differential equations for I_i, i.e. the master−

equation, for SIS model is

$$\frac{dI_i}{dt} = -\rho I_i + \lambda \sum_{j \in \partial i} [S_i I_j], \tag{2.1}$$

where ∂i are neighbors of i, and $[S_i I_j]$ is the joint probability that i is Susceptible and j is Infectious. The first term is the rate at which i independently recovers from infection, whereas the second term refers to the rate at which if i is Susceptible, it transition to become Infectious by getting infected from its neighbors.

However, although this system is exact, it is not closed, i.e. the marginal probability I_i depends on the two-point marginals $[S_i I_j]$, whose time derivatives, as we review below, depend in turn on three-point marginals, and so on.

2.2 First-order moment closure

In the first-order moment closure method, we assume (approximate) that neighboring nodes are uncorrelated, i.e.,

$$[S_i I_j] \approx S_i I_i. \tag{2.2}$$

Thus the ordinary differential equation system via this first-order moment closure becomes

$$\frac{dI_i}{dt} = -\rho I_i + \lambda S_i \sum_{j \in \partial i} I_j. \tag{2.3}$$

In some cases, this can be a reasonable assumption. For instance in a large complete network, where every pair of nodes is connected by an edge, the state of the system is essentially driven the fraction of individuals in a given compartmental state and is equivalent to a fully mixed mean-field model.

But in some instances, we cannot ignore correlations. Consider, for example, a simple but pathological case of the SI model where there are only two nodes in the

Figure 2.1: Two simple, yet illustrative, cases of networks, where the darker node is initially Infectious.

graph, i and j, with an edge between them as shown in Fig. 2.1. If the transmission rate is λ, and if we assume the nodes are independent (i.e., if we use first-order moment closure) we obtain the following differential equations,

$$\begin{aligned} \frac{dI_i}{dt} &= \lambda S_i I_j \\ \frac{dI_j}{dt} &= \lambda S_j I_i \,, \end{aligned} \tag{2.4}$$

where $S_i(t) = 1 - I_i(t)$ and similarly for j.

Now suppose that j is initially Infectious with probability δ, and that i is initially Susceptible, i.e., $I_j(0) = \delta$ and $I_i(0) = 0$. Since in the SI model nodes never recover, the infection will eventually spread from j to i, but only if i was Infectious in the first place. Thus the marginals $I_i(t)$ and $I_j(t)$ should tend to δ as $t \to \infty$.

However, integrating Eq. (2.4) gives a different result. Once I_i becomes positive, dI_j/dt becomes positive as well, allowing i to infect j with the infection that it received from j in the first place. As a result, $I_j(t)$ approaches 1 as $t \to \infty$. Thus the "echo chamber" between i and j leads to the absurd result that j eventually becomes Infectious, even though with probability $1 - \delta$ there was no initial infection in the system.

2.3 Second-order moment closure, i.e. the pair approximation

One way to account for correlations between neighboring states, such as the echo-chamber effects we discussed in the previous section, is by tracking the dynamics of neighboring pairs $[S_i I_j]$ as well. The ordinary differential equation for the pair $[S_i I_j]$ in the SIS model is given by

$$\frac{d[S_i I_j]}{dt} = -(\rho + \lambda)[S_i I_j] + \rho[I_i I_j] + \lambda \sum_{k \in \partial j \setminus i} [S_i S_j I_k] - \lambda \sum_{\ell \in \partial i \setminus j} [I_\ell S_i I_j]. \tag{2.5}$$

But as we can see, the derivative of the pair $[S_i I_j]$ depends on the three-point marginals $[S_i S_j I_k]$ and $[I_\ell S_i I_j]$. In principle, we could again write the derivative of these three-point marginals in terms of four-point marginals, and those of four-point marginals in terms of five-point marginals, and so on for any finite network. But the number of variables increases exponentially, as a result of which, we need to close this hierarchical dependence of correlations at some level at the cost of some accuracy.

In the second-order moment closure, also known as the pair approximations, we factor the triplet distribution $[S_i S_j I_k]$ as

$$[S_i S_j I_k] \simeq \frac{[S_i S_j][S_j I_k]}{S_j}. \tag{2.6}$$

Therefore, the system of ordinary differential equations for the SIS model via this pair approximations becomes

$$\frac{d[S_i I_j]}{dt} = \rho[I_i I_j] - (\rho + \lambda)[S_i I_j] + \frac{\lambda[S_i S_j]}{S_j} \sum_{k \in \partial j \setminus i} [S_j I_k] - \frac{\lambda[S_i I_j]}{S_i} \sum_{\ell \in \partial i \setminus j} [S_i I_\ell]. \tag{2.7}$$

Note that the evolution of $[S_i I_j]$ depends on other two-point marginals like $[I_i I_j]$ and $[S_i S_j]$. But, by definition we have

$$\begin{aligned} S_i &= [S_i I_j] + [S_i S_j] \\ I_i &= [I_i S_j] + [I_i I_j], \end{aligned} \tag{2.8}$$

which, along with Eq. (2.1), closes the system at the level of pairs.

In the two-node example of Fig. 2.1 (left), of course, the pair approximation is exact, since it maintains separate variables such as $[S_j I_k]$ for each of the joint states of the two nodes. Now, consider a case with three nodes, as in Fig. 2.1 (right), where j is the common neighbor of i and k. The pair-approximation assumes that given j is Susceptible (as in the figure), the pairs $[S_i S_j]$ and $[S_j I_k]$ are independent. In non-recurrent models like SI, nodes do not revert back to the Susceptible state once they transition out of it. So, knowing that j — being the only node between i and k — is Susceptible is a sufficient condition to block all correlations between i and k.

However, in a recurrent epidemic model, i and k could be correlated, for instance if j infected them both and then returned to the Susceptible state. As a result, the pair approximation is vulnerable to a distance-two echo chamber, where i and k infect each other through j.

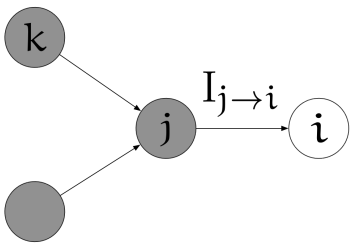

Figure 2.2: We define messages on the directed edges of a network to carry causal information of the flow of contagion, e.g. $I_{j\to i}$ is the probability that j is Infectious because it received the infection from a neighbor k other than i. This prevents effects from immediately backtracking to the node they came from, and avoids "echo chamber" infections.

2.4 Message-passing approach: a prelude

Well, it seemed to me quite evident that the idea that a particle acts on itself is not a necessary one. And so I suggested to myself that electrons cannot act on themselves; they can only act on other electrons.

- Richard Feynman

To prevent echo chamber effects, we consider messages that are defined on the *directed* edges of the network. The idea is to prevent infection signals from backtracking to the node that they immediately came from and thus avoid an echo chamber of information flow.

This idea of making the flow of information directional in dynamic message-passing is similar to belief propagation [24, 25], where we use the network structure to update posterior probabilities of the vertices' states. However, unlike belief propagation where we update posterior distributions according to Bayes' rule, the causal structure of information flow is captured directly by the time evolution of DMP. In

other words, in DMP we write differential equations for the messages over time. This approach takes correlations between neighboring nodes into account while preventing causal signals from backtracking to their immediate source, and thus avoids "echo chamber effects" where a pair of adjacent nodes each amplify the probability that the other is infectious.

In the next two chapters, we present in detail and in a self-contained way, the message-passing approach to models of epidemics in networks.

Chapter 3

Message-passing for non-recurrent threshold models of epidemics

The work described in this chapter is a result of collaboration with Cristopher Moore and is published in Phys. Rev. E 89, 022805 (2014) [27]

Mathematical modeling of epidemics has attracted the interest of researchers from diverse academic disciplines [1, 2, 16, 10, 31, 32, 33, 34, 35, 36, 37, 38, 39, 5, 11, 40, 41]. Epidemics range from outbreaks of infectious disease to the contagion of social behaviors such as trends, memes, fads, political opinions, rumors, innovations, financial decisions, and so on. In an early study, Granovetter [3, 4] proposed a threshold model, where individuals adopt a behavior when they are informed by at least T of their neighbors.

We consider a stochastic model similar to Granovetter's with a trend propagating on a network. At each time, an individual has integer valued awareness of a trend ranging from 0 to T. Each time an individual is informed by one of its neighbors, this awareness is incremented until it reaches the threshold T. At that point, that individual adopts the trend, and starts informing its neighbors about it. We will

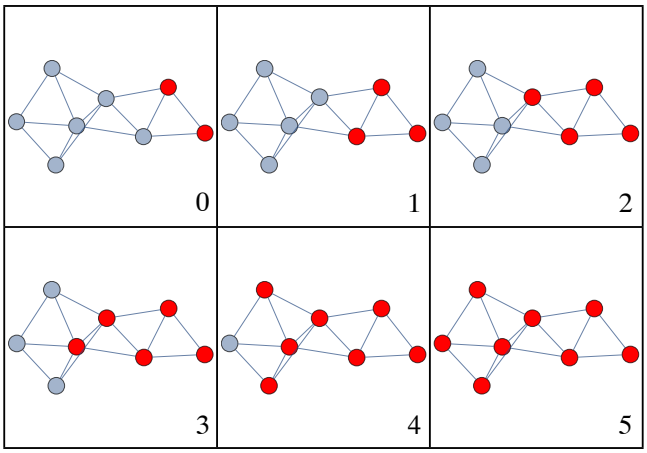

Figure 3.1: Schematic illustration of the evolution of contagion in a finite network, where individuals have threshold $\mathsf{T} = 2$. Adopters are colored red. For simplicity, we assume adopters inform their neighbors one time step after they become adopters. In fact we study a stochastic continuous-time version of this model.

assume that the network topology is fixed, but our model of information flow (or "contagion") is probabilistic. Each adopter informs each of its neighbors at a rate $\mathsf{r}(\tau)$, where τ is the time elapsed since it became an adopter. Since $\mathsf{r}(\tau)$ may depend on τ, the resulting dynamics can be non-Markovian. In an illustrative example, we show a schematic evolution of the contagion in Figure 3.1, where $\mathsf{T} = 2$.

Given an initial condition, where some individuals have already become adopters, or have done so with some probability, our goal in this chapter is to calculate the probability that any given individual i is an adopter (or not an adopter) as a function of time. More generally, we calculate the probability $\mathsf{P}^{\mathsf{i}}_{\mathsf{a}}(\mathsf{t})$ that i has awareness a at time t. Then the probability that i is an adopter is $\mathsf{P}^{\mathsf{i}}_{\mathsf{T}}$.

Calculating the time evolution of the probability $\mathsf{P}^{\mathsf{i}}_{\mathsf{a}}(\mathsf{t})$ is non-trivial as a result of intrinsic nonlinearities in the dynamics. The heterogeneous network interactions between individuals make it even harder. One simple way to estimate these probabilities is to put on a computational-frequentist hat, simulate the model many times

independently by a Monte Carlo agent-based method, and measure in what fraction of these runs each vertex becomes an adopter. Doing this is computationally costly, however, as we are required to perform many independent runs of the simulation

We thus consider the dynamic message passing algorithm (DMP), where we evolve the probabilities $P_a^i(t)$ directly according to certain update equations. Compared to a Monte Carlo simulation that requires many independent runs, we only need to run the DMP algorithm once. In the special case where $T = 1$, DMP was recently formulated by Karrer and Newman [42] to analytically study non-Markovian dynamics of the Susceptible, Infected, Recovered (SIR) epidemic model of the networks. In an analogy with the SIR model, we sometimes refer to a vertex as *susceptible* if it is not yet an adopter, *infected* if it is an adopter, and *recovered* if it is an adopter but the rate $r(\tau)$ at which it informs its neighbors has dropped to zero.

The underlying idea of dynamic message passing is similar to belief propagation [24, 25], where we use the network structure to update posterior probabilities of the vertices' states. However, unlike belief propagation where we update posterior distributions according to Bayes' rule, the causal structure of information flow is captured directly by the time iteration of DMP. As in belief propagation, the DMP algorithm assumes that the neighbors of each vertex are conditionally independent of each other. As a result, like belief propagation, DMP is exact on trees and approximate on networks with loops, where the conditional independence assumption cannot capture higher order correlations.

However, as we will see, DMP gives good approximations to the probabilities even on real networks with many loops. We will show this by implementing it in a real social network, specifically Zachary's karate club network [43]. Although the Zachary's club network contains many loops, the probabilities computed by DMP compare well with those from the Monte Carlo simulation. We present this in Section 3.2.

In the limit of large random networks in the Erdős-Rényi model, or networks with a given degree distribution, DMP is asymptotically exact because these networks are locally treelike. In Section 3.3, we use DMP to obtain the exact results for such random networks in the thermodynamic limit.

3.0.1 Related Work

There are many related studies that consider what fraction of vertices eventually become adopters if each neighbor informs them with probability p. The set of eventual adopters are the ones who have at least T neighbors who are also adopters. This is reminiscent of the model commonly studied in statistical physics as k-core (or bootstrap) percolation. The k-core is the maximal induced subgraph in the network, such that each vertex has at least k other neighbors in the subgraph.

By deleting each edge with probability $1-\mathsf{p}$ independently, we can ask whether the resulting diluted network in the thermodynamic limit contains an extensive k-core in the ensemble of similarly prepared networks. Interestingly for $\mathsf{k} \geqslant 3$, the emergence of a k-core in random networks is a first-order (discontinuous) phase transition in the sense that when it first appears it covers a finite fraction of the network [44]. An early work on k-core percolation was on the Bethe lattice in the context of magnetic systems [45]. Recently, it has been used in studies of the Ising model and nucleation [46, 47], analysis of zero temperature jamming transitions [48], and in a bootstrap percolation model in square lattices and random graphs [49, 50, 51, 39, 52].

3.1 Message-passing approach

We now formulate the dynamic message passing (DMP) technique for the threshold models. We define the message $u_{i\leftarrow j}(t)$ as the probability that vertex j has *not* informed i about the trend by time t. If we have $u_{i\leftarrow j}(t)$ for all neighboring pairs i, j, we will be able to calculate the marginal probability $p^i_a(t)$ that i has awareness a at time t, i.e. that it has been informed by a of its neighbors. We focus on initial conditions where each vertex is either an adopter or has awareness zero. So given that i is not an initial adopter,

$$p^i_a(t) = \sum_{\substack{\Theta\subseteq\partial i\\ |\Theta|=a}} \prod_{j\in\Theta}(1-u_{i\leftarrow j}(t)) \prod_{j\in\partial i\setminus\Theta} u_{i\leftarrow j}(t). \tag{3.1}$$

Here, ∂i is the set of i's neighbors, and Θ ranges over all subsets of ∂i of size a. Note the conditional independence assumption in Equation (3.1), where we assume that the events that j has informed (or not informed) i are independent. That is, we assume that the probability that i has been informed by a given set of neighbors Θ is the product over $j \in \Theta$ and $j \notin \Theta$ of the probability that j has or has not informed i respectively.

Given that i is not an initial adopter, the probability $p^i_S(t)$ that the vertex i is susceptible at time t, i.e. its awareness is less than T at time t, is then

$$p^i_S(t) = \sum_{a=0}^{T-1} p^i_a(t). \tag{3.2}$$

Equivalently,

$$p^i_S(t) = \sum_{\substack{\Theta\subseteq\partial i\\ |\Theta|<T}} \prod_{j\in\Theta}(1-u_{i\leftarrow j}(t)) \prod_{j\in\partial i\setminus\Theta} u_{i\leftarrow j}(t). \tag{3.3}$$

We can see that this expression is easy to generalize to the case where each individual has its own threshold T_i. For instance, we could set T_i to some fraction of i's degree.

We could also assume a probabilistic threshold T_i for each i drawn from some distribution $P(T_i)$ and take an average over the threshold in Equation (3.1). We can also capture the case where i initially has awareness a_i by setting $T_i = T - a_i$. However, for simplicity, we assume that every individual has the same threshold, and everyone starts with an initial awareness of 0 or T.

Given $P_S^i(t)$, we note that i is an adopter if it is at the root of a T-ary tree, whose nodes are mapped onto the vertices of the network, such that 1) the leaves of the tree are initial adopters, 2) the T children of each tree node are mapped to distinct vertices, 3) none of the paths from the root to the leaves backtracks; that is, an edge (u, v) cannot be immediately followed by the edge (v, u), and 4) the trend is successfully transmitted along each edge of this tree.

To capture the information flow that the message $U_{i\leftarrow j}(t)$ represents, we define $P_S^{j\setminus i}(t)$, which is the probability that j would be susceptible at time t if i were absent from the network. Alternately, this is the probability that j is susceptible at time t if we ignore the possibility of j being informed of the trend by i. In removing the vertex i (or ignoring the flow of information to j from i), we bring the information flow to i based on the information or messages that neighbor j receives from j's other neighbors. We thus avoid the "echo-chamber" effect, where i informs j, and j informs i back, and so on.

In an analogy with the cavity method of statistical physics, we call $P_S^{j\setminus i}(t)$ the cavity probability that j is susceptible given that i is in a noninteracting "cavity state". Hence, using Equation (3.3), if j was not an initial adopter, then $P_S^{j\setminus i}$ can be written as

$$P_S^{j\setminus i}(t) = \sum_{\substack{\Theta\subseteq \partial j\setminus i \\ |\Theta|<T}} \prod_{\ell\in\Theta}(1 - U_{j\leftarrow\ell}(t)) \prod_{\ell\in\partial j\setminus\{\Theta,i\}} U_{j\leftarrow\ell}(t). \tag{3.4}$$

Note that initially $P_S^{j\setminus i}(0) = P_S^j(0)$, since the initial probability that j is an an adopter

does not depend on i. Similarly, the cavity rate $p_I^{j\setminus i}(t)$ at which j becomes an adopter at time t, if it was not an adopter initially, is then

$$p_I^{j\setminus i}(t) = -\frac{dP_S^{j\setminus i}(t)}{dt}. \tag{3.5}$$

It is convenient to define $f(\tau)$ as the rate at which j *first* informs i at time t, if j became an adopter at time $t' = t - \tau$. In particular, if j informs i at a rate $r(\tau)$, then $f(\tau) = r(\tau)e^{-\int_0^\tau d\tau' r(\tau')}$ is the rate at which j informs i for the first time at time t. Note that $f(\tau)$ might not be normalized, since the probability $p = \int_0^\infty d\tau f(\tau)$ that j ever informs i may be less than 1. By letting $f(\tau)$ depend arbitrarily on the time τ since j became an adopter, we can handle both Markovian and non-Markovian models. In particular, if an adopter inform its neighbors at some constant rate β, we have

$$f(\tau) = \beta e^{-\beta\tau}. \tag{3.6}$$

If an adopter "recovers" with rate γ as in the SIR model, after which it no longer informs its neighbors about the trend, then $f(\tau)$ becomes

$$f(\tau) = \beta e^{-(\gamma+\beta)\tau}, \tag{3.7}$$

where $e^{-\gamma\tau}$ is the probability that an adopter has itself not recovered up to the elapsed time τ. Note in general we can let $f(\tau)$ depend on i and j, giving arbitrary inhomogeneous rates at which individuals inform each other; we do not pursue this here.

Although we have defined the messages and shown how they allow us to calculate the probabilities $P_a^i(t)$, we have not yet shown how to calculate the messages themselves.

So, let us now calculate the messages $U_{i\leftarrow j}(t)$. The rate at which $U_{i\leftarrow j}(t)$ decreases at time t is the rate at which j informs i for the first time at time t. This

happens in two ways. If j was an initial adopter, it informs i for the first time at time t at the rate $f(t)$. Or, if j was initially susceptible, j becomes an adopter at some time $t' = t - \tau$, and informs i for the first time at the rate $f(t - t')$ at time t. Integrating this over t' up to time t, we see that j will inform i for the first time at the rate $\int_0^t dt' f(t - t') p_I^{j\setminus i}(t')$. Combining these two cases with Equation (3.5), the rate at which the message $U_{i\leftarrow j}(t)$ decreases at time t is thus given by

$$\begin{aligned} -\frac{dU_{i\leftarrow j}(t)}{dt} &= f(t)[1 - P_S^j(0)] + P_S^j(0) \int_0^t dt' f(t - t') p_I^{j\setminus i}(t') \\ &= f(t)[1 - P_S^j(0)] - P_S^j(0) \int_0^t dt' f(t - t') \frac{dP_S^{j\setminus i}(t')}{dt'}. \end{aligned} \tag{3.8}$$

Integrating by parts gives

$$\begin{aligned} \frac{dU_{i\leftarrow j}(t)}{dt} &= -f(t) + f(0) P_S^j(0) P_S^{j\setminus i}(t) \\ &\quad + P_S^j(0) \int_0^t dt' P_S^{j\setminus i}(t') \frac{df(t - t')}{dt}. \end{aligned} \tag{3.9}$$

One may check that the solution of (3.9) is

$$U_{i\leftarrow j}(t) = 1 - \int_0^t d\tau f(\tau) + P_S^j(0) \int_0^t d\tau f(\tau) P_S^{j\setminus i}(t - \tau). \tag{3.10}$$

We can explain this expression, as in [42], as follows. The term $1 - \int_0^t d\tau f(\tau)$ is the probability that the elapsed time τ, after which j informs i for the first time, is greater than the absolute time t, i.e. $\tau > t$. In this case, i is not informed by j, even if j became an adopter before time t. The second term is the probability that i would have been informed at time t if j had been an adopter at time $t - \tau$, but that j was not yet an adopter at that time.

Note however that Equation (3.9) is an integro-differential equation, so numerically integrating it can be computationally costly. It is possible to numerically integrate (3.10), or, for particular functions $f(\tau)$, we can transform (3.9) into an

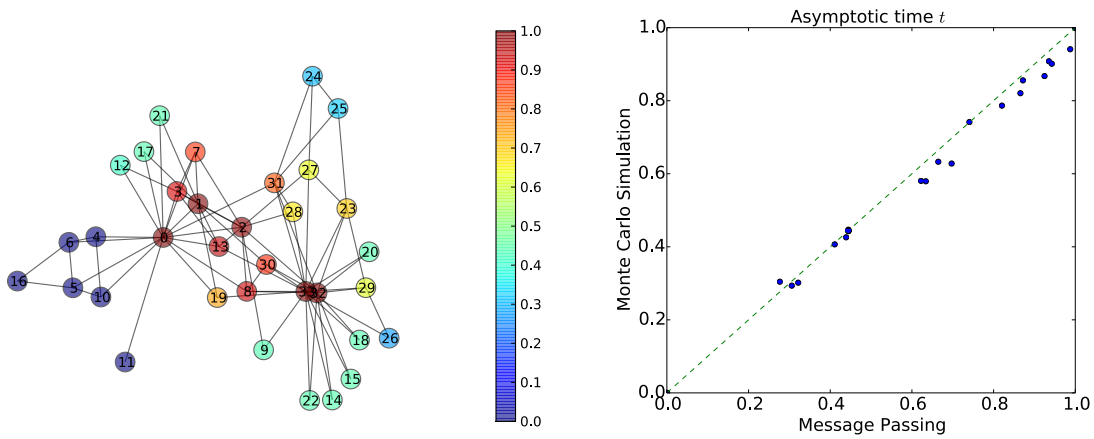

Figure 3.2: Comparison (right) with a scatter plot of individuals eventual infection probability in the Zachary club (left), where threshold $T = 2$. Horizontal axis is the eventual infection probability calculated by the DMP, whereas vertical axis is the result from the Monte Carlo simulation. Each point refers to the eventual infection probability of one of the individuals in the club. Here, four initially infected individuals are $\{0, 1, 32, 33\}$. Simulation is averaged over 10^5 runs. Transmission rate $\beta = 0.6$, and recovery rate $\gamma = 0.3$. Vertices on the left are colored according to their eventual infection probability from the DMP.

ordinary differential equation. For example if we plug $f(\tau)$ from (3.7) and integrate the last term in (3.9) by parts, we obtain

$$\frac{dU_{i\leftarrow j}(t)}{dt} = -\beta U_{i\leftarrow j}(t) + \gamma(1 - U_{i\leftarrow j}(t)) + \beta P_S^j(0) P_S^{j\setminus i}(t) \tag{3.11}$$

So, given the initial conditions $U_{i\leftarrow j}(0)$ and $P_S^i(0)$, we numerically integrate this or (3.9) to compute $P_a^i(t)$, $P_S^i(t)$, and $P_T^i(t)$ using (3.1) and (3.3) respectively.

3.2 Message passing vs Monte Carlo simulation in real networks

The message passing formulation in Section 3.1 is exact only on trees, since we assumed that the probabilities $P_a^i(t)$ are independent. However, typical networks

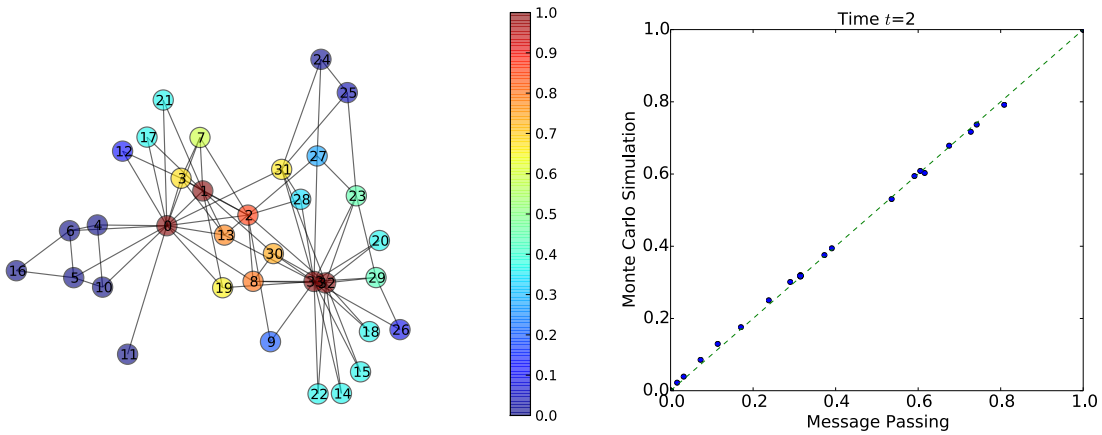

Figure 3.3: Same parameters and initial conditions as Fig. 3.2, except that we are comparing the infection probability at time $\mathbf{t} = 2$.

contain many loops. Thus, the independence assumption of the message passing approach is an approximation in real networks. Our goal in this section is to see how accurate DMP is in real networks by comparing it with Monte Carlo simulations of the actual stochastic process.

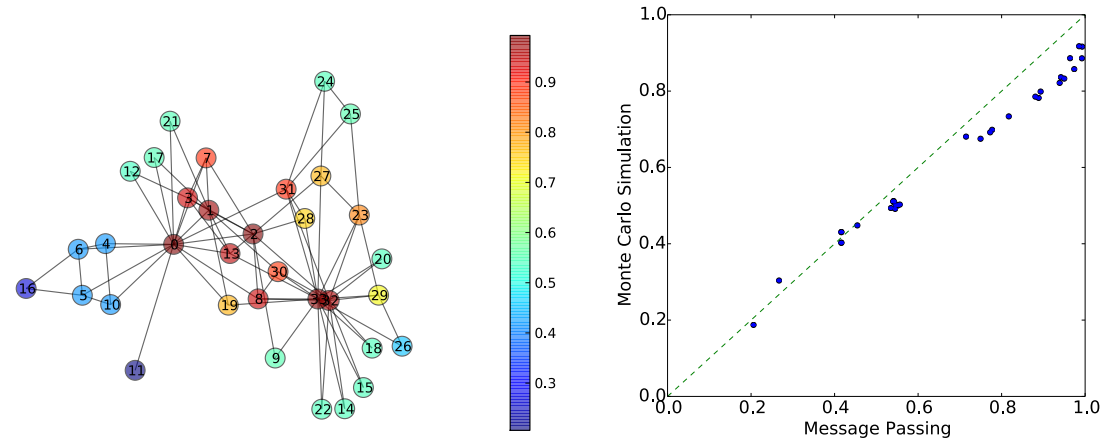

Figure 3.4: Same as Fig. 3.2, where we compare individuals probability of eventually getting infected. Here the initial condition is such that each is infected with probability 0.2.

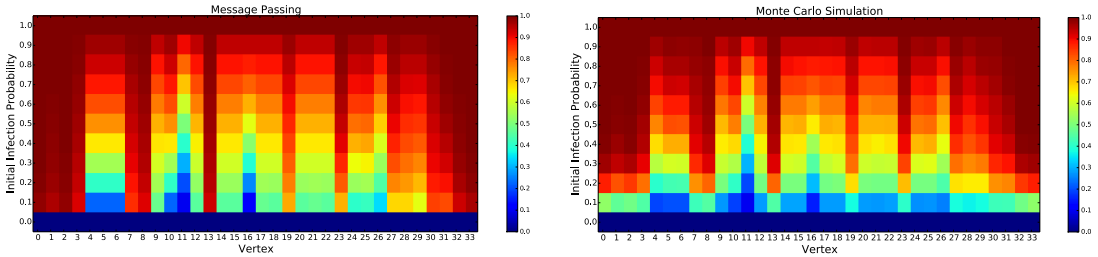

Figure 3.5: We show the eventual infection probability of each individual (horizontal axis) in the Zachary karate club network at increasing uniform probability (vertical axis) of getting infected initially. Here, threshold $T = 2$, transmission rate $\beta = 0.6$, and recovery rate $\gamma = 0.3$. On the left is the result calculated through the DMP. Whereas, on the right, we show the result from the Monte Carlo simulations, where the probabilities are averaged over 10^5 runs for each initial infection probability.

To compare the results between DMP and Monte Carlo simulations, we show the infection probability of each individual calculated through both methods in a scatter plot. In Fig. 3.2, we compare the eventual infection (adoption) probability of each individual in Zachary's karate club network. Each point in the scatter plot refers to the eventual infection probability of an individual in the club. If the DMP were exact, all points in the figure would lie exactly on the dotted diagonal line.

Here, each individual's threshold T is set to 2. Four vertices labeled $\{0, 1, 32, 33\}$ in Fig. 3.2 (left) are the initially infected individuals. We assume $f(\tau) = \beta e^{-(\gamma+\beta)\tau}$ with a transmission rate $\beta = 0.6$ and a recovery rate $\gamma = 0.3$. We simulate the actual stochastic process using a continuous-time Monte Carlo method algorithm. Events are maintained in a priority queue using a heap data structure to sort the events in the model: specifically, sort the edges (i, j) according to the time at which j will inform i. The probabilities are then averaged over 10^5 independent runs.

In Fig. 3.3, using the same parameters and initial conditions as Fig. 3.2, we compare the infection probability of each individual at a particular finite time $t = 2$. We chose this time because this is when the average number of infected individuals

is at its maximum.

In Fig. 3.4, we again use the same parameters as Fig. 3.2, but with different initial conditions. Each individual is initially infected with probability 0.2. There are now two sources of randomness in the model: the dynamics and the set of initial adopters. This again forces us to do many independent runs of the Monte Carlo simulation to estimate the infection probabilities. By setting $\mathsf{P}_{\mathsf{S}}^{\mathsf{i}}(0) = 0.8$ in Equation (3.3) however, we can calculate the infection probability with the same computational cost as before where the initial infectors were fixed. Accordingly in Fig. 3.5, we show the density plot of the probability that each individual (horizontal axis) is eventually infected, when each of them is initially infected with increasing probability (vertical axis).

3.2.1 A note on correlations for threshold models

Checking the scatter plot of the results computed from DMP and Monte Carlo simulation in Figures 3.2 - 3.4, we first see that the results computed from DMP do not match perfectly with those from the simulation. As pointed out in [42], where $\mathsf{T} = 1$ the probability estimated by DMP is always an upper bound on the true probability, since the events that two or more neighbors become infected are positively correlated.

However, for $\mathsf{T} > 1$ the situation is more complicated, and DMP does not necessarily give an upper bound on the infection probability. Indeed, in Figs. 3.2–3.4, we see several cases when DMP underestimates the infection probability rather than overestimating it. This includes the vertices labeled $\{26\}$ in Fig. 3.2, $\{12, 26, 27, 28\}$ in Fig. 3.3, and $\{5, 6, 16\}$ in Fig. 3.4.

To see why this happens, suppose i has two neighbors, j and k. Let $\mathsf{P}[\mathsf{i}]$ denote the probability that i becomes infected, and let $\mathsf{P}[\mathsf{j}]$ and $\mathsf{P}[\mathsf{k}]$ denote the probabilities

that j and k inform i respectively. If $\mathsf{T} = 1$, then

$$\mathsf{P}[\mathsf{i}] = \mathsf{P}[\mathsf{j} \vee \mathsf{k}] = \mathsf{P}[\mathsf{j}] + \mathsf{P}[\mathsf{k}] - \mathsf{P}[\mathsf{j} \wedge \mathsf{k}].$$

Let's assume that DMP computes the right marginals, so that $\mathsf{P}_{\mathrm{DMP}}[\mathsf{j}] = \mathsf{P}[\mathsf{j}]$ and $\mathsf{P}_{\mathrm{DMP}}[\mathsf{k}] = \mathsf{P}[\mathsf{k}]$. However, DMP ignores correlations, and assumes that these events are independent. Thus

$$\mathsf{P}_{\mathrm{DMP}}[\mathsf{i}] = \mathsf{P}[\mathsf{j}] + \mathsf{P}[\mathsf{k}] - \mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}].$$

However, j and k are positively correlated if they have a common neighbor that may have infected them both, or if they are neighbors of each other. That is,

$$\mathsf{P}[\mathsf{j} \wedge \mathsf{k}] > \mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}].$$

Then $\mathsf{P}[\mathsf{i}] < \mathsf{P}_{\mathrm{DMP}}[\mathsf{i}]$, and DMP overestimates $\mathsf{P}[\mathsf{i}]$. On the other hand, if $\mathsf{T} = 2$, then

$$\mathsf{P}[\mathsf{i}] = \mathsf{P}[\mathsf{j} \wedge \mathsf{k}] > \mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}] = \mathsf{P}_{\mathrm{DMP}}[\mathsf{i}],$$

and DMP underestimates $\mathsf{P}[\mathsf{i}]$.

Similarly, suppose i has three neighbors, j, k, and ℓ. Again taking $\mathsf{T} = 2$, we have

$$\mathsf{P}[\mathsf{i}] = \mathsf{P}[\mathsf{j} \wedge \mathsf{k}] + \mathsf{P}[\mathsf{j} \wedge \ell] + \mathsf{P}[\mathsf{k} \wedge \ell] - 2\mathsf{P}[\mathsf{j} \wedge \mathsf{k} \wedge \ell],$$

whereas, DMP gives

$$\mathsf{P}_{\mathrm{DMP}}[\mathsf{i}] = \mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}] + \mathsf{P}[\mathsf{j}]\mathsf{P}[\ell] + \mathsf{P}[\mathsf{k}]\mathsf{P}[\ell] - 2\mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}]\mathsf{P}[\ell].$$

In this case, DMP can either underestimate or overestimate $\mathsf{P}[\mathsf{i}]$, depending on the strength of the correlations between its neighbors. For example, if ℓ is independent of j and k, then

$$\begin{aligned}\mathsf{P}[\mathsf{i}] &= \mathsf{P}[\mathsf{j} \wedge \mathsf{k}] + \mathsf{P}[\mathsf{j}]\mathsf{P}[\ell] + \mathsf{P}[\mathsf{k}]\mathsf{P}[\ell] - 2\mathsf{P}[\mathsf{j} \wedge \mathsf{k}]\mathsf{P}[\ell] \\ &= \mathsf{P}[\mathsf{j} \wedge \mathsf{k}](1 - 2\mathsf{P}[\ell]) + (\mathsf{P}[\mathsf{j}] + \mathsf{P}[\mathsf{k}])\mathsf{P}[\ell].\end{aligned}$$

If j and k are positively correlated so that $\mathsf{P}[\mathsf{j} \wedge \mathsf{k}] > \mathsf{P}[\mathsf{j}]\mathsf{P}[\mathsf{k}]$, then DMP underestimates $\mathsf{P}[\mathsf{i}]$ if $\mathsf{P}[\ell] < 1/2$ and overestimates it if $\mathsf{P}[\ell] > 1/2$.

3.3 Exact solution in networks with arbitrary degree distributions

In this section, we consider the message passing approach in the ensemble of random networks in the thermodynamic limit. Our goal is to show that DMP can be applied to large random networks just as well as to a particular finite network.

In random networks, we are interested in the expected behavior of the dynamics rather than the dynamics in a single realization of the network. So, instead of computing messages for individual vertices, we assume that these messages are drawn from some probability distribution, and update this distribution based on their average behavior. We can then compute the distribution of marginals as well.

We consider random networks with a given degree distribution, specifically an ensemble of networks called the *configuration* model [53]. Each of n vertices is first assigned an integer degree from a specified degree distribution, say p_k. We think of a vertex with degree k as having k "spokes" or half-edges coming out of it. We then choose a uniformly random matching of these $2m$ spokes with each other, where m is the number of edges in the network. The key fact is then that, in the thermodynamic limit, i.e. $n \to \infty$, following an edge from any given vertex connects with a vertex of degree k with probability proportional to kp_k. Strictly speaking, this model generates random multigraphs. But, the average size of such graphs is a constant as $n \to \infty$, as a result of which the density of self-loops and multiple edges vanishes when n is large.

Now, consider the message $U_{i \leftarrow j}(t)$ from Equation (3.10). Recall that this is the probability that j has not informed i by time t. In the configuration model however, different individuals j are connected to i in different realizations of the network. But, edges are now statistically identical in the sense that each edge identically connects

to a vertex based on its degree. So, we consider a single average message $U(t)$.

This average message $U(t)$ then has the following interpretation. It is the average probability that by following a random edge, the neighbor we reach has not informed the vertex we came from by time t. This in turn will tell us the probability $P_a(t)$ that a randomly chosen vertex has awareness a at time t. However, this probability depends on the degree of the vertex: specifically, if it has degree k, then

$$P_a(k,t) = P_S(0)\binom{k}{a}U(t)^{k-a}(1-U(t))^a. \tag{3.12}$$

Averaging over p_k, we get

$$P_a(t) = P_S(0)\sum_k^\infty p_k\binom{k}{a}U(t)^{k-a}(1-U(t))^a. \tag{3.13}$$

It is useful to write this in terms of the generating function $G(x)$ of the degree distribution and its derivatives:

$$G(x) = \sum_k p_k x^k, \tag{3.14}$$

$$G^{(a)}(x) = \frac{d^a G(x)}{dx^a}. \tag{3.15}$$

Then $P_a(t)$ can be written as

$$P_a(t) = P_S(0)\frac{(1-U(t))^a}{a!}G^{(a)}(U(t)). \tag{3.16}$$

Thus the probability $P_S(t)$ that a randomly chosen vertex is susceptible at time t is

$$P_S(t) = \sum_{a=0}^{T-1} P_a(t). \tag{3.17}$$

Equivalently,

$$P_S(t) = P_S(0)\sum_{a=0}^{T-1}\frac{(1-U(t))^a}{a!}G^{(a)}(U(t)). \tag{3.18}$$

So, we see that given $U(t)$, computing $P_a(t)$ and $P_S(t)$ in the configuration model reduces to knowing $G^{(a)}$ to some order.

To capture the information flow that $U(t)$ represents in the configuration model, we define the cavity probability $Q(t)$ by simplifying Equation (3.4). This is the probability that a randomly chosen edge leads to a vertex that has not been infected by time t, if the vertex we came from is assumed to be absent from the network. Equivalently, $Q(t)$ is the probability that if we follow a random edge from a vertex i, the vertex j it leads to has been informed by at most $T-1$ of its neighbors other than i. This probability also depends on j's degree. Namely, if it has degree $k+1$, then

$$Q(k,t) = \sum_{a=0}^{T-1} \binom{k}{a} U(t)^{k-a}(1-U(t))^a, \tag{3.19}$$

where k is the number of neighbors that j has other than i. As discussed above, a random edge leads to a vertex with degree k with probability proportional to kp_k. Therefore, the probability that j has k neighbors other than i is

$$q_k = \frac{(k+1)p_{k+1}}{\sum_k k p_k} = \frac{(k+1)p_{k+1}}{G^{(1)}(1)}. \tag{3.20}$$

Averaging $Q(k,t)$ over q_k, we obtain

$$Q(t) = \sum_k q_k \sum_{a=0}^{T-1} \binom{k}{a} U(t)^{k-a}(1-U(t))^a. \tag{3.21}$$

Similar to Equation (3.18), we can write $Q(t)$ in terms of the generating function as

$$Q(t) = \frac{1}{G^{(1)}(1)} \sum_{a=0}^{T-1} \frac{(1-U(t))^a}{a!} G^{(a+1)}(U(t)). \tag{3.22}$$

We now calculate $U(t)$ by simplifying (i.e. averaging) Equation (3.10) for the configuration model. But, note the right-hand side of (3.10) consists of products of $U(t)$, and the average of products is not always the product of averages. In the limit $n \to \infty$ however, the network is locally treelike in the sense that the typical size of the shortest loops diverges as $O(\log n)$. As a result, $U(t)$ is asymptotically

independent, and the average of products is equal to the product of averages. So, the self-consistent relation for $U(t)$ becomes

$$U(t) = 1 - \int_0^t d\tau f(\tau) + P_S(0) \int_0^t dt' f(t-t')Q(t'). \tag{3.23}$$

To numerically integrate this equation in time, we differentiate it with respect to t,

$$\frac{dU(t)}{dt} = -f(t) + P_S(0)f(0)Q(t) + P_S(0) \int_0^t dt' Q(t') \frac{df(t-t')}{dt}. \tag{3.24}$$

It is also possible to get this from Equation (3.9). We can further simplify this to an ordinary differential equation in some cases. For example, if $f(\tau) = \beta e^{-(\beta+\gamma)\tau}$, we can write it as

$$\frac{dU(t)}{dt} = -\beta U(t) + \gamma(1 - U(t)) + \beta P_S(0)Q(t). \tag{3.25}$$

So, given the initial conditions $U(0) = 1, P_S(0)$, and $G^{(a)}(x)$, we can calculate $P_S(t)$ using Equation (3.18). Similarly, the fraction of infected and recovered vertices at time t can be calculated. Note that, in general, we can let $f(\tau)$ depend on the degree of the vertex by following a degree dependent transmission method formulated by Newman [16]. Similarly, we can allow for the case where the probability $P_T(0) = 1 - P_S(0)$ of getting initially infected depends on the degree of the vertex.

In Fig. 3.6 (left), we show the time evolution of the fraction of susceptible (blue), infected (red), and recovered (green) vertices in the configuration model, where the degrees are drawn from the Poisson distribution with mean c, or equivalently the Erdős-Rényi graphs $G(n, p = c/n)$. For Poisson distribution, $G^{(a)}(x)$ are given by $c^a e^{-c(1-x)}$. We take $c = 9$, $T = 3$, $f(\tau) = \beta e^{-(\beta+\gamma)\tau}$, where $\beta = 0.8$ and $\gamma = 0.2$, and the initial fraction of adopters/infecteds is $P_T(0) = 0.1$.

Continuous lines in Fig. 3.6 (left) are obtained by numerically integrating Equation (3.25), whereas dots are from Monte Carlo simulations with 10^4 vertices aver-

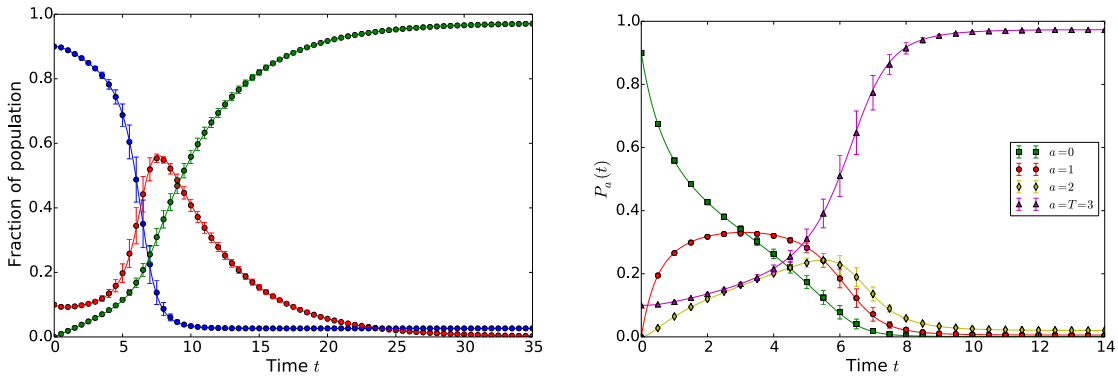

Figure 3.6: On the left is the dynamics in the Erdős-Rényi graphs $G(n, p = c/n)$ where individuals have threshold $T = 3$, average degree $c = 9$, initial fraction of adopters/infecteds $P_T(0) = 0.1$. The fractions of infected, recovered and susceptible vertices are red, green, blue respectively. Continuous lines are analytic results calculated using our DMP approach, by numerically integrating Equation (3.25), whereas dots are from the Monte Carlo based simulations with 10^4 vertices averaged over 100 runs. Transmission rate $\beta = 0.8$, and recovery rate $\gamma = 0.2$. On the right is the time evolution of $P_a(t)$, where continuous lines are calculated using Equation (3.16). Root Mean Square deviations in the simulation are provided when they are larger than the markers.

aged over 100 runs. Similarly, Fig. 3.6 (right) gives the fraction $P_a(t)$ of vertices with awareness a, where the continuous lines are obtained by using Equation (3.16).

In Fig. 3.7, we show the fraction $P_T(t)$ of adopters as a function of time for the same parameter values as Fig. 3.6, except where T is 1 (green square), 2 (blue circle), 3 (magneta triangle), and 4 (black diamond). Root Mean Square deviations in the simulation are provided when they are larger than the the markers.

Using the same framework, we can calculate the asymptotic probability $u = U(\infty)$ that the infection has not been transmitted along a random edge. This in turn will tell us the asymptotic probability that a randomly chosen vertex ever becomes infected.

We can think of the long time behavior as k-core percolation. Either the edge is

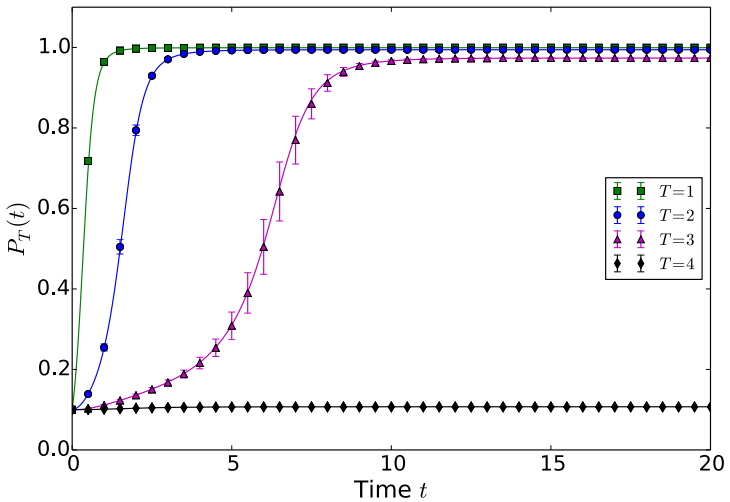

Figure 3.7: Same parameters and initial conditions as Fig. 3.6, except we are computing the fraction $P_T(t)$ of adopters, i.e. either infected or recovered vertices, as a function of time when the threshold T is 1 (green square), 2 (blue circle), 3 (magneta triangle), and 4 (black diamond).

closed in the sense that its other endpoint fails to inform the vertex we came from, which happens with the probability $1-p = 1-\int_0^\infty f(\tau)d\tau$. In this case, it does not matter if the neighbor gets infected by its other neighbors, since it fails to inform the vertex we came from. Or, it can be the case that the edge is open (with probability p), but the vertex we reach is itself not infected eventually by its other neighbors. This happens when the neighbor we reach by randomly following the edge is informed by at most $T-1$ other neighbors, provided it was not initially infected. Summing up both cases, we arrive at the following self-consistent relation for u:

$$\begin{aligned} u &= 1-p+pP_S(0)\sum_k^\infty q_k \sum_{a=0}^{T-1}\binom{k}{a}u^{k-a}(1-u)^a \\ &= 1-p+\frac{pP_S(0)}{G^{(1)}(1)}\sum_{a=0}^{T-1}\frac{(1-u)^a}{a!}G^{(a+1)}(u). \end{aligned} \tag{3.26}$$

Note that we could have written this equally by taking the limit $t\to\infty$ in Equation (3.23). Similarly, the probability P_S that a randomly chosen vertex never gets

infected, i.e. the fraction of susceptible vertices is

$$P_S = P_S(0) \sum_{a=0}^{T-1} \frac{(1-u)^a}{a!} G^{(a)}(u). \tag{3.27}$$

For Erdős-Rényi networks $G(n, p = c/n)$, or equivalently the Poisson distribution with average degree c, we have the following self-consistent relation for u:

$$u = 1 - p + pP_S(0)e^{-c(1-u)} \sum_{a=0}^{T-1} \frac{c^a(1-u)^a}{a!}. \tag{3.28}$$

We can also obtain this expression by following [52]. Similarly, P_S in Erdős-Rényi networks is

$$P_S = P_S(0)e^{-c(1-u)} \sum_{a=0}^{T-1} \frac{c^a(1-u)^a}{a!}. \tag{3.29}$$

Equations (3.26) and (3.27) have a nice interpretation in terms of well-studied problems in random graphs, including percolation and the emergence of the k-core. We say that Equation (3.26) is the generating function in $P_S(0)$ of the size of the connected component of susceptible vertices by following a random edge in the long time limit. Similarly, Equation (3.27) is the generating function of the size of the connected susceptible component of a randomly chosen vertex.

Chapter 4

Message-passing for recurrent epidemics

The work described in this chapter is a result of collaboration with Samuel V. Scarpino and Cristopher Moore is submitted for publication in Phys. Rev. E.

For many epidemic models, such as SI (susceptible-infectious), SIR (susceptible-infectious-recovered) and SEIR (susceptible-exposed-infectious-recovered), only one-way state changes can occur. For example, in the SIR model, once an individual has left the Susceptible class and become Infectious, they cannot return to being Susceptible; once they become Recovered, they are immune to future infections, and might as well be Removed. For these non-recurrent models, DMP is known to be be an efficient algorithm to estimate $\mathtt{I}_{\mathtt{i}}(\mathtt{t})$, and it is exact on trees [26]; it can also be applied to threshold models [27, 29, 30] and used for inference [19].

However, for many real-world diseases individuals can return to previously inhabited states. In these *recurrent* models, such as SIS (susceptible-infectious-susceptible) , SIRS (susceptible-infectious-recovered-susceptible), and SEIS (susceptible-exposed-infectious-susceptible), individuals can cycle through the states multiple times, giving

multiple waves of infection traveling through the population. The most obvious examples of recurrent models are seasonal influenza, where due to the evolution of the virus individuals are repeatedly infected during their lifetime [70], vaccination where protective immunity wanes over time [71], and diseases curable by treatment which does not result in antibody-mediated immunity, such as gonorrhea [76]. In all three cases, individuals leave the Susceptible class, only to return at some point in the future (although for influenza, it is worth mentioning that if the evolutionary rate of the virus is functionally related to the number of susceptible individuals, then the recovery rate may not be independent from the state of one's neighbors.) Unfortunately, the DMP approach of [26] cannot be directly extended to recurrent models, since their equations for messages only track the first time an individual makes the transition to a given state.

The purpose of this chapter is to develop a novel DMP algorithm for recurrent models of epidemics on networks, which we call rDMP. We will show that rDMP gives very good approximations for marginal probabilities on networks, and is often more accurate than the pair approximation. Moreover, whereas the pair approximation requires keeping track of $\mathtt{m}\mathtt{k}^2$ variables, if there are $\mathtt{m}$ edges and $\mathtt{k}$ states per node, rDMP requires just $2\mathtt{m}\mathtt{k}$ variables. For complex models where $\mathtt{k}$ is large—for instance, for diseases with multiple stages of infection or immunity, or multiple-disease epidemics where one disease makes individuals more susceptible to another one—this gives a substantial reduction in the computational effort required. Finally, the rDMP approach is conceptually simple, making it easy to write down the system of differential equations for a wide variety of epidemic models.

4.1 Message-passing and preventing the echo chamber effect

As shown in Fig. 2.2, the variables of rDMP are messages along directed edges of the network (in addition to one-point marginals). For instance, $I_{j\to i}$ is the probability that j is Infectious because it was infected by one of its other neighbors k. The intuition behind this is the following, where we take the SIS model as an example. If i is Susceptible, the rate at which j will infect i is proportional to the probability I_j that j is infected. But when computing this rate, we only include the contribution to I_j that comes from neighbors other than i. In other words, we deliberately neglect the event that j receives the infection from i, and immediately passes it back to i, even if i has become Susceptible in the intervening time.

This choice avoids a kind of "echo chamber" effect, where neighboring nodes artificially amplify each others' probability of being Infectious which we discussed in chapter 2.1.

In other words, consider again the two-node case as shown in Fig. 2.1. Here, with DMP, we fix the effect of "echo chamber" by replacing I_i and I_j with the messages they send each other,

$$\frac{dI_i}{dt} = \lambda S_i I_{j\to i}\,,$$
$$\frac{dI_j}{dt} = \lambda S_j I_{i\to j}\,,$$

so that i can only infect j if i received the infection from some node other than j. (Below we give the equations on a general network, including the time derivatives of the messages.) In this example, there are no other nodes, so if $I_{j\to i}(0) = \delta$ and $I_{i\to j}(0) = 0$, then $I_j(t) = \delta$ for all t as it should be.

Note that we do not claim that rDMP is exact in this case. In particular, as

in (2.4), $I_i(t)$ tends to 1 as $t \to \infty$. This is because, unlike the system of [26], rDMP assumes that the events that j infects i at different times are independent.

Preventing backtracking completely may seem like a strong assumption, and in recurrent models it is *a priori* possible, for instance, for a node to re-infect the neighbor it was infected by. Despite the well-documented importance of recurrent infections for diseases including (but certainly not limited to) seasonal influenza [70], Plasmodium malaria [78], and urinary tract infections [77], little is known about the source of recurrent infections. For certain sexually transmitted diseases such as gonorrhea [76] and repeated ringworm infections [79], there is evidence that backtracking plays a significant role; on the other hand, it may be that recurrent infections are caused by different strains, each of which is acting essentially without backtracking. Thus while our non-backtracking assumption is clearly invalid in some cases, we believe it is a reasonable approach for most recurrent state infections.

4.2 The rDMP equations for the SIS, SIRS, and SEIS models

In this section, we illustrate the rDMP approach for several recurrent epidemic models. We start with the simplest one: in the SIS model, each node is either Infectious (I) or Susceptible (S). Infectious nodes infect their Susceptible neighbors at rate λ, and their infections wane back into the Susceptible state at rate ρ. We denote the probability that that node i is Infectious or Susceptible by I_i and S_i respectively. The objective then is to efficiently and accurately compute these probabilities as a function of time t.

We define variables or "messages" that live on the directed edges (i, j) of the network. The directed nature of these messages prevent infection from backtracking

from an Infectious node back to its infection source, e.g., if node i infects node j, then we prevent j from re-infecting i. In addition to tracking the one-point marginal I_j, we define a message $I_{j\to i}$ from j to i as the probability that j is in the Infectious state as a result of being infected from one of its neighbors other than i. Given these incoming messages, the rate at which I_i evolves in time is given by

$$\frac{dI_i}{dt} = -\rho I_i + \lambda S_i \sum_{j\in\partial i} I_{j\to i}, \tag{4.1}$$

where ∂i denotes the neighbors of i. Similarly, the rate at which $I_{j\to i}$ evolves in time is given by

$$\frac{dI_{j\to i}}{dt} = -\rho I_{j\to i} + \lambda S_j \sum_{k\in\partial j\setminus i} I_{k\to j}, \tag{4.2}$$

where $k \in \partial j \setminus i$ denotes the neighbors of j excluding i.

For the SIRS model, we let ρ and γ denote the transition rates from Infectious to Recovered and from Recovered to Susceptible respectively. Then the rDMP system for the SIRS model is given by

$$\frac{dI_{j\to i}}{dt} = -\rho I_{j\to i} + \lambda S_j \sum_{k\in\partial j\setminus i} I_{k\to j}, \tag{4.3}$$

which is coupled with the one-point marginals through

$$\begin{aligned}
\frac{dS_i}{dt} &= \gamma R_i - \lambda S_i \sum_{j\in\partial i} I_{j\to i} \\
\frac{dI_i}{dt} &= -\rho I_i + \lambda S_i \sum_{j\in\partial i} I_{j\to i} \\
\frac{dR_i}{dt} &= \rho I_i - \gamma R_i \,.
\end{aligned} \tag{4.4}$$

In the SEIS model, upon becoming exposed to an infected neighbor, Susceptible nodes first go through a latent period called the Exposed state. In this state, individuals are infected but not yet Infectious. Exposed nodes become Infectious at the

rate ε, and Infectious nodes again wane back to Susceptible at rate ρ. The rDMP system for the SEIS model is

$$\begin{aligned}
\frac{dE_{j\to i}}{dt} &= -\varepsilon E_{j\to i} + \lambda S_j \sum_{k\in\partial j\setminus i} I_{k\to j},\\
\frac{dI_{j\to i}}{dt} &= -\rho I_{j\to i} + \varepsilon E_{j\to i}\,,
\end{aligned} \tag{4.5}$$

which is coupled with the one-point marginals as

$$\begin{aligned}
\frac{dS_i}{dt} &= \rho I_i - \lambda S_i \sum_{j\in\partial i} I_{j\to i}\\
\frac{dI_i}{dt} &= -\rho I_i + \varepsilon E_i\\
\frac{dE_i}{dt} &= -\varepsilon I_i + \lambda S_i \sum_{j\in\partial i} I_{j\to i}\,.
\end{aligned} \tag{4.6}$$

Note that here we track messages for the Exposed state, in addition to one-point marginals, since they act as precursors for the Infectious messages. There is no need to track messages for the Susceptible state, since it does not cause state changes in its neighbors.

Generalizing these equations to more complex epidemic models with k different states, as opposed to three or four, is straightforward. Even in a model where every state can cause state changes in its neighbors—for instance, where having Susceptible neighbors speeds up the rate of recovery, or where Exposed nodes can also infect their neighbors at a lower rate—the total number of variables we need to track in a network with n nodes and m edges is at most $2mk$ in addition to the nk one-point marginals. In contrast, the pair approximation requires mk^2 states to keep track of the joint distribution of every neighboring pair.

4.3 Experiments in real and synthetic networks

In this section we report on numerical experiments for rDMP for the SIS and SIRS models on real and synthetic networks. As a performance metric, we use the average L_1 error per node between the marginals computed from rDMP and the true probabilities computed (up to sampling error) using continuous-time Monte Carlo simulations. That is,

$$L_1^{\mathrm{rDMP}}(t) = \frac{1}{n} \sum_i \left| I_i^{\mathrm{MC}}(t) - I_i^{\mathrm{rDMP}}(t) \right| , \tag{4.7}$$

We use this metric to compare the performance of rDMP with the independent-node approximation and the pair approximation, or equivalently first- and second-order moment closure [18, 58]. As we will see, for a wide range of parameters, rDMP is more accurate than either of these approaches, even though it is computationally easier than the pair approximation.

In Fig. 4.1, we show results for the SIS model on Zachary's Karate Club [63]. On the left, we show the marginal probability that a particular node is Infectious as a function of time, estimated by rDMP and by first- and second-order moment closure, and compared with the true marginals given by Monte Carlo simulation. On the right, we show the average L_1 error for the three methods. Here $\lambda = 0.1$, $\rho = 0.05$, and the initial condition consists of a single infected node (shown in red in the inset). The Monte Carlo results were averaged over 10^5 runs. We see that rDMP is significantly more accurate than the other two, except at some early times when the pair approximation marginally outperforms rDMP.

As a further illustration, in Fig. 4.2 we show the steady-state marginal I_i for each node i (measured by running the system until $t = 50$, at which point $I_i(t)$ is nearly constant), with the same parameters and initial condition as in Fig. 4.1. We show the true marginal of each node on the y-axis, and the marginals estimated by rDMP and the pair approximation on the x-axis. If the estimated marginals were perfectly

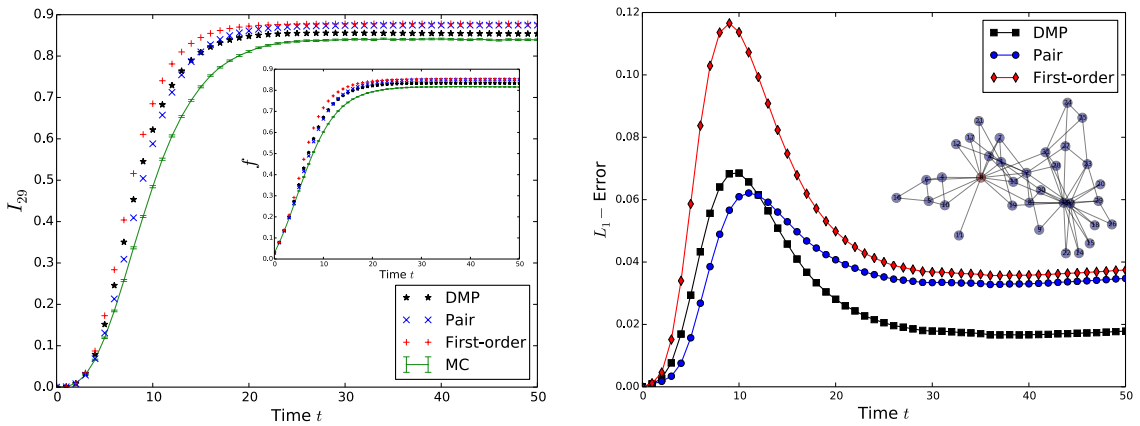

Figure 4.1: Results on the SIS model. On the left, the marginal probability that node 29 in Zachary's Karate club (see inset on right) is Infectious as a function of time. We compare the true marginal derived by 10^5 independent Monte Carlo simulations with that estimated by rDMP, the independent node approximation, and the pair approximation. In the inset, we show the fraction f of Infectious nodes as a function of time. On the right is the L_1 error, averaged over all nodes; we see that rDMP is the most accurate of the three methods. Here the transmission rate is $\lambda = 0.1$, the waning rate is $\rho = 0.05$, and vertex 0 (colored red) was initially infected.

accurate, the points would fall on the line $y = x$. Both methods overestimate the marginals to some extent, but rDMP is more accurate than the pair approximation on every node. Thus rDMP makes accurate estimates of the marginals on individual nodes, as opposed to just the average across the population.

To investigate how rDMP compares with the pair approximation across a broader range of parameters, in Fig. 4.3 we vary the ratio between waning rate ρ and the transmission rate λ. Since we can always rescale time by multiplying λ and ρ by the same constant, we do this by holding $\lambda = 0.1$ as before, and varying ρ. We then measure the difference in the L_1 error of the two methods, $L_1^{\text{rDMP}} - L_1^{\text{pair}}$.

In the blue region, rDMP is more accurate than the pair approximation; in the red region, it is less so. We see that rDMP is more accurate except at early times (as in Fig. 4.1) or when ρ is small compared to λ, i.e., if the model is close to the SI

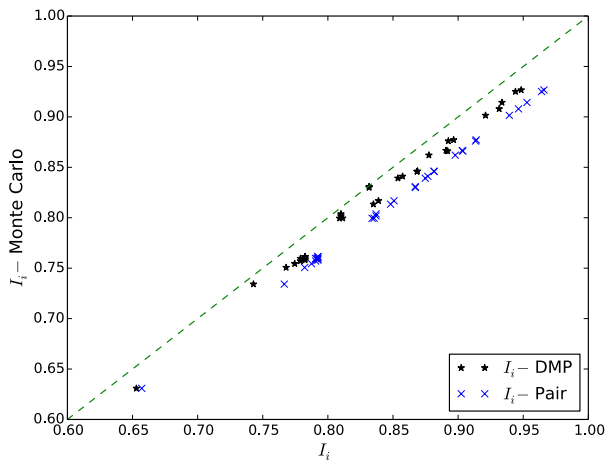

Figure 4.2: A scatterplot of the steady-state marginals I_i for the $n = 33$ nodes in Zachary's Karate Club, with the same parameters as in Fig. 4.1. The vertical axis is the true marginal computed by Monte Carlo simulations; the horizontal axis is the estimated marginals from rDMP (black $\star$) and the pair approximation (blue $\times$). Both methods overestimate the marginal, but rDMP is closer to the true value (the line $y = x$) for every node.

model where Infectious nodes rarely become Susceptible again.

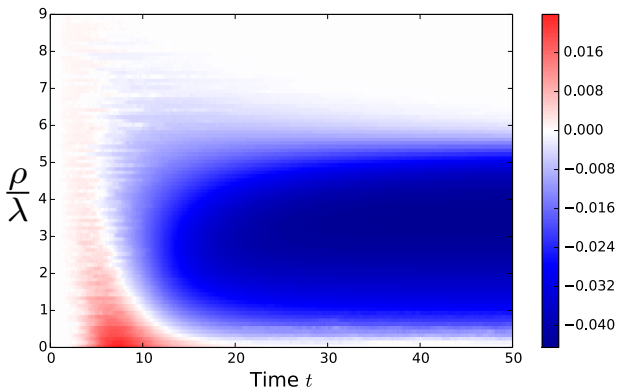

Figure 4.3: The difference between L_1^{rDMP} and L_1^{pair} on Zachary's Karate Club for various values of the ratio ρ/λ. We rescale time so that $\lambda = 0.1$ as before. In the blue region, $L_1^{\mathrm{rDMP}} < L_1^{\mathrm{pair}}$ and rDMP is more accurate; in the red region, $L_1^{\mathrm{rDMP}} > L_1^{\mathrm{pair}}$. We see that rDMP is more accurate except at early times or when ρ/λ is small.

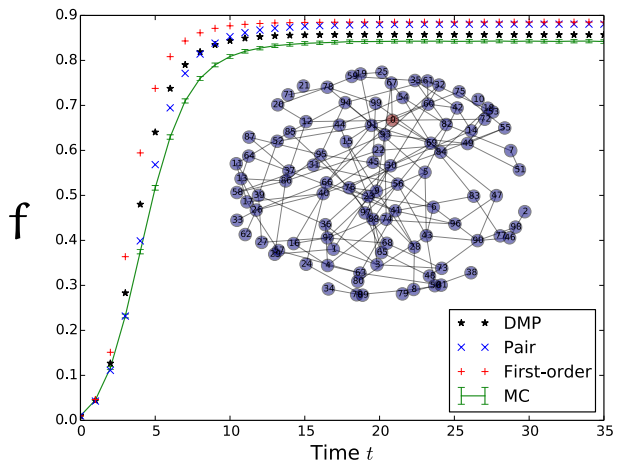

Figure 4.4: The fraction f of Infectious nodes as a function of time in the SIS model on an Erdős-Rényi graph (inset) with $n = 100$ and average degree 3. Here $\lambda = 0.4$, $\rho = 0.1$, and the initial condition consists of a single Infectious node (colored red). Monte Carlo results were averaged over 10^3 independent runs. Except at early times, rDMP tracks the true trajectory more closely.

In Fig. 4.4, we simulate the SIS model on an Erdős-Rényi graph with $n = 100$ and average degree 3, with $\lambda = 0.4$, $\rho = 0.1$, and a single initially Infectious node. As with the Karate Club, rDMP does a better job of tracking the true fraction of Infectious nodes, except at early times when the pair approximation is superior; in particular, it does a better job of computing the steady-state size of the epidemic.

In Fig. 4.5 we show results for the SIRS model on Zachary's Karate Club. As in Fig. 4.1, on the left we show the marginal probability I_{29} that node 29 is Infectious; on the right, we show the L_1 error for I_i averaged over the network. In the insets, we show the marginal probability R_{29} for the Recovered state and the corresponding average L_1 error. Here the transmission rate is $\lambda = 0.1$, the waning rate from Infectious to Recovered is $\rho = 0.05$, and the rate from Recovered to Susceptible is $\gamma = 0.2$. The initial condition consisted of a single infected node, and Monte Carlo results were averaged over 10^5 runs. As for the SIS model, rDMP is significantly more accurate than the independent node approximation, and is more accurate than

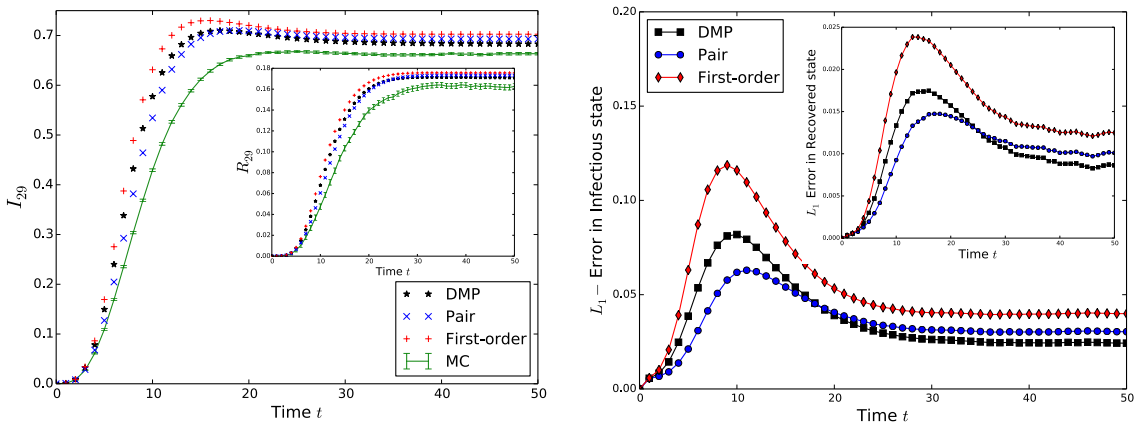

Figure 4.5: The SIRS model on the Karate Club. On the left, we show the true and estimated marginal probability that a node 29 is Infectious (main figure) or Recovered (inset) as a function of time. On the right is the average L_1 error for the Infectious and Marginal states. The transmission rate is $\lambda = 0.1$, and the transition rates from Infectious to Recovered and from Recovered to Susceptible are $\rho = 0.05$ and $\gamma = 0.2$ respectively. Node 0 (colored red) was initially infected. Monte Carlo results were averaged over 10^5 runs. As for the SIS model, rDMP is significantly more accurate than the first-order model where nodes are independent, and is more accurate than the pair approximation except at early times.

the pair approximation except at early times.

We found similar results on many other families of networks, including random regular graphs, random geometric graphs, scale-free networks, Newman-Watts-Strogatz small world networks, and a social network of dolphins [59]. Namely, rDMP outperforms the first-order approximation where nodes are independent, and outperforms the pair approximation across a wide range of parameters and times.

4.4 Linear stability, epidemic thresholds, and related work

Systems of differential equations for rDMP, such as (4.2), do not appear to have a closed analytic form due to their nonlinearities. On the other hand, we can compute quantities such as epidemic thresholds by linearizing around a stationary point, such as $\{I^*_{j\to i} = 0\}$ where the initial outbreak is small. Given a perturbation $\epsilon_{j\to i} = I_{j\to i} - I^*_{j\to i}$, the linear stability of the system, i.e., whether or not $\epsilon_{j\to i}$ diverges in time, is governed by the eigenvalues of the Jacobian matrix $\mathbf{J}$ of the right hand side of (4.2) at the stationary point I^*_i. The Jacobian for (4.2) at $\{I^*_{j\to i}\}$ is

$$\mathbf{J}_{(j\to i),(k\to j')} = -\delta_{kj}\delta_{ij'}\rho + \lambda(1 - I^*_j)\mathbf{B}_{(j\to i),(k\to j')}\,. \tag{4.8}$$

where

$$\mathbf{B}_{(j\to i),(k\to j')} = \delta_{jj'}(1 - \delta_{ik})\,. \tag{4.9}$$

This definition of $\mathbf{B}$ is another way of saying that the edge $k \to j$ influences edges $j \to i$ for $i \neq k$, but does not backtrack to k. This corresponds to our assumption that infections, for instance, do not bounce from k to j and back again and create an echo chamber effect. For this reason, $\mathbf{B}$ is also known in the literature as the non-backtracking matrix [65] or the Hashimoto matrix [61].

Now, for a small perturbation $\vec{\epsilon}$ away from a stationary point $\{I^*_{j\to i}\}$, the linearized system of (4.2) becomes

$$\frac{d\vec{\epsilon}}{dt} = \mathbf{J}\vec{\epsilon}, \tag{4.10}$$

If $\mathbf{J}$ has any eigenvalues with positive real part, then $\|\vec{\epsilon}(t)\|$ grows exponentially in time. So, the fixed point $\{I_{j\to i}\}$ is stable as long as the leading eigenvalue J_1 of $\mathbf{J}$ has negative real part.

One trivial, but important, stationary point to test is $I^*_{j\to i} = 0$ for all edges. A small perturbation around $\vec{0}$ corresponds to a small initial probability that each node is infected. From (4.8), $\mathbf{J}$ becomes

$$\mathbf{J} = \lambda \left(\mathbf{B} - \frac{\rho}{\lambda} \mathbb{1} \right) , \tag{4.11}$$

where $\mathbb{1}$ is the $2m \times 2m$ identity matrix. So, the leading eigenvalue of $\mathbf{J}$ becomes positive when the largest eigenvalue B_1 of $\mathbf{B}$ is greater than ρ/λ. In other words, if

$$R_0 = \frac{\lambda}{\rho} B_1 \geqslant 1 , \tag{4.12}$$

where R_0 is the reproductive number, even a small initial probability of infection will lead to a widespread endemic state, where the infection becomes extensive. If (4.12) does not hold, a small initial probability of infection will instead decay back to an infection-less state.

Since $\mathbf{B}$ is not symmetric, not all its eigenvalues are real. However, by the Perron-Frobenius theorem, it's leading eigenvalue is real; moreover, it is upper bounded by A_1, the leading eigenvalue of the adjacency matrix $\mathbf{A}$. Interestingly, if we examine the linear stability of the first-order approximation where nodes are independent, [18], the epidemic threshold for the SIS model is given by

$$\frac{\lambda}{\rho} A_1 \geqslant 1 . \tag{4.13}$$

Since $B_1 \leqslant A_1$, the threshold (4.12) gives a better upper bound for the true epidemic threshold than we would get from the first-order approximation. A similar threshold for the SIR model in sparse networks, or equivalently for percolation, using B_1 was recently demonstrated in [66]. (We note that when backtracking is allowed, it has important consequences for epidemic thresholds on power-law networks [67].)

Whereas the leading eigenvector of $\mathbf{B}$ governs the epidemic threshold, the spectral gap between $\mathbf{B}$'s top two eigenvectors governs how quickly the epidemic converges to the leading behavior (at least until we leave the linear regime). Qualitatively, this

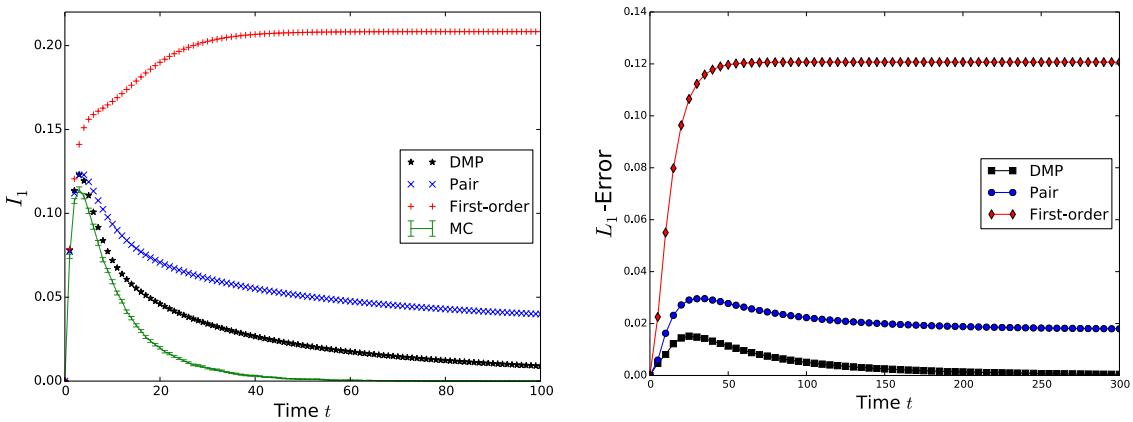

Figure 4.6: Same as in Fig. 4.1, but with transmission rate $\lambda = 0.1$ and waning rate $\rho = 0.54$. A well known upper bound on the epidemic threshold of the SIS model can be computed from the leading eigenvalue Λ_1 of the adjacency matrix (the Jacobian matrix of first-moment-closure approach) of a network. In other words, if $\frac{\rho}{\lambda} < \Lambda_1$, it is known from the first-moment-method that an infection-free state becomes unstable and epidemics become widespread and endemic. Here we show the results from SIS model in Zachary's Karate Club, where $\Lambda_1 \approx 6.7$. Even though $\frac{\rho}{\lambda} = 5.4 < \Lambda_1$ which is well below the threshold from the first-moment method, the contagion fades away eventually, which is correctly captured by our DMP approach.

depends on bottlenecks in the network such as those due to community structure, where an epidemic spreads quickly in one community but then takes a longer time to cross over into another. Indeed, the second eigenvector of the non-backtracking matrix **B** was recently used to detect community structure [65].

Similarly, just as the leading eigenvector of **B** was recently shown to be a good measure of importance or "centrality" of a node [69], it may be helpful in identifying "superspreaders"—nodes where an initial infection will generate the largest outbreak, and be the most likely to lead to a widespread epidemic.

Chapter 5

Stability analysis of financial contagion due to overlapping portfolios

The work described in this chapter is a result of collaboration with Fabio Caccioli (first-author), Cristopher Moore, J. Doyne Farmer and is published in Journal of Banking & Finance Volume 46, (2014).

5.1 Introduction

The 2007–2009 financial crisis highlighted the complex interconnections between financial institutions and made it clear that we need a better understanding of how financial contagion propagates and the circumstances under which it is amplified[80, 10, 81, 82, 83, 84, 85, 86, 87]. Financial contagion comes through different channels, including *(i)* counterparty risk, *(ii)* roll-over risk, and *(iii)* common asset holdings, i.e. *overlapping portfolios*. Of these the first two have so far received the most atten-

tion, even though the primary problem is believed by many to have been due to the third. Our goal in this chapter is to remedy this by gaining a better understanding of the problem of overlapping portfolios. To do this we develop a method of computing the stability of financial networks under contagion due to overlapping portfolios. To understand the factors that determine network stability, we develop and study a stylized model, and suggest how it can be extended to be more realistic. This model can be regarded as a multiple asset extension of the single asset model developed in reference [88][1].

Inter-institutional lending drives the problem of counterparty and roll-over risk. Counterparty risk occurs when a bankrupt institution is unable to pay its debts and consequently causes other institutions to fail [89]. Roll-over risk occurs when financial institutions depend on short term lending for liquidity and their creditors stop lending because they fail or are under stress, so that they are no longer able to borrow and consequently fail or become under stress [83]. These have now been extensively studied and we are rapidly developing better insight into the circumstances where interbank lending causes problems (see for instance [80, 84, 89]).

Financial contagion due to overlapping portfolios is driven by common asset holdings [84, 90]. In the event that an asset price fluctuation causes an institution to fail, the resulting "fire sale" of assets by that institution further depresses prices, which in turn may cause other institutions to fail, causing a spiral of selling and further asset price decreases. This also induces correlations between different assets that further exacerbate the problem [91].

The problem of overlapping portfolios is very general. It occurs even without inter-institutional lending, and applies to any institutions that manage money. Although this can occur even without leverage[2], the use of leverage makes it particularly

[1] Reference [88] considered the properties of leveraged single asset portfolios. It was shown that under deleveraging market impact can cause bankruptcy if leverage is too large.

[2] In this chapter we assume that institutions sell assets only when they become

acute. We are particularly interested in the banking system, where it is not uncommon for investments to be leveraged by a factor of 30 or more, but our analysis applies equally well to hedge funds or any other financial institutions that make leveraged investments. For convenience we will use the word *bank* to refer to institutions in general, but the reader should bear in mind that our model applies equally well to any leveraged financial institution.

The problem of overlapping portfolios has previously been considered in references [80, 84, 92, 93][3]. In these papers, however, liquidation effects were considered on top of counterparty or roll-over risk. Here we are interested in the situation in which shocks can propagate between different financial institutions through a pattern of local portfolio overlaps (e.g. bank i has assets in common with bank j, that has other assets in common with bank k, etc.). The model is simple: we assume that banks own a portfolio of assets, that when a bank goes bankrupt due to a loss in the value of its portfolio it sells its assets, and that this in turn causes these assets to be devalued according to a simple market impact function relating the size of the sale to the change in price.

The model we consider is purely mechanistic, i.e. we do not attempt to describe decision-making processes by banks. The underlying assumption is that, during the development of a crisis, banks do not have time to deleverage or rebalance their portfolios before failing. Thus we consider portfolios fixed until default occurs, and assume that they are fully liquidated when it occurs. We then perform a macroprudential stress test by applying localized shocks affecting either a single bank or a single asset. After the initial shock is applied we test to see whether it causes any

bankrupt, in which case the problem of financial contagion occurs only when leverage is used. In general asset sales may be triggered by losses that are less severe, for example if investment funds are forced to liquidate even when they are solvent, as occurred during the stat-arb meltdown in 2007.

[3]After completing these results we received reference [94], whose independent results are complementary to ours.

bank failures; if so we iterate the process as needed until either there are no more failures or all banks have failed. The only trades during the course of the dynamics are fire sales of the assets of insolvent banks.

Our focus in this chapter is therefore in understanding the specific role of market impact and portfolio overlaps as a contagion mechanism between leveraged financial institutions. To this end, we consider a network of banks and assets, and we test how the average level of diversification in bank portfolios, the ratio of the number of banks to the number of assets (crowding), and the leverage attained by banks impact the stability of the system with respect to an initial shock affecting a single asset or bank.

The stability of the system will be measured in terms of the probability of observing a global cascade of failures, with a smaller probability being associated with a higher stability. A global cascade of failures, in this context, refers to the failure of a significant fraction of the banks: that is, a non-zero fraction in the limit of infinite network size. By mapping our model onto a generalized branching process, we show analytically that there is a region in parameter space where global cascades of failures occur. One advantage of this mechanistic approach is that it can in principle be calibrated against real data and used to perform stress tests on real financial systems.

We find that, as the diversification of the banks' portfolios increases, the system undergoes two phase transitions, with a region in between where global cascades occur. Below the first transition, banks are not interconnected enough for shocks to propagate in the network. Above the second transition, banks are robust to devaluations in a few of their assets. In between these two transitions, banks are both vulnerable to shocks in their asset prices, and interconnected enough for these shocks to spread. We also find that more leverage increases the overall instability of the network and that the system exhibits a "robust yet fragile" behavior, with regions of parameter space where contagion is rare but the whole system is brought

down whenever it occurs.

The chapter is organized as follows. In the next section we introduce the model. In Section 5.3 we map the model into a generalized branching process and present the analytical approach that allows us to identify the region of phase space where global cascades occur. In Section 5.4 we report results from numerical simulations exploring how stability of banking systems depends on parameters and network properties. In section 5 we compare the results of numerical simulations to those of stability analysis, and we present our conclusions in the last section.

5.2 The model

5.2.1 Banks, assets, and cascades of bankruptcies

We consider a representation of a financial system given in terms of a network of N banks and M assets. Whenever a bank invests in an asset, we draw a link in the network connecting that bank to that asset. The resulting network is bipartite (see Figure 5.1), meaning that there are two groups of nodes (banks and assets) and that there are links only between these two groups.

The number of assets in the portfolio of bank i, i.e. the number of links of the corresponding node, is its degree k_i. The *average diversification*, i.e. the average degree of banks in the network, is then

$$\mu_b = \frac{1}{N}\sum_{i=1}^{N} k_i \,, \tag{5.1}$$

where the sum runs over all N banks. Conversely, the number of banks that hold asset j in their portfolio is its degree ℓ_j, and the average degree of the assets is

$$\mu_a = \frac{1}{M}\sum_{j=1}^{M} \ell_j \,. \tag{5.2}$$

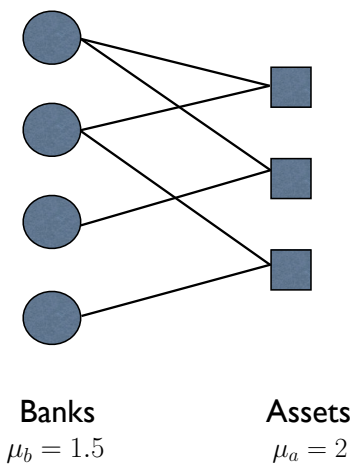

Figure 5.1: Graphical representation of the bipartite network of banks and assets. Banks are denoted by circles, assets by squares. Links connect banks to the assets they have in their portfolios. In this example $N = 4$, $M = 3$, the average banks' degree is $\mu_b = 1.5$ and the average assets' degree is $\mu_a = 2$.

Since each link connects a bank to an asset, the total degree of the banks must equal the total degree of the assets, so

$$\mu_b N = \mu_a M \,. \tag{5.3}$$

Although a complete description of the network's topology would require more information, a rough characterization can be given in terms of two parameters, μ_b and $n = N/M$. The *crowding parameter* n is a measure of the density of institutions choosing their investments from the same pool of assets.

Each solvent bank i holds a portfolio $\{Q_{i,1}, \ldots, Q_{i,M}\}$. Its value at time t is

$$A_i^t = \sum_{j=1}^{M} Q_{ij} p_j^t \,,$$

where Q_{ij} is the number of shares of asset j held by bank i and p_j^t the price of asset j at time t. In our dynamics a bank holds on its portfolio as long as it is solvent, so

Q_{ij} is independent of time. Notice that, given that bank i invests in k_i assets, only k_i of the M portfolio weights Q_{ij} will be non-zero for bank i.

Each solvent bank also holds cash C_i, and we denote by L_i its total liabilities; neither of these quantities depend on time. If A_i^0 is the initial value of bank i's portfolio, its initial equity (or capital) is therefore $E_i^0 = A_i^0 + C_i - L_i$. The leverage of a bank is the ratio between the amount of risky assets on its balance sheet and its equity. Assuming no risk associated with cash holdings, the initial leverage of bank i is $\lambda_i = A_i^0/E_i^0$.

The condition for bank i to be solvent at time t is

$$\sum_{j=1}^{M} Q_{ij} p_j^t + C_i \geqslant L_i \,. \tag{5.4}$$

Given that $E_i^0 = A_i^0 + C_i - L_i$, the above condition can be expressed as

$$A_i^0 - \sum_{j=1}^{M} Q_{ij} p_j^t \leqslant E_i^0 \,. \tag{5.5}$$

The left hand side represents the loss with respect to the initial investment. If such a loss happens to be greater than the initial capital of the bank, the bank is out of business.

Note that leverage is a necessary condition for banks to fail. A bank investing only its own capital always satisfies condition (5.5), since its maximal loss is equal to its equity. We can write (5.5) as a condition on the leverage,

$$\lambda_i \leqslant \frac{\sum_{j=1}^{M} Q_{ij} p_j^t}{E_i^0} + 1 \,. \tag{5.6}$$

Even in the worst case scenario where $p_j^t = 0$ for all assets, this condition can be violated only if $\lambda_i > 1$, i.e. if the bank is leveraged.

Whenever a bank does not satisfy the solvency condition (5.5), we assume its portfolio undergoes a fire sale, i.e. all its assets are immediately liquidated. The fire

sale causes the price of the assets in the bank's portfolio to drop. If x_j is the fraction of asset j that has been liquidated, the price is updated as

$$p_j \to p_j f_j(x_j) \tag{5.7}$$

We are interested in the response of the system to an initial shock. We consider two kinds of initial shocks:

- **Presence of a toxic asset.** We select a random asset j and devalue it at time 0.
- **Initial failure of a bank.** We select a random bank i and cause it to go bankrupt.

In each case we follow the chain of events caused by the initial shock. The dynamics we consider is very simple: after shocking the system at time $t = 0$, at each time step $t = 1, 2, \ldots$ the solvency condition (5.5) is checked for each bank, the portfolios of newly insolvent (bankrupted) banks are liquidated, and new prices are computed for each asset. The dynamics stops when no new bankruptcies occur between two consecutive time steps. This can be expressed with the following algorithm:

1. introduce the initial shock in the system;
2. liquidate the portfolio of insolvent banks;
3. recompute prices of assets;
4. if new banks are insolvent go to step 2, otherwise end.

Note that we don't allow for new banks to enter the system, so that once a bank has gone bankrupt it remains in this state for the rest of the process.

In the limit of large systems, when $N, M \to \infty$ while the parameters μ_b and $n = N/M$ remain finite, the initial shock we consider only affects an infinitesimal (of order $O(1/N)$) fraction of the banking system. We are interested in understanding if and when such infinitesimal shocks can trigger global cascades of failures. A *global cascade* of failures is defined as a cascade affecting a finite fraction of banks in the infinite system. In the following we will measure the probability and the average extent of contagion. We define the *probability of contagion* as the probability that a global cascade of failures occurs, and the *average extent of contagion* as the average size of a global cascade.

5.3 Stability analysis

In this section we develop a theoretical approach that allows us to compute a bound on stability, which as we will show is a good estimate, for when cascading bank failures are likely to occur. We show how this can be applied to understand the stability of specific banking networks (i.e. a given set of banks and their balance sheets), and we also show how it can be used to understand how stability depends on the parameters of the network, such as diversification, crowding, and leverage.

Let us start by discussing what happens if there is an external shock that causes a particular bank to go bankrupt. Through the combination of leverage and impact, this failure can trigger the failure of other banks investing in the same assets. If the parameters of the system amplify shocks, this can generate a cascading failure that propagates through the system. One of our main points is that, while the likelihood of the first bank failure depends on the nature of the shocks, whether or not this propagates depends on whether the financial system is stable, which in turn depends on parameters such as the leverage, market impact and network structure. We begin with a general discussion of branching processes. We then discuss how it can be

applied to understand a given banking network, and make some specific assumptions that allow us to demonstrate how the stability of the banking system depends on parameters.

5.3.1 The Galton-Watson process

The scenario of cascading failures for banks closely resembles the branching process introduced by Galton and Watson to study the survival probability of family names over generations [95]. This process is formulated in terms of a progenitor that gives rise to $\mathfrak{x}$ children, where $\mathfrak{x}$ is a non-negative integer drawn from a probability distribution $\mathfrak{g}(\mathfrak{x})$. Each of the children, in turn, independently generates a number of offspring distributed according to $\mathfrak{g}(\mathfrak{x})$, and the same process is repeated at each generation. The question is whether such a process is doomed to extinction or not, i.e. if the population drops to zero after a finite number of generations, so that the total number of descendants is finite. A fundamental result in the theory of branching processes states that such a process goes extinct with probability one if $\mathbb{E}[\mathfrak{x}] < 1$, where $\mathbb{E}[\mathfrak{x}]$ is the expected number of offspring per individual.

For our purposes it is essential to consider a generalized Galton-Watson process with individuals of different types $\mathfrak{i} \in 1, 2, \ldots, \mathfrak{q}$. The key quantities are then, for each pair of types $\mathfrak{i}, \mathfrak{j}$, the expected number of offspring of type $\mathfrak{i}$ produced by an individual of type $\mathfrak{j}$. We denote these as a $\mathfrak{q} \times \mathfrak{q}$ matrix $\mathcal{N}_{\mathfrak{i}\mathfrak{j}}$. The condition for extinction is then that the largest eigenvalue ξ_1 of $\mathcal{N}$ is smaller than one [96]. Conversely, if this eigenvalue is greater than one, then with positive probability this process lasts forever, producing an infinite number of offspring. We say that the branching process is *subcritical* or *supercritical* if this eigenvalue is less than or greater than one, respectively.

In the context of our model, we are interested in computing the expected number

of banks that go bankrupt because of the previous failure of another bank. Consider for example the case in which a random bank $\mathtt{i}$ receives a shock at time $\mathtt{t} = 0$ that causes it to become bankrupt. This bank is equivalent to the progenitor of the Galton-Watson process, and banks whose bankruptcy is triggered by that of $\mathtt{i}$ are equivalent to its offspring. In the language of branching processes, banks failing at time $\mathtt{t}$ correspond to individuals in the $\mathtt{t}$-th generation. We are interested in understanding when there is a non-zero chance that financial contagion keeps on spreading over time, which is equivalent to asking whether shocks will be amplified rather than dying out. If the branching process is supercritical, then this initial shock results in a global cascade with non-zero probability, affecting a non-zero fraction of all the banks in the limit of infinite system size.

Note that in our model we have in principle banks with different properties (degree, leverage, size...) that can be considered as individuals of different types in the generalized Galton-Watson process. Thus $\mathcal{N}_{\mathtt{hk}}$ is the expected number of banks of type $\mathtt{h}$ that fail because of the failure of a bank of type $\mathtt{k}$. There is obviously considerable flexibility in how we classify banks into types, which at its most fine-grained extreme allows the "types" to correspond to individual banks.

It is important to stress at this point that the process here considered is more complex than the usual generalized Galton-Watson process. In particular, in our case, the ultimate fate of bank $\mathtt{i}$ depends not only on its properties, but also on those of all the other banks whose portfolio overlaps with $\mathtt{i}$. This happens because the price drop that follows the fire sale liquidation of an asset depends on the fraction of total shares of that asset being liquidated, which changes from bank to bank.

A second important difference is that the Galton-Watson process occurs on a tree, so that individuals of a given generation are independent. For banks the failure process is not necessarily a tree, but is rather a more general graph which may have loops. To see this, consider a simple example of three banks $\mathtt{i}$, $\mathtt{j}$, and $\mathtt{k}$ with one

asset in common. Let us suppose that $\mathtt{i}$ is robust with respect to the failure of $\mathtt{j}$ by itself, but not with respect to the failure of both $\mathtt{j}$ and $\mathtt{k}$ together. Now, if the failure of $\mathtt{j}$ is enough to trigger that of $\mathtt{k}$, then $\mathtt{i}$ is effectively vulnerable to the failure of $\mathtt{j}$. Such situations, which occur whenever assets have degree higher than 2, are neglected under the analytical calculation that we perform here. Therefore, our analytical treatment gives a sufficient but not a necessary condition for global cascades to occur, and gives only an upper bound on the stability of the banking system. We will see, however, that it is nonetheless a good approximation, in rough agreement with the results of numerical simulations.

It is in principle possible to improve this approximation to account for the non-linearities induced by loops in the branching process by considering multiple time-step dynamics. This method is commonly used in dynamical system theory: the $\mathtt{t}$-th iteration of the dynamics converts cycles of length $\mathtt{t}$ into fixed points. For instance, to properly treat triplets one can compute a two-step matrix that counts the average number of banks of type $\mathtt{i}$ whose failure is triggered by the bankruptcy of a bank of type $\mathtt{j}$ within two time steps of the dynamics. Comparing to the example given above, if an initial shock causes $\mathtt{j}$ to fail, $\mathtt{k}$ will fail after one iteration, and since both $\mathtt{j}$ and $\mathtt{k}$ have now failed, $\mathtt{i}$ will fail in the second time step. While our one-step approximation is already quite accurate, this approach provides a path for systematically improving the degree of approximation, which deserves further investigation.

5.3.2 Stability of a given system

If we have complete information about the banking system, i.e. if we know the portfolios $\mathtt{Q}_{\mathtt{ba}}$ of all the banks and, in addition, the market impact function for their assets, then we can describe the stability of the system through a matrix $\mathcal{B}$, where $\mathcal{B}_{\mathtt{ij}}$ is the probability that bank $\mathtt{i}$ will fail under the failure of bank $\mathtt{j}$. Bank $\mathtt{i}$ becomes

insolvent when bank j fails if and only if the market impact due to the sale of their overlapping assets causes a loss to bank i that exceeds its equity E_i. As described above, we focus for now on the direct effect on bank i of the failure of bank j. Since we assume that bank j's entire portfolio $\{Q_{ja}\}$ is liquidated, the new price for asset a is $p_a(1 - f_a(Q_{ja}))$. Using the shorthand $\text{Prob}(x)$ to indicate the probability that condition x is satisfied, the stability matrix $\mathcal{B}_{ij}$ is defined as

$$\mathcal{B}_{ij} = \text{Prob}\left[\sum_{a=1}^{M} Q_{ia} p_a \left(1 - f_a\left(Q_{ja}\right)\right) - E_i > 0\right] . \tag{5.8}$$

In order to understand whether a cascade of failures will spread, we compute $\mathcal{B}_{ij}$ in the case where the assets shared by banks i and j have not yet been devalued, and still have their initial prices. That is, we focus on the "boundary" of the cascade, with failures and devaluations spreading outward through the network through banks and assets that have not yet been touched by the crisis. In that case, since the dynamics themselves are deterministic, $\mathcal{B}_{ij}$ depends only on the initial structure of the banks' portfolios, and in particular on the network structure. The stability of the banking system can then be estimated by simply computing the largest eigenvalue ξ_1 of $\mathcal{B}$ and determining whether ξ_1 is greater than or less than one.

Note that, rather than using the simplifying approximation that the market impact function is deterministic, one could more realistically use a stochastic market impact function as in [88]. Similarly, imperfect knowledge about bank portfolios and equity can be coped with using probabilities to represent uncertainties in their values. In either case, we can still bound the stability of the network by computing $\mathcal{B}$'s largest eigenvalue.

5.3.3 Simplifying assumptions

The approach described above makes it possible to estimate the stability of the banking system when it is in a particular state, corresponding to a particular config-

uration of the balance sheets of each bank. One of our main goals here, however, is to understand more generically how the stability of the banking system depends on its network properties. To make a high-level characterization it is necessary to think in terms of ensembles of networks, and to understand how stability varies as properties of the ensemble are varied. As a first step in this direction we will make some specific assumptions in order to simplify the problem and gain intuition. While these assumptions are rather arbitrary, the basic method used here is easily generalized, as discussed later.

- **Network topology:** We will consider random networks with Poisson degree distributions for both banks and assets. Specifically, for each possible bank-asset pair a link is drawn with probability μ_b/M. The resulting network is drawn from the bipartite Erdős-Renyi ensemble of random networks with average degrees μ_b and $\mu_a = \mu_b N/M$ for the banks and assets respectively.

- **Structure of balance sheets:** We will assume all banks have the same amount of money $A_i^0 = A^0$ available for investment, and that each bank uniformly splits its investment in the assets that are in its portfolio. The asset side of bank's balance sheets will be composed of 80% assets and 20% cash. For bank i each link thus corresponds to an investment of $0.8A^0/k_i$, where k_i is the number of assets in i's portfolio. Unless otherwise stated, we assume for each bank an initial equity $E_i^0 = E^0$ corresponding to 4% of its total assets. This corresponds to all banks having initial leverage $\lambda = A^0/E^0 = 20$.

- **Market impact function:** We will assume that the market impact function has the form $f_j(x_j^t) = e^{-\alpha x_j^t}$, where x_j^t is the fraction of asset j liquidated up to time t. The parameter α is chosen such that the price drops by 10% when 10% of the asset is liquidated, i.e. $\alpha = 1.0536$. All prices are set to $p_j^0 = 1$ at time 0. This choice corresponds to linear market impact for log-prices, as originally used to describe price dynamics in [97, 98]. It should be noted

that recent empirical and theoretical evidence indicates that market impact for large trades is a concave function of the number of traded shares, which under normal conditions impact is well approximated by a square-root function [99]. By normal conditions we mean that execution is slow enough for the order book to replenish between successive trades. Under extreme conditions, like those of a fire sale, market impact is expected to become less concave and even linear or super-linear [100], which motivates our choice of functional form here.

Altering these assumptions does not change the qualitative behavior of the system. In particular, our methods generalize easily to degree distributions other than Poisson, e.g. power laws, and also to multiple types of banks with different sizes, portfolio structures, and amounts of leverage, or multiple types of assets with different initial prices and market impact functions.

5.3.4 Explicit calculation of the stability matrix

In order to understand how stability depends on network properties, we lump banks into equivalence classes according to their degree, equating their degree with their type in the generalized Galton-Watson process. We define the following notation:

- N_{h} is the number of banks of degree h.
- $\mathcal{P}(\mathsf{h}, \mathsf{k}|\mathsf{a})$ is the probability that a given bank of degree h and a given bank of degree k share a given asset a, i.e., are both connected to a in the network.
- $\mathsf{F}(\mathsf{h}|\mathsf{k}, \mathsf{a})$ is the probability that a bank of degree h fails given that it is connected to a failed bank of degree k through asset a.

Under the assumption that we are in the limit where $\mathsf{M} \to \infty$, $\mathsf{N} \to \infty$ while μ_{b} and $\mathfrak{n} = \mathsf{N}/\mathsf{M}$ are finite, the network is sparse, and we can easily compute the

probability $\mathcal{B}_{ij}$ by summing over each asset one at a time. If i has degree h and j has degree k, the probability that the failure of bank j causes bank i to fail can be written

$$\mathcal{B}_{ij} = \sum_{a} \mathcal{P}(h, k|a)\, F(h|k, a)\,. \tag{5.9}$$

Summing over all banks of degree h, the expected number of failures of banks of degree h caused by the failure of a bank of degree k, is

$$\mathcal{N}_{hk} = N_h \sum_{a=1}^{M} \mathcal{P}(h, k|a) F(h|k, a)\,. \tag{5.10}$$

This is the matrix defining the branching process, i.e. the expected number of offspring of type h from an individual of type k.

We can now compute each of the entires of $\mathcal{N}_{hk}$ in turn. Since the degree distribution of our network ensemble is Poisson, the number of banks of degree h is simply $N_h = NP_b(h)$ where

$$P_b(h) = \frac{e^{-\mu_b}\mu_b^h}{h!} \tag{5.11}$$

is the probability that a bank has degree h. A given bank of degree h is connected to a given asset a with degree ℓ_a with probability $h\ell_a/(\mu_b N)$, where $\mu_b N$ is the total number of edges in the network. The probability that a failed bank of degree k is also connected to the same asset a is $h(k-1)\ell_a(\ell_a-1)/(\mu_b N)^2$, where the factor of $k-1$ comes from the fact that one of the k edges of the failed bank is already connected to the asset that caused its failure. This gives

$$\mathcal{P}(h, k|a) = \frac{h\ell_a(k-1)(\ell_a-1)}{\mu_b^2 N}\,. \tag{5.12}$$

We now compute the probability $F(h|k, a)$ that a bank i of degree h fails due to failure of a bank j of degree k given that they share an asset a. The shift in price when a fraction x_a of an asset is sold is $(1 - f_a(x_a))$. (Recall that the initial price is

set to one for convenience). Thus the condition for a bank of degree k to fail because bank j sells a fraction x_a of asset a is

$$\frac{A^0}{k}\left(1 - f_a(x_a)\right) > E^0. \tag{5.13}$$

If $\nu(a)$ denotes the set of banks investing in asset a, the fraction of a that is liquidated when j fails is

$$x_a = \frac{A^0/k}{\sum_{m\in\nu(a)} A^0/k_m} = \frac{1/k}{\sum_{m\in\nu(a)} 1/k_m} = \frac{1/k}{1/h + 1/k + \sum_{m\in\nu'(a)} 1/k_m}, \tag{5.14}$$

where $\nu'(a)$ denotes the set of banks, *other than* i *and* j, that invest in a and k_m the degree of bank m.

To compute $F(h|k, a)$ we must add up the probability of failure for each possible configuration of banks that are compatible with the condition of choosing a specific pair of banks of degrees h and k that are connected through asset a. If a has degree ℓ_a, there are $\ell_a - 2$ remaining banks. Letting i index these banks, we must average over the possible configurations $\{m_1, \ldots, m_{\ell_a-2}\}$. Fortunately the degrees of the banks are independent. The probability that bank i has degree m_i is the ratio of the number of edges for banks of degree m to the total number of edges. Since $N_m = NP_b(m)$, the number of edges for banks of degree m is mN_m, and the total number of edges in the network is $\mu_b N$. Thus each bank has degree m with probability $mN_m/(\mu_b N) = mP_b(m)/\mu_b$ and the probability of any given configuration of bank degrees is

$$\prod_{i=1}^{\ell_a-2} \frac{m_i P_b(m_i)}{\mu_b}. \tag{5.15}$$

Combining equations (13 - 15) and summing over all the possible configurations $\{m_1, \ldots, m_{\ell_a-2}\}$ gives

$$F(h|k,a) = \sum_{m_1=1}^{\infty} \cdots \sum_{m_{(\ell_a-2)}=1}^{\infty} \prod_{i=1}^{\ell_a-2} \frac{m_i P_b(m_i)}{\mu_b} \tag{5.16}$$

$$\times \Theta\left[\frac{A^0}{h}\left(1 - f_a\left(\frac{1/k}{1/h + 1/k + \sum_{i=1}^{\ell_a-2} 1/m_i}\right)\right) - E^0\right], \tag{5.17}$$

where Θ is the Heaviside step function, $\Theta(x) = 1$ if $x > 0$ and zero otherwise.

After summing over assets equation (5.10) becomes

$$\mathcal{N}_{hk} = \frac{e^{-\mu_b}\mu_b^h}{h!}\frac{h(k-1)}{\mu_b^2 n}\sum_{\ell}\frac{e^{-\mu_a}\mu_a^\ell}{\ell!}\ell(\ell-1)F(h,k,\ell), \tag{5.18}$$

where we have used the fact that the number of assets with given degree ℓ is $MP_a(\ell)$ and explicitly introduced the Poisson degree distributions of banks and assets.

The form of the matrix $\mathcal{N}$ confirms that the independent parameters of the model are μ_b, n, λ and α. We can see in particular that, although leverage has a similar effect on stability to the market impact constant α, the two are not related through a simple relation that allows us to eliminate one of the two dependencies. However, if we had used a market impact that was linear in the price, instead of the log-price, i.e. of the form $f_a(x) = \alpha' x$, then the stability would depend only on the product $\alpha'\lambda$ and not on the two parameters separately.

For networks in which all the banks have the same degree k we can compute the largest eigenvalue of $\mathcal{N}$ in closed form. In this case the matrix $\mathcal{N}$ reduces to the scalar quantity

$$\mathcal{N} = \xi_1 = (k-1)\mu_b n \frac{\Gamma(l^* - 1, \mu_b n)}{\Gamma(l^* - 1)}, \tag{5.19}$$

where

$$l^* = \frac{1}{\log\left(\frac{\lambda}{\lambda - k}\right)}, \tag{5.20}$$

$\Gamma(x)$ is the gamma function, and $\Gamma(x, z)$ is the incomplete Gamma function.

If we make the approximation $\frac{1}{k_b} \to \mathbb{E}\left[\frac{1}{k}\right]$ we can obtain a closed expression for $\mathcal{N}_{h,k}$. However, given that this approximation is uncontrolled, we do not give an explicit form for the matrix elements, but rather compute them exactly via Montecarlo methods.

5.4 Dependence on leverage and network properties

We now explore how the stability of the banking network depends on parameters. We first show results based on numerical simulations and then compare them to results based on the stability matrix $\mathcal{N}$.

In numerical simulations N and M are both finite, and global cascades can be defined as cascades for which the fraction of bankrupted banks exceeds a fixed threshold. For consistency with previous work on counterparty loss [80, 101, 102], we set this threshold to 5%. The contagion probability is then measured as the fraction of runs in which a global cascade results from the initial shock. The conditional average extent of contagion is the fraction of failed banks, averaged *only* over those runs where a global cascade occurs.

5.4.1 Effect of diversification and crowding

We begin with the case where the initial shock consists of devaluing a random asset., and examine the dependence on diversification and crowding. In the left panel of Figure 5.2, we plot the probability and conditional extent of contagion measured for a system of $N = 10^4$ banks and $M = 10^4$ assets as a function of the average

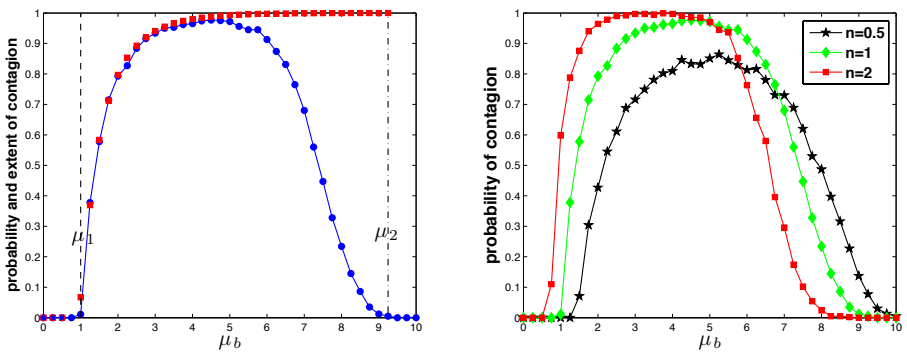

Figure 5.2: **Left panel:** contagion probability (blue dots, the solid line is a guide for the eye) and conditional extent of contagion (red squares) measured from 1000 simulations of a system with $N = M = 10^4$. In each run, the initial shock consists of dropping the price of a random asset by 35% at the beginning of the simulation. We vary the average degree of diversification $\mu_b = \mu_a$. The two vertical dashed lines mark our numerical estimates for the critical values μ_1 and μ_2 where phase transitions occur, and show the existence of a contagion window between these transitions where global cascades occur with non-zero probability. The system also displays a "robust yet fragile" behavior for μ_b slightly below μ_2: the probability of a global cascade is small, but when one occurs it affects almost all the banks. **Right panel**: contagion probability for systems with $N = 10^4$ and $M = 5 \times 10^3$ (red squares), $M = 10^4$ (green diamonds) and $M = 2 \times 10^4$ (black stars) as a function of the average banks' degree. Solid lines are a guide for the eye. The boundaries μ_1, μ_2 of the contagion window depend on the value of the crowding parameter $n = N/M$: for larger n both phase transitions are shifted to the left.

banks' degree μ_b. Results refer to 1000 runs in which a random asset is initially devalued by 35%. We observe phase transitions at two critical values μ_1, μ_2 of μ_b, with a *contagion window* in between where global cascades of failures occur with non-zero probability. Above and below this window, where $\mu_b > \mu_2$ or $\mu_b < \mu_1$, global cascades do not occur.

The existence of a contagion window, and the nonmonotonicity of the contagion probability as a function of μ_b, can be understood with the following arguments. On one hand, for sufficiently low values of μ_b, stress cannot propagate through the system because the network is poorly connected; there is not enough overlap between

the banks' portfolios to spread the cascade. In particular, for small enough μ_b the network of banks and assets consists of small components disconnected from one another, so even if every bank is extremely vulnerable to collapse, an initial shock will only affect one of these components. Thus there is a critical μ_1 below which the cascade cannot propagate; the initial shock might affect a few nearby banks, but the cascade quickly dies out.

On the other hand, if the banks' portfolios are sufficiently diverse, they are robust with respect to devaluing any single asset in their portfolio. Moreover, a larger average bank degree μ_b also implies a larger average asset degree μ_a, so each institution typically holds a smaller fraction of the shares of any given asset. As a consequence, each bank failure has a relatively small effect on asset prices, and most banks remain solvent even if some of their assets are devalued. Thus there is a critical μ_2 above which cascades quickly die out even though the network is highly connected.

The left panel of Figure 5.2 also shows that the system displays a "robust yet fragile" behavior [80] for some values of the parameters. Specifically, if μ_b is slightly less than μ_2, just inside the upper end of the contagion window, the probability of a global cascade is very small, tending continuously to zero as μ_b approaches μ_2 from below. But when a global cascade does occur, it affects almost all the banks: the conditional extent of the contagion is almost 1.

In the right panel of Figure 5.2 we plot the contagion probability for different values of the crowding parameter $\pi = N/M$. As π increases, the contagion window shifts to the left, decreasing both μ_1 and μ_2.

The shift in μ_1 can be understood in terms of the appearance of a giant connected component in the network. In the ensemble of random networks considered here, the emergence of the giant component corresponds to the situation where the average number of banks to which a given bank b is exposed, i.e. the average number of

banks whose portfolios share at least one asset with $\mathfrak{b}$, is one. Equivalently, this is the average degree of the *projected network* where two banks are connected if they share an asset. For this ensemble (essentially the bipartite version of the Erdős-Rényi model) this degree is $\mu_\mathfrak{a}\mu_\mathfrak{b} = \mu_\mathfrak{b}^2\mathfrak{n}$, giving $\mu_1 = 1/\sqrt{\mathfrak{n}}$.

To explain the shift in the second transition point μ_2, we note that the drop in price of an asset caused by the liquidation of a portfolio is a decreasing function of $\mathfrak{n}$. This is because the average number of banks investing in a given asset is $\mu_\mathfrak{a} = \mu_\mathfrak{b}\mathfrak{n}$. If each bank owns a smaller fraction of an asset, the market impact of a fire sale on that asset is smaller. When $\mathfrak{n}$ is larger, this effect takes over at a smaller value of $\mu_\mathfrak{b}$.

Note that, as a result, changing the crowding parameter $\mathfrak{n}$ has different effects on the network's stability depending on the value of $\mu_\mathfrak{b}$. If $\mu_\mathfrak{b}$ is close to μ_1, increasing $\mathfrak{n}$ while keeping $\mu_\mathfrak{b}$ fixed increases the instability of the system, moving it into the contagion window by increasing the connectivity of the network. The opposite is true if $\mu_\mathfrak{b}$ is close to μ_2, where increasing $\mathfrak{n}$ moves us outside the contagion window by making assets and banks more robust. Thus the contagion probability is not a monotonic function of $\mathfrak{n}$.

5.4.2 Dependence on shocks

The above simulations started with an initial shock consisting of devaluing a random asset. We now consider the case where we begin with the failure of a random bank. Figure 5.3 shows a comparison between simulations with shocks of these two types. We observe that the probability of contagion depends on the type of shock, but the contagion window and the conditional extent of contagion are the same for both types of shock. The reason is simple: while the initial conditions of these two processes are different, their dynamics are the same. Once a cascade has begun, it doesn't matter

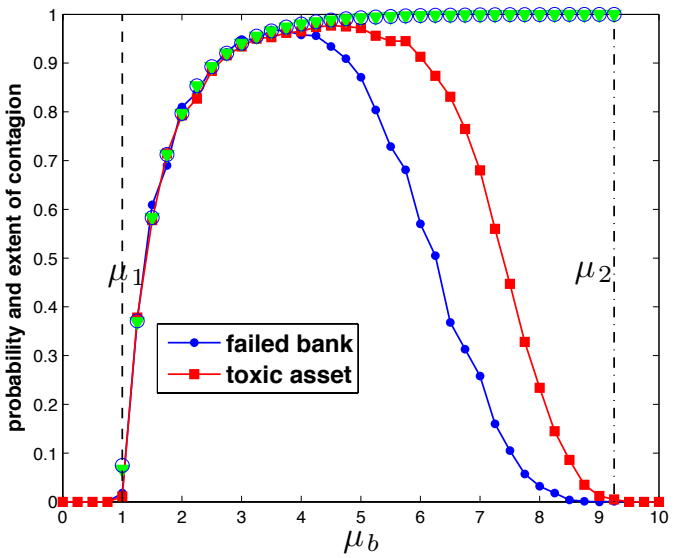

Figure 5.3: The probability of contagion, and the average conditional extent of contagion, as a function of μ_b for the two types of initial shock (failed asset vs. failed bank). Red squares: contagion probability where a random asset is devalued by 35%. Blue dots: the contagion probability when a random bank fails. Blue circles and green triangles: conditional extent of contagion for asset shocks and bank shocks respectively. We see that while the probability of contagion differs between the two types of shocks, the window $\mu_1 < \mu_b < \mu_2$ in which they occur with non-zero probability is the same. Moreover, when a global cascade does occur, its average size is the same for both types of shocks. Results refer to 1000 simulations of systems with $N = M = 10^4$.

what kind of shock began it. Thus the region where the dynamics cause a cascade to spread rather than die out is the same in both cases, as is the eventual size of a global cascade if one occurs.

5.4.3 Leverage

We now show what happens for different values of initial leverage. In the left panel of Figure 5.4 we plot the contagion probability for different values of μ_b as a function of λ. We observe that, for each μ_b, there is a critical value of λ above which global

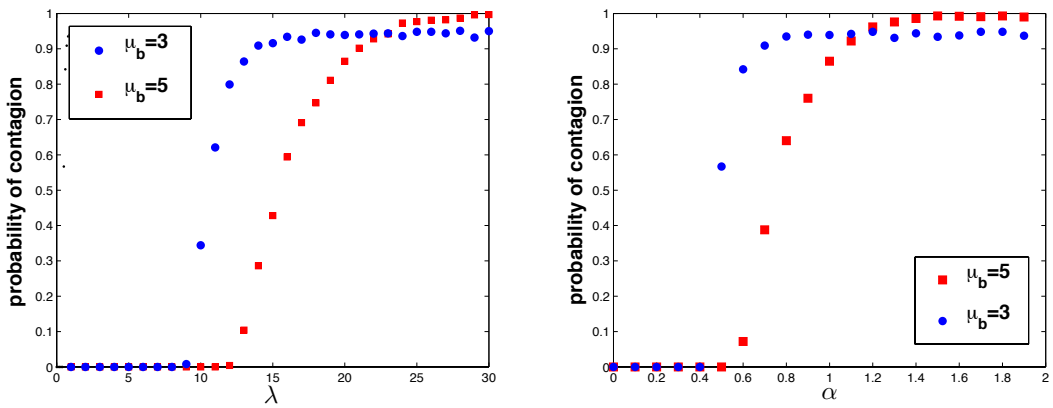

Figure 5.4: **Left panel:** Contagion probability as a function of leverage measured from 1000 simulations of a system with $N = 10000$ for different values of μ_b.The initial shock considered is the initial failure of a random bank. Contagion probability is a monotonic function of leverage, and a phase transition separates a regime where no global cascades are observed from one where they occur with non-zero probability. **Right panel:** Contagion probability as a function of the market impact parameter α. Increasing market impact has a similar effect as increasing leverage.

cascades occur with non-zero probability, and below which they do not. This is of interest for regulatory purposes, since it implies the existence of a critical level of leverage below which systemic stability is guaranteed. In addition, the critical value of λ increases as μ_b increases: in other words, increasing diversification allows for a greater degree of leverage without creating systemic events.

In the right panel of Figure 5.4, we show that a similar behavior occurs as we change the parameter α that appears in the market impact function while keeping the leverage fixed. That is, for a given value of μ_b and λ, there is a critical value of α above which contagion occurs. This is not unexpected, since under the assumptions specified in Section 5.3.3 the solvency condition for bank i can be written as

$$\lambda_i \leqslant \frac{\sum_{j=1}^{M} Q_{ij} e^{-\alpha x_j^t}}{E_i^0} + 1, \tag{5.21}$$

where x_j^t is the fraction of shares of asset j liquidated up to time t. When α is larger,

the market impact of a fire sale is greater, causing a sharper drop in asset prices. On the other hand, increasing diversification μ_b increases this critical value of α, showing that diversification allows banks to survive a larger price impact.

Summarizing, we presented in this section results of numerical simulations for bipartite networks with Poisson degree distributions for both banks and assets. The probability and the average extent of contagion have been measured for two different types of shocks, namely the initial depreciation of a random asset or the initial failure of a random bank. Our simulations suggest that:

- As a function of the average diversification of banks' portfolios, represented by their average degree μ_b, the system is characterized by two phase transitions that define a contagion window where global cascades occurs with non-zero probability.

- Changing the crowding parameter n, i.e. the ratio of the number of banks to the number of assets available for investment, can increase or decrease the contagion probability depending on which of these transitions we are close to.

- Although the contagion probability is different for the two types of initial shocks, the contagion window within which global cascades occur, and the average extent of these cascades when they occur, are the same.

- The system displays a "robust yet fragile" behavior, with regions in parameter space where global cascades are very unlikely, but where almost the entire system is affected if one occurs.

- For each fixed μ_b and n, there is a critical value of the leverage λ above which the system becomes unstable. This critical value of λ increases with μ_b.

5.5 Comparison to predictions from stability analysis

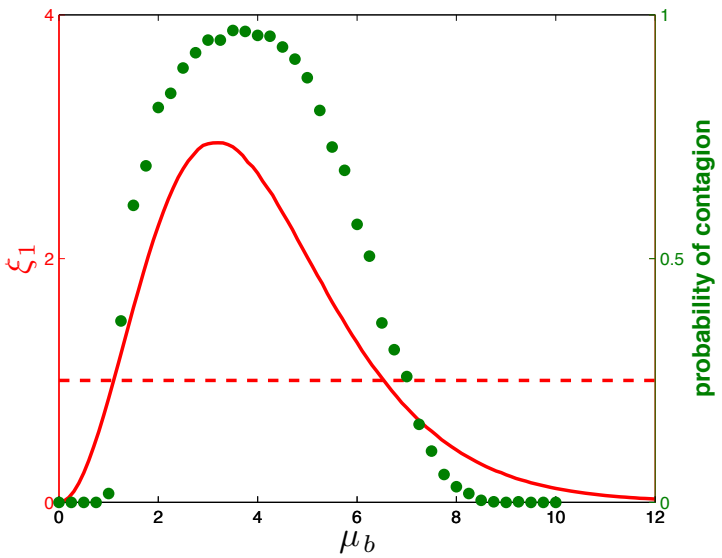

Figure 5.5: Contagion probability (green dots, right axis) as computed from numerical simulations of a system of size $\mathsf{N} = \mathsf{M} = 10^4$. The red solid line (left axis) represent the largest eigenvalue ξ_1 of the matrix $\mathcal{N}$. The dashed horizontal line is in correspondence to $\xi_1 = 1$. If $\xi_1 > 1$ global cascades are observed in numerical simulations. The theory underestimates the width of the contagion window, as it only gives a sufficient condition for global cascades to occur. However, the discrepancy between theory and numerical results is partly due to finite size effects (see Figure 5.6).

We now compare the numerical results presented in the previous section to those based on stability analysis. The stability analysis depends on two assumptions that are not necessarily well-satisfied in the simulation. The first is that M and N are both infinite (even though their ratio $\mathfrak{n} = \mathsf{N}/\mathsf{M}$ is finite), and the second is that the failure process can be described through a branching process on a tree.

We estimated $\mathsf{F}(\mathsf{h}, \mathsf{k}, \ell)$ through monte-carlo methods and assumed that the contribution coming from banks with degree higher than 200 is negligible. We then

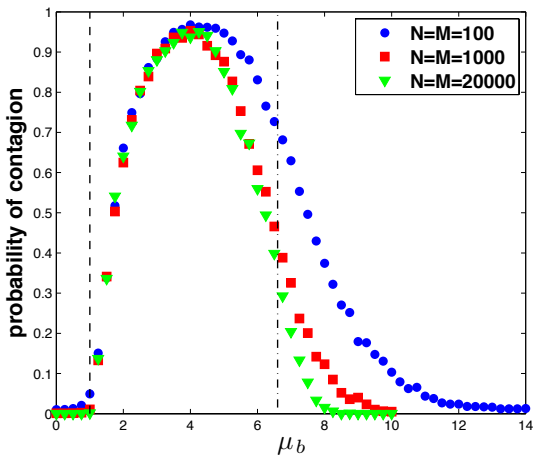

Figure 5.6: Simulation results for $\mathsf{N} = 100$ (blue circles), $\mathsf{N} = 1000$ (red squares), $\mathsf{N} = 20000$ (green triangles), $\mathfrak{n} = 1$. The vertical dashed lines are drawn in correspondence to the phase transitions predicted by the analytic calculation. As the size of the system increases the agreement between theory and simulations improves. Finite size effects are expected given that the theory is valid in the limit $\{\mathsf{N}, \mathsf{M}\} \to \infty$.

numerically diagonalized the 200 by 200 matrix $\mathcal{N}$. We discuss in the following the results obtained in the case where $\mathsf{P}_a(\ell)$ and $\mathsf{P}_b(\mathfrak{h})$ are Poisson distributions.

In Figure 5.5 we plot for $\mathfrak{n} = 1$ the largest eigenvalue of $\mathcal{N}$ and we compare it with the contagion probability as computed from numerical simulations. As expected, when the largest eigenvalue of $\mathcal{N}$ is greater than 1 global cascades are observed. We see from the figure that the analytic calculation underestimates the size of the contagion window. This is partly due to finite size effects, as observed for instance in [103]. We plot in Figure 5.6 the contagion probability as measured from numerical simulations for $\mathfrak{n} = 1$ and different values of N. From the figure we clearly see that by increasing the size of the system the discrepancy between analytic and numerical calculations gets smaller, and that the analytic solution, although giving only a sufficient condition for global cascades to occur, produces a reasonable estimate of the contagion window when N and M are large.

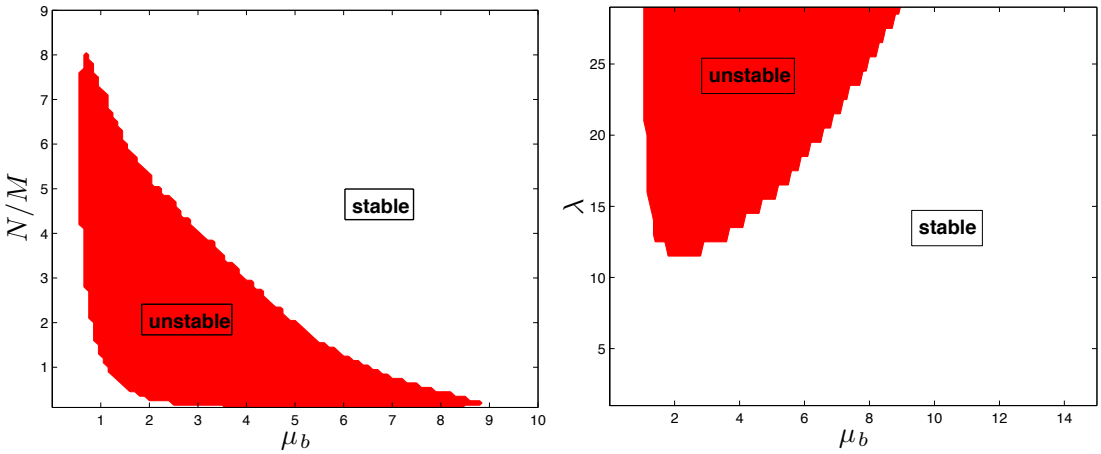

Figure 5.7: **Left panel**: The red region is the region of phase space where global cascades occur for a system with $\lambda = 20$ as a function of μ_b and n. **Right panel**: the red region is the region of phase space where global cascades occur for a system with $n = 1$ as a function of μ_b and λ. Points refer to the phase transition as measured from the largest eigenvalue of $\mathcal{N}$. Lines are a guide for the eye.

We finally plot in Figure 5.7 the phase diagram obtained with our analytic approach. The region within the solid line in the left panel represent the region where global cascades occur for $\lambda = 20$. From the figure we can see the features already observed in numerical simulations. In particular, for fixed n, we clearly see the existence of two phase transitions that define a window of connectivities where global cascades occur with non-zero probability. As we change n, the analytic calculation also predicts the shift in the transition points, that tend to move to higher values of μ_b as n is decreased. In the right panel of Figure 5.7, we depict the phase diagram for $n = 1$ as a function of leverage λ and average diversification μ_b. As expected, we see that the contagion window widens as λ increases. A three dimensional visualization of the phase diagram is reported in Figure 5.8, where the red region denotes the unstable region of parameter space. Interestingly, we observe the existence of a minimum level of leverage ($\lambda \simeq 12$) that leads to the occurrence of global cascades of failure. This feature is of potential interest for regulators, since it is equivalent

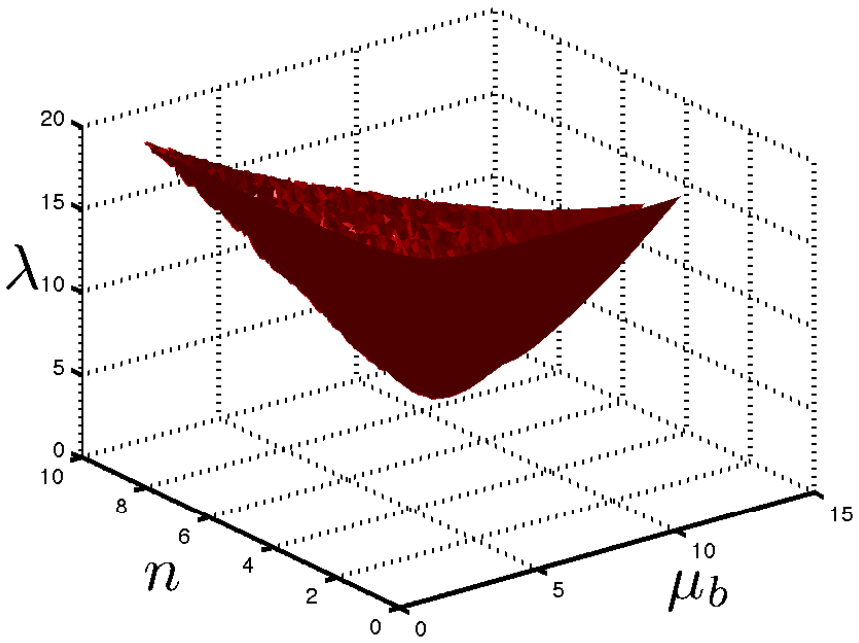

Figure 5.8: 3D visualization of the region of parameter space where global cascades occur with non-zero probability as predicted with the analytical approach. Global cascades are observed within the cone-shaped colored region. We note in particular that there is critical value of leverage below which global cascades do not occur for any values of diversification and crowding parameter.

to the existence of a maximum level of leverage below which the system is overall stable.

Another point of potential interest for regulators is the eigenvalue of the matrix $\mathcal{N}$. In a dynamic setting in which banks operate under no stress circumstances, one expects ξ_1 to change over time as banks trade to rebalance their portfolios. By monitoring the time behavior of ξ_1, a regulator would notice if the system is approaching a dangerous regime as ξ_1 gets closer to 1, and could act to increase the stability of the system.

Chapter 6

Concluding remarks and future work

In this thesis, we first consider in chapter 3 the dynamic message-passing (DMP) technique to study a simple threshold model of behavior in networks. In doing so, we are able capture how each individual's probability of becoming an adopter evolves in time in an arbitrary network with far less computational cost than Monte Carlo simulations. Although DMP is exact only on trees, we observe that it compares well with simulations even in a real social network where there are many loops. Interestingly, unlike in the SIR model, or equivalently the case $\mathsf{T} = 1$, there are cases where DMP can either underestimate or overestimate the probability of infection.

In addition, we apply DMP equations to give analytical results in the thermodynamic limit of large random networks. We provide an exact analytic result for calculating the time dependence of the probabilities, thereby learning something about the dynamics of bootstrap percolation.

The message-passing dynamics we consider here can be generalized in many ways, including letting the transmission probability and the threshold vary arbi-

trarily across edges and vertices. Because the transmission rate $r(\tau)$ may depend on the elapsed time τ since an individual became an adopter, our study can be implemented in networks where some non-Markovian assumptions are warranted, as we pointed out in Section 3.1.

We can include so-called "rumor spreading" models where, rather than setting $r(\tau) = 0$ until an individual's awareness reaches a threshold as we have done here, an individual starts telling its neighbors about the rumor even if it has only heard about it once. Such models were recently applied to the diffusion of microfinance [41]. We can also let the rate at which an individual receives new information depend on its own awareness. An interesting case is to consider a unimodal function.

We can also consider a model where j can transmit repeatedly to i, raising i's awareness each time. We simply replace each directed edge (j, i) with T multi-edges. So, each message $U_{i \leftarrow j}(t)$ would now be mapped to T identical copies of itself. The update equations and expressions are the same as above, but now we sum over all these multi-edges accordingly.

In Section 3.3, we focus on random networks in the configuration model. However, DMP equations can be easily generalized to many other families of random graphs, including interdependent networks [40], scale-free networks [54], small-world networks [33, 32], and bipartite networks [11] to name a few. In some cases this is a matter of plugging in a different degree distribution, and allowing for a finite number of types of vertices. However, for preferential attachment networks the topology is correlated with the vertices' ages, so we would have to let the messages $U(t)$ depend on the age of the vertices sending them.

We can also extend this study to a network that has community structures such as the stochastic block model. We can then study how trends move through communities, and how the distribution of initial adopters (for instance, whether they are

concentrated in one community, or are spread across many communities) affects the eventual fraction of the network that adopts the trend. Community structures can be driven by socio-economic, ethnic, religious and linguistic separations. So, it would be useful to gain some perspective on how the structures of communities contribute to the norms and social preferences that prevail in real populations, and in turn how differences in these norms drive the division of social networks into communities.

However, DMP has been exclusively applied to non-recurrent models, where the states of the nodes evolve in one-way or irreversibly. Modern epidemiological studies often require recurrent models, where nodes can return to their previous inhabited states multiple times. For example, consider diseases such as influenza where individuals are infected multiple times throughout their lives, or whooping cough where vaccine effectiveness wanes over time; in both cases, individuals return to the Susceptible class.

In chapter 4, we extend DMP to recurrent epidemic models which we call rDMP. Our rDMP approach defines messages on the directed edges of a network in such a way as to prevent signals, such as the spread of infection, from backtracking immediately to the node that they came from. By preventing these "echo chamber effects," rDMP obtains good estimates of the time-varying marginal probabilities on a wide variety of networks, estimating both the fraction of infectious individuals in the entire network, and the probabilities that individual nodes become infected.

Like the pair approximation, rDMP takes correlations between neighboring nodes into account. However, our experiments show that rDMP is more accurate than the pair approximation for a wide variety of network structures and parameters. Moreover, rDMP is computationally less expensive than the pair approximation, especially for complex epidemic models with a large number of states, using $O(mk)$ instead of $O(mk^2)$ variables for models with k states on networks with m edges.

Finally, rDMP is conceptually simple, allowing the user to immediately write down the system of differential equations for a wide variety of epidemic models, such as those with multiple stages of infection or immunity [72, 73], or those with multiple interacting diseases [74, 75]. We expect that given its simplicity and accuracy, it will be an attractive option for future epidemiological studies.

In an applied side of this thesis, we develop a framework for thinking about the stability properties of banking networks due to overlapping portfolios. This framework emphasizes that the key property is stability: If the system is stable, shocks will not propagate; if it is unstable, a shock can be amplified and trigger cascading bankruptcies. This can be discussed in terms of a branching process that gives insight into the dynamics of failure. While we call these "banking networks" for simplicity, the basic ideas are relevant for any leveraged financial institutions.

To understand how the stability of banking networks might depend on parameters such as diversification, leverage and crowding, we formulate a stylized model of a financial system in which N banks with average diversification μ_b invest in a common pool of M assets. The system can be conveniently described in terms of a bipartite network, with banks being connected through links to the assets in their portfolios. Links have a two-fold role in such a network. On one hand, they allow individual banks to diversify their investment and reduce their exposure to a specific asset. On the other hand, they are channels for the propagation of financial contagion. We characterize the response of such system to initial shocks affecting a single asset or bank.

The relevant parameters for the model are the average diversification μ_b, the crowding parameter $n = N/M$ (that measures the proportion of banks to assets), and the initial leverage λ. By means of numerical simulations we show the existence of phase transitions separating a region in parameter space where global cascades occur from a region where global cascades never occur. In particular, the double

role played by links in the bipartite network representing the system is reflected in a non-monotonic behavior of the contagion probability as a function of μ_b, where we observe the existence of two phase transitions at $\mu_b = \mu_1$ and $\mu_b = \mu_2$ that define a window of connectivities such that global cascades occur if $\mu_1 \leqslant \mu_b \leqslant \mu_2$. Changing the crowding parameter π has the effect of shifting the location of the phase transitions. Finally, our model shows that increasing leverage increases the overall instability of the system, but that there is a critical level of leverage below which global cascades do not occur for any value of diversification or crowding.

Using an analytical approach based on generalized branching processes on networks, we are able to analytically estimate the region of parameter space where global cascades occur. This branching process is different from standard ones in the fact that the fate of a node depends on its degree and on the degree of all its neighbors. This greatly increases the difficulty of the problem. We are nonetheless able to solve it by generalizing existing methods. Thus, apart from their specific application to financial contagion, our methods can be applied to a wide variety of contagion models, where susceptibility and transmission probabilities depend on node degrees.

The mechanistic model we consider in this chapter can be extended in several directions. First of all, it would be interesting to relax some of the specific assumptions considered in this chapter (homogeneity of banks' balance sheets, Poisson degree distributions, market impact function) in order to understand how different choices for the network topology or the statistical properties of balance sheets impact the stability of the system. Although we do not expect different results from a qualitative point of view, it should nonetheless be possible to asses the relative stability of systems with different properties, similarly to what has been done for counter-party loss in [87]. In particular, it would be very useful to empirically characterize real systems and calibrate the model with real data. This could potentially make it possible to test the effectiveness of new policies aimed at reducing systemic risk.

A further direction we plan to pursue in the future is to go beyond the mechanistic model by considering a more realistic price dynamics and allowing banks to react to price fluctuations by rebalancing their portfolios. This should allow the system to develop endogenous crisis similar to the ones observed in [104], and to generate the systemic instabilities induced by leverage and mark-to-market accounting practices discussed in [88].

References

[1] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications*. Hafner Press, New York (1975).

[2] R. M. Anderson and R. M. May, *Infectious Diseases of Humans*. Oxford University Press, Oxford (1991).

[3] M. Granovetter, Threshold models of collective behavior, *American Journal of Sociology* **83(6)**, 1420-1443(1978).

[4] M. Granovetter, The strength of weak ties *American Journal of Sociology* **78(6)**, 1360-1380(1973).

[5] J.H. Miller and S.E. Page, The standing ovation problem, *Complexity* **9**, 8-16 (2004).

[6] B. Gonçalves, N. Perra, A. Vespignani, Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number. *PLoS ONE* **6 (8)**, e22656 (2011).

[7] P. Bak, K. Chen, and C. Tang, A forest-fire model and some thoughts on turbulence. *Phys. Lett. A,* **147**, 297-300 (1990).

[8] B. Drossel, and F. Schwabl, Self-organized critical forest-fire model. *Phys. Rev. Lett.* **69**, 1629-1632 (1992).

[9] P. Grassberger, Critical behaviour of the Drossel-Schwabl forest fire model. *New J. Phys,* **4**, 17 (2002).

[10] R. M. May and A. G. Haldane, Systemic risk in banking ecosystems. *Nature* **469**, 351–355 (2011).

[11] F. Caccioli, M. Shrestha, C. Moore, and J. D Farmer, Stability analysis of financial contagion due to overlapping portfolios. *Journal of Banking & Finance* **46**, 233-245 (2014).

[12] M. Mézard and A. Montanari, *Information, Physics, and Computation.* Oxford University Press (2009).

[13] C. Moore and S. Mertens, *The Nature of Computation.* Oxford University Press (2011).

[14] R. Morris, Zero-temperature Glauber dynamics on Z^{d}. *Prob. Theory Rel. Fields,* **149**, 3-4 (2011).

[15] R.I.M Dunbar, Neocortex size as a constraint on group size in primates. *Journal of Human Evolution* **22 (6)**, 469-493 (1992)

[16] M.E.J. Newman, The spread of epidemic disease on networks, *Phys. Rev. E* **66**, 016128 (2002).

[17] L. A. Meyers, Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, *Bulletin of the American Mathematical Society* **44** 63-86 (2007).

[18] M. E. J. Newman, *Networks: An Introduction.* Oxford University Press (2010).

[19] A.Y. Lokhov, M. Mézard, H. Ohta, and L. Zdeborovà, Inferring the origin of an epidemic with dynamic message-passing algorithm. *Phys. Rev. E* **90**, 012801 (2014)

[20] F. Altarelli, A. Braunstein, L. Dall'Asta, A. Ingrosso, and R. Zecchina, The zero-patient problem with noisy observations. *J. Stat. Mech* P10016 (2014)

[21] F. Altarelli, A. Braunstein, L. Dall'Asta, J.R. Wakeling, and R. Zecchina, Containing epidemic outbreaks by message-passing techniques. *Phys. Rev. X* **4** 021024 (2014)

[22] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Optimizing spread dynamics on graphs by message passing. *J. Stat. Mech* P09011 (2013)

[23] F. Altarelli, A. Braunstein, A. Ramezanpour, and R. Zecchina, Stochastic optimization by message passing. *J. Stat. Mech* P11009 (2011)

[24] J. Pearl, Reverend Bayes on inference engines: a distributed hierarchical approach, *AAAI Proceedings* **82**, (1982).

[25] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications, *Phys. Rev. E* **84**, 066106 (2011).

[26] B. Karrer and M.E.J. Newman, Message passing approach for general epidemic models. *Phys. Rev. E* **82**, 016101 (2010)

[27] M. Shrestha and C. Moore, Message passing approach for threshold models of behavior in networks. *Phys. Rev. E* **89**, 022805 (2014)

[28] Joel C. Miller, Anja C. Slim and Erik M. Volz, Edge-based compartmental modelling for infectious disease spread. *Journal of the Royal Society Interface [Internet].* **9** 890-906 (2010).

[29] F. Altarelli, A. Braunstein, L. Dall'Asta, and R. Zecchina, Large deviations of cascade processes on graphs. *Phys. Rev. E* **87** 062115 (2013)

[30] A.Y. Lokhov, M. Mézard, and L. Zdeborovà, Dynamic message-passing equations for models with unidirectional dynamics. *Phys. Rev. E* **91**, 012811 (2015)

[31] S. H. Strogatz, Exploring complex networks. *Nature* **410**, 268–276 (2001).

[32] C. Moore and M. E. J. Newman, Epidemics and percolation in small-world networks. *Phys. Rev. E* **61**, 5678–5682 (2000).

[33] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).

[34] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Åberg, The web of human sexual contacts. *Nature* **411**, 907–908 (2001).

[35] M. Kuperman and G. Abramson, Small world effect in an epidemiological model. *Phys. Rev. Lett.* **86**, 2909–2912 (2001).

[36] R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203 (2001).

[37] D. J. de S. Price, Networks of scientific papers. *Science* **149**, 510–515 (1965).

[38] A. L. Lloyd and R. M. May, How viruses spread among computers and people. *Science* **292**, 1316–1317 (2001).

[39] D. Watts, A simple model of global cascades on random networks. *PNAS* **99-9**, 57665771. (2001)

[40] S. Buldyrev, R. Parshani, G. Paul, H. Stanley, and S. Havlin, Catastrophic cascade of failures in interdependent networks, *Nature Physics* **464**, 1025-1028(2010).

[41] A. Banerjee, A. Chandrasekhar, E. Duflo, M. Jackson, The Diffusion of Microfinance, *Science* **341-6144**, 888-893(2013).

[42] B. Karrer and M.E.J. Newman, A message passing approach for general epidemic models, *Phys. Rev. E* **82**, 016101 (2010)

[43] W. W. Zachary An information flow model for conflict and fission in small groups, *Journal of Anthropological Research* **33 (4)**, 452-473 (1977)

[44] B. Pittel, J. Spencer, N. Wormald, Sudden Emergence of a Giant k-Core in a Random Graph. *Journal of Combinatorial Theory Series B* **67-1**,111-151(1991).

[45] J. Chalupa, Bootstrap percolation on a Bethe lattice. *J. Phys. C* **12**, L31-35(1979).

[46] R. Cerf and F. Manzo, Nucleation and growth for the Ising model in d dimensions at very low temperatures. Preprint. Available at arXiv:1102.1741.

[47] R. Cerf and F. Manzo, A d-dimensional nucleation and growth model. Preprint, arXiv:1001.3990.

[48] J. M. Schwarz, A. J. Liu, and L. Q. Chayes, The onset of jamming as the sudden emergence of an infinite k-core cluster, *Europhys. Lett.* **73**, 560-566 (2006).

[49] M. Aizenman, and J. L. Lebowitz, Metastability effects in bootstrap percolation. *J. Phys. A* **21**, 38013813. MR0968311 (1988)

[50] S. Janson, T. Luczak, T. Turova, and T. Vallier, Bootstrap percolation on the random graph $\mathsf{G}_{\mathsf{N},\mathsf{P}}$. *The Annals of Applied Probability* **22-5**, 1989-2041 (2012)

[51] H. Amini, Bootstrap percolation and diffusion in random graphs with given vertex degrees, *Electron. J. Combin.* Research Paper 25, 20, 1077-8926 (2010)

[52] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Bootstrap Percolation on Complex Networks, arXiv:1003.5583 (2010)

[53] M.E.J. Newman, S.H. Strogatz, and D.J.Watts, Random graphs with arbitrary degree distributions and their applications, *Phys. Rev. E* **64**, 026118 (2001)

[54] A. Barabási, R. Albert, Emergence of scaling in random networks, *Science* **286 (5439)** 509512 (1999).

[55] M. Granovetter, Threshold models of collective behavior. *American Journal of Sociology* **83(6)**, 14201443(1978).

[56] M. Granovetter, The strength of weak ties. *American Journal of Sociology* **78(6)**, 13601380(1973).

[57] J.H. Miller and S.E. Page, The standing ovation problem. *Complexity* **9**, 8-16 (2004).

[58] M. A. Porter and J. P. Gleeson, Dynamical systems on networks: A tutorial. arXiv:1403.7663 (2014).

[59] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology* **54**, 396-405 (2003).

[60] P. Zhang, and C. Moore, Scalable detection of statistically significant communities and hierarchies: message-passing for modularity. *Proceedings of the National Academy of Sciences* **111** (51), 18144-18149

[61] K. Hashimoto, Zeta functions of finite graphs and representations of p-adic groups. *Advanced Studies in Pure Mathematics* **15** 211-280 (1989).

[62] J. Pearl, Reverend Bayes on inference engines: a distributed hierarchical approach. *AAAI Proceedings* **82**, (1982).

[63] W. W. Zachary, An information flow model for conflict and fission in small groups. *Journal of Anthropological Research* **33 (4)**, 452-473 (1977).

[64] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*. Princeton and Oxford: Princeton University Press (2008).

[65] F. Krzakala, C. Moore, E. Mossel, J. Neeman, A. Sly, L. Zdeborová, and P. Zhang, Spectral redemption in clustering sparse networks. *Proceedings of the National Academy of Sciences* **110 (52)**, 20935-20940 (2013).

[66] B. Karrer, M. E. J. Newman, and L. Zdeborová, Percolation on sparse networks. *Phys. Rev. E* **113**, 208702 (2014).

[67] S. Chatterjee and R. Durrett, Contact processes on random graphs with power law degree distributions have critical value 0. *The Annals of Probability* **37**, 2332–2356 (2009).

[68] H W Watson, and Francis Galton, On the Probability of the Extinction of Families *Journal of the Anthropological Institute of Great Britain*, **4**, 138-144, (1875).

[69] T. Martin, X. Zhang, M. E. J. Newman, Localization and centrality in networks. *Phys. Rev. E* **90**, 052808 (2014).

[70] D. J. D Earn, J Dushoff, S. A Levin, Ecology and evolution of the flu. *Trends in ecology & evolution* **17**, 334–340 (2002).

[71] M. G. M Gomes, L. J White, G. F Medley, Infection, reinfection, and vaccination under suboptimal immune protection: epidemiological perspectives. *Journal of Theoretical Biology* **228**, 539–549 (2004).

[72] S. Melnik, J. A. Ward, J. P. Gleeson, and M. A. Porter, Multi-stage complex contagions. *Chaos* **23**, 013124 (2013).

[73] J. C. Miller and E. M Volz, Incorporating Disease and Population Structure into Models of SIR Disease in Contact Networks. *PLoS ONE* **8**, (8) e69162 (2013).

[74] B. Karrer and M. E. J. Newman, Competing epidemics on complex networks, *Phys. Rev. E* **84**, 036106 (2011).

[75] J. C. Miller, Cocirculation of infectious diseases on networks, *Phys. Rev. E* **87**, 060801 (2013).

[76] M. R. Golden, W. L. H. Whittington, H. H. Handsfield, J. P. Hughes, W. E. Stamm, M. Hogben, A. Clark, C. Malinski, J. R. L Helmers, K. K. Thomas, and K. K Holmes, Effect of expedited treatment of sex partners on recurrent or persistent gonorrhea or chlamydial infection. *New England Journal of Medicine* **352**, 676–685 (2005).

[77] P. H. Conway, A. Cnaan, T. Zaoutis, and B. V. Henry, R. W. Grundmeier, and R. Keren, Recurrent urinary tract infections in children: risk factors and association with prophylactic antimicrobials. *Journal of the American Medical Association* **2**, 179–186 (2007).

[78] G. M. Jeffery, Epidemiological significance of repeated infections with homologous and heterologous strains and species of Plasmodium. *JBulletin of the World Health Organization* **35**, 873 (1966).

[79] L. M. Drusin, B. G. Ross, K. H. Rhodes, A. N. Krauss, R. A. Scott, Nosocomial Ringworm in a Neonatal Intensive Care Unit A Nurse and Her Cat. *Infection Control* **21**, 605–607 (2000).

[80] P. Gai and S. Kapadia, "Contagion in financial networks," Proc. R. Soc. A, vol. 466, no. 2120, pp. 2401–2423, 2010.

[81] F. Allen, A. Babus, and E. Carletti, "Asset commonality, debt maturity and systemic risk," *Journal of Financial Economics*, vol. 104, no. 3, pp. 519–534, 2012.

[82] R. Ibragimov, D. Jaffee, and J. Walden, "Diversification disasters," *Journal of Financial Economics*, vol. 99, pp. 333–348, 2011.

[83] P. Gai, A. Haldane, and S. Kapadia, "Complexity, concentration and contagion," *Journal of Monetary Economics*, vol. 58, no. 5, pp. 453 – 470, 2011.

[84] R. M. May and N. Arinaminpathy, "Systemic risk: the dynamics of model banking systems," *Journal of The Royal Society Interface*, vol. 7, no. 46, pp. 823–838, 2010.

[85] H. Amini, R. Cont, and A. Minca, "Resilience to contagion in financial networks," *Available at SSRN: http://ssrn.com/abstract=1865997*, 2010.

[86] C.-P. Georg, "The effect of the interbank network structure on contagion and financial stability," Global Financial Markets Working Paper Series 12-2010, Friedrich-Schiller-University Jena, Oct. 2010.

[87] F. Caccioli, T. A. Catanach, and J. D. Farmer, "Heterogeneity, correlations and financial contagion," Advances in Complex Systems, vol. 15, no. supp. 2, p. 1250058, 2012.

[88] F. Caccioli, J.-P. Bouchaud, and J. Doyne Farmer, "A proposal for impact-adjusted valuation: Critical leverage and execution risk," *ArXiv e-prints*, Apr. 2012.

[89] J. C. Staum, "Counterparty contagion in context: Contributions to systemic risk," *Available at SSRN :http://ssrn.com/abstract=1963459*, 2011.

[90] N. Beale, D. G. Rand, H. Battey, K. Croxson, R. M. May, and M. A. Nowak, "Individual versus systemic risk and the regulator's dilemma," *Proceedings of the National Academy of Sciences*, vol. 108, no. 31, pp. 12647–12652, 2011.

[91] R. Cont and L. Wagalath, "Running for the exit: Distressed selling and endogenous correlation in financial markets," *Mathematical Finance*, pp. no–no, 2012.

[92] E. Nier, J. Yang, T. Yorulmazer, and A. Alentorn, "Network models and financial stability," *Journal of Economic Dynamics and Control*, vol. 31, no. 6, pp. 2033–2060, 2007.

[93] N. Arinaminpathy, S. Kapadia, and R. May, "Size and complexity in model financial system," *Bank of England Working Paper No. 465*, 2012.

[94] X. Huang, I. Vodenska, S. Havlin, and H. E. Stanley, "Cascading Failures in Bi-partite Graphs: Model for Systemic Risk Propagation," *ArXiv e-prints*, Oct. 2012.

[95] H. W. Watson and F. Galton, "On the probability of the extinction of families," *The Journal of the Anthropological Institute of Great Britain and Ireland*, vol. 4, pp. pp. 138–144, 1875.

[96] C. J. Mode, *Multitype branching processes; theory and applications*. American Elsevier Pub. Co., New York,, 1971.

[97] J.-P. Bouchaud and R. Cont, "A langevin approach to stock market fluctuations and crashes," *European Physics Journal B*, vol. 6, pp. 543–550, 1998.

[98] J. D. Farmer, "Market force, ecology and evolution," *Industrial and Corporate Change*, vol. 11, no. 5, pp. 895–953, 2002.

[99] J.-P. Bouchaud, J. D. Farmer, and F. Lillo, "How markets slowly digest changes in supply and demand," in *Handbook of Financial Markets: Dynamics and Evolution* (T. Hens and K. Schenk-Hoppe, eds.), pp. 57–156, Elsevier, 2009.

[100] J. Gatheral, "No-dynamic-arbitrage and market impact," *Quantitative Finance*, vol. 10, pp. 749–759, 2010.

[101] J. Gleeson, T. R. Hurd, S. Melnik, and A. Hackett, "Systemic risk in banking networks without monte carlo simulation," in *Advances in Network Analysis and its Applications* (E. Kranakis, ed.), pp. 57–156, Springer Verlag, 2011.

[102] T. R. Hurd and J. P. Gleeson, "A framework for analyzing contagion in banking networks," *ArXiv e-prints*, Oct. 2011.

[103] D. Watts, "A simple model of global cascades on random networks," PNAS, no. 99, pp. 5766–5771, 2002.

[104] S. Thurner, G. Geanakoplos, and J. D. Farmer, "Leverage causes fat tails and clustered volatility," *Quantitative Finance*, vol. 12, no. 5, pp. 695–707, 2012.