A Branching Process Approximation of Complex Contagion on Random Networks

by

Michael P. Lynch

Under the supervision of:
Prof. James P. Gleeson

A thesis presented for the degree of
Master of Science

Department of Mathematics and Statistics
University of Limerick
August 2021

Abstract

Watts' threshold model (Watts, 2002) is reproduced numerically to examine cascade size distributions for a range of values of mean degree on Poisson random networks. We propose a branching process approximation to model cascade size distribution with a numerical stochastic model and a theoretical asymptotic analysis.

The numerical branching process approximation shows excellent agreement with results from the numerical Watts' threshold model at the lower end and middle of the cascade window. The location of upper transition differs between results from the numerical Watts' threshold model and the numerical branching process approximation. This is shown to be caused by the independent, identically distributed (*iid*) assumption of the branching process. The practical implication of this is that the vulnerability condition is limited to the contribution of, at most, one active neighbour. Therefore, nodes of higher degree above a certain threshold cannot be activated. In Watts' threshold model, multiple active neighbours may contribute to the activating of a node. The numerical branching process also shows excellent agreement with the theoretical branching number derived by Watts (2002) using a similar assumption of only one active neighbour contributing to the vulnerability condition.

An asymptotic analysis of the branching process approximation is carried out to derive the long-term distribution of large cascade sizes at, and near, the lower critical point. At the critical point of the lower transition, we derive analytically the $-\frac{3}{2}$ power-law distribution that Watts (2002) observed numerically, showing excellent agreement between the asymptotics results and numerical Watts' threshold model results. Moving into the cascade window, the near-critical asymptotic approximation of the branching process model predicts the cascade size distribution qualitatively, showing bimodal distribution, and also quantitatively, albeit with decreasing accuracy moving further from the critical point.

Acknowledgements

I would like to acknowledge the wisdom, guidance, patience and encouragement of my supervisor Prof. James Gleeson in both matters of research and writing. I am sincerely grateful for the time and energy he invested in helping me realise this thesis. I would like to give special thanks to my parents, siblings and close friends for their support, optimism and encouragement through a year which brought with it many challenges. Finally, I would like to express my thanks to the colleagues I have been so fortunate to work with over the last number of years who have instilled in me a confidence to pursue an unconventional path in education.

Contents

List of Figures

Chapter 1

Introduction

Networks can be used to describe a large number of systems in biology, physics, sociology and many other disciplines (Newman, 2018, 2003). Studying the structure of a network may give insight into how systems operate on that network. For example, one may examine the shortest path between any two nodes on a network representing human relationships and see that it explains the Small World phenomenon demonstrated by Travers and Milgram (1969). In the context of sexually transmitted disease, the connectivity of the network explained why some individuals, who were otherwise associated with high risk factors for infection, did not acquire the disease (Darrow et al., 1999). To complement the study of network structure, it is also possible to study network dynamics, whereby a phenomenon which follows a predefined rule acts on the network. The speed and extent to which the phenomenon spreads can then be observed for different graph structures and parameters.

Cascades occur when an initial change of state of a node, which we will call a shock, triggers a change of state of many nodes with which it is directly or indirectly connected. Global cascades are cascades which are of a size which exceeds a certain fraction of the network, although no fixed definition exists. It may be useful to think of cascades as a chain reaction of any size, regardless of how small or large, and global cascades as filling most of the available network. Shocks which cause global cascades are often indistinguishable from those that do not, and on systems which were previously shown to be stable after similar shocks. Such phenomena occur in many applications including financial markets (Doherty, 2018), power transmission (Schäfer et al., 2018) and ischemic attacks (Felberg et al., 2000).

Watts (2002) proposed a model for the diffusion of innovation,[1] whereby each node

[1]The model also applies to rumours or fads.

could have adopted, or not adopted, the technology. Each node is assigned a threshold which can be thought of as a resistance to uptake of the technology. A node with a low threshold is considered to have a lower resistance to uptake, and a node with a high threshold is considered to have a high resistance to uptake. An initial seed node or innovator is placed randomly in the network[2] and adoption by a node is determined by whether the fraction of active neighbours is greater then the node's personal threshold. Watts showed that there exists an interval in the control parameter mean degree of the graph outside of which global cascades do not occur. We define a transition point as one where global cascades begin or cease to occur in terms of the parameter mean degree. A continuous transition in maximum cascade size (number of nodes activated) was noted at the lower end of the cascade window and a discontinuous transition in size was noted at the upper end. We defer a more in-depth introduction to Watts' threshold model (WTM) until Chapter 3, where we also introduce the branching process (BP) approximation. By this time we will have introduced some concepts of network science which are key to its understanding. As this thesis is based on the WTM from Watts (2002), for the sake of brevity we will often refer to its contents as Watts' paper, Watts' results etc.

The remainder of the thesis is structured as follows: Chapter 2 gives an introduction to graph theory and some of the measures and statistics that will be used in later sections. Chapter 3 introduces the WTM in greater detail as well as some of Watts' results most relevant to the thesis. At this time, we also conceptually introduce the branching process approximation upon which most of the novel results in this thesis are based. Chapter 4 details the analytical model used to describe cascade size distribution, including theory on probability generating functions, branching process and the asymptotic analysis. Chapter 5 details the numerical implementation of creating Poisson random graphs, studying dynamics on networks as applied to the WTM, and our branching process approximation. The numerical implementation of the branching process model is also detailed here. Chapter 6 refreshes the main results of Watts (2002), adds to these results with the numerical WTM and compares results to the numerical BP approximation. Finally, the cascade size distributions from the asymptotic analysis are compared with the numerical WTM at and near the lower critical point. In Chapter 7, we conclude the thesis by summarising the main findings and we make propositions for how the model may be generalised.

As there is a large numerical aspect to this thesis, a supplementary MATLAB script is available which includes the functions detailed in Chapter 5. [3]

[2]This is equivalent to a shock in the context of diffusion of innovation.

[3]`https://github.com/mikelyynch/A-Branching-Process-Approximation-of-Complex-Contagion-on-Random-Networks`

Chapter 2

Network Theory

In this section, the foundation concepts of how networks are described mathematically are introduced. This provides a basis upon which much of the subsequent chapters are based.

Network theory studies agents, their relationships with other agents, and how phenomena spread across a system of agents through these relationships. Agents, otherwise called nodes, may represent people, computers, websites or cities, to name but a few examples. Relationships may take the form of a friendship, internet connection, hyperlink or road. Regardless of the system being studied, the underlying network structures are described in the same way and many of the metrics used to describe their structure remain applicable regardless of application. Henceforth, the term nodes will be used to refer to agents in a network. The term edges will be used to describe the connections between nodes in a network.

2.1 Network Representation

The symbol n is generally used to refer to the number of nodes present in a network. Relationships between nodes can be directed or undirected. The structure of a network, its nodes and edges, are most efficiently contained in an $n \times n$ matrix referred to as the adjacency matrix, denoted A. Where $A_{ij} = 0$, no edge exists between node i and j. Generally, an edge between nodes i and j is defined as $A_{ij} = 1$, although there are cases where a non-zero value other than 1 is used. These cases are known as weighted matrices, but are not relevant to WTM.

Networks can be both directed or undirected. Undirected networks represent mutual or reciprocated connections, for example a Facebook friend request. Directed networks

represent non-reciprocal relationships, for example a social media influencer may have millions of followers but is unlikely to follow the majority back. In some cases, mutual connection is not possible. An example of this is citations in published academic papers.

WTM is based on undirected, unweighted networks and so an emphasis will be placed on those graph types in this introduction. For the undirected case, because edges are reciprocal, $A_{ij} = A_{ji}$. Consequently, the undirected adjacency matrix is always symmetric around the leading diagonal. Taking the simplified network

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.1}$$

we can plot the network structure as in Figure 2.1. This adjacency matrix represents a network with three nodes, with an undirected, unweighted connection between nodes 1 & 2. Self-loops are also possible, for $A_{ij} = 1$ where $i = j$, but these are not a feature of WTM.

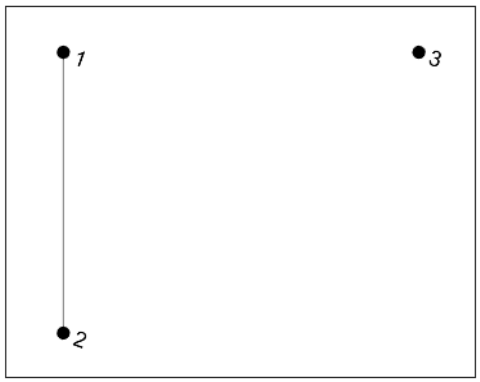

Figure 2.1: Simplified network corresponding to Equation 2.1.

The degree is the term used for the number of edges a particular node has, each of the nodes which make up the degree are then considered a neighbour of that particular node. Summing across a particular row or column of the adjacency matrix gives the degree of the respective node. The degree of node i, denoted k_i, is written mathematically as

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

While it is possible to construct a graph in such a way that each node is of equal degree,

usually there is some variability in the degree of nodes in a network. Finding the degree of each node in a network and plotting their probability maps the degree distribution p_k for that network. For well-known degree distributions, the probability of occurrence of a node of degree k can be calculated explicitly. The mean degree of the network, denoted z, is the average number of connections per node and is defined as

$$z = \frac{1}{n}\sum_{j}^{n}\sum_{i}^{n} A_{ij} = \frac{\text{total edges}}{\text{total nodes}}. \tag{2.3}$$

2.2 Random Graphs

Prior to the mainstream use of computers, information on the real structure of networks had to be collected manually, involving either questioning people on who was in their social network or by direct observation. Given the tedious nature of both of these methods, the sizes of the networks under study were limited to tens of people. With the development of computers came the potential to study much larger networks. Actual real-world network structures can be used including those based on Twitter interactions, for example De Domenico et al. (2013).

It is also possible to randomly generate graphs with certain statistical properties, producing a particular degree distribution. This allows us to study multiple realisations of a particular degree distribution, as well as work with arbitrary degree distributions. So, when we refer to a random graph, it is an ensemble of potential realisations of that graph type, rather than a unique structure.

2.2.1 Erdos-Renyi Random Graph

Random graphs, where all edges may exist with equal probability, are most closely associated with Erdos and Renyi after whom this type of graph is often referred. There are two slightly differing methods for creating an Erdos-Renyi random graph.

- The $G(n, p)$ method creates edges between n nodes with probability p. The mean number of edges is not fixed but rather follows a distribution. As n increases, the mean number of edges converges to $p \times (n-1)$.
- The $G(n, m)$ method creates m edges between n nodes uniformly at random. In this case the number of edges is fixed, unlike the $G(n, p)$ method where it is a random variable.

Both of these methods result in a graph with similar statistical properties and both are straightforward to implement in MATLAB, however, some important differences exist in their computational efficiency which is discussed in Chapter 5.

As all edges are created with equal probability, this method of graph creation represents a binomial process, and so the degree distribution will follow a binomial distribution. It can be shown that for large n, the binomial distribution tends towards the Poisson distribution. The following derivation is taken from Newman (2018). We begin by defining p_k, the probability of a node being connected to exactly k others as

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}, \tag{2.4}$$

following the binomial distribution. We also define p as

$$p = \frac{z}{n-1}. \tag{2.5}$$

Most large networks are sparse, so the mean degree z increases slower than n. This allows us to write

$$ln[(1-p)^{n-1-k}] = (n-1-k)ln(1-\frac{z}{n-1}) \simeq -z, \tag{2.6}$$

and taking exponentials of both sides, we get

$$(1-p)^{n-1-k} = e^{-z}. \tag{2.7}$$

For large n,

$$\binom{n-1}{k} = \frac{(n-1)!}{(n-1-k)!k!} \simeq \frac{(n-1)^k}{k!}, \tag{2.8}$$

and putting these together gives

$$p_k = \frac{(n-1)^k}{k!} p^k e^{-z} = \frac{(n-1)^k}{k!} (\frac{z}{n-1})^k e^{-z} = e^{-z}\frac{z^k}{k!}. \tag{2.9}$$

Here, the mean degree z has been used in place of the more commonly used rate parameter λ. For this reason, Erdos-Renyi random graphs are also known as Poisson graphs. For consistency, they will be referred to as Poisson graphs henceforth.

2.2.2 Configuration Model

Alternatively, one can specify the desired degree distribution of the network. Each node is assigned a degree from a distribution. For a node of degree k, k half stubs are assigned. These half stubs are connected uniformly at random. It can be shown that a configuration type model, when assigned a Poisson degree distribution produces a graph with similar statistical properties to those discussed previously. This is the widely used method for creating non-Poisson graphs, including power-law graphs. The probability of a node being connected to another node of degree k is given by

$$q_k = \frac{k}{z} p_k, \tag{2.10}$$

which is known as the excess degree distribution. In the Poisson case, where $p_k \sim \text{Poiss}$, the excess degree distribution reduces to

$$p_{k-1} = e^{-z} \frac{z^{(k-1)}}{(k-1)!}. \tag{2.11}$$

The probability of a node being chosen is weighted by how well connected it is, meaning high degree nodes are more likely chosen than low degree nodes. Conveniently, and intuitively, this removes the possibility of a node being connected to a zero degree node. We use the excess degree distribution outside of the context of generating configuration type random graphs to derive the branching process offspring in Chapter 5. As, for this thesis, we focus on a simplified network structure, the degree distribution on which all results are based is Poisson, so we do not use the configuration type model to produce random graphs.

2.3 Largest Connected Component

For the purposes of this thesis, we will limit ourselves to a mainly conceptual understanding of the largest connected component of a network. For a more rigorous examination of this measure, we refer the reader to Newman (2018). In a randomly generated graph, depending on the desired degree distribution and control parameters, there may exist a non-zero probability that at least some nodes are not connected to any others. Naturally, poorly connected graphs are more likely to have isolated nodes than highly connected graphs. Isolated nodes cannot be influenced by dynamics occurring on the rest of the graph. Similarly, a phenomenon seeded at an isolated node cannot propagate to the rest

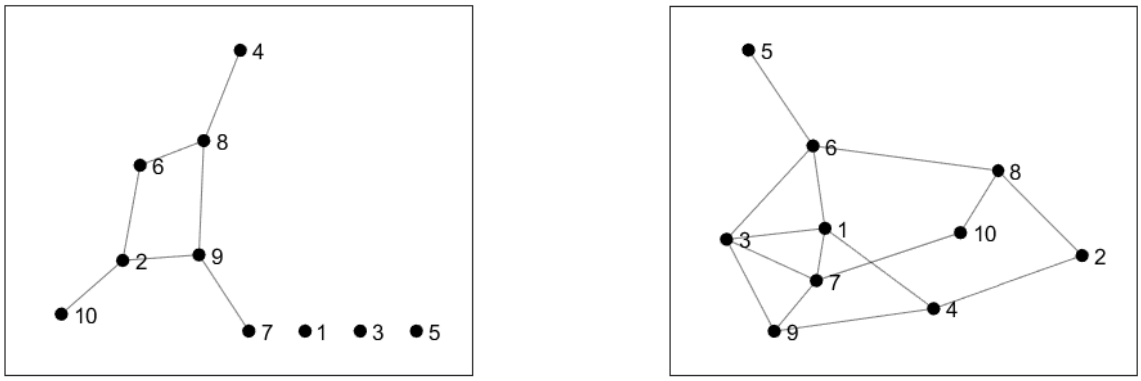

Figure 2.2: Single realisations of connected components for Erdos-Renyi random graph at $z = 1.4$ and $z = 4$.

of the network.

We show two realisation of a small Erdos-Renyi random graph to illustrate this concept. In Figure 2.2 (left) a poorly connected graph with mean degree $z = 1.4$ has several isolated nodes. A phenomenon seeded in the largest connected component will never be able to access nodes $1, 3$ & 5 and so, can at most affect 70% of the network. In Figure 2.2 (right), the network is better connected and there are no isolated nodes. So, the potential exists for a phenomenon to affect 100% of the network. As, when we turn to numerical simulation, we deal with much larger network sizes (usually $n = 10,000$), it is not uncommon for isolated nodes to occur even in well connected networks, but they occur less and less frequently as the network connectivity increases.

Watts (2002) showed that the maximum cascade size in the WTM was roughly equivalent to the largest connected component of the graph. Understanding this concept is a prerequisite in understanding Watts' own results introduced in Chapter 3 and the results presented from the numerical WTM in Chapter 6.

2.4 Clustering Coefficient

Again, as with the largest connected component, we are content with a mainly conceptual understanding of the clustering coefficient of a graph and refer the reader to Newman (2018) for further information. The term clustering refers to the tendency to which edges in a graph form triangles. A tree-like structure is one with no triangles, or we may more generally think of it as there being a maximum of one path between any two nodes. As

most graphs are not strictly tree-like, the clustering coefficient measures the extent to which this condition does or does not apply. A graphical comparison of tree-like and non tree-like structure is given in Figure 2.3. Different degree distributions and network types are associated with different levels of clustering. For the Poisson random graph, the clustering coefficient is defined as

$$c = \frac{z}{n-1}, \tag{2.12}$$

which goes to zero as $n \to \infty$ This metric, and the concept of tree-like structures becomes relevant when we apply the Discrete Time Galton-Watson branching process approximation to the WTM, as this assumes independent and identically distributed (*iid*) offspring.

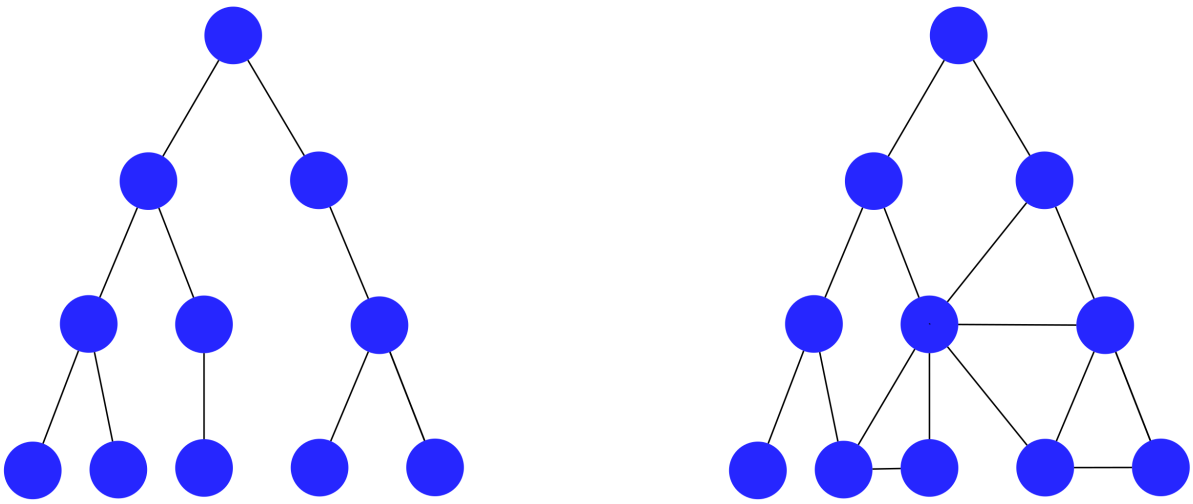

Figure 2.3: Tree-like and non tree-like network structure.

Chapter 3

Watts' Threshold Model and Branching Processes

In this chapter we first introduce some network dynamics terminology. Following this, we introduce Watts' threshold model (WTM) in more depth and the concept of a branching process.

3.1 Dynamics on Networks

Most interest in network science lies in the study of how a phenomenon travels across a network of a given structure. The range of application is vast and spans multiple disciplines. At this point we will introduce the terminology for different types of contagion or spreading, whether that be for a disease, idea or new technology. A simple contagion is one where a single exposure is sufficient to infect a node, like in a disease spread model. The probability of infection then increases linearly with number of exposures. In a complex contagion, like the WTM, a single exposure does not guarantee a non-zero probability of activation.

3.2 Watts' Threshold Model

The WTM is an example of a complex contagion model with binary externalities. The binary externalities refer to the states which a node can take, either active or inactive and this is governed by the node's neighbours, i.e. external to the node. Activated nodes in the WTM also display a permanently active property, meaning once a node becomes active, it cannot deactivate. Most intuitively this differs from a disease spread model, where

infectiousness generally only lasts for a finite time before the body immunises. More recently, the WTM has been adapted to include a time-dependent feature to represent a limited attention span (Karimi and Holme, 2021) but we restrict our study to the original WTM.

We introduced the concept of a personal threshold in Chapter 1 as a measure of resistance to activation. Typically, threshold values range from $0 - 1$ but values outside this range are also possible. A node with threshold less than zero will immediately adopt the technology once it becomes available regardless of the state of its neighbours, whereas a node with threshold greater than one will never adopt. In the context of diffusion of innovation, we refer to the first node[1] seeded as being active as the innovator. Nodes can either have adopted the technology (active) or not have adopted the technology (inactive). Adoption occurs when the active fraction of a node's neighbours exceed his or her personal threshold and we shall refer to this going forward as the vulnerability condition, such that the state S of node i is

$$S_i = \begin{cases} 1, & \text{if active fraction of neighbours} > \text{node's personal threshold.} \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

Again, we remind the reader that once a node becomes active, it remains active indefinitely.

3.2.1 Relevance of Watts' Threshold Model

On initial inspection, the reduction of a complex decision to one based purely on the uptake of their peers seems like a gross simplification. It would not be unreasonable to expect that one's decision to adopt a new technology would be impacted by a large number of variables. However, there is a strong body of evidence to suggest that a) peer pressure plays a large role in our decision making, or in other words, when faced with a complex decision, we defer to the experience of those around us (Dearing and Cox, 2018) and b) personal preferences can largely be captured in the threshold distribution (Valente, 1996).

Watts' paper and the following thesis examine the occurrence of global cascades outside of the context of application. We deliberately study a simplified model with homogeneous threshold on Poisson random networks, which still displays the behaviour of interest. It

[1] We assume infinitely small seed size for this thesis, equivalent to a seed size of one in a large network. The impact of finite seed size was investigated by Gleeson and Cahalane (2007)

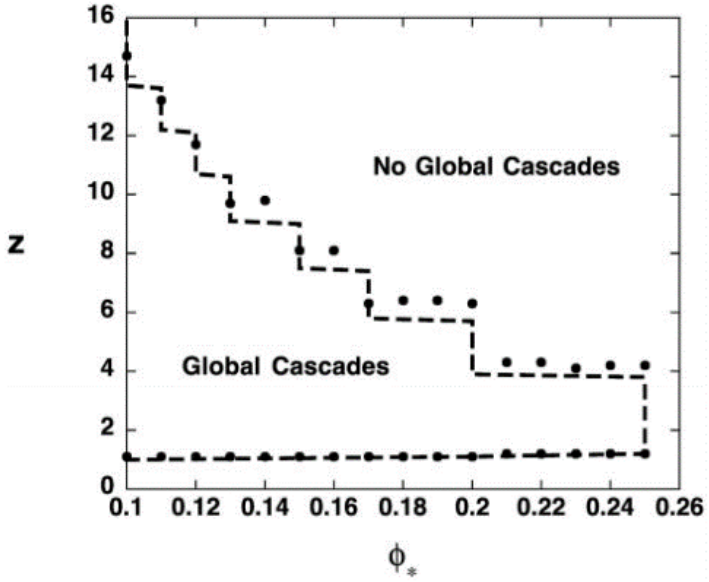

Figure 3.1: Cascade window from Watts (2002) (dashed lines represent theoretical cascade window, full circles represent global cascades for numerical model).

is essential to understand these behaviours in their simplest form before we can extend this understanding to real-world systems.

3.2.2 Behaviour

The Watts (2002) paper focused on the occurrence of global cascades which occurred for a subset of values of z (which we already defined as the mean degree) and ϕ, the homogeneous threshold. The range of z and ϕ for which global cascades occur is called the cascade window and is shown in Figure 3.1. We will now also define the lower transition point as where global cascades begin to occur, and the upper transition point as where global cascades cease to occur, in terms of the parameter z. Although the criteria for what makes a cascade global is not explicitly defined in Watts' paper, a value of 10% of the network size reproduces similar results. Intuitively, if the graph is too poorly connected, global cascades do not occur. Potentially less intuitively, if the graph is *too* well connected, cascades also do not occur. This upper transition is a direct consequence of the adoption criteria. If all neighbours of the seed node are of sufficiently high degree such that they cannot be activated, then no cascade can occur. Alternatively, the seed node may be able to activate one or more neighbours, but their neighbours may be of too high a degree to be activated, and so on.

For the region where global cascades do occur, Watts showed the maximum cascade

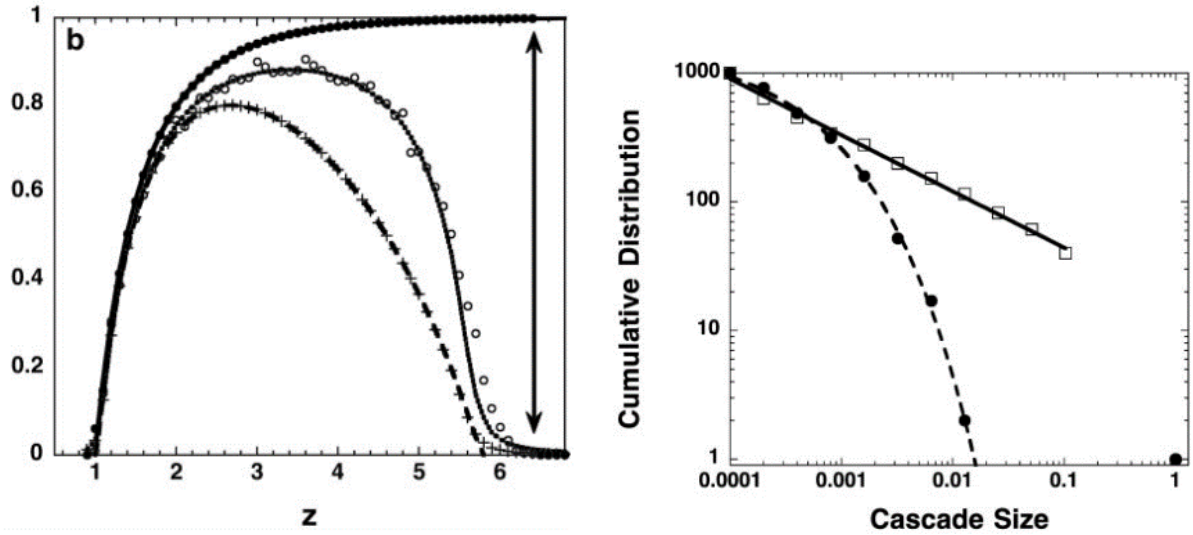

Figure 3.2: Mean cascade (crosses) and mean global cascade (filled circles) size and cascade frequency (open circles) from Watts (2002) (left). Cascade size distribution near lower ($z = 1.05$) (solid line) and upper ($z = 6.14$) (dashed line) transition points from Watts (2002) (right).

size was well approximated by the largest connected component of the graphs. This resulted in a continuous transition in terms of maximum cascade size at the lower transition point, and a discontinuous transition in terms of maximum cascade size at the upper transition point shown in Figure 3.2 (left). Both global cascade frequency and mean cascade size undergo continuous transitions at both the upper and lower transition point. We show results for cascade size distribution near the lower and upper transition points in Figure 3.2 (right). Near the lower transition point, cascade sizes follow a power-law distribution with exponent of $-\frac{3}{2}$, a result which we reproduce with both the numerical and asymptotic branching process approximations in Chapter 6. Near the upper transition point, the cascade sizes are bimodal, with most dying out. The remainder then fill most or all of the available graph. The branching process approximation does not accurately model the upper transition point of the WTM and we expand on the reasons for this in Chapters 6 & 7 in greater detail.

3.3 Branching Process Approximation

Branching processes were originally developed to model extinction probabilities of family names (Watson and Galton, 1875). The main premise of the Discrete Time Galton-Watson branching process is the following: A single individual lives for one unit of time, after

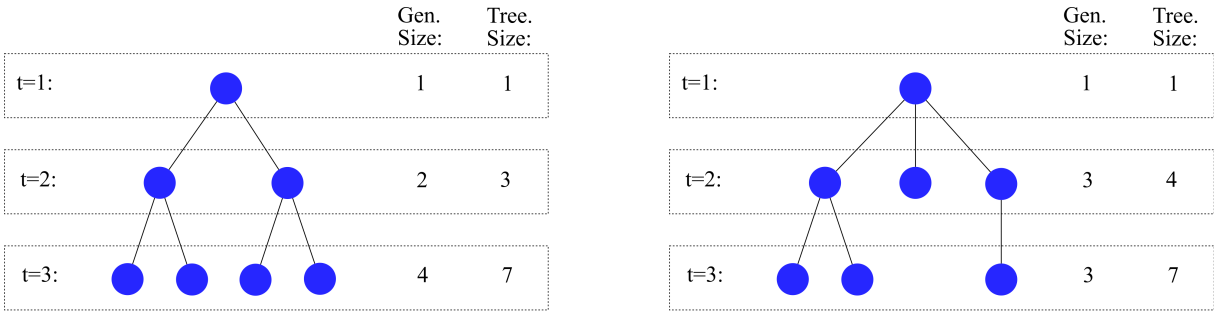

Figure 3.3: Branching process for deterministic (left) and probabilistic (right) offspring.

which they die and produce a number of offspring $k = 0, 1, 2...$ drawn from a probability distribution. Each of those offspring then live for one unit of time, after which they too produce offspring from the same distribution. The offspring distributions of each of the children are independent and identically distributed (*iid*). In the case of evolution, the size of a generation at time a, Z_a, is the result of interest, as a generation size of zero results in extinction. In other applications, the entire tree size, not just the generation size, are of interest. The applications of branching processes have diversified significantly since their first use. Their application in network dynamics include sexually transmitted disease spread (Lashari and Trapman, 2018), epidemiology (Jacob, 2010; Levesque et al., 2021) and information spread (Gleeson et al., 2014; Kapsikar et al., 2021).

In contrast to other deterministic growth models, branching processes and probability generating functions (PGFs) have the useful property of providing detail on the probability of different outcomes with the same control parameters. Figure 3.3 compares a deterministic branching process (comparable to a deterministic exponential growth model) to a single realisation of a probabilistic branching process. Here, each node in the deterministic process will always produce two offspring. For the probabilistic branching process, as long as there is a non-zero probability of producing zero offspring, there also exists a non-zero probability of the process dying out in the first few generations, although the average number of children defined by the offspring distribution may be larger than one.

This is perhaps an early indicator that a branching process may suitably model the behaviour observed in WTM. A key difference between WTM and a branching process approximation is that the branching process can have a maximum of one parent, so at most, one node can contribute to the active fraction of neighbours of the node. We consider the implications of this assumption and quantify them in Chapter 6.

Chapter 4

Analytical Model

We introduced the motivation for probability generating functions (PGFs) and branching processes in Chapter 3, as well as some applications in epidemiology and information spread. This chapter is devoted to introducing the theory of generating functions and branching processes. We conclude the chapter with an asymptotic analysis of the large η long-term solution to the tree size [1] equation.

4.1 Probability Generating Functions

We denote probability generating functions as $f(x)$, to maintain consistency with the probability generating function notation used for offspring distributions in the branching process approximation

$$f(x) = \sum_{k=0}^{\inf} p_k x^k, \tag{4.1}$$

with some useful results of $f(x)$ including the sum of all probabilities equalling 1, written as

$$f(1) = 1, \tag{4.2}$$

the expected value of the distribution is

$$f'(1) = E(X) = \mu, \tag{4.3}$$

[1] This is analogous to cascade size in the WTM.

and we can write an expression for the variance as

$$f''(1) + f'(1) - (f'(1))^2 = V(X). \tag{4.4}$$

These results are equally valid when we apply PGFs to branching processes.

We can recover the probability of a particular event from the PGF in a number of ways. The probability of event η can be written as

$$p_\eta = \frac{1}{\eta!} \frac{d^\eta}{dx^\eta} f(x)|_{x=0}\,. \tag{4.5}$$

For some rare cases, direct differentiation of the PGF is possible, but this is generally not the case. Numerical differentiation is possible, but becomes unstable for large values of η.[2] For more accurate results, Fast Fourier Transform methods can be used (Cavers, 1978). Each of these methods also requires an explicit form of the probability generating function.[3]

4.2 Branching Processes

In Chapter 3, we introduced the Discrete Time Galton Watson process as starting from a single seed which survives for one unit of time, after which it dies and spawns offspring from a probability distribution. Each of these offspring will then survive for one unit of time after which they too will each produce offspring from the same distribution. It is assumed that the offspring distributions are independent of all others in the generation (*iid* assumption).

For the Discrete-time Galton-Watson process, the generation size for generation a is

$$Z_a = f(Z_{a-1}), \tag{4.6}$$

where f is the offspring distribution, with total tree size from Gleeson et al. (2021) as

$$G_a(x) = xf(G_{a-1}(x)). \tag{4.7}$$

[2] In the case of WTM, the probability of a cascade size of 10,000, for example, would require the $10{,}000^{\text{th}}$ derivative.

[3] In our case, we do not have an explicit form of the PGF so we will use an asymptotic analysis to recover the large η, long term cascade size distribution.

4.2.1 Criticality of Branching Processes

Although the mean μ of the branching process tells us little about the size distribution of the trees, it is an important marker of the probability of extinction of the process. The complement of this, the probability that the process never goes extinct, can be considered the equivalent to global cascades in WTM. The mean cascade size at generation a can be written as

$$E(f(a)) = \mu^a, \tag{4.8}$$

where μ is the mean of the offspring distribution, also called the branching number.

- For $\mu < 1$, extinction is definite. Note, $\mu^a \to 0$ as $a \to \infty$. This is known as a sub-critical branching process.
- For $\mu = 1$, extinction is definite, except for the case where $P(\#\text{ of offspring} = 1) = 1$. This is known as a critical branching process.
- For $\mu > 1$, the probability of extinction is less than one. Note, $\mu^a \to \infty$ as $a \to \infty$. This is known as a super-critical branching process. Here, the probability that the process never goes extinct is approximately equal to the probability of global cascades.

It must be noted, when using branching processes to model dynamics on networks, that the degree distribution of the node does not always equal the offspring distribution. With the exception of the seed node, if a node (in a graph) is of degree k, in the branching process representation, one of its k neighbours is from the parent generation. Therefore, it will pass on $k - 1$ offspring to the next generation. We expand on this idea further when introducing the numerical BP in Chapter 5.

4.3 Asymptotic Analysis

Certain assumptions are used to determine the long-term behaviour of cascade sizes. This approximation assumes a constant offspring distribution, and this is not strictly true in our case. We elaborate on the reasons for this in Chapter 5 and discuss its implications in Chapter 6. The asymptotic approximation is also intended for large values of η. The asymptotic analysis of the long-term distribution of large cascade sizes is carried out in a similar manner to that used by Gleeson et al. (2014).

Beginning with Equation 4.7, we assume that

$$G(\infty, x) = xf(G(\infty, x)), \tag{4.9}$$

where $G(\infty, x) = \lim_{a\to\infty} G(a, x)$ and the function f is the offspring distribution.

We make use of Lemma 5.3.2 from Wilf (1994). Let $\psi(x)$ be the PGF for the distribution π_k and suppose ψ has the following asymptotic series as $x \to 1$:

$$\psi(1-w) \sim \textit{analytic part} + \sum_{m=1}^{\infty} c_m w^{\alpha_m} \quad \text{as} \quad w \to 0, \tag{4.10}$$

where $w = 1 - x$ and $\alpha_1 < \alpha_2 < \ldots$ are positive, non-integer powers. The leading order asymptotic behaviour of π_k is

$$\pi_k \sim \frac{c_1}{\Gamma(-\alpha_1)} k^{-\alpha_1 - 1} \quad \text{as} \quad k \to \infty. \tag{4.11}$$

Returning to the tree-size equation, an asymptotic expansion of $G(\infty, x)$ is sought in the form of

$$G(\infty, 1-w) \sim \sum_{m} c_m w^{\alpha_m} \quad \text{as} \quad w \to 0. \tag{4.12}$$

Two cases will be examined, at and near the lower transition point. To begin with, we expand $f(G)$ as a Taylor series around $w = 0$ to give

$$1 - \mu\psi(w) + \frac{f''(1)}{2}\psi(w)^2 \ldots, \tag{4.13}$$

noting that $f(1) = 1$ (Equation 4.2) and $f'(1) = \mu$ (Equation 4.3). We take this opportunity to reiterate that μ is the mean of the offspring distribution, which differs from z, the mean degree of the network.

4.3.1 Case 1: Critical Branching Process

The first case to be examined is the critical branching process which occurs at $\mu = 1$. We begin with the equation for tree size

$$G_\infty(x) - xf(G_\infty(x)) = 0 \tag{4.14}$$

and substitute in the Taylor expansion of $f(G)$ from Equation 4.13, $(1-w)$ for x and $(1-\psi(w))$ for G. This gives

$$1-\psi(w)-[(1-w)(1-\mu\psi(w)+\frac{f''(1)}{2}(\psi(w))^2)]=0. \tag{4.15}$$

At the critical point, where $\mu=1$, we have

$$(w-1)\frac{f''(1)}{2}(\psi(w))^2-w\psi(w)+w=0. \tag{4.16}$$

Substituting in $-c_1w^\alpha$ for $\psi(w)$, we balance the leading order terms of w with $\alpha_1=\frac{1}{2}$, then find $c_1=(\frac{f''(1)}{2})^{-\frac{1}{2}}$.

Using Lemma 5.3.2 from Wilf (1994), the probability distribution π_η can then be recovered as

$$\pi_\eta \sim B\eta^{-\frac{3}{2}}, \tag{4.17}$$

where B is given as

$$B=\frac{c_1}{\Gamma(-\alpha)}=\frac{(0.5f''(1)\pi)^{-\frac{1}{2}}}{2}. \tag{4.18}$$

This analytically reproduces the power-law distribution with $-\frac{3}{2}$ exponent Watts' observed numerically, introduced in Figure 3.2 (right).

4.3.2 Case 2: Near-critical Branching Process

Next we will look at the case of a near-critical process. We now examine a point offset by a small distance δ from the critical point such that $\mu=1+\delta$. Taking Equation 4.15 and expanding, we form a quadratic equation in ψ, retaining terms of order ψ, ψ^2 and w but not ψw, with solution

$$\psi(w)=\frac{-\delta+\sqrt{\delta^2-4w(\frac{-f''(1)}{2})}}{-f''(1)}, \tag{4.19}$$

with a branch point in the complex x plane at

$$\beta=1+\frac{\delta^2}{2f''(1)}. \tag{4.20}$$

Using the result from Gleeson et al. (2014), we obtain an expression for large η, long term asymptotic form of the cascade size distribution as

$$\pi_\eta \sim B\eta^{-\frac{3}{2}} e^{-\frac{\eta}{\kappa}}, \tag{4.21}$$

where B is defined as in Equation 4.18 and κ is defined as

$$\kappa = \frac{2f''(1)}{\delta^2}. \tag{4.22}$$

To find the actual cascade size distributions, we must find the corresponding value of δ for the near-critical case, and $f''(1)$ for the critical and near-critical cases. At the lower critical point of the branching process (near $z = 1$), almost all nodes in the network become vulnerable when a single neighbour is active. As a consequence of this, the offspring distribution at the lower critical point is well approximated by the Poisson excess degree distribution defined in Equation 2.11. For the near-critical case, as the mean degree increases, the assumption that the offspring are approximately Poisson distributed breaks down. We must then calculate δ and $f''(1)$ from the numerical model using Equations 4.3 and 4.4 respectively.

Chapter 5

Numerical Models

This chapter will deal solely with the numerical implementation of WTM and the branching process approximation. A brief introduction to network-specific functions in MATLAB is also given.

5.1 Introduction to Networks in Matlab

With the exception of the three MATLAB functions described later in the chapter, there are two network-specific built-in functions we make use of. The *graph()* function creates a sparse graph object from a number of input types including adjacency matrix and node/edge lists. We also make use of the *conncomp()* in-built function, for calculating the size of the connected component each node is a member of. From this we can extract the largest connected component of the graph. We refer the reader to the accompanying MATLAB script for a more thorough introduction to working with network structures in MATLAB, if they so wish. [1]

5.2 Numerical Watts' Threshold Model

Numerical simulation allows for the analysis of stochastic processes in the event where an analytical solution is not available or not possible. The programming of a stochastic process such as the WTM is conceptually straightforward to implement but faces limitations in terms of computational cost which will be discussed in context below. A deterministic simulation is one where, for a given set of inputs, there is no variation in the outputs.

[1] `https://github.com/mikelyynch/A-Branching-Process-Approximation-of-Complex-Contagion-on-Random-Networks`

For stochastic simulation or a probabilistic model, repeated trials with the same control parameters will result in the output taking different values which, when repeated a sufficient number of times, will form a probability distribution of outputs. As the number of repetitions increases, the observed distribution of the outputs tends towards the actual distribution of the outputs. In the numerical WTM we have two sources of randomness. The creation of a given random graph (e.g. $G(n, m)$) for a given set of parameters does not produce a unique graph and so will produce different dynamics to some extent. Similarly, the placement of the initial seed node may have a large impact on the dynamics. [2] If the initial seed is placed on a node of degree zero, the phenomenon will not spread.

For all of the results presented, a focus will be given to homogeneous thresholds, whereby all nodes in the network have the same threshold[3] although it is straightforward to adapt the script to produce many different threshold distributions. Again, as in Section 3.2.1, we reiterate that this simplification is deliberate and a prerequisite of understanding more complex models.

5.2.1 Random Graphs

In terms of MATLAB implementation, there are roughly equivalent implementations for the $G(n, p)$ and $G(n, m)$ realisations of the Erdos-Renyi random graph.

Two important differences lie in how these commands are executed. In the first case $G(n, p)$, the function binornd() cycles through each element of the $n \times n$ matrix, assigning 1 or 0 with probability p. In the $G(n, m)$ case, two vectors of length $\frac{zn}{2}$ are created. So, the computational cost in the $G(n, p)$ case scales as $O(n^2)$ while the computational cost in the $G(n, m)$ case scales as $O(n)$. We compare computational costs for each method in Figure 5.1.

As well as wall time, the $G(n, m)$ method scales sufficiently well in terms of system memory requirements. The $G(n, p)$ method reaches system memory capacity at approximately $n = 20,000$, whereas the $G(n, m)$ method can produce sparse graphs of over $1,000,000$ nodes. This allows for suitable testing of the impact of finite graph size on the results. We present the MATLAB implementation of the $G(n, m)$ method overleaf.

[2] The placement of the seed only matters when the graph structure is unchanged between repetitions.

[3] This is consistent with Watts who restricted most of his analysis to the homogeneous threshold case of $\phi = 0.18$.

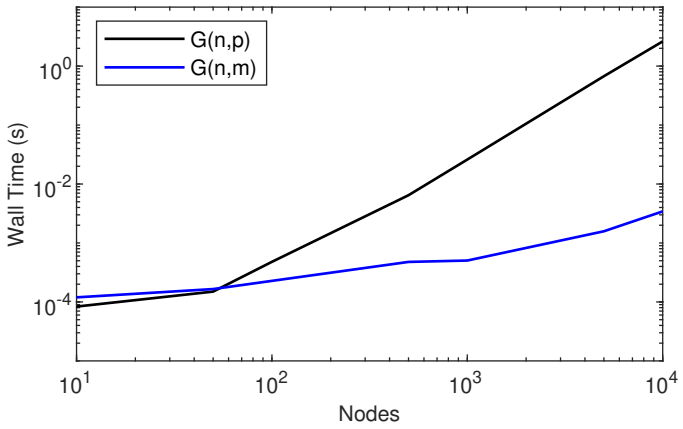

Figure 5.1: Computational cost for $G(n,p)$ and $G(n.m)$ Poisson random graphs.

```
function [A]=ER(n,z)
%Create two vectors of length 0.5*z*n uniformly from all possible nodes
v1=randi([1 n],1,round(z*n/2));
v2=randi([1 n],1,round(z*n/2));
%Loop to remove self loops-not strictly neccessary but reasonably low ...
    cost
i=1;
while i≤n*z/2
    if v2(i)==v1(i)
        v2(i)=randi([1 n]);
    else
        i=i+1;
    end
end
%Create graph from node lists v1 and v2
%Note: ones() vector in position 3 below creates unweighted graph
G=graph(v1,v2,ones(1,length(v1)),n);
A=adjacency(G);
end
```

5.2.2 Updating Process

As inputs, we require information on the adjacency matrix upon which the model is applied, the number of nodes and the threshold. The number of active neighbours can be

found, as illustrated by Porter and Gleeson (2014), with the following expression

$$AN = A \times S, \tag{5.1}$$

where the adjacency matrix A, multiplied by a vector S which contains the states of each node ($S_i = 0$ if node i is inactive, $S_i = 1$ if node i is active) returns the number of active neighbours of node i in the vector AN_i. This removes the need to individually check each node, avoiding a costly $O(n)$ loop. We then simply test the vulnerability condition of each node, updating to active if vulnerable, then repeating for each time-step until a steady state is reached. Below we present the MATLAB implementation of WTM.

```
%Function inputs adjacency matrix, network size and homogeneous threshold
function[SUM, tfin]=WattsTh(A,n,phi)
%Randomly assign seed node
S=zeros(n,1);seed=[randi(n,1,1)];
S(seed)=1;
%Assign thresholds from distribution (heterogeneous) or constant ...
    (homogeneous) as in this case
T=ones(n,1)*phi;
%We wish to iterate until dynamics reach steady state ie no new ...
    adopters in a generation, so define conditions for loop as new ...
    sum of adopters being greater than sum of adopters from last ...
    iteration
Snew=1;Sold=0;t=0;

while Snew>Sold
    AN=A*S;          %active neighbours
    TN=sum(A);       %total neighbours
    Supd=AN./TN'> T;%Boolean operation for vulnerability conditon
    S=(S+Supd)>0;    %Boolean operation to remove duplicate active ...
        nodes, or combine active nodes from previous generation with ...
        current generation
    Sold=Snew;       %update values for next time step
    Snew=sum(S);     %update values for next time step
    t=t+1;           %increase time index by 1
    SUM(t)=sum(S);   %assign total number of active nodes at time=t
end

tfin=t;              %total time to steady state
end
```

5.3 Numerical Branching Process Approximation

As discussed in Chapter 4, branching processes that lend well to full analytical derivation are the exception, rather than the rule, so a first step in understanding the dynamics of the branching process is a numerical stochastic model.[4] As with the WTM, each time the model is run represents one potential realisation. Repeated realisations give an idea of the probability and distribution of different outcomes.

5.3.1 Offspring Distribution

We demonstrate the process of generating the offspring distribution for the numerical BP approximation with the aid of Figure 5.2, which represents a single realisation of the first two generations. The seed node (node 1 on the diagram) is chosen uniformly at random. So, the probability of node 1 having degree k and k offspring is equal to the degree distribution of the network. In our example, node 1 has a randomly generated degree of 3, and so has three potentially active offspring labelled 2, 3 and 4. We must now randomly generate the degree for each of these nodes to assess whether they are vulnerable and become active or not vulnerable and are discarded. Unlike the seed node, which was an exterior node, all subsequent nodes are interior nodes and so their degree is conditioned on the fact that they are already connected to a node. So, their degree distribution follows the excess degree distribution defined in Equation 2.11.

- For node 2, we have randomly generated a degree of four. This means node 2 meets the vulnerability criteria and so will become active (filled in blue). Node 2 is of degree four but one of its neighbours is from the parent generation so it will produce $k-1$ offspring. These $k-1$ children are retained to check their vulnerability in the next generation.

- For node 3, we have randomly generated a degree of one. This means node 3 meets the vulnerability criteria and so will become active. Node 3 is of degree one but its only neighbour is from the parent generation so will produce no offspring.

- For node 4, we have randomly generated a degree of six. This means node 4 does not meet the vulnerability criteria and will not become active. Its $k-1$ children are discarded.

[4] We have derived an analytical solution for the long term large cascade size, but a stochastic model provides more information about the exact process, as well as provides a benchmark against which the asymptotic result can be compared.

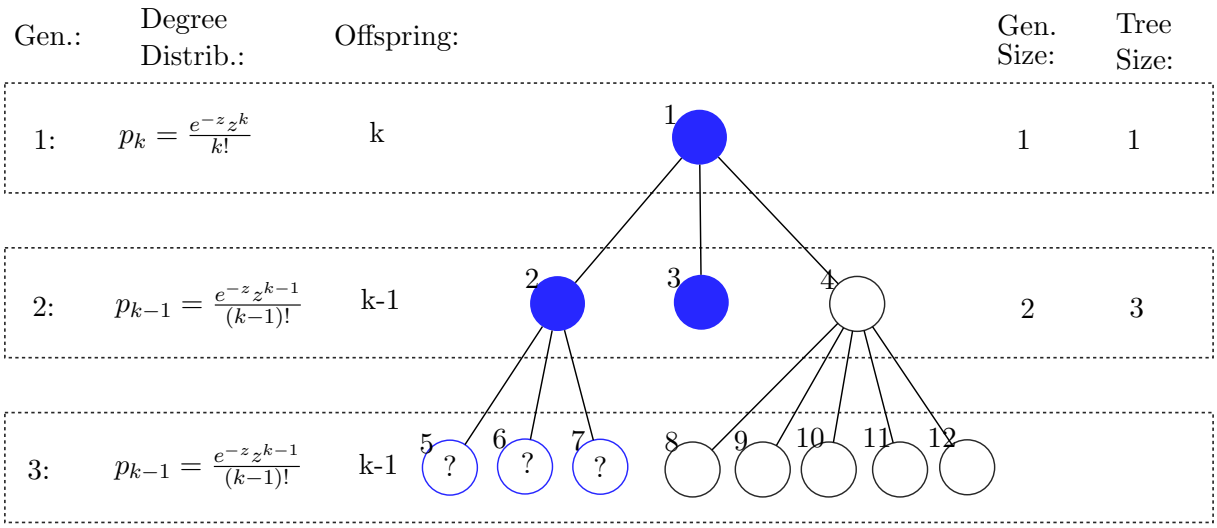

Figure 5.2: Offspring distribution for branching process (active=filled blue,inactive=unfilled black, yet unknown=unfilled blue).

We then repeat this process for the next generation with nodes $5, 6, 7$, noting that nodes $8 - 12$ will not become active because their parent was not vulnerable.

As a consequence of the offspring being *iid*, only one branch of the tree can contribute to vulnerability of a node. In WTM, this is not the case. In a graph, especially after a large fraction of the graph has been activated, it is possible that two or more active nodes could contribute to the activating of a high degree node whose vulnerability condition cannot be satisfied by the contribution of a single active node. The implications of this are discussed in Chapters 6 & 7.

5.3.2 Updating Process

Unlike the numerical implementation of WTM, the BP model is not naturally restricted by a finite graph size. Consequently, in the case of a super-critical branching process, the tree size will grow indefinitely. To ensure reasonable computational time, the maximum tree size is restricted to a cut-off point sufficiently large so that it can be assumed, if reached, the probability of extinction is negligibly small. Consistent with this requirement, and to make comparison with WTM easier, a maximum tree size of $10,000$ was chosen as a sensible cut-off. We present the MATLAB implementation of the numerical BP approximation below, noting that the offspring for the seed node is calculated outside

the of the time-step loop.[5]

```
function [treesize,gensize]=bpa(z,phi,upper_limit,tmax)
T=1/phi;  %Convert homogeneous threshold to limit for highest ...
    accessible node degree
%Defining generation 1 and vector sizes
gensize=1; gensize(2:tmax)=0;
treesize=1; treesize(2:tmax)=0;
%Exterior node offspring distribution
CoA=poissrnd(z);%CoA=children of active ndoes
%loop in time for interior nodes
for t=2:tmax
    if max(treesize)>upper_limit
        treesize(t:end)=upper_limit;
        break
    else
    end
    %Excess degree distribution
    k=poissrnd(z,CoA,1); k=k+1;
    %Tree size: tree size up to previous generation+current generation
    treesize(t)=treesize(t-1)+sum(k<T);
    %Children of active nodes for current generation
    CoA=sum(nonzeros(k.*(k<T))-1);%Boolean restricts contribution to ...
        nodes below threshold
    gensize(t)=sum(k<T);
end
end
```

5.4 Model Sensitivity

The main limitation of numerical stochastic models is the computational cost. This is particularly true when we wish to compare with analytical results which assume infinite graph size. As we have seen in Section 5.2.1, the efficiency of algorithms is of utmost importance. A similar attention to efficiency is required when programming the dynamics on the graph. The two mains drivers of computational cost in the context of stochastic models on random graphs are the system size and number of repeated realisations.

[5]For the numerical BP, the use of Boolean operations makes for efficient scripting. If we were to extend the model to a heterogeneous threshold case, the script would need to be altered.

5.4.1 Finite System Size

Most of the theoretical arguments based on the Watts model assume infinite network size. As this is physically not possible to implement in a numerical model, the aim is to find an optimum balance between maximising network size (and minimising impact of finite network size) and maintaining a reasonable computational cost. The majority of simulations shown in Chapter 6 were carried out with system sizes of $10,000$ nodes, in line with Watts (2002). There are a number of instances where discrepancies in results are observed, particularly in the tails of distributions. In these cases, an effort is made to quantify the impact of network size on the results. we refer the reader to Appendix A for a sensitivity analysis.

The numerical implementation of the BPM is also susceptible to finite system size but to a much lesser extent than the numerical WTM. It is expected that this is mainly due to the assumption of *iid* offspring. The main limitation on the branching process numerical model is that the upper limit is placed sufficiently high so that it falls outside of the tail of the distribution of cascades which go extinct.

5.4.2 Repeated Realisations

The basis behind numerical stochastic models is to repeatedly simulate the process such that the sample distribution is representative of the true distribution. This can be recognised as when we see little change in the distribution of results from one set of simulations to the next. We take care to ensure a sufficient number of repetitions are carried out for each set of results so that the results presented are representative of the true distribution.

Chapter 6

Results

We begin the chapter by refreshing some results introduced in Chapter 3, as well as extending them, using the numerical Watts' threshold model (WTM). These will then be compared to results from the numerical branching process (BP) model and branching process asymptotic analysis.

6.1 Numerical Watts' Threshold Model

Beginning with Watts' cascade window for values of ϕ (homogeneous threshold) and z (mean degree) shown in Figure 3.1, we notice a clear step-like behaviour. This is an intuitive consequence of the homogeneous threshold. Taking a homogeneous threshold of $\phi = 0.18$, the highest degree node which can be activated is five, but this is also true for thresholds of 0.17 & 0.19 etc. In Figure 6.1, superimposing the fraction of active neighbours for nodes of degree $4, 5, 6...10$ with one active neighbour, we see good agreement between these superimposed lines and the locations of the steps.[1] More importantly, we do not see any additional steps which may be formed by two neighbours contributing to the activating of a node.[2] This is the first suggestion that a branching process approximation may suitably approximate WTM.

For the remainder of the chapter, we restrict our analysis to the case where $\phi = 0.18$, in line with Watts' paper. Watts describes two transition points at either end of the cascade window, where we have already defined a transition point as the point in terms of

[1] Recall the vulnerability condition introduced in Chapter 3.

[2] We take this to mean that the dynamics are dominated by events where a single neighbour contributes to the activating of a node. This does not mean that events do not occur where two neighbours contribute to the activating of a node, simply that it occurs in a minority of cases. We discuss this further when studying the results of the branching process approximation.

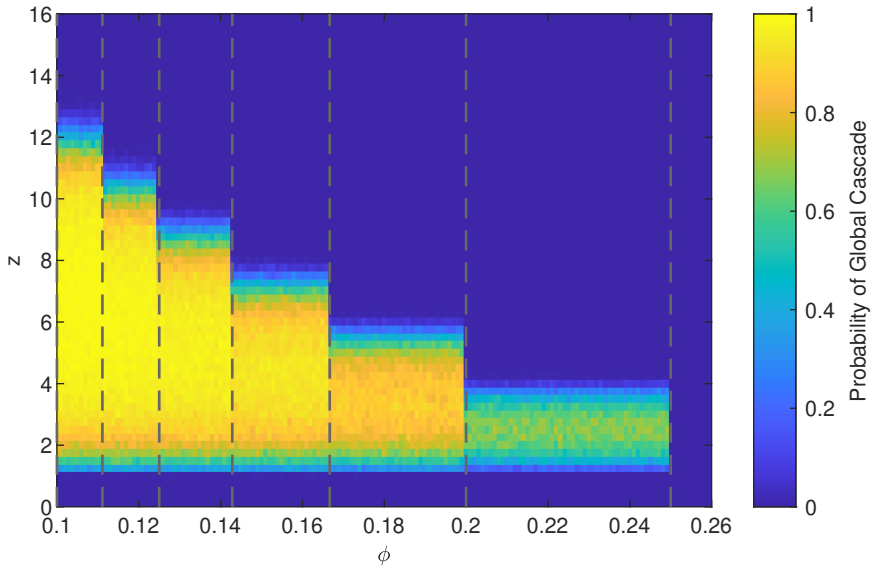

Figure 6.1: Cascade window with fraction of active neighbours with one active neighbour superimposed.

z where global cascades begin or cease to occur. Focusing on cascade size and frequency originally introduced in Figure 3.2 (left), we reproduce the data from Watts' paper and plot the global cascade probability and cascade size graphs separately in Figure 6.2. Both global cascade probability and mean cascade size undergo continuous transitions at the upper and lower end of the cascade window. In terms of maximum cascade size, there is a continuous transition at the lower transition (near $z = 1$) and a discontinuous transition at the upper transition (near $z = 6.5$). We also note the agreement between maximum cascade size and largest connected component, as observed by Watts.

We now display some additional plots to emphasise some features of the cascade size distributions. Watts limited plotting of cascade frequency and size distribution to global cascade frequency, maximum global cascade size and mean size (of all cascades) across a range of values for z. We can, to get a more complete understanding, plot the probability of a given cascade size as a third graph dimension either using a two dimensional heatmap or surface plot. We may think of this as individual histograms plotting cascade size probability, for a range of z, being placed consecutively.

Figure 6.3 (left) and (right) essentially show the same information in a different format. We have reduced the resolution in Figure 6.3 (left) intentionally, as the actual cascade size distribution is of such a low variance around the zero point or largest connected component that they become difficult to read. As with Figure 6.2, one can see a continuous transition

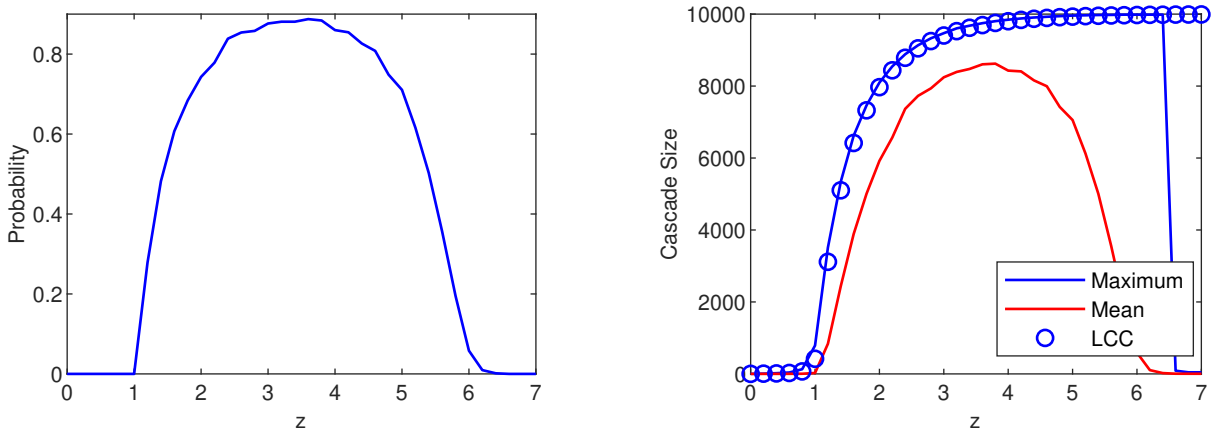

Figure 6.2: Cascade size and probability for a range of z. Global cascade probability (left). Maximum and mean cascade size and largest connected component (right).

in global cascade size at low z and discontinuous transition in global cascade size at high z. This format highlights the continuous nature of the transition in global cascade probability at both lower and upper transitions. Also notable from these plots is the purely bimodal nature of the size distribution inside the cascade window. Both Figure 6.3 (left) and (right) show the cascade sizes either very small, or filling the giant component with low variance. There are no intermediate cascade sizes. This "all or nothing" type behaviour is the second suggestion that branching processes may suitably approximate the WTM.

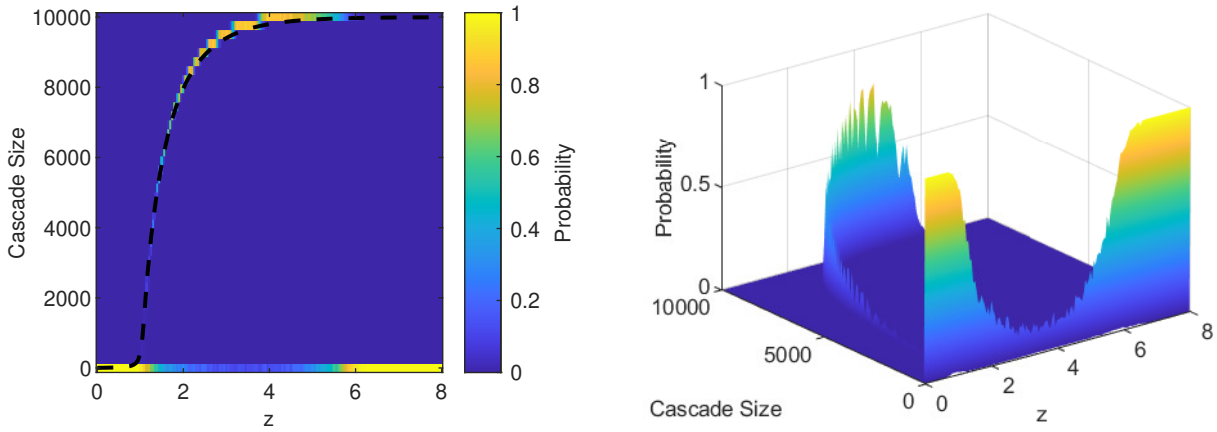

Figure 6.3: Distribution of cascade sizes and probability for numerical WTM shown as heatmap (LCC=dashed line) (left) and surface plot (right).

6.2 Numerical Branching Process

At this point it is worth noting some key differences between results of the numerical WTM and numerical BP approximation. In the WTM we see variability within cascade size for a given z which is quite clearly bimodal for most of the cascade window. Inside the lower end of the cascade window, the largest cascade sizes are limited to a fraction of the graph significantly smaller than the total graph size, and we have already discussed how this is limited by the largest connected component. We must understand the distinction between a phenomenon that has died out and one that has filled near to or all of the available graph. As we approach the lower transition point near $z = 1$ from the upper side, it becomes more difficult to distinguish these in the numerical WTM. In the branching process model, the spreading phenomenon is not restricted by a finite graph or connected component size so will either die out or tend to infinity.[3] When viewed like this, it may be arguable that the BP model is discontinuous in global cascade size at both the lower and upper end of the cascade window, as shown in Figure 6.4. Again, we have reduced the resolution in Figure 6.4 (left) intentionally, as the actual distribution of cascade sizes are of such low variance around the zero point or largest connected component that they become difficult to read.

In terms of criticality of the branching process, the branching number can be calculated for each value of z. In a method consistent with that used by Gleeson et al. (2021), if $Y_{m,a}$ is defined as offspring m in generation a where $m = 1, 2, 3....M$, and Z_a is the sum

[3] In our numerical implementation, we have placed a maximum limit of $10,000$ on the cascade size due to computational limitations.

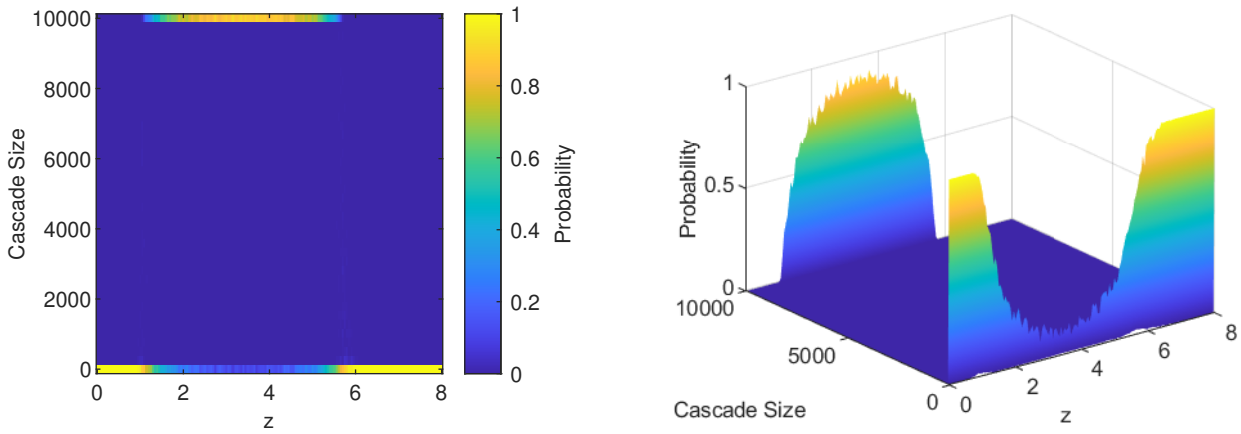

Figure 6.4: Distribution of cascade size and probability for branching process model shown as heatmap (left) and surface plot (right).

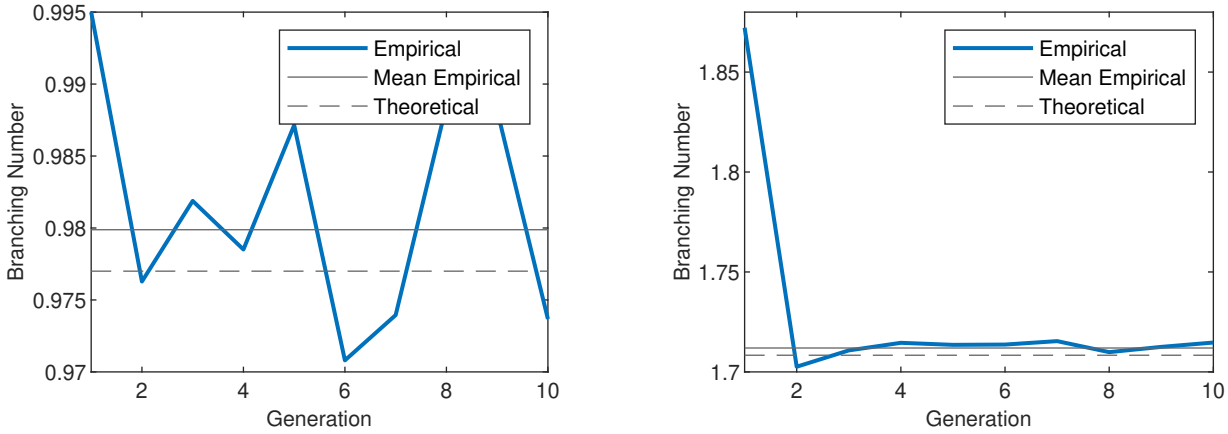

Figure 6.5: Branching number for each generation for $z = 1$ (left) and $z = 2$ (right), noting the difference in y-axis scales.

of all nodes in generation a, written as

$$Z_a = \sum_{m=1}^{M} Y_{m,a}, \tag{6.1}$$

then the empirical branching number for generation a is given as

$$\mu_a = \frac{Z_{a+1}}{Z_a}. \tag{6.2}$$

From this we can construct a cascade window for the branching process model.

Plotting the branching number for the first ten generations in Figure 6.5, we see the anomalous branching number for generation one in the $z = 2$ case, which is the branching number between first and second generations. As shown in Figure 5.2, the first generation of the branching process, made up of a single exterior node, has a different offspring distribution to all subsequent generations as interior nodes. Consequently, this is removed from the calculation of the average branching number for the process. Interestingly, we do not see this occurring in the $z = 1$ case.

Figure 6.6 shows there is excellent agreement between the empirical branching number from the simulated branching process and the theoretical branching number derived by Watts (2002) for infinitely small seed size.[4] In a manner consistent with the BP approximation we propose, the theoretical branching number here only consider one ac-

[4] A theoretical branching number was later derived by Gleeson and Cahalane (2007) for finite seed size.

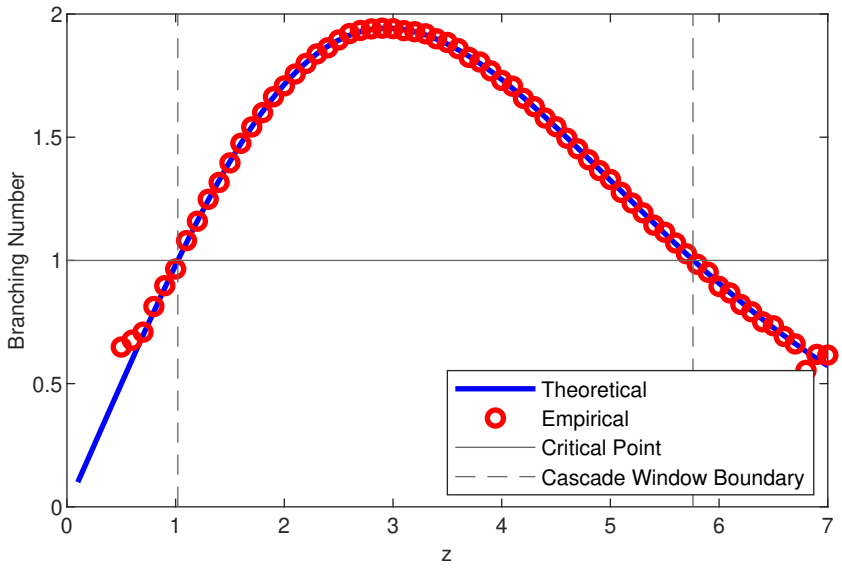

Figure 6.6: Empirical and theoretical branching number cascade window.

tive neighbour contributing to the vulnerability condition of a node. We also note here that unlike the Poisson graph, which undergoes a transition of connectivity at $z = 1$, the branching process does not become critical until just after $z = 1$, at approximately $z = 1.02$. We do not investigate this observation thoroughly, apart from noting that the criteria for a giant component (Molloy and Reed, 1995) in the Poisson graph differs from the cascade condition derived by Watts (2002). We return to this observation briefly in Chapter 7.

Comparing the numerical WTM with the numerical BP approximation in Figure 6.7, excellent agreement is observed near the lower transition. Some differences are observed between the models at the tail of the distribution, but these differences are amplified by the logarithmic scaled axes. We also refer the reader to Appendix A where we show that the results for the numerical WTM and numerical BP approximation converge as the graph size and number of repetitions increase. In terms of the upper transition point, observing cascade window sizes for the $\phi = 0.18$ case for the numerical WTM (Figure 6.2) and the numerical BP approximation (Figure 6.6), we notice the two models differ in terms of location of the upper transition point. The upper transition for the numerical BP approximation occurs at $z \approx 5.76$, compared with $z \approx 6.5$ for the numerical WTM. For clarity, we will refer to these as the BP upper transition and the WTM upper transition points respectively.

We have already discussed that one of the most notable differences between the WTM

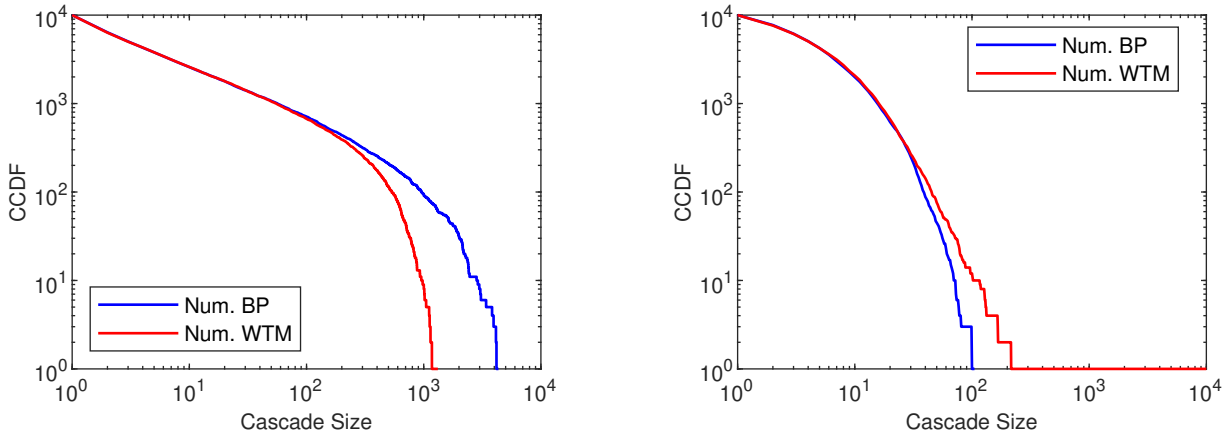

Figure 6.7: Complementary cumulative distribution (CCDF) of numerical WTM and numerical BP approximation at $z = 1$ (left), and $z = 6.5$ (right).

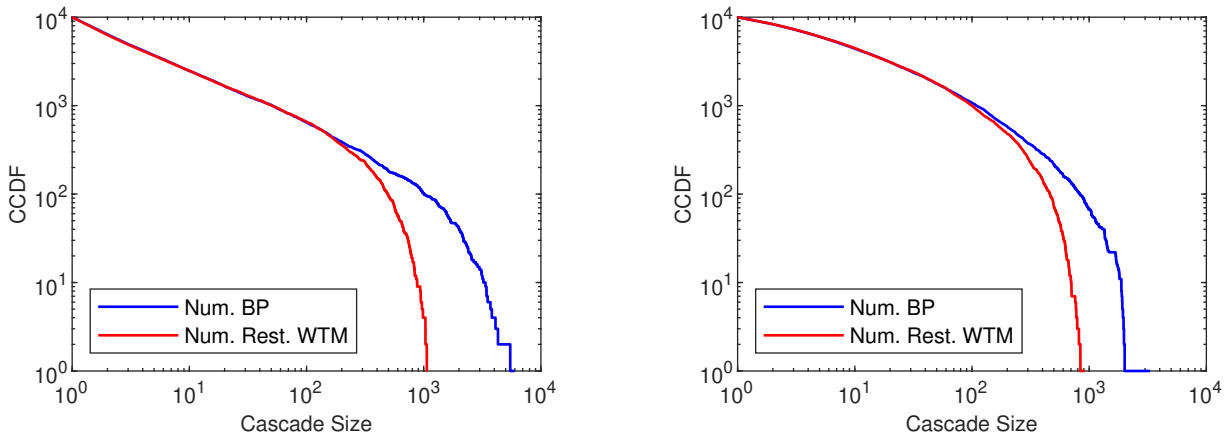

Figure 6.8: Complementary cumulative distribution of restricted numerical WTM and numerical BP model at $z = 1$ (left), and $z = 5.76$ (right).

and the BP approximation is the *iid* assumption of the BP approximation. In practice, where a high degree node in the WTM may be activated by multiple active neighbours, only one active neighbour may ever contribute to the activating of a node in the BP approximation.[5] Graphs of higher mean degree (z) contain more high degree nodes that can never be activated by a single active neighbour. It follows from this that the impact of this *iid* assumption will be greater at higher values of z and it would be reasonable to assume that this would explain the discrepancies in the results.

Conveniently, it is straightforward to restrict the numerical WTM so that, at most, only one active neighbour can contribute to the activating of a node, effectively allowing

[5]Again, recall the vulnerability condition introduced in Chapter 3.

this hypothesis to be tested. Figure 6.8 shows good agreement at both the lower and BP upper transition point. This effectively confirms the hypothesis that at the lower end and middle of the cascade window, the numerical BP approximation shows good agreement with the WTM because the majority of activation events are due to the contribution of a single active neighbour. This no longer holds after approximately $z = 5.5$ (not shown), after which events where multiple active neighbours contribute to the activation of a node begin to play a significant role in causing global cascades.

6.3 Branching Process Asymptotic Analysis

In this section, we restrict our comparisons to those between the asymptotic BP approximation and the numerical WTM, as we have already determined that the numerical BP approximation shows excellent agreement with the numerical WTM at, and near, the lower transition point. We have also shown that the numerical BP approximation does not show good agreement with the numerical WTM at the upper end of the cascade window, so we restrict our analysis further to the lower transition point and lower end of the cascade window. When presenting results, to make comparison with the numerical WTM and BP approximation easier, the probability of a global cascade (the complement of the probability that a cascade which will become extinct in a finite number of generations) is assigned to the position of $\eta = 10,000$.

6.3.1 Critical Branching Process

The following cascade size distribution is the analytical equivalent to Watts' result introduced in Figure 3.2 (right). We begin by looking to solve Equation 4.17 with constant B from Equation 4.18. Here, $f''(1)$ is the only undefined variable, and may be calculated analytically, taking certain assumptions into account, or numerically, from Equation 4.4. For the critical case, at $z \approx 1$, with $\phi = 0.18$, most nodes in the network become vulnerable when a single neighbour is active. So, the offspring distribution is approximately Poisson, with the result that $f'(1) = \mu$ and $f''(1) = \mu^2$. As this occurs at the critical point, $\mu \approx 1$, and so are $f'(1)$ and $f''(1)$.

Figure 6.9 shows excellent agreement between the numerical WTM and the asymptotic approximation. This is consistent with the $-\frac{3}{2}$ power law distribution of cascade sizes that Watts observed at the lower transition point. In Figure 6.9 (left), we observe some noise at the tail of the distribution. This is a regular feature of a simulated power-law distribution

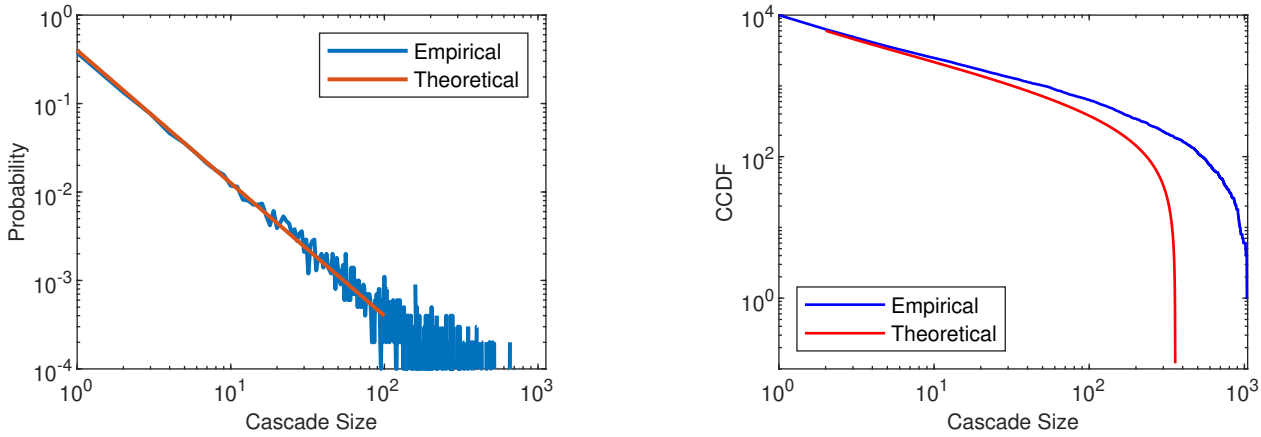

Figure 6.9: Probability density (left) and CCDF (right) for numerical WTM and asymptotic BP approximation at the critical point.

(Newman, 2005) and may be made more aesthetically pleasing by using logarithmically scaled bins to plot the probability of outcomes. Here we intentionally leave it as is, as a reminder of this limitation of numerical stochastic models. In Figure 6.9 (right), we plot the power-law distribution in an equivalent format to Figure 3.2 (right) we introduced from Watts' paper for direct comparison. The error at large cascade sizes is emphasised by the logarithmic scaling, but good agreement is observed for most cascade sizes. We note that the result is sensitive to small changes in $f''(1)$.

6.3.2 Near-critical Branching Process

For the near-critical branching process, the model assumes that δ is a small parameter. We plot two scenarios to assess how quickly the accuracy of the solution diminishes as δ increases. Again, we restrict our comparisons to those between the numerical WTM and the asymptotic BP approximation.

For Figure 6.10 (left), we see good agreement between the numerical WTM and asymptotic BP approximation at a value of $z = 1.12$. In Figure 6.10, at a value of $z = 2.02$ we see a larger error between the numerical WTM and asymptotic BP approximation. Here, most of the error is incurred in the first generation. An assumption of our asymptotic analysis was that the approximation holds for large cascade sizes and this may impact the result here. In Figure 6.5 (right), we showed that the branching number in the first generation is different to the mean of successive generations. We noted that the magnitude of this anomaly increased as we moved away from the critical point and this effect may also impact the results here. As with the critical branching process, we note the results

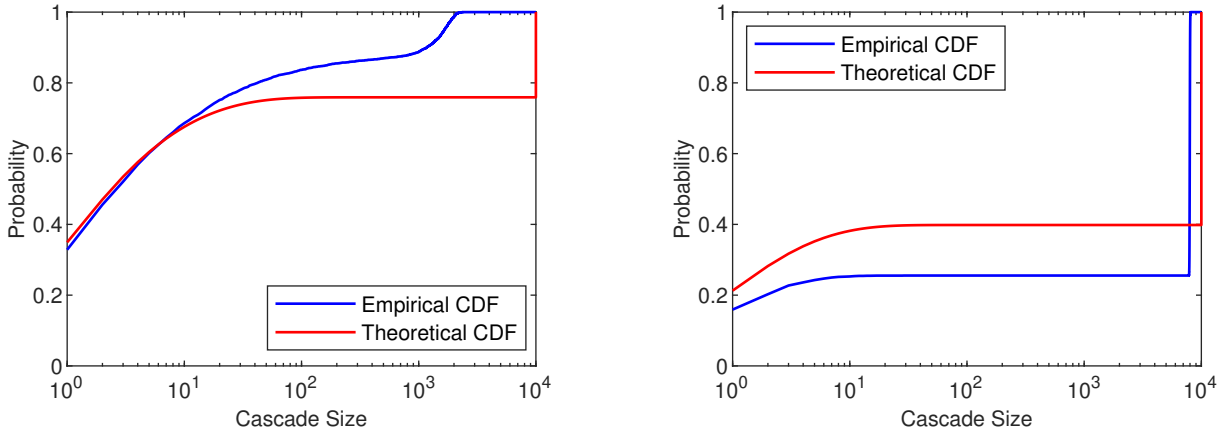

Figure 6.10: Cumulative cascade size probability distributions for $z = 1.12$ and $z = 2.02$.

are highly sensitive to changes in $f''(1)$, and some error is likely to have been incurred in the calculation of this parameter.

Chapter 7

Conclusion

We developed a numerical and theoretical novel branching process (BP) approximation to Watts' threshold model (WTM) complex contagion (Watts, 2002), modelling the size distribution of cascades for the homogeneous threshold case, acting on Poisson random networks.

By implementing a numerical WTM, we expanded on Watts' results to show that a) we see the step change in the cascade window occurring at values of $\phi = \frac{1}{4}, \frac{1}{5}...\frac{1}{10}$ (Figure 6.1), indicating that the contagion process may be dominated by events where a single active neighbour is sufficient to activate a particular node, and b) inside the cascade window, cascade sizes are bimodal; the cascade either dies out very quickly or fills all or almost all of the largest connected component of the graph (Figure 6.3). These two factors indicated that a branching process model may approximate the WTM on at least some of the cascade window.

Beginning with a numerical BP approximation, we conceptually derived the offspring distribution in Figure 5.2. We see excellent agreement between the numerical WTM and numerical BP approximation at the lower end and middle of the cascade window. At the upper end, the results begin to diverge and the numerical BP approximation undergoes an upper transition at $z \approx 5.76$ compared with $z \approx 6.5$ for the numerical WTM.

Hypothesising that the most prominent difference between the numerical WTM and numerical BP approximation is the independent and identically distributed (*iid*) assumption of the BP approximation, we restricted the numerical WTM in such a way that, at most, one neighbour could contribute to the activation of a node. This then shows excellent agreement in cascade size distribution at the upper end of the BP cascade window and BP upper transition point location (Figure 6.8).

Finally, we carried out an asymptotic analysis of the long-term large cascade size

distribution at, and near, the lower critical point. Again, we see excellent agreement for cascade size distribution at the lower transition (critical) point between the numerical WTM and asymptotic BP approximation (Figure 6.9). For the near-critical case, we see good agreement close to the critical point ($z = 1.12$), which diminishes as we move away from the critical point ($z = 2.02$) (Figure 6.10).

We noted that the transition point for connectivity of the Poisson graph and the critical point of the branching process were close, but not exactly coincident, resulting in a continuous or near-continuous transition at the lower boundary of the cascade window of the numerical WTM. This raises the question of whether a different network structure might display a larger gap in terms of mean degree z between the transition point for connectivity of the graph and onset of global cascades, thus forcing a discontinuous transition in global cascade size at the lower transition point also.

We propose adapting the numerical BP approximation to consider the probability of two branches joining to activate a high degree node, acknowledging that this would violate the *iid* assumption of the current BP approximation. We speculate that if this change were implemented, and we calculated the empirical branching number (Equation 6.2) near the WTM upper transition point, we would see that the branching number would not remain constant for successive generations, as was the case in the current model[1] and would in fact increase as the probability of joining branches with another active node increased also, changing the BP from sub-critical or critical to super-critical at some point in time. As we have already restricted the numerical WTM to better match the numerical BP approximation, and show excellent agreement between these models across the lower and upper BP transition points (Figure 6.8) it seems a worthwhile exercise to see whether an adapted numerical BP approximation could provide good agreement with the numerical WTM at the WTM upper transition point. We are unsure at this time how this mechanism could be transferred to an analytical BP approximation.

Finally, we restricted our study to Poisson graphs, which have a low clustering coefficient for large n (Chapter 2). Having demonstrated that the BP approximation shows good agreement with the numerical WTM on Poisson graphs for most of the cascade window, a natural progression from this would be to test the agreement on different graph structures. Melnik et al. (2011) showed that tree-like theories, like those presented in this thesis, can work well even on networks with clustering, so we remain cautiously optimistic that this approximation will transfer to real-world network structures.

[1]The branching number did differ for generation one, the exterior node, but was constant for successive generations, see Figure 5.2 and associated explanation.

Bibliography

Cavers, J. K. (1978), 'On the Fast Fourier Transform Inversion of Probability Generating Functions', *IMA Journal of Applied Mathematics* **22**(3), 275–282.

Darrow, W. W., Potterat, J. J., Rothenberg, R. B., Woodhouse, D. E., Muth, S. Q. and Klovdahl, A. S. (1999), 'Using Knowledge of Social Networks to Prevent Human Immunodeficiency Virus Infections: The Colorado Springs Study', *Sociological Focus* **32**(2), 143–159.

De Domenico, M., Lima, A., Mougel, P. and Musolesi, M. (2013), 'The Anatomy of a Scientific Rumor', *Scientific Reports 2013 3:1* **3**(1), 1–9.

Dearing, J. W. and Cox, J. G. (2018), 'Diffusion Of Innovations Theory, Principles, And Practice', *Health Affairs* **37**(2), 183–190.

Doherty, O. (2018), 'Informational Cascades in Financial Markets: Review and Synthesis', *Review of Behavioral Finance* **10**(1), 53–69.

Felberg, R. A., Burgin, W. S. and Grotta, J. C. (2000), 'Neuroprotection and the Ischémie Cascade', *CNS Spectrums* **5**(3), 52–58.

Gleeson, J. P. and Cahalane, D. J. (2007), 'Seed size strongly affects cascades on random networks', *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **75**(5).

Gleeson, J. P., Onaga, T., Fennell, P., Cotter, J., Burke, R. and O'Sullivan, D. J. (2021), 'Branching Process Descriptions of Information Cascades on Twitter', *Journal of Complex Networks* **8**(6), 1–29.

Gleeson, J. P., Ward, J. A., O'Sullivan, K. P. and Lee, W. T. (2014), 'Competition-Induced Criticality in a Model of Meme Popularity', *Physical Review Letters* **112**(4).

Jacob, C. (2010), 'Branching Processes: Their Role in Epidemiology', *International Journal of Environmental Research and Public Health* **7**(3), 1186–1204.

Kapsikar, S., Saha, I., Agarwal, K., Kavitha, V. and Zhu, Q. (2021), 'Controlling Fake News by Collective Tagging: A Branching Process Analysis', *IEEE Control Systems Letters* **5**(6), 2108–2113.

Karimi, F. and Holme, P. (2021), 'A Temporal Network Version of Watts' Cascade Model', *Understanding Complex Systems* **3**, 315–329.

Lashari, A. A. and Trapman, P. (2018), 'Branching Process Approach for Epidemics in Dynamic Partnership Network', *Journal of Mathematical Biology* **76**(1), 265.

Levesque, J., Maybury, D. W. and Shaw, R. H. (2021), 'A Model of COVID-19 Propagation based on a Gamma Subordinated Negative Binomial Branching Process', *Journal of Theoretical Biology* **512**, 110536.

Melnik, S., Hackett, A., Porter, M. A., Mucha, P. J. and Gleeson, J. P. (2011), 'The Unreasonable Effectiveness of Tree-based Theory for Networks with Clustering', *Physical Review E* **83**(3), 036112.

Molloy, M. and Reed, B. (1995), 'A critical point for random graphs with a given degree sequence', *Random Structures & Algorithms* **6**(2-3), 161–180.

Newman, M. (2003), 'The Structure and Function of Complex Networks', *SIAM Review* **45**(2), 167–256.

Newman, M. (2005), 'Power Laws, Pareto Distributions and Zipf's Law', *Contemporary Physics* **46**(5), 323–351.

Newman, M. (2018), *Networks*, Vol. 1, Oxford University Press.

Porter, M. A. and Gleeson, J. P. (2014), *Dynamical Systems on Networks: A Tutorial*, Springer.

Schäfer, B., Witthaut, D., Timme, M. and Latora, V. (2018), 'Dynamically Induced Cascading Failures in Power Grids', *Nature Communications 2018 9:1* **9**(1), 1–13.

Travers, J. and Milgram, S. (1969), 'An Experimental Study of the Small World Problem', *Sociometry* **32**(4), 425–443.

Valente, T. W. (1996), 'Social Network Thresholds in the Diffusion of Innovations', *Social Networks* **18**(1), 69–89.

Watson, H. and Galton, F. (1875), 'On the Probability of the Extinction of Families', *The Journal of the Anthropological Institute of Great Britain and Ireland* **4**, 138–144.

Watts, D. J. (2002), 'A Simple Model of Global Cascades on Random Networks', *Proceedings of the National Academy of Sciences* **99**(9), 5766–5771.

Wilf, H. S. (1994), *Generating Functionology: Second Edition*, Elsevier Inc.

Appendix A-Sensitivity Analysis

We present the results of a sensitivity analysis below (Figure 7.1) and overleaf (Figure 7.2). Each graph is the result of 25,000 repeated simulations. We take the case of the complementary cumulative distribution at $z = 1$ to study, comparing the numerical WTM with numerical BP. We produce results at graph sizes of $n = 1,000$, $10,000$, $50,000$ & $100,000$. We see as the graph size increase, the numerical WTM results tend towards those of the numerical BP. However, this comes at significant computational cost and so is not feasible to produce all results with this graph size. We are content that to show convergence in one case is sufficient to show that the numerical BP does approximate the numerical WTM well, for large graph sizes in the numerical WTM.

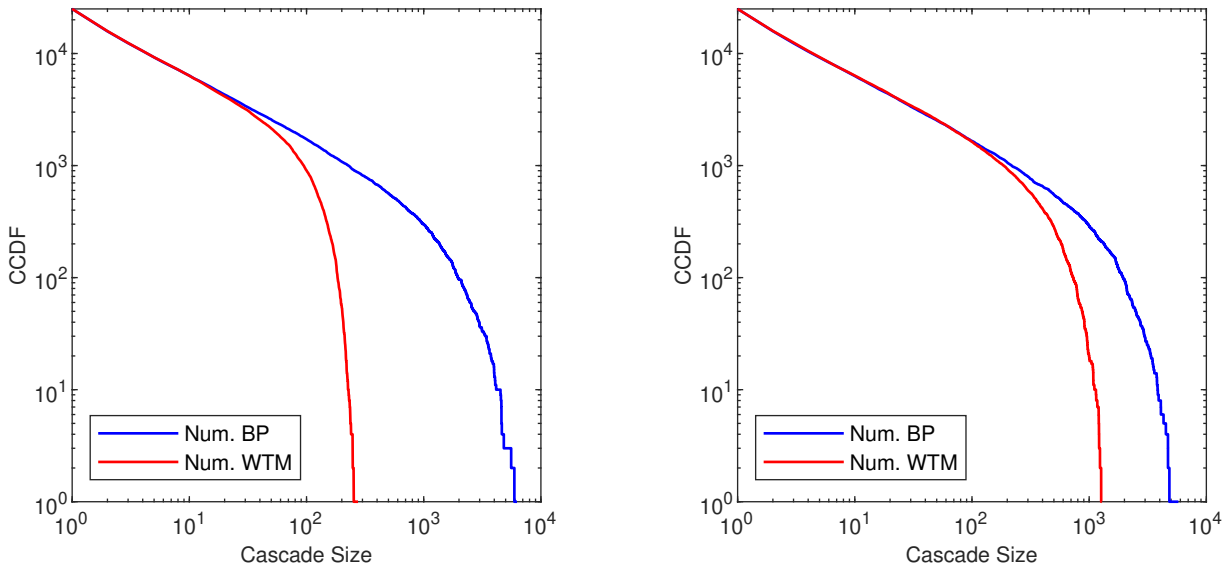

Figure 7.1: Sensitivity analysis for finite graph size at $n = 1,000$ (left) and $n = 10,000$ (right).

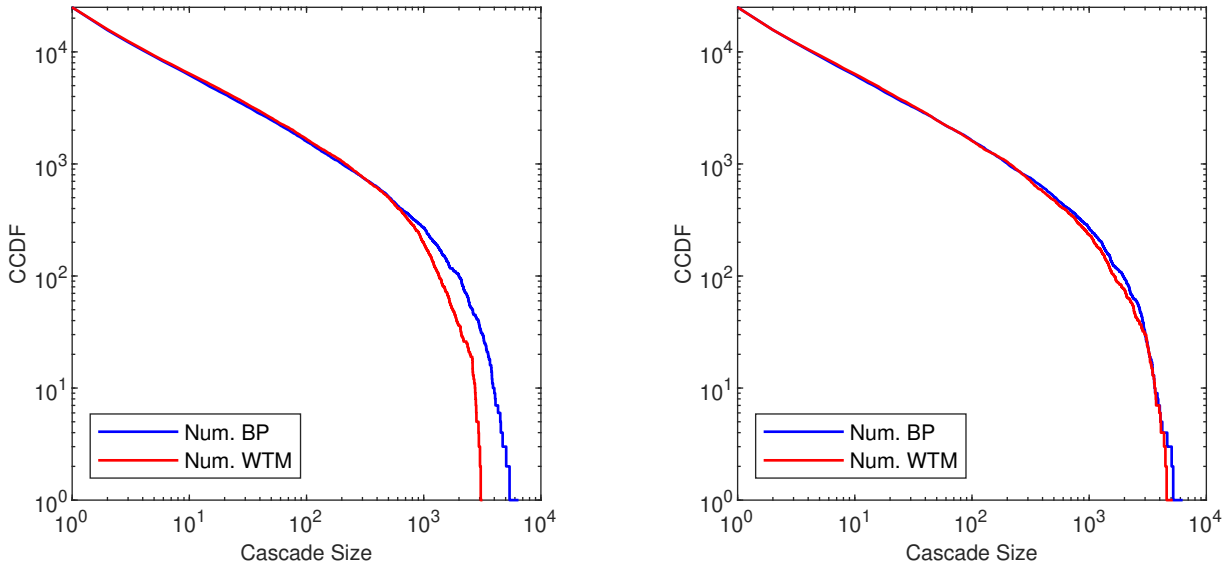

Figure 7.2: Sensitivity analysis for finite graph size at $n = 50,000$ (left) and $n = 100,000$ (right).