Effect of Network Topology on Cascade Behaviour

Verhoeven, Lisa S.

Award date:
2024

Eindhoven University of Technology
Department of Mathematics and Computer Science

Effect of Network Topology on Cascade Behaviour

Authors:
Lisa Verhoeven
(1576607)
l.s.verhoeven@student.tue.nl

Supervisor:
Fiona Sloothaak
f.sloothaak@tue.nl
Agnieszka Janicka
a.j.janicka@tue.nl

Abstract

This thesis examines the influence of network topology on cascading failures within power grids through a combination of theoretical and numerical analyses. By employing a high voltage Direct Current (DC) approximation as described in [22], this thesis looks into the behaviour of cascades across different network structures. The different network structures which will be covered in this thesis are tree and cycle configurations, as well as the Erdős–Rényi (ER) and lattice topologies. We have the following key findings.

Firstly, tree structures limit cascades to the initial failure, whereas cycles typically conclude after two edge failures. Secondly, in ER models, the analysis identifies a specific interval for edge probability p where we see a relative increase in edge failures and disconnection size. Additionally, for large values of p the components that disconnect are either single nodes or small line components. Lastly, for lattice topologies, this thesis highlights cross-shaped patterns of disconnection caused by the cascade and the approximation of emergency edge limits.

March 10, 2024

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Contents

1 Introduction

Modern society is deeply intertwined with the reliability of power grids, which form the backbone of our infrastructure. However, in a power grid, the damage of one transmission line could trigger an overload on other parts of the grid and cause more transmission lines to fail, also called a cascade of failures [9, 4]. These failures, like a series of falling dominos, have the capability to propagate through a system, bringing in danger even the most robust and resilient power grids.
Cascading failures in power grids can lead to flow loss of electricity. Even more so, when the cascade is not contained, it may lead to a major blackout. The satellite image of the 2003 North-Eastern blackout in the United States and Canada shown in Figure 1.1 exemplifies that blackouts can have devastating effects. This event affected over 50 million people and

Figure 1.1: A satellite image from a blackout in North America [18].

resulted in economic losses for the US of up to 10 billion dollars [27]. It highlights the severe consequences that can arise from blackouts, and underscores the critical importance of research into blackouts and cascading failures in power grids.
When it comes to the analysis of blackouts, the challenge lies in accurately quantifying the extent of them, which has proven to be difficult through analytical and simulation approaches. Nonetheless, the distribution of blackout sizes is recognized as scale-free [3, 5, 13, 10]. This means that the probability of a blackout S exceeding a certain size, x, for some constants $\alpha, C > 0$, is given by

$$\mathbb{P}(S > x) \sim Cx^{-\alpha},$$

where $\sim$ means that the ratio of the two quantities turns to one as $x \to \infty$. This is a significant problem, as this distribution implies that large instances are not virtually impossible. In [22] they present a model to explain the emergence of scale-free blackouts. It is explained that the constant C is dependent on the topology of the network, thus the topology influences the blackout size. The impact of topology is so pronounced that, should the cascade stop without disconnecting the power grid, the blackout size is relatively small. Therefore, the primary aim of this thesis is to analyze how network topology impacts the cascade model established in [22].

1.1 High-level description

The aim of this thesis is highly dependent on the cascade model as described in [22] using the high voltage DC approximation. Therefore, in this section, we give an overview of that specific cascade model.
The model starts with a power grid modelled as a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ with the set of nodes $\mathcal{N}$ representing cities and the set of edges $\mathcal{E}$ representing transmission lines. Before the cascading failures process is initialized, the emergency edge limits are defined for each edge, which denote the maximum amount of flow that can pass over that edge. Next to that, every node has a demand and generation. The generation per node is determined by the solution to the DC optimal power flow (DC-OPF) model, which is an optimization problem that determines the most cost-efficient way to distribute generation among nodes in the network. The demand is equal to the population X_i at node i, where we assume the X_i's are identically and independently Pareto distributed, meaning that $\mathbb{P}(X > x) \approx Kx^{-\alpha}$ for some constants $K, \alpha > 0$. After the cascade is invoked with an initial random failure flow has to be redistributed over other edges. This redistribution potentially leads to more edges exceeding their emergency edge limits, resulting in more failures. After several edge failures, disconnection can occur, leading to load shedding, meaning that the disbalance of generation and demand in the subnetworks results in a power shortage in a subnetwork. The load shedding occurs due to the disconnected subnetworks not being able to transfer their electricity to the component where there is a power shortage. Figure 1.2 illustrates the scenario in which load shedding occurs as the node on the left side becomes disconnected from the network. Consequently, depending on the decision of demand and generation either, the generation capacity at the single node is insufficient to meet the demands of the larger network, or similarly, the larger network cannot supply adequate electricity to meet the demand at the single node.
The distribution of demand and generation play a pivotal role in determining the amount of load shed. Recall that the demand is Pareto distributed. Whenever the total sum of Pareto distributed random variables is large, this is due to the contribution of a single random variable, while the others are relatively small.
This phenomenon is known as the "principle of a single big jump" in extreme value theory [21]. It states that for certain distributions, particularly those that are heavy-tailed like the

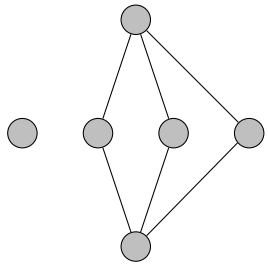


Figure 1.2: Two disconnected subnetworks

Pareto distribution, the largest value dominates the sum. Consequently, the presence of a large city is often a key factor in the occurrence of widespread blackouts.
Recognizing that the node with the largest demand often dictates the scale of a blackout due to the Pareto distribution's heavy-tailed nature, this thesis will delve into a specialized case within the model: the unit demand case, where only the largest city possesses significant demand (normalized to 1), while the remaining cities have negligible demand of zero. This theoretical boundary case approximates the conditions observed during major blackouts, where the size and impact are overwhelmingly influenced by the largest city's power demand. As such, this justifies our focus on the unit demand case within this thesis.

1.2 Related literature

With a clearer understanding of this thesis' objectives, we now turn our attention to existing research on power grid modelling and its role in cascades within power grids. Additionally, we explore the results from the field of cascade failures that are pertinent to the questions posed at the beginning of this thesis: how does topology impact the cascade?
We start with exploring different models that are used to represent the topology of a power grid. Firstly, the scale-free network approach, particularly the Barabási-Albert (BA) model [6], represents grids where the degree of the nodes have a power-law degree distribution, although its node connection strategy might not fully mimic real-world grid geography. Furthermore, the Small-World model, introduced by Watts and Strogatz [28], combines local clustering with short path lengths, reflecting power grids' structure of local and long-range connections, i.e. low and high voltage networks [28]. Lastly, the Configuration Model, from Bollobás [2] and Molloy and Reed [20], offers flexibility in simulating various network types based on specific degree sequences. It is useful for analyzing non-random features in power grids when these are known.
It is crucial to acknowledge these models are approximations. As noted in [7], they often overlook aspects like electrical connectivity, making the evaluation of different topology models essential to obtain a more comprehensive understanding of power grids. Hence, in this thesis, we focus on simple (random) graphs to get an initial understanding of the impact of topology, which can then provide a set-up for the more complex and realistic graphs. We

explore two main types: Erdős–Rényi (ER) [11] and lattices.
Besides the graph models relevant to this thesis, another critical consideration is the method of modelling a high-voltage transmission system for analysis. The DC power flow model is the model used in [22], which is employed in this thesis. The DC model is most attractive to use in the theoretical analysis because of its relative simplicity by neglecting reactive power and being an approximation of the AC power flow model [17, 24], which is a model closer to real life.
As stated, this thesis delves deeper into the effect of the topology on the cascade for the Erdős–Rényi and lattice model. In both of these models, percolation theory is a well-established research field and describes the behaviour of a graph when edges are uniformly at random added or removed. Percolation theory is crucial in understanding cascades in specific graphs, particularly in identifying thresholds necessary for forming certain structures that influence these cascades. This theory has been extensively explored, notably in the context of Erdős–Rényi graphs, where research has delved into aspects like the percolation threshold, which defines the threshold on the probability of node connections at which most nodes get clustered together [12, 11]. Additionally, percolation in square lattices represents another well-examined area within this field [1].
These important characteristics like percolation thresholds can be used in defining different properties of the topology. To get a broader understanding of these characteristics, we also look at topological metrics and characteristics that can influence the cascade model as described in [22]. While this has not been done for the specific model, there are results for topological metrics and their influence on the cascade for different models. For instance, the edge probability in the Erdős–Rényi graph, which defines the probability that two nodes are connected, is shown to be of possible influence in a cascade [25]. Next to that, average degree and average path length [8] are shown to be topological metrics that can impact the propagation of failures.

1.3 The contribution of this thesis

Having reviewed the research conducted on cascade failures, we outline the scope of this thesis. The impact of the topology of particular graphs in the model described in [22] is still unknown. Therefore, this thesis will primarily examine the influence of the topology of simple (random) graphs on the cascade behaviour in the cascade model of [22]. To gain insight into the impact of the topology on the cascade behaviour, we concentrate on the following key aspects:

- Analytical results for the cascade in cycle and tree network structures,
- Numerical results for the cascade in the ER model, specifically looking at the ratio of edge failures and the size of the components that disconnect,

- Numerical results for the cascade in the lattice model, specifically looking at the ratio of edge failures and the size of the components that disconnect,

- An initial analytical study on the cascade in square lattices.

For the remainder of this thesis, we introduce the cascade model [22] in Section 2.1. Additionally, the ER and lattice models are presented in Section 2.2. Secondly, the work on particular graph structures and their cascade behaviour are studied in Section 3. Furthermore, the results for the ER model are described in Section 4. Lastly, the numerical results and an initial analytical study for lattices are given in Section 5.

2 Model description

In this section, we explain the cascade model introduced in [22] in more detail, as well as the network topologies studied in this thesis.

2.1 Cascading failure model

The cascade model used in this thesis is the same as [22] and the supplementary material [23]. In this section, the model is explained to the extent it is used in this thesis, but for more details, one can consult [22] and [23].

2.1.1 The flow network modelling

The electric power grid can be formally represented as a connected graph, denoted as $\mathcal{G} = (\mathcal{N}, \mathcal{E})$. Here, $\mathcal{N}$ represents nodes, which are the entities that either consume or generate electricity. On the other hand, $\mathcal{E}$ represents the edges over which electricity is transmitted. The number of nodes and edges are represented by the cardinalities n and m respectively.
To represent the generation and demand at each node, vectors $\mathbf{g}$ and $\mathbf{d}$ are introduced, both belonging to $\mathbb{R}^n$. These vectors denote the relative electricity generation and demand for each respective node. For modelling purposes, the Direct Current (DC) power flow representation is adopted [24]. This allows the relationship between power injections, represented as $\mathbf{p} = \mathbf{g} - \mathbf{d}$, and power flows, denoted by $\mathbf{f} \in \mathbb{R}^m$. This relationship is expressed by the equation $\mathbf{f} = \mathbf{V}(\mathbf{g} - \mathbf{d})$, where $\mathbf{V} \in \mathbb{R}^{m \times n}$ corresponds to the Power Transfer Distribution Factors (PTDF) matrix, which describes the flow of electricity over the edges.

PTDF construction

To define the Power Transfer Distribution Factors (PTDF) matrix, we first introduce the edge-vertex incidence matrix, denoted by $\mathbf{C} \in \mathbb{R}^{m \times n}$. This matrix is characterized as

$$C_{l,i} = \begin{cases} 1 & \text{if } l = (i,j), \\ -1 & \text{if } l = (j,i), \\ 0 & \text{otherwise} \end{cases}$$

with $l \in \mathcal{E}$.
Denote $\mathbf{B} \in \mathbb{R}^{m \times m}$ as the matrix with on the diagonal the resistance of the transmission edge. In this thesis we consider the $\mathbf{B}$ to be the identity matrix, therefore we assume that every edge l has an equal resistance of one, denoted by $\beta_l = 1$. Moreover, for our specific scenario, we examine the Laplacian matrix $\mathbf{L} \in \mathbb{R}^{n \times n}$ associated with graph $\mathcal{G}$. It is expressed as $\mathbf{L} = \mathbf{C}^\top \mathbf{B} \mathbf{C}$, equivalent to

$$L_{(i,j)} = \begin{cases} -1 & \text{if } (i,j) \text{ or } (j,i) \in \mathcal{E} \text{ and } i \neq j, \\ \sum_{k \neq j} \beta_{(i,k)} & \text{if } i = j. \end{cases}$$

Note from this formulation that every row in $\mathbf{L}$ adds up to zero, therefore $\mathbf{L}$ is a singular matrix. Consequently, we can determine the Moore-Penrose pseudo-inverse, denoted as $\mathbf{L}^+$. The PTDF is then constructed as follows:

$$\mathbf{V} = \mathbf{C}\mathbf{L}^+.$$

Demand Vector

In the study by [22], it is demonstrated that city populations are modelled as independent and identically distributed (i.i.d.) from a Pareto distribution. In this thesis, we model each node as a representation of a city. For every inhabitant in a city, an additional unit of demand is accounted for, as discussed in Section 1.1.
For the purposes of our model, we move to the limiting case as explained in Section 1.1, therefore, the node with the highest demand is given a value of 1 with all other nodes being set to 0. Given the *a priori* nature of the i.i.d. Pareto distribution, before any populations are realized, each node in our set $\mathcal{N}$ has an equal probability of $\frac{1}{|\mathcal{N}|}$ of drawing the largest value. Consequently, we can uniformly randomly select a node and assign its demand as 1.

Generation vector

In the initialization phase, we must optimize generation per node to minimize system generation costs while ensuring demand and supply constraints are met. To achieve this, we employ the DC optimal power flow (DC-OPF) equations minimizing costs while maintaining

the constraints and balance between generation and demand[23]. Simplifications arise due to the specific constraints on the demand vector given previously, yielding a direct solution for the generation vector $\mathbf{g}$. The only assumption to be met is $\mathbf{V}\mathbf{e_1} \geq \mathbf{0}$, which can be done w.l.o.g. That is, in [23] it is shown that this requirement can always be satisfied by choosing orientation correctly. Generation then only depends on the loading factor $\lambda \in [0, 1]$ which is given by the following closed-form expression with all-ones vector $\mathbf{e}$,

$$\mathbf{g}(\lambda) = \lambda \frac{1}{n}\mathbf{e} + (1 - \lambda)\mathbf{e_1}. \tag{1}$$

Edge Limits

A cascade can only be triggered if the transmission edges have set limits that can potentially be exceeded. Adopting the modelling choices in [22], we set the edge limit for edge $j \in \mathcal{E}$ as

$$F_j = |(\mathbf{V}\mathbf{d})_j|.$$

If the edge limits would not exist the flow over edge j would equal F_j.

Cascade dynamics

Now that the flow network is defined, the cascade can be initiated. The cascade is triggered by removing an edge $l \in \mathcal{E}$ uniformly at random. The failed edge is removed from the graph. If the resulting graph is disconnected, redistribution takes place. Moreover note that for every component disconnected from the demand node, the disconnected component can not disintegrate any further because there is no demand to which the generated electricity can flow and therefore, no edge limits will be exceeded.

Load shedding rule after disconnection

If a disconnection occurs, the total generation and demand in each component are no longer balanced. The balance is restored via the uniform shedding rule [23].

Definition 2.1. *(Uniform Shedding Rule when $\mathbf{d} = \mathbf{e_1}$) Suppose that the removal of edges $l^{(1)}, \ldots, l^{(m)}$, $m \geq 1$ yields the graph $\mathcal{G}^{(m)} = (\mathcal{N}^{(m)}, \mathcal{E}^{(m)})$ in which the demand node is present. We define the power imbalance in this component as*

$$Y_{\mathcal{G}^{(m)}} = \sum_{k \in \mathcal{N}^{(m)}} (g_k^{(m)} - d_k^{(m)})$$

where $g^{(m)}$ and $d^{(m)}$ are the generation and demand vectors after $m-1$ edge failures and $g^{(1)} = g$ and $d^{(1)} = d$ with g and d the initial demand and generation vectors. In order to achieve power balance we have for $k \in \mathcal{N}^{(m)}$

$$d_k^{(m+1)} = \left(1 + \frac{Y_{\mathcal{G}^{(m)}}}{\sum_{i \in \mathcal{N}^{(m)}} d_i^{(m)}}\right) d_k^{(m)},$$
$$g_k^{(m+1)} = g_k^{(m)}.$$

After the flow redistribution of demand and generation, the PTDF $\mathbf{V}$ and flow $\mathbf{f}$ have also changed, therefore flow redistribution is also needed.

Definition 2.2. *We assume that the removal of edges $l^{(1)}, \ldots, l^{(m)}$, $m \geq 1$ disconnects the grid in components leaving us with a new network $\mathcal{G}^{(m)} = (\mathcal{N}^{(m)}, \mathcal{E}^{(m)})$ in which the demand node present. Then the new edge flows in $\mathcal{G}^{(m)}$ are*

$$f_{\mathcal{E}^{(m)}}^{(m+1)} = \mathbf{V}^{(m+1,\mathcal{G}^{(m)})}(\mathbf{g}_{\mathcal{N}^{(m)}}^{(m+1)} - \mathbf{d}_{\mathcal{N}^{(m)}}^{(m+1)})$$

where $\mathbf{V}^{(m+1,\mathcal{G}^{(m)})}$ is the PTDF matrix for the graph $\mathcal{G}^{(m)}$ and $\mathbf{g}_{\mathcal{N}^{(m)}}^{(m+1)}$ and $\mathbf{d}_{\mathcal{N}^{(m)}}^{(m+1)}$ are defined in Definition 2.1. No flow exists in the component that becomes disconnected from the component that contains the demand node.

Edge failure

After the first edge failure, the cascade continues when edge limits are exceeded. We first define the set of edges that exceed their edge limit as $\mathcal{A}^{(m)} = \{j \in \mathcal{E}^{(m)} | f_j^{(m)}| \geq F_j\}$. Then, the next or $m\text{-}th$ edge failure occurs at the edge where the relative exceedance it highest, i.e.

$$l^{(m)} = \arg\max_{j \in \mathcal{A}^{(m)}} \left\{ \begin{cases} \infty & \text{if } F_j = 0, \\ \frac{|f_j^{(m)}|}{F_j} & \text{otherwise} \end{cases} \right\}.$$

When $\mathcal{A}^{(m)}$ is non-empty, a subsequent edge failure takes place, which in turn requires another disconnection check and potential redistribution. This cycle continues until there are no further edge failures.

Algorithm

In Algorithm 1, we outline the steps we previously discussed. Lines 2-7 detail the initialization, as explained in Section 2.1.1. Note that the adjacency matrix A has an entry at position (i, j) if nodes i and j are connected by an edge. Furthermore, after the switching in line 3, the rest of the algorithm assumes that the demand node is at position 0. Starting

from line 10, as outlined in Equation 2.1.1, the cascade begins by uniformly selecting the initial edge to fail. Subsequently, the algorithm checks for any disconnections in the network and, if found, adjusts both demand and generation. The PTDF matrix and the flow are then updated. Next, the system checks for edges exceeding their edge limit to find the next failure. From line 19 onwards, the procedure mirrors that of lines 10-15: the system examines network disconnections, and recalibrates demand, generation, PTDF matrix, and flow. This iterative process concludes once no additional edge failures are detected.

Algorithm 1: Cascade Algorithm

1 **Input:** matrix $\mathbf{A}$, double λ, int **demandNode**
2 $\mathbf{d} \leftarrow \mathbf{e_1}$
3 Switch rows and columns 0 and **demandNode** in adjacency matrix $\mathbf{A}$.
4 $\mathbf{g} \leftarrow \lambda \frac{1}{n}\mathbf{e} + (1-\lambda)\mathbf{e_1}$
5 Construct PTDF matrix $\mathbf{V}$ and correct orientation ensuring that $\mathbf{V}e_1 \geq 0$.
6 $F_j \leftarrow |(\mathbf{V}\mathbf{d})_j|$ *// Setting the edge limits*
7 $\mathbf{f} \leftarrow \mathbf{V}(\mathbf{g} - \mathbf{d})$ *// Computing the initial flow*
8 *// Initiate cascade by failing an edge.*
9 Randomly select and fail edge l; subsequently remove l from $\mathbf{A}$.
10 **if** *network becomes disconnected* **then**
11 Remove rows and columns of nodes in $\mathbf{A}$, $\mathbf{d}$, and $\mathbf{g}$ that are disconnected from node 0.
12 Adjust demand and generation based on Definition 2.1, updating both $\mathbf{d}$ and $\mathbf{g}$.
13 **end**
14 Re-evaluate the PTDF matrix $\mathbf{V}$ using updated $\mathbf{A}$.
15 Modify the flow, $\mathbf{f}$, based on Definition 2.2 using the new $\mathbf{V}$, $\mathbf{d}$, and $\mathbf{g}$.
16 *// Continue cascade until no further edge failures.*
17 **while** *a edge l fails* **do**
18 Omit the failed edge l from $\mathbf{A}$.
19 **if** *network becomes disconnected* **then**
20 Remove rows and columns in $\mathbf{A}$, $\mathbf{d}$, and $\mathbf{g}$ that are disconnected from node 0.
21 Readjust demand and generation based on Definition 2.1, updating $\mathbf{d}$ and $\mathbf{g}$.
22 **end**
23 Update PTDF matrix $\mathbf{V}$ based on current network structure $\mathbf{A}$.
24 Modify flow $\mathbf{f}$ using Definition 2.2 with the updated $\mathbf{V}$, $\mathbf{d}$, and $\mathbf{g}$
25 **end**

Computations and notation for the emergency edge limits

In this section we start by showing that to calculate the emergency edge limits, only the first column of a PTDF **V** matrix is used. Firstly, the emergency edge limits are given by the equation $|\mathbf{Vd}|$ with $\mathbf{d} = \mathbf{e_1}$. Additionally, due to the rows of **V** adding up to 0, computing the flow subsequent to any number of edge failures, requires us to only use the first column of **V** necessary. This requirement stems from the sum of the rows of **V** and the specific structure of the demand and generation vectors **d** and **g**. More specifically, for $\lambda \in [0, 1]$ and all-ones vector $\mathbf{e} \in \mathbb{R}^n$, by using the closed form for **g** given in Equation 1 it holds that

$$\begin{aligned}\mathbf{g} - \mathbf{d} &= \lambda\frac{1}{n}\mathbf{e} + (1-\lambda)\mathbf{e_1} - \mathbf{e_1} \\ &= \lambda\frac{1}{n}\mathbf{e} - \lambda\mathbf{e_1}.\end{aligned}$$

Therefore as the rows of **V** add up to zero,

$$\mathbf{V}(\mathbf{g} - \mathbf{d}) = \mathbf{V}\left(\frac{\lambda}{n}\mathbf{e} - \lambda\mathbf{e_1}\right) = \mathbf{V}(-\lambda\mathbf{e_1}). \tag{2}$$

Consequently, only the first row of the PTDF **V** is necessary for these calculations. Hence, we introduce the following notation.

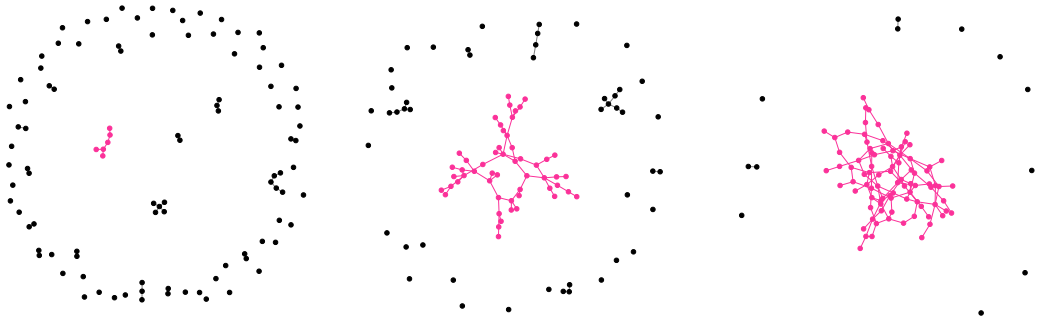
Notation 1. *(Operational and emergency edge limits) For the emergency edge limits we use F^0 after 0 edge failures. We use F^i for the first column of the PTDF matrix after the i-th edge failure.*

2.2 Network models

This thesis focuses on both the Erdős–Rényi and lattice models. In the following part, both these models are explained in detail, including their specific use and implementation for this thesis.

2.2.1 Erdős–Rényi

Consider the Erdős–Rényi model $\mathrm{ER}_n(p)$ defined by n vertices and edge probability p. In this model, each pair of vertices is connected with a probability p independent of other pairs. The ER model does not guarantee a connected graph. Therefore, to implement a cascade, the largest connected component is utilized. The ER model has a distinct threshold known as the percolation threshold, denoted p_c. If the edge probability $p > p_c$, the graph predominantly contains one large component of size $O(n)$ [1]. For the ER model with n vertices, the percolation threshold is approximately $p_c \approx \frac{1}{n}$ [1].



(a) (b) (c)

Figure 2.1: Two instances of Erdős–Rényi random graphs are illustrated, each with 100 nodes. The first instance (a) has an edge probability of 0.005, so $p = \frac{0.5}{n}$, and the second instance (b) has an edge probability of 0.015, so $p = \frac{1.5}{200}$. Finally, the last figure (c) is constructed by an edge probability of 0.025, so $p = \frac{2.5}{200}$. The pink colour illustrates the connected component with the largest number of nodes.

In Figure 2.1, two instances of the ER model with $n = 100$ are depicted. Figure 2.1 (a) has $p = 0.005$, below the threshold $p_c = \frac{1}{100}$. Meanwhile, Figure 2.1 (b) has $p = 0.015$, exceeding the threshold. The latter exhibits a dominant component, called a "giant" component, whereas the former primarily comprises smaller components. To implement the cascade in the ER model, we specify a value for n and select a probability $p > \frac{1}{n}$. In this thesis, we focus on the case $p > \frac{1}{n}$. This configuration ensures that the resulting graph has a giant component, denoted as $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ stands for the set of nodes and $\mathcal{E}$ for the set of edges in the giant (component).
Another foundational aspect of the Erdős-Rényi model arises as the critical parameter np approaches 1 (i.e. as p approaches $\frac{1}{n}$ for a fixed n) then the average degree K of the giant component—calculated as $K = \frac{2|\mathcal{E}|}{|\mathcal{N}|}$-tends towards 2. This is the minimum average degree necessary for cycle formation within the network. For a fixed n, as p increases, so does K, indicating a gain in the giant's structural complexity, which is mostly shown in a higher likelihood of cycles. This principle is illustrated in Figure 2.1. Note that as p increases from (a) to (c) the number of cycles increases in the giant component, which is highlighted in pink. The analytical relationship between the number of cycles and p is further explored in [26]. Essentially, a giant near the critical threshold ($p = \frac{1}{n}$) may display minimal complexity in the form of more tree structures, but as p grows compared to a fixed n, its tendency for connectivity patterns, such as cycles, is increased.
Lastly, as referenced in Section 2.1, in the context of the limiting behaviour of the Pareto distribution, we uniformly and independently select a node from $\mathcal{G}$ and assign it a demand of 1.

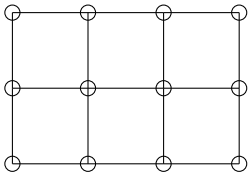


Figure 2.2: 3 by 4 lattice

2.2.2 Lattice

Consider the n by m square lattice $\mathcal{L}$ with nm amount of nodes. In Figure 2.2, an example of such a lattice is given. As referenced in Section 2.1, in the context of the limiting behaviour of the Pareto distribution, we can uniformly and independently select a node from $\mathcal{L}$ and assign it a demand of 1. However, the unit demand node might also be fixed. This fixed placement could be the centre or a corner, for instance, and due to the highly structured lattices can give more insights.
We choose to place the unit demand node in the centre, using an n by n matrix where n is odd to make sure there is always a node in the centre. Placing the unit demand node in the centre allows us to look at situations with larger n, as every node is the centre node when $n \to \infty$.

3 Results of cascade on particular graph structures

As noted in Section 2.2.1, three distinct structures are commonly observed in the giant component of the ER model: trees, cycles, and a combination of both. These structures are presented in an order reflecting their increasing likelihood of occurrence as the edge probability p rises. Therefore, this section aims to provide insights into how these various graph structures, including trees, cycles, and their combinations, impact the cascade's behaviour.

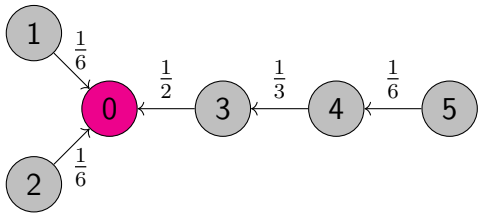
3.1 Cascade on tree

Before we start with our focus on the behaviour of a cascade in a tree we define a path which will be used in the rest of this chapter.

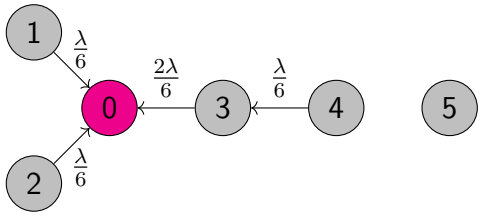
Definition 3.1. *(Path) For a graph G with unit demand node 0 and size n, a path S_j is the ordered set of edges from node j to node 0 of minimal cardinality m_j, i.e. the shortest path. $S_j(i)$ gives the i-th edge in the path, where $S_j(1)$ is the first edge and $S_j(m_j)$ is the last edge. Such a path is not always unique, therefore S_{j_k} is used for different k when there is more than one shortest path.*

For this section, we consider a tree $\mathcal{T}$ of size n with as demand node 0 and set of nodes and edges $\mathcal{N}$ and $\mathcal{E}$. Let S_j be a path in $\mathcal{T}$ for every node $j \in \mathcal{N}$. The purpose of this

section is to demonstrate that the cascade will stop immediately after the initial edge failure, independently of the location of this first failure. An example of this behaviour is given in Figure 3.1.



(a) emergency edge limits



(b) absolute flows after edge failure

Figure 3.1: Example of a tree network

The relatively straightforward configuration of a tree structure eases the explanation of the emergency edge limits and the flows. In a tree structure, each node is capable of transmitting its output exclusively along a single path to the demand node. Hence, the edge limits on any given edge are dependent on the number of nodes that pass the edge on their path to the demand node.
Illustrating this concept, consider the edge between nodes 4 and 5 in Figure 3.1; this edge only needs to support the generation from node 5, which leads to an emergency edge limit of $\frac{1}{6}$. Conversely, the edge between nodes 3 and 4 must carry the generation from both nodes 4 and 5, thus, the emergency edge limit is $\frac{2}{6}$ or simplified, $\frac{1}{3}$.
The flows after an initial edge failure follow from the same line of argument, however, there are fewer nodes because of the disconnection and therefore flows are either reduced or stay the same. Hence, the flows after an initial edge failure never exceed the emergency edge limits because there are fewer nodes and therefore less or equal flow for every edge compared to the emergency edge limits.
In Figure 3.1 (b), the flows are given after the first edge failure and disconnection and none of them exceed their emergency edge limit for any $\lambda \in [0, 1]$, and we observe that indeed the cascade stops in the tree after an initial failure. The rest of this section shows this behaviour rigorously by describing the emergency edge limit and the flows, finally showing that the flows do not exceed the emergency edge limits. In other words, the cascade stops after the initial edge failure.

3.1.1 Emergency edge limits

As noticed in the example given in Figure 3.1, the edge limits of an edge are determined by the number of paths of nodes that use the edge. Therefore, note that there is a unique path for every node to the demand node.

Lemma 3.1. *(Uniqueness of path in Tree) The path S_j for every node $j \in \mathcal{N}$ is unique.*

Proof. Let $j \in \mathcal{N}$. There exists at least one path $S_j = \{(j, a_1), (a_1, a_2), ..., (a_{m_j-1}, 0)\}$ of size m_j because of connectivity. Assume there to be a second path $S'_j = \{(j, b_1), (b_1, b_2), ..., (b_{k_j-1}, 0)\}$ of size k_j. However, then there exists at least one cycle starting in j, reaching node 0 over path S_j and going back to node j over path S'_j, contradicting the tree property. □

Having established the uniqueness of every path in a tree, we can next support our initial claim that the emergency edge limits for an edge, denoted as F^0, are determined by counting the number of nodes that utilize this edge in their route to the demand node, i.e. the amount nodes that are in the subtree below that edge.

Lemma 3.2. *(Emergency edge limits in a tree) Let S_j of length m_j be the path of node $j \in \mathcal{N}$. For all edges (k, l) of $\mathcal{T}$ such that they are oriented towards the demand node,*

$$F^0_{(k,l)} = \sum_{j \in \mathcal{N}} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (k, l)\}}{n}.$$

Proof. The PTDF matrix describes how the power flow is distributed according to Kirchhoff's Law [29]. Therefore, we prove that Kirchhoff's law gives a unique solution and holds given the emergency edge limit in the lemma.
In Section 2.1 we noted that the resistance at every edge is 1 given by the matrix $\mathbf{B}$, therefore, by Ohm's law $V = I \cdot R$ implies that the current equals the voltage [19]. Then by Kirchhoff's Current Law, for every node a_l with power injection $g(a_l)$ (with $\lambda = 1$) and outgoing edge (a_l, a_{l+1}), we need to show that the outgoing flow equals the sum of the incoming and the generation at that node [14]. Let the edges be oriented towards the demand node over the shortest path. Then Kirchhoff's Current Law (KCL) states

$$F^0_{(a_l,a_{l+1})} = \sum_{h \in \mathcal{N}} \mathbb{I}\{(h, a_l) \in \mathcal{E}\} F^0_{(h,a_l)} + g(a_l). \tag{3}$$

Define the set N_0 as all the leaf nodes, i.e. nodes of degree 1, and define

$$N_l = \left\{ j \in \mathcal{N} : \left| (j, i) \in \mathcal{E} : i \in \mathcal{N} \setminus \bigcup_{k=0}^{l-1} N_k \right| = 1 \right\},$$

which is a set of nodes of distance l from the furthest leaf node j such that the nodes are on the path S_j. Note that the incoming flow of a node in N_l can be from any node in the set $\bigcup_{k=0}^{l-1} N_k$.

Unique solution. We show that there is a unique solution to the KCL equations given by Equation 3. From the KCL equations (3) for the leaf nodes, it holds that for every leaf node $j \in N_0$, node $i \in N_1$, and edge $(j,i) \in \mathcal{E}$, $F^0_{(j,i)} = \frac{1}{n}$. Now assume for $a_l \in \bigcup_{k=0}^{l} N_k$ and $a_{l-1} \in \bigcup_{k=0}^{l-1} N_k$ that $F^0_{(a_{l-1},a_l)}$ is known and unique. Then for $a_{l+1} \in N_{l+1}$, by the KCL equations (3)

$$F^0_{(a_l,a_{l+1})} = \sum_{h\in\mathcal{N}} \mathbb{I}\{(h,a_l) \in \mathcal{E}\} F^0_{(h,a_l)} + g(a_l).$$

Because all incoming edges $\sum_{h\in\mathcal{N}} \mathbb{I}\{(h,a_l) \in \mathcal{E}\} F^0_{(h,a_l)}$ are already known, this is again uniquely determined. Then by induction, all emergency edge limits are known for all edges. By the uniqueness of path the Kirchhoff equation for the demand node $\sum_{e\in\mathcal{E}} F^0_e + g(0) = \frac{|\mathcal{E}|-1}{n} + \frac{1}{n} = 1$ holds as well. Therefore, there is a unique solution to Kirchhoff's equality.

The solution. We show by induction on l that Kirchhoff's equality holds for the given solution in the lemma. Note that by 3.1, m_j is the length of path S_j. For all leaf nodes $a_0 \in N_0$ and neighbouring nodes a_1 of a_0 we have,

$$\begin{aligned}
F^0_{(a_0,a_1)} &= \sum_{j\in\mathcal{N}} \sum_{i=1}^{m_j} \frac{\mathbb{1}\{S_j(i) = (a_0,a_1)\}}{n} \\
&= \sum_{i=1}^{m_{a_0}} \frac{\mathbb{1}\{S_{a_0}(i) = (a_0,a_1)\}}{n} + \sum_{j\in\mathcal{N}\setminus\{a_0\}} \sum_{i=1}^{m_j} \frac{\mathbb{1}\{S_j(i) = (a_0,a_1)\}}{n} \\
&= \frac{\mathbb{1}\{S_{a_0}(1) = (a_0,as_1)\}}{n} + \sum_{i=2}^{m_{a_0}} \frac{\mathbb{1}\{S_{a_0}(i) = (a_0,a_1)\}}{n} + \sum_{j\in\mathcal{N}\setminus\{a_0\}} \sum_{i=1}^{m_j} \frac{\mathbb{1}\{S_j(i) = (a_0,a_1)\}}{n} \\
&= \frac{1}{n} + 0 + 0 \\
&= g(a_0).
\end{aligned}$$

Now assume for $a_l \in \bigcup_{k=0}^{l} N_k$ and $a_{l-1} \in \bigcup_{k=0}^{l-1} N_k$ that

$$F^0_{(a_{l-1},a_l)} = \sum_{h\in\mathcal{N}} \mathbb{I}\{(h,a_{l-1}) \in \mathcal{E}\} F^0_{(h,a_{l-1})} + g(a_{l-1}).$$

Then for $a_{l+1} \in N_{l+1}$

$$\begin{aligned}
F^0_{(a_l,a_{l+1})} &= \sum_{j\in\mathcal{N}} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (a_l, a_{l+1})\}}{n} \\
&= \sum_{j\in\bigcup_{k=0}^{l} N_k} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (a_l, a_{l+1})\}}{n} \\
&= \sum_{j\in\bigcup_{k=0}^{l-1} N_k} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (a_l, a_{l+1})\}}{n} + \sum_{j\in N_l} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (a_l, a_{l+1})\}}{n} \\
&= \sum_{j\in\bigcup_{k=0}^{l-1} N_k} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (a_l, a_{l+1})\}}{n} + \frac{1}{n} \\
&= \sum_{h\in\mathcal{N}} \mathbb{I}\{(h, a_l) \in \mathcal{E}\} F^0_{(h,a_l)} + g(a_l)
\end{aligned}$$

□

Given the known emergency edge limits, the initial flow on all edges can also be determined, as detailed in the following corollary.

Corollary 3.1. *(Flows before initial edge failure in tree) Let S_j of length m_j be the path of node $j \in \mathcal{N}$. For all edges (k, l) of $\mathcal{T}$ such that they are oriented towards the demand node,*

$$f_{(k,l)} = -\lambda \sum_{j\in\mathcal{N}} \sum_{i=1}^{m_j} \frac{\mathbb{I}\{S_j(i) = (k, l)\}}{n}.$$

Proof. This follows from Lemma 3.2 and the flow computations in Equation 2. □

3.1.2 Flows after the initial edge failure

Now that the emergency edge limits are determined, the next step is to determine the flow after the initial edge failure. As an example in Figure 3.1, the failure of edge $(4, 3)$ affects the flow on $(3, 0)$ and $(5, 4)$, the latter not being important due to the disconnection from the demand node. However, the other two edges on the left, $(1, 0)$ and $(2, 0)$, are not affected by this edge failure and disconnection. In the general case, the set of edges that are affected by the initial edge failure will contain all edges that were carrying the flow of the nodes that are now disconnected, excluding the disconnected nodes themselves. In Figure 3.2, this set of edges is depicted by A, Next to that, E shows the nodes that are disconnected.

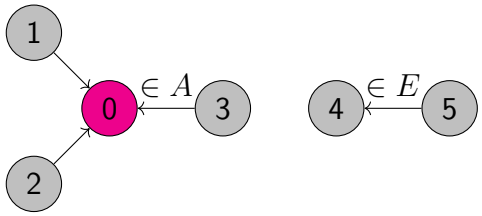


Figure 3.2: Tree example after disconnection

Definition 3.2. *(Set A and E) Assume edge (k,l) is removed from $\mathcal{T}$ which disconnects the set of edges E and nodes N from the demand node such that the tree $\mathcal{T}' = (\mathcal{N}', \mathcal{E}')$ is the subgraph with demand node $0 \in \mathcal{N}$. Then, $E = \mathcal{E} \setminus \mathcal{E}'$. Let S_j be the path such that there exists $i \in (0, m_j]$ for which $S_j(i) = (k,l)$ then*

$$A_{(k,l)} = \{S_j(h) \quad \forall h > i\}.$$

Hence, the flow distribution is only changed for those edges in the set $A_{(k,l)}$. However, because of the disconnection all flows are lowered a bit due to the redistribution of the demand vector. These factors are combined when finding the flow of the edges after the first edge failure in the next lemma.

Lemma 3.3. *(Flow on a tree after the first edge failure) Let f_m^0 be the flow for edge m of $\mathcal{T}$. We assume edge (k,l) is removed which disconnects the set of edges E and nodes N from the demand node 0.*
For a given λ it holds that the flow f'_m after edge failure (k,l) is given by

$$f'_m = \begin{cases} f_m - f_{(k,l)} & m \in A_{(k,l)} \\ f_m & m \notin A_{(k,l)} \\ 0 & m \in E \end{cases}.$$

Proof. Firstly note that due to the disconnection, there is a redistribution of flow. Therefore, assume we have initial demand and generation vectors, $g^{(0)}$ and $d^{(0)}$. After the removal of edge (k,l) the nodes N and edges E are disconnected from the component with the unit demand node $\mathcal{T}' = (\mathcal{N}', \mathcal{E}') = (\mathcal{N} \setminus N, \mathcal{E} \setminus E)$. According to Definition 2.2, $g^{(0)}$ (Equation 1) and $d^{(0)}$ in Section 2.1.1 we then have

$$Y_{\mathcal{T}'} = \sum_{j \in \mathcal{N}'} g_j^{(0)} - d_j^{(0)} = (1-\lambda) + \frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda - 1$$

and thus for $j \in \mathcal{N}'$ and the all-ones vector $\mathbf{e} \in \mathbb{R}^{|\mathcal{N}'|}$ and first unit vector $\mathbf{e_1} \in \mathbb{R}^{|\mathcal{N}'|}$

$$d_j^{(1)} = \begin{cases} \left(1 + \frac{Y_{\mathcal{T}'}}{\sum_{i \in \mathcal{N}'} d_i^{(0)}}\right) d_j^{(0)} & \text{if } j = 1, \\ 0 & \text{else} \end{cases}$$
$$= \begin{cases} \left(1 - \lambda + \frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right) & \text{if } j = 1, \\ 0 & \text{else} \end{cases}$$

and

$$g^{(1)} = g^{(0)} = (1-\lambda)\mathbf{e_1} + \frac{\lambda}{|\mathcal{N}|}\mathbf{e}.$$

Consequently, we have that

$$g^{(1)} - d^{(1)} = (1-\lambda)\mathbf{e_1} + \frac{\lambda}{|\mathcal{N}|}\mathbf{e} - \left(1 - \lambda + \frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right)\mathbf{e_1} = \frac{\lambda}{|\mathcal{N}|}\mathbf{e} - \left(\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right)\mathbf{e_1}.$$

Note that as described in Section 2.1.1 we can now compute the flow f' for each edge $m \in \mathcal{E}'$ after the edge failure (k,l) with the first column or the PTDF matrix for the graph $\mathcal{T}'$ which is denoted as $F^{(1)}$. Thus, as the rows of the PTDF matrix sum up to zero, and for the PTDF matrix $V^{(1)}$ for the graph $\mathcal{T}'$, we compute f'_m as follows

$$\begin{aligned} f'_m &= V^{(1)}\left(g^{(1)} - d^{(1)}\right) \\ &= V^{(1)}\left(\frac{\lambda}{|\mathcal{N}|}\mathbf{e} - \left(\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right)\mathbf{e_1}\right) \\ &= V^{(1)}\left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right) && \text{this equality holds because } \mathbf{e} \in Ker(V^{(1)}) \\ &= F_m^{(1)}\left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|}\lambda\right). \end{aligned}$$

With this set up we can start to prove the first case.
Case 1. First, consider the case of f'_m with $m \in A_{(k,l)}$, implying that before disconnection the flow over edge (k,l) was also carried by edge m. The first step is to find the relationship between $F_m^{(0)}$ and $F_m^{(1)}$. Assume $F_m^{(0)}$ is of the form $\frac{a}{|\mathcal{N}|}$ for some positive $a \in \mathbb{N}$. Then as $|N|$ nodes are disconnected and their flow was going over the edge m we now have,

$$F_m^{(1)} = \frac{a - |N|}{|\mathcal{N}| - |N|}.$$

Then the relation between $F_m^{(0)}$, the first row of the PTDF matrix of $\mathcal{T}$, and $F_m^{(1)}$, the first row of the PTDF matrix of $\mathcal{T}'$, is

$$\left(\frac{a - |N|}{|\mathcal{N}| - |N|}\right)\left(\frac{|\mathcal{N}|}{a}\right) F_m^{(0)} = F_m^{(1)}. \tag{4}$$

Furthermore, we want to find the relationship between $F_m^{(0)}$ and $F_{(k,l)}^{(0)}$, where $F_{(k,l)}^{(0)}$ is the first row of the PTDF for $\mathcal{T}$ and the emergency edge limit for edge (k,l). As (k,l) was the edge that connected $|N|$ to the demand node

$$F_{(k,l)}^{(0)} = \frac{|N|}{|\mathcal{N}|}$$

and thus

$$\frac{|N|}{a} F_m^{(0)} = F_{(k,l)}^{(0)}. \tag{5}$$

With this information, we can now compute f'_m as follows

$$\begin{aligned}
f'_m &= F_m^{(1)} \left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|} \lambda \right) \\
&= \left(\frac{a - |N|}{|\mathcal{N}| - |N|} \right) \left(\frac{|\mathcal{N}|}{a} \right) F_m^{(0)} \left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|} \lambda \right) && \text{by Equation 4} \\
&= F_m^{(0)} \left(\frac{a - |N|}{|\mathcal{N}|} \right) \left(\frac{|\mathcal{N}|}{a} \right) (-\lambda) \\
&= F_m^{(0)} \left(1 - \frac{|N|}{a} \right) (-\lambda) \\
&= -\lambda F_m^{(0)} - (-\lambda) F_{(k,l)}^{(0)} && \text{by Equation 5} \\
&= f_m - f_{(k,l)}.
\end{aligned}$$

This proves the first case and we can continue to the next.

Case 2. Secondly, when $m \notin A_{(k,l)}$ we have that none of the flow is coming from the disconnected component used to travel over m. Therefore, if we again assume that $F_m^{(0)}$ is of the form $\frac{a}{|\mathcal{N}|}$ for some positive $a \in \mathbb{N}$, then

$$F_m^{(1)} = \frac{a}{|\mathcal{N}| - |N|}.$$

Therefore we have the following relation

$$F_m^{(0)} \left(\frac{|\mathcal{N}|}{|\mathcal{N}| - |N|} \right) = F_m^{(1)}. \tag{6}$$

Consequently, we can finish the proof by computing the flow f'_m,

$$\begin{aligned}
f'_m &= F_m^{(1)} \left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|} \lambda \right) && \text{by Equation 6} \\
&= \left(\frac{|\mathcal{N}|}{|\mathcal{N}| - |N|} \right) F_m^{(0)} \left(-\frac{|\mathcal{N}| - |N|}{|\mathcal{N}|} \lambda \right) \\
&= -\lambda F_m^{(0)} \\
&= f_m.
\end{aligned}$$

This concludes our second case and brings us to the last part of this proof.
Case 3. Lastly, suppose that edge $m \in E$, which means that m belongs to the component $\mathcal{T} \setminus \mathcal{T}'$. The total demand in this component is 0. Hence, by Definition 2.2 the new generation in this component is also set to 0. As a result, there is no flow and we conclude that $f'_m = 0$. □

3.1.3 Results of the cascade

Now that the emergency edge limits and the flows are given, the final step is to show that flows will never exceed the emergency edge limits. Hence, the cascade stops after the first edge failure.

Theorem 3.1. *(The cascade stops immediately in a tree after the first edge failure) The cascade on $\mathcal{T}$ as defined in Section 2.1 stops after one edge removal.*

Proof. Let F_m^0 be the emergency edge limit for edge m of $\mathcal{T}$, initial demand and generation vector d and g as in Section 2.1 and $\lambda \in [0, 1]$. Remove the edge (k, l) from $\mathcal{T}$, which disconnects the set of nodes N. We have multiple cases.
Case 1. Firstly, we take the case for all edges $m \notin A_{(k,l)}$. Then by Lemma 3.3, we have that

$$f'_m = f_m - f_{(k,l)} = -\lambda F_m^0 - \frac{-\lambda |N|}{n} = -\lambda \left(F_m^0 - \frac{|N|}{n} \right).$$

We have that $\left| \left(F_m^0 - \frac{|N|}{n} \right) \right| < F_m^0$ and $|\lambda| < 1$. Therefore $|f'_m| < |F_m^0|$ and the cascade stops in this case.
Case 2. Now we look at the edges $m \in A_{(k,l)}$, by Lemma 3.3

$$f'_m = f_m = -\lambda F_m^0$$

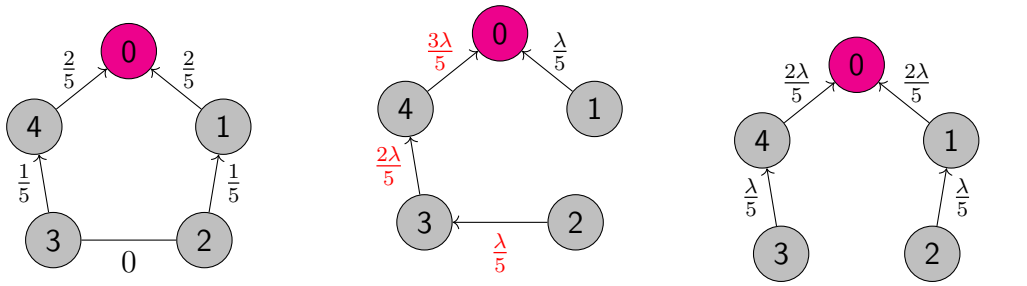
We have that $\lambda < 1$, therefore $|f_m| < |F_m^0|$ and the cascade stops in this case as well.
Case 3. The case where the edge is the removed one or part of the disconnected set is trivial as the flow is 0.
This concludes the proof as there does not exceed an edge for which the flow exceeds the emergency edge limit. □

In summary, this demonstrates that in tree-structured power networks, cascades are contained after the initial failure of a single edge. The analysis reveals that due to the absence of cycles in a tree structure, the flow on any edge is only affected by the number of nodes reliant on that edge for transmission to the demand node. Consequently, when one edge is randomly removed, the resulting redistribution of flows does not exceed the remaining edges' capacities, thus ending the cascade.

3.2 Cascade on odd cycle

In this section consider $\mathcal{O}$ to be a connected cycle of odd size n with as unit demand node 0 and set of nodes and edges $\mathcal{N}$ and $\mathcal{E}$. In this section, we show that after the initial edge failure $m \in \mathcal{E}$

- if $F_m^0 = 0$, the cascade stops,
- if $F_m^0 > 0$ there is second edge failure at the edge $e \in \mathcal{E}$ with $F_e^0 = 0$ and then the cascade stops.



(a) emergency edge limits (b) absolute flows (case 1) (c) absolute flows (case 2)

Figure 3.3: The edge limits and flow after first edge failure in an odd cycle

In Figure 3.3, an odd cycle is shown. Here, we clearly see that the removal of an edge with a positive emergency edge limit results in the second failure in edge $(2, 3)$. However, also note that a removal of the edge $(3, 2)$ which has an emergency edge limit equal to zero, implies the cascade does not continue. Lastly, in the case that there is a second edge failure, there will not be a third because the removal of the edge $(2, 3)$ reduces the flow for the edges that also exceed their emergency edge limit. Finally, we show that the expected size for the number of nodes that will be disconnected after the cascade stops is $\frac{(n-1)(n+1)}{4n}$. The rest of this section shows the illustrated pattern in the cascade by describing the emergency edge limit and the flows and finally showing that the flows do exceed the emergency edge limits in the given cases. Additionally, the expected size of the disconnected components is described.

3.2.1 Emergency edge limits

When considering the example Figure 3.3 (a) we see two distinct paths S_3 and S_2 in which the emergency edge limit increases for each node that we get closer to the demand node. Thus, for any odd cycle $\mathcal{O}$, consider two paths S_j, where each j represents the node at the maximum distance from the demand node. On these paths, the value of F^0 is incremented by $\frac{1}{n}$ at each subsequent edge. This behaviour is defined in the following lemma.

Lemma 3.4. *(Emergency edge limit in odd cycle) Define D to be the set of the two nodes that have the furthest distance from the demand node 0. Define for both $j \in D$ the path S_j from j to the demand node, then*

$$F^0_{(S_j(i))} = \frac{i}{n} \qquad \forall \quad 0 < i < \frac{n-1}{2}.$$

This covers all edges except for the edge m between the two nodes in D, for which the emergency edge limit is

$$F^0_m = 0.$$

Proof. The PTDF matrix describes how the power flow is distributed according to Kirchhoff's Law [29]. Therefore, we prove that Kirchhoff's laws give a unique solution and hold given the emergency edge limit in the lemma.
In Section 2.1 we noted that the resistance at every edge is 1 given by the matrix **B**, therefore, by Ohm's law $V = I \cdot R$ implies that the current equals the voltage [19]. Then by Kirchhoff's Current Law, for every node a_l with power injection $g(a_l)$ (with $\lambda = 1$), we need to show that the sum incoming flow with the power injection equals to outgoing flow [14]. Note that Kirchhoff's Voltage Law does apply because there is a cycle, this states that the sum of all flows over the cycle should equal zero [14].
Define the two nodes with the furthest distance from the unit demand node by l and r and paths S_l and S_r. Let the edge (l, r) be oriented towards l. Then Kirchhoff's Current Law (KCL) states

$$F^0_{S_l(1)} = F_{(l,r)} + \frac{1}{n} \tag{7}$$

$$F^0_{S_r(1)} = -F_{(l,r)} + \frac{1}{n} \tag{8}$$

and for all $1 < i \leq \lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}$,

$$F^0_{S_l(i)} = F^0_{S_l(i-1)} + \frac{1}{n} \tag{9}$$

$$F^0_{S_r(i)} = F^0_{S_r(i-1)} + \frac{1}{n}. \tag{10}$$

To simplify notation set $F_{(l,r)} = x$. Note that the explicit form for the recursive formulas Equation 9 and Equation 10 are

$$F^0_{S_l(i)} = x + \frac{i}{n}$$

$$F^0_{S_r(i)} = -x + \frac{i}{n}.$$

Now we can use Kirchhoff's Voltage Law (KVL) to find x. By KVL

$$\begin{aligned}
\sum_{i=1}^{\frac{n-1}{2}} F^0_{S_l(i)} - \left(\sum_{i=1}^{\frac{n-1}{2}} F^0_{S_r(i)}\right) + F_{(l,r)} &= 0 \\
\sum_{i=1}^{\frac{n-1}{2}} \left(x + \frac{i}{n}\right) - \left(\sum_{i=1}^{\frac{n-1}{2}} \left(-x + \frac{i}{n}\right)\right) + x &= 0 \\
\frac{(n-1)x}{2} + \frac{1}{n}\sum_{i=1}^{\frac{n-1}{2}} i + \frac{(n-1)x}{2} - \frac{1}{n}\sum_{i=1}^{\frac{n-1}{2}} i + x &= 0 \\
\frac{2(n-1)x}{2} + x &= 0 \\
x &= 0.
\end{aligned}$$

This equality holds if $x = 0$. Therefore, we also find that for $1 \leq i \leq \frac{n-1}{2}$

$$F^0_{S_l(i)} = \frac{i}{n} \text{ and } F^0_{S_r(i)} = \frac{i}{n}.$$

□

Given the known emergency edge limits, the initial flow on all edges can also be determined, as detailed in the following corollary.

Corollary 3.2. *(Flows before initial edge failure in odd cycle) Define D to be the set of the two nodes that have the furthest distance from the demand node 0. Define for both $j \in D$ the path S_j from j to the demand node, then*

$$f_{(S_j(i))} = -\lambda\frac{i}{n} \qquad \forall \quad 0 < i < \frac{n-1}{2}.$$

This covers all edges except for the edge m between the two nodes in D, for which the emergency edge limit is

$$f_m = 0.$$

Proof. This follows from Lemma 3.2 and the flow computations in Equation 2. □

3.2.2 Flows after the initial edge failure

Now that the initial flow and emergency edge limits are determined, the next step is to determine the flow after the initial edge failure. In Figure 3.3 (b), the remaining edge on the right side has a reduced flow, while all the other nodes have an increased flow of $\frac{\lambda}{5}$ which is the edge limit of the removed edge $(2, 1)$. This behaviour is generalized in the following lemma.

Lemma 3.5. *(The flow of an odd cycle after first edge failure) Let j, r be the nodes of the furthest distance from the demand node. Assume we remove an edge $(k,l) \in S_j$ from $\mathcal{O}$ such that the orientation on (k,l) is pointed towards l. Then,*

$$f_m = \begin{cases} -\lambda \left(F^0_m - F^0_{(k,l)} \right) & m \in S_j \\ -\lambda \left(-F^0_{(k,l)} \right) & m = (j,r) \\ -\lambda \left(F^0_m + F^0_{(k,l)} \right) & m \in S_r \end{cases}$$

Proof. The flow is given by $f = -\lambda F^1$ since no disconnection occurs. All that is left show is

$$F^1_m = \begin{cases} F^0_m - F^0_{(k,l)} & m \in S_j \\ -F^0_{(k,l)} & m = (j,r) \\ F^0_m + F^0_{(k,l)} & m \in S_r \end{cases}$$

given the set-up in the lemma. We perform a case distinction.
Case 1. First consider the case where edge $m \in S_l \subset S_j$. The flow on these edges was flowing away from the edge (k,l) and was receiving the incoming flow from (k,l). Therefore the flow on m is now reduced by the amount of flow on (k,l), and, $F^1_m = F^0_m - F^0_{(k,l)}$.
Case 2. Now assume $m \in S_j \setminus S_l$. We show by induction that the lemma holds. Let a_i be the node with distance i from k that is part of path S_j, such that a_1 is the neighbour of k other than l and a_2 the neighbour of a_1 etc. We take the base case $m = (a_1, k)$. Because there are no incoming edges at k anymore, $F^1_{(a_1,k)} = -\frac{1}{n}$, where the minus sign occurs because of the change direction. The base case holds, because,

$$\begin{aligned} F^1_{(a_1,k)} &= F^0_{(a_1,k)} - F^0_{(k,l)} \\ &= F^0_{(k,l)} - \frac{1}{n} - F^0_{(k,l)} \\ &= -\frac{1}{n} \end{aligned}$$

Now assume $F^1(a_{i+1}, a_i) = F^0_{(a_{i+1},a_i)} - F^0_{(k,l)}$. Then,

$$\begin{aligned} F^1(a_{i+2}, a_{i+1}) &= F^1(a_{i+1}, a_i) + \frac{1}{n} \\ &= F^0_{(a_{i+1},a_i)} - F^0_{(k,l)} + \frac{1}{n} \\ &= F^0_{(a_{i+2},a_{i+1})} - F^0_{(k,l)}. \end{aligned}$$

This concludes our induction and finishes the proof for the case $m \in S_j \setminus S_l$.
Case 3. Assume $m = (r, j)$ such that $F^0_m = 0$. Let edge m be orientated towards j and that

j is closer to edge (k,l) than r. Assume k was at distance β from node j, then $F^0_{(k,l)} = \frac{\beta+1}{n}$ by Lemma 3.4. However, after the edge failure, node k is now the new starting point of the flow that is directed in the opposite direction. Therefore $F^1_m = -\frac{\beta+1}{n} = -F^0_{(k,l)}$.
Case 4. Let $m \in S_r$. We show by induction that the lemma holds for all edges of this path. Consider the path of the form $(r,a_1),(a_2,a_3),...(a_{m_r-1},0)$ up to the demand node such that $0 = a_{m_r}$. For the base case,

$$\begin{aligned} F^1_{(r,a_1)} &= -F^1_{(r,j)} + \frac{1}{n} \\ &= F^0_{(k,l)} + \frac{1}{n} \\ &= F^0_{(k,l)} + F^0_{(r,a_1)}. \end{aligned}$$

This proves our base case and we can now continue with the induction step. Assume $F^1(a_i,a_{i+1}) = F^0_{(a_i,a_{i+1})} + F^0_{(k,l)}$. Then,

$$\begin{aligned} F^1(a_{i+1},a_{i+2}) &= F^1(a_i,a_{i+1}) + \frac{1}{n} \\ &= F^0_{(a_i,a_{i+1})} + F^0_{(k,l)} + \frac{1}{n} \\ &= F^0_{(a_{i+1},a_{i+2})} + F^0_{(k,l)}. \end{aligned}$$

This concludes our induction and finishes the proof for the case $m \in S_r$.

□

3.2.3 Results of the cascade

As shown in Figure 3.3, the cascade process terminates when the first edge failure occurs opposite the demand node. Additionally, a secondary edge failure happens when an initial removal involves any edge with a positive emergency edge limit other than the one across from the demand node.

Theorem 3.2. *(The cascade may propagate in an odd cycle component) The cascade on $\mathcal{O}$ after one uniform random edge removal (k,l) as defined in Section 2.1*

- *if $F^0_{(k,l)} > 0$, continues by failing the edge m for which $F^0_m = 0$ and then stops,*
- *if $F^0_k = 0$, stops.*

Proof. Let F^0_m for edge m of $\mathcal{O}$, and $\lambda \in [0,1]$. Let j,r be the nodes of the furthest distance from the demand. Remove the edge $(k,l) \in S_j$ from $\mathcal{O}$.

Case 1. Firstly, consider the case $F_m^0 = 0$. Then by Lemma 3.5,

$$f_m = \begin{cases} -\lambda\left(F_m^0 - F_{(k,l)}^0\right) & m \in S_j \\ -\lambda\left(-F_{(k,l)}^0\right) & m = (j,r) \\ -\lambda\left(F_m^0 + F_{(k,l)}^0\right) & m \in S_r \end{cases}$$
$$= \begin{cases} -\lambda\left(F_m^0\right) & m \in S_j \\ 0 & m = (j,r)\,. \\ -\lambda\left(F_m^0\right) & m \in S_r \end{cases}$$

Clearly, there is no edge m for which $|f_m| > |F_m^0|$. Therefore, the cascade stops.
Case 2. Now consider the case $F_m^0 > 0$. Then by Lemma 3.5,

$$f_m = \begin{cases} -\lambda\left(F_m^0 - F_{(k,l)}^0\right) & m \in S_j \\ -\lambda\left(-F_{(k,l)}^0\right) & m = (j,r)\,. \\ -\lambda\left(F_m^0 + F_{(k,l)}^0\right) & m \in S_r \end{cases}$$

Clearly, (j,r) is the next edge to fail because $F_{(j,r)}^0 = 0$ which the exceedance is defined as ∞ as described in Section 2.1.
All that is left to show is that the cascade will stop afterwards after this second edge failure. Note that after the first edge failure, the cycle is broken up. Removing the edge (j,r) will result in the edges on the path S_r being reduced by the flow on the edge (j,r) because after the first edge failure, this has reduced to an edge component. When the second edge failure occurs there is a disconnection of edges $T \subset S_j$. Therefore, the flow after the second disconnection becomes

$$f_m = \begin{cases} -\lambda\left(F_m^0 - F_{(k,l)}^0\right) & m \in S_j \setminus T \\ -\lambda\left(F_m^0 + F_{(k,l)}^0 - \left(-F_{(k,l)}^0\right)\right) & m \in S_r \end{cases}$$
$$= \begin{cases} -\lambda\left(F_m^0 - F_{(k,l)}^0\right) & m \in S_j \setminus T \\ -\lambda\left(F_m^0\right) & m \in S_r \end{cases}.$$

Clearly, there is no edge m anymore for which $|f_m| > |F_m^0|$ therefore, the cascade stops. □

With this theorem, it is possible to find the number of nodes that get disconnected because the pattern of edge failures is defined after the first edge failure.

Corollary 3.3. *(Size of disconnected component) Let S be the set of nodes that are disconnected from the demand node after the cascade has stopped. Then,*

$$\mathbb{E}(|S|) = \frac{(n-1)(n+1)}{4n}.$$

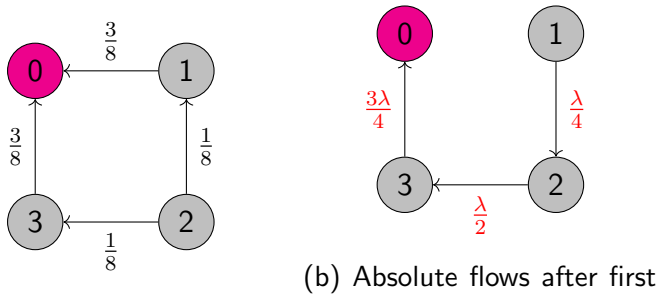
Proof. The probability of an edge being the first failure is $\frac{1}{n}$. The maximum size that gets disconnected is in the case when the first edge failure is one of the edges closest to the demand node. The second edge removed will then be the one farthest from the demand node, resulting in $|S| = \lfloor \frac{n}{2} \rfloor$. Note that $\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}$ as n is odd. The minimum size of the disconnected component is 0 in the case that the first edge failure is the one with an emergency edge limit of zero. However, the minimum in the case that this edge is not the first edge failure, is $|N| = 1$. This happens when the first edge failure is one of the edges that neighbours the edge with an emergency edge limit of zero. Then the second failure is the edge next to it, resulting in one node being disconnected. Then by symmetry

$$\begin{aligned}\mathbb{E}(|S|) &= 0 \cdot \frac{1}{n} + \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \frac{i}{n} + \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \frac{i}{n} = \sum_{i=1}^{\frac{n-1}{2}} \frac{2i}{n} = \frac{2}{n} \left(\frac{\left(\frac{n-1}{2}\right)\left(\frac{n-1}{2}+1\right)}{2} \right) \\ &= \frac{\left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right)}{n} = \frac{(n-1)(n+1)}{4n}\end{aligned}$$

□

3.3 Cascade on even cycle

In this section, consider $\mathcal{C}$ to be a connected cycle of even size n with unit demand node 0 and set of nodes and edges $\mathcal{N}$ and $\mathcal{E}$, respectively. In this section, we show that after the initial edge failure, $m \in \mathcal{E}$ the cascade may continue and the second edge failure is the edge furthest away from both the demand node and the edge that was removed. However, we also show that the second edge failure only occurs when $\lambda > K$ for a certain threshold K. In Figure 3.4, the edge to fail is $(2, 3)$, but only if $\frac{\lambda}{2} > \frac{1}{8}$ implying $\lambda > \frac{1}{4}$ for the cascade to



(a) Emergency edge limits

(b) Absolute flows after first edge failure

Figure 3.4: Even cycle with edge limits and flow after edge failure

continue. Thus, we see a similar pattern compared to the odd cycle, however, in this case, there exists a bound, so the cascade does not always continue after the first edge failure. This is partly because there is no edge with an emergency edge limit of 0 compared to the odd cycle. However, the node j at the farthest distance $\frac{n}{2}$ from the demand node splits its

generation into two equal parts, resulting in lower emergency edge limits over those edges. Therefore, in the remainder of this section, we consider two paths S_{j_l} and S_{j_r} which both cover a different path from node j to the demand node. Together they cover all edges in the cycle and have no overlap.

3.3.1 Emergency edge limits

As illustrated by Figure 3.4 there is a shortest path for almost every edge, except for the one which is the furthest from the demand node. This is shown in the following lemma.

Lemma 3.6. *(None-unique path in even cycle) Every node $j \in \mathcal{N}$ has a unique shortest path, except for the node at distance $\frac{n}{2}$, which has two shortest paths.*

Proof. Let $j \in \mathcal{N}$ such that the distance to 0 is less than $\frac{n}{2}$. Again there are two paths from j to 0, one of size k and the other of size $n-k$. Since $k \neq \frac{n}{2}$. Either $k > n-k$ or $k < n-k$, therefore there is a unique shortest path.
However, in the case that the distance to node 0 equals $\frac{n}{2}$ we have that both paths have equal length and therefore this node has two shortest paths. □

The emergency edge limits in an even cycle are very similar to those in an odd cycle. However, the two longest paths originate from just one node. Therefore, this initial node splits the generation into two fractions of $\frac{1}{2n}$. After which at each step closer to the demand node an additional capacity of $\frac{1}{n}$ is added to the edge compared to the edge at one distance further from the demand node. In the following lemma, this is specified.

Lemma 3.7. *(Emergency edge limits in even cycle) Let j be the node at distance $n/2$ from the demand node and consider the two paths S_{j_l} and S_{j_r} from node j to the demand node. Given this setup, the emergency edge limits in an even cycle are given by*

$$F_{(S_{j_l}(i))} = F_{(S_{j_r}(i))} = \frac{i-1}{n} + \frac{1}{2n} \qquad \forall \quad 0 < i \leq \frac{n}{2}.$$

Proof. The flow in the DC-OPF model follows the Kirchoff's Law [29]. Therefore, we prove that Kirchhoff's laws are given by a unique solution and hold for the emergency edge limit in the lemma.
In Section 2.1 we noted that the resistance at every edge is 1 given by the matrix $\mathbf{B}$, therefore, by Ohm's law $V = I \cdot R$ implies that the current equals the voltage [19]. Then by Kirchhoff's Current Law, for every node a_l with power injection $g(a_l)$ (with $\lambda = 1$), we need to show that the sum incoming flow with the power injection equals the outgoing flow [14]. Note that Kirchhoff's Voltage Law also applies because there is a cycle, and states that the sum of all flows over the cycle should equal zero [14].
Then Kirchhoff's Current Law (KCL) states

$$F^0_{S_{j_l}(1)} + F^0_{S_{j_r}(1)} = \frac{1}{n} \tag{11}$$

and for all $1 < i \leq \frac{n}{2}$,

$$F^0_{S_{j_l}(i)} = F^0_{S_{j_l}(i-1)} + \frac{1}{n}, \tag{12}$$

$$F^0_{S_{j_r}(i)} = F^0_{S_{j_r}(i-1)} + \frac{1}{n}. \tag{13}$$

Set x to be the fraction of the generation of node j that is sent over the path S_{j_l}. Then $F^0_{S_{j_l}(1)} = \frac{x}{n}$ and by Equation 11, $F^0_{S_{j_r}(1)} = (1-x)\frac{1}{n}$. Note that the explicit form for the recursive formulas Equation 12 and Equation 13 are

$$F^0_{S_l(i)} = F^0_{S_l(1)} + \frac{i-1}{n},$$
$$F^0_{S_r(i)} = F^0_{S_r(1)} + \frac{i-1}{n}$$

with starting points

$$F^0_{S_l(1)} = \frac{x}{n},$$
$$F^0_{S_r(1)} = (1-x)\frac{1}{n}.$$

Now we can use Kirchhoff's Voltage Law (KVL) to find x. By KVL

$$F^0_{S_{l_j}(1)} + \sum_{i=2}^{\frac{n}{2}} F^0_{S_{j_l}(i)} - \left(F^0_{S_{j_r}(1)} + \sum_{i=2}^{\frac{n}{2}} F^0_{S_{j_r}(i)} \right) = 0$$

$$F^0_{S_{l_j}(1)} + \sum_{i=1}^{\frac{n}{2}} F^0_{S_{j_l}(i)} - F^0_{S_{l_j}(1)} - \left(F^0_{S_{j_r}(1)} + \sum_{i=1}^{\frac{n}{2}} F^0_{S_{j_r}(i)} - F^0_{S_{j_r}(1)} \right) = 0$$

$$\sum_{i=2}^{\frac{n}{2}} \left(\frac{x}{n} + \frac{i-1}{n} \right) - \sum_{i=1}^{\frac{n}{2}} \left(\frac{1-x}{n} + \frac{i-1}{n} \right) = 0$$

$$\frac{n}{2} \left(\frac{x}{n} - \frac{1-x}{n} \right) = 0$$

$$\frac{x}{2} - \frac{1-x}{2} = 0$$

$$x = \frac{1}{2}.$$

This equality holds if $x = \frac{1}{2}$ Therefore, we find that for $1 \leq i \leq \frac{n}{2}$

$$F^0_{S_l(i)} = F^0_{S_r(i)} = \frac{1}{2n} + \frac{i-1}{n}.$$

□

Given the known emergency edge limits, the initial flow on all edges can also be determined, as detailed in the following corollary.

Corollary 3.4. *(Flows before initial edge failure in even cycle) Let j be the node at distance $n/2$ from the demand node and consider the two paths S_{j_l} and S_{j_r} from node j to the demand node. Given this setup, the flows of an even cycle is given by*

$$f_{(S_{j_l}(i))} = -\lambda \left(\frac{i-1}{n} + \frac{1}{2n} \right) \qquad \forall \quad 0 < i \leq \frac{n}{2}$$

and

$$f_{(S_{j_r}(i))} = -\lambda \left(\frac{i-1}{n} + \frac{1}{2n} \right) \qquad \forall \quad 0 < i \leq \frac{n}{2}.$$

Proof. This follows from Lemma 3.2 and the flow computations in Equation 2. □

3.3.2 Flows after the initial edge failure

Now that, the initial flow and emergency edge limits are determined, the next step is to determine the flow after the initial edge failure. The next step in the cascade is a failure. In the example in Figure 3.4, the flow after the edge failure for edge $(1, 2)$ is $\frac{\lambda}{4}$ because the flow is reduced over the edge by $\frac{\lambda 3}{8}$. The remaining edges have a flow increased by $\frac{\lambda 3}{8}$ which is the initial flow of the removed edge $(0, 1)$. This behaviour is generalized in the following lemma.

Lemma 3.8. *(Flow of an even cycle after the first edge failure) Consider the node j that has the furthest distance $\frac{n}{2}$ from the demand node 0. Assume we have two paths S_{j_l} and S_{j_r} which both cover a different path from node j to the demand node. Assume all edges are oriented towards the demand node on the shortest path. Now assume we remove an edge (k, l), with an orientation towards l, from $\mathcal{C}$. Assume w.l.o.g. that the edge that was removed was part of the patch S_{j_l}. Then,*

$$f_m = -\lambda \left(F^0_m - F^0_{(k,l)} \right) \quad m \in S_{j_l}.$$

For the other edges, which are on the other path we have to show that

$$f_m = -\lambda (F^0_m + F^0_{(k,l)}) \quad m \in S_{j_r}.$$

Proof. The flow is given by $f = -\lambda F^1$ since no disconnection occurs. All that is left to show is

$$F^1_m = F^0_m - F^0_{(k,l)} \quad m \in S_{j_l}.$$

For the other edges, which are on the other path we have

$$F^1_m = F^0_m + F^0_{(k,l)} \quad m \in S_{j_r},$$

given the set-up in the lemma. We perform a case distinction.
Case 1. Let $m \in S_{j_l}$. First consider $m \in S_l \subset S_{j_l}$ The flow on these edges was flowing away from the edge (k,l) and was receiving the incoming flow from (k,l), therefore, the flow on m is now reduced by the amount of flow on (k,l), hence, $F^1_m = F^0_m - F^0_{(k,l)}$.
Now consider the other edges $m \in S_{j_l} \setminus S_l$. We show by induction that the lemma holds. Let a_i be the node with distance i from k that is part of path S_{j_l}, such that a_1 is the neighbour of k other than l and a_2 the neighbour of a_1 etc. Assume j is at distance β from k then $j = a_\beta$. We take the base case $m = (a_1, k)$. Because there are no incoming edges at k anymore, $F^1_{(a_1,k)} = -\frac{1}{n}$, where the minus sign occurs because of the change direction. The base case holds, because

$$F^1_{(a_1,k)} = F^0_{(a_1,k)} - F^0_{(k,l)} = \frac{\beta - 1}{n} + \frac{1}{2n} - \left(\frac{\beta}{n} + \frac{1}{2n}\right) = -\frac{1}{n}.$$

Now assume $F^1(a_{i+1}, a_i) = F^0_{(a_{i+1},a_i)} - F^0_{(k,l)}$. Then,

$$F^1_{(a_{i+2},a_{i+1})} = F^1_{(a_{i+1},a_i)} + \frac{1}{n} = F^0_{(a_{i+1},a_i)} - F^0_{(k,l)} + \frac{1}{n} = F^0_{(a_{i+2},a_{i+1})} - F^0_{(k,l)}.$$

This concludes the proof of our first case.
Case 2. Let $m \in S_{j_r}$. We show by induction that the lemma holds for all edges of this path. Consider the path of the form $(j, a_1), (a_2, a_3), ..., (a_{m_{j_r}-1}, 0)$ up to the demand node such that $0 = a_{m_{j_r}}$ is the demand node. For the base case,

$$F^1_{(r,a_1)} = -S_{j_l}(1) + \frac{1}{n} = \frac{\beta}{n} + \frac{1}{n} = \left(\frac{\beta}{n} + \frac{1}{2n}\right) + \frac{1}{2n} = F^1_{(k,l)} + F^0_{(r,a_1)}.$$

Assume $F^1(a_i, a_{i+1}) = F^0_{(a_i,a_{i+1})} + F^0_{(k,l)}$. Then,

$$F^1(a_{i+1}, a_{i+2}) = F^1(a_i, a_{i+1}) + \frac{1}{n} = F^0_{(a_i,a_{i+1})} + F^0_{(k,l)} + \frac{1}{n} = F^0_{(a_{i+1},a_{i+2})} + F^0_{(k,l)}.$$

This shows the induction for our second case and finalizes the proof. □

3.3.3 Results of the cascade

As mentioned previously, the likelihood of a cascade in the network is contingent on a threshold value for λ and the initial location of an edge failure, however, we also see that even if a second edge failure occurs, the cascade will always cease afterwards. In the following, we focus on determining the threshold condition necessary for this second edge failure to occur. Additionally, we establish that the cascade is limited to a maximum of two edge failures. The expected number of nodes that become disconnected as a result of these cascading failures will also be determined.

Theorem 3.3. *(The cascade may propagate in an even cycle component) Let the first edge failure m be such that either $S_{j_r}(\alpha) = m$ or $S_{j_l}(\alpha) = m$ for $\alpha \in \mathbb{N}$. Assume $m \in S_{j_l}$ or $m \in S_{j_r}$ then this second edge failure is the edge $S_{j_r}(1)$ or $S_{j_l}(1)$ respectively if*

$$\alpha > \frac{1}{2\lambda}.$$

Otherwise, the cascade terminates after the first failure. Furthermore, the cascade will stop after at most two edge failures.

Proof. Consider the node j that has the furthest distance $\frac{n}{2}$ from the demand node 0. Assume we have two paths S_{j_l} and S_{j_r} which both cover a different path from node j to the demand node. Now assume we remove an edge (k,l) with the orientation towards the node l from $\mathcal{C}$. Let S_l be the path from the node l to the demand node. Assume w.l.o.g. that the edge that was removed was part of path S_{j_l} such that $S_{j_l}(\alpha) = (k,l)$. We show that $S_{j_r}(1)$ is the next edge that fails. We consider three cases for the edges.
Case 1. Let edge $m_i \in S_l$ such that $\alpha < i$ and $S_{j_l}(i) = m_i$. Note that $\alpha < i$ because the edge (k,l) is further from the demand node than any edges on the path starting from node l to the demand node. By Lemma 3.8,

$$f_{m_i} = -\lambda \left(F^0_{m_i} - F^0_{(k,l)}\right).$$

By Lemma 3.7 $F^0_{m_i} = \frac{i-1}{n} + \frac{1}{2n} > \frac{\alpha-1}{n} + \frac{1}{2n} = F^0_{(k,l)}$ as $\alpha < i$ for every i. Therefore the flow on edge m_i does not exceed the edge limit. Note that the set S_l can also be the empty set when (k,l) is an edge directly connected to the demand node, in which case there is also no edge with any exceedance.
Case 2. Now take $m_i \in S_{j_l} \setminus (S_l \cup (k,l))$ such that $S_{j_l}(i) = m_i$. Note that $\alpha > i$ if $S_{j_l} \setminus (S_l \cup (k,l))$ is not empty because the edge (k,l) is closer to the demand node than any edge between the nodes j and k on the path S_{j_l}. By Lemma 3.8,

$$f_m = -\lambda \left(F^0_{m_i} - F^0_{(k,l)}\right).$$

By Lemma 3.7

$$|f_m| = \left|-\lambda\left(\frac{i-1}{n} + \frac{1}{2n} - \left(\frac{\alpha-1}{n} + \frac{1}{2n}\right)\right)\right| = \left|-\lambda\left(\frac{i-\alpha}{n}\right)\right| = -\lambda\left(\frac{i-\alpha}{n}\right).$$

As $\alpha > i$, there exists i such that $|f_m| > \frac{i-1}{n} + \frac{1}{2n} = |F^0_m|$ so some emergency edge limits could be exceeded. Hence, the relative edge exceedance for m_i is

$$\begin{aligned}\frac{|f_{m_i}|}{|F^0_{m_i}|} &= \frac{\left|-\lambda\left(F^0_{m_i} - F^0_{(k,l)}\right)\right|}{|F^0_{m_i}|} = \frac{\left|-\lambda\left(\frac{i-\alpha}{n}\right)\right|}{\frac{i-1}{n} + \frac{1}{2n}} = \frac{-\lambda\left(\frac{i-\alpha}{n}\right)}{\frac{i-0.5}{n}} = \frac{\lambda(\alpha-i)}{i-0.5} \\ &= \frac{-\lambda(i-0.5) + \lambda(\alpha+0.5)}{i-0.5} = -\lambda + \frac{\lambda(\alpha+0.5)}{i-0.5}.\end{aligned}$$

Note that in this case the the relative exceedance is maximal for $i = 1$. Also, note again that $S_{j_l} \setminus (S_l \cup (k,l))$ can be empty when (k,l) with $k = j$. In that case, no edges exist and therefore, there is no edge with any exceedance.
Case 3. Now take $m_i \in S_{j_r}$ such that $S_{j_r}(i) = m_i$. Note that $0 < i < \frac{n}{2}$ and that S_{j_r} is never the empty set. By Lemma 3.8,

$$f_{m_i} = -\lambda \left(F^0_{m_i} + F^0_{(k,l)}\right)$$

$F^0_{(k,l)} > 0$, so the emergency edge limits are exceeded. Therefore, the relative edge exceedance for m_i is

$$\begin{aligned}\frac{|f_{m_i}|}{|F^0_{m_i}|} &= \frac{\left|-\lambda\left(F^0_{m_i} + F^0_{(k,l)}\right)\right|}{|F^0_{m_i}|} = \frac{\left|-\lambda\left(\left(\frac{i-1}{n} + \frac{1}{2n}\right) + \left(\frac{\alpha-1}{n} + \frac{1}{2n}\right)\right)\right|}{\frac{i-1}{n} + \frac{1}{2n}} \\ &= \frac{\left|-\lambda\left(\left(\frac{i+\alpha-1}{n}\right)\right)\right|}{\frac{i+0.5}{n}} = \frac{\lambda\left(\frac{i+\alpha-1}{n}\right)}{\frac{i-0.5}{n}} = \frac{\lambda(i+\alpha-1)}{i-0.5} = \frac{\lambda(i-0.5)+\lambda(\alpha+0.5)+1}{i-0.5} \\ &= \lambda + \frac{\lambda(\alpha+0.5)+1}{i-0.5}\end{aligned}$$

Note that in this case the the relative exceedance is maximal for $i = 1$.
Largest relative exceedance. We have two candidate edges to fail, namely $S_{j_l}(1)$ and $S_{j_r}(1)$, assuming $S_{j_l} \setminus S_l$ is not empty. Their relative exceedance is respectively

$$\frac{|f_{S_{j_l}(1)}|}{|F^0_{S_{j_l}(1)}|} = \frac{\lambda(\alpha-1)}{1-0.5},$$

and

$$\frac{|f_{S_{j_r}(1)}|}{|F^0_{S_{j_r}(1)}|} = \frac{\lambda(1+\alpha-1)}{i-0.5} = 2\lambda\alpha.$$

Clearly,

$$\frac{|f_{S_{j_r}(1)}|}{|F^0_{S_{j_r}(1)}|} > \frac{|f_{S_{j_l}(1)}|}{|F^0_{S_{j_l}(1)}|},$$

therefore the next edge to fail is the edge $S_{j_r}(1)$. Note that the cardinality of $S_{j_l} \setminus (S_l \cup (k,l))$ therefore does not matter.
Threshold on α. Now we show the threshold on α. We find that

$$\begin{aligned}|F^0_{S_{j_r}(1)}| &< |f_{S_{j_r}(1)}| \\ 1 &< 2\alpha\lambda \\ \frac{1}{2\lambda} &< \alpha.\end{aligned}$$

Note that $\alpha > 0$ therefore the threshold is

$$\alpha > \frac{1}{2\lambda}.$$

There are at most 2 edge failures. Note that after the second edge fails there is a disconnection and the only edges that are left are $S_{j_r} \setminus S_{j_r}(1)$ and S_l. Note that removing the edge $S_{j_r}(1)$ only affects the flow in edges $m_i \in S_{j_r} \setminus S_{j_r}(1)$ such that $S_{j_r}(i) = m_i$. Take such an m_i then the flow f', disregarding the decrease caused by the redistribution of the demand and generation vector, after the second edge failure for these edges is

$$\begin{aligned} f'_{m_i} &= f_{m_i} - f_{S_{j_r}(1)} \\ &= -\lambda \left(F^0_{m_i} + F^0_{(k,l)}\right) - \lambda \left(\frac{1}{2n} + F^0_{(k,l)}\right) \\ &= -\lambda \left(F^0_{m_i} - \frac{1}{2n}\right). \end{aligned}$$

Note that $|f'_{m_i}| < |F^0_{m_i}|$ even without the redistribution of the demand and generation vector, thus the cascade stops. □

With this, it is possible to find the number of nodes that get disconnected because the pattern of edge failures is defined after the first edge failure.

Corollary 3.5. *(Size of disconnected component) Let the first edge failure m be such that either $S_{j_r}(\alpha) = m$ or $S_{j_l}(\alpha) = m$. Let S be the set of nodes that are disconnected from the demand node after the cascade has stopped then*

$$\mathbb{E}[|S|] = \frac{2}{n} \sum_{\alpha=\lfloor \frac{1}{2\lambda} \rfloor + 1}^{\frac{n}{2}} \alpha.$$

Proof. Let K be the shortest distance to the first edge failure to the demand node. Recall that the first edge to fail is chosen uniformly at random. By symmetry of the cycle, there are always exactly two edges at distance $\frac{n}{2} - \alpha$ from the demand node for $\alpha \in [1 \dots \frac{n}{2}]$. Hence $K \sim \text{Unif}(\{1, ..., \frac{n}{2}\})$ and $\mathbb{P}(K = \alpha) = \frac{2}{n}$. Due to Theorem 3.3, given that $K = \alpha > \frac{1}{2\lambda}$, then $|S| = \alpha$. This is because there are α nodes between the edges that fail. Hence, by conditioning on the value K we obtain

$$\mathbb{E}[|S| \mid K = \alpha] = \alpha \mathbb{I}\left\{\alpha > \frac{1}{2\lambda}\right\}.$$

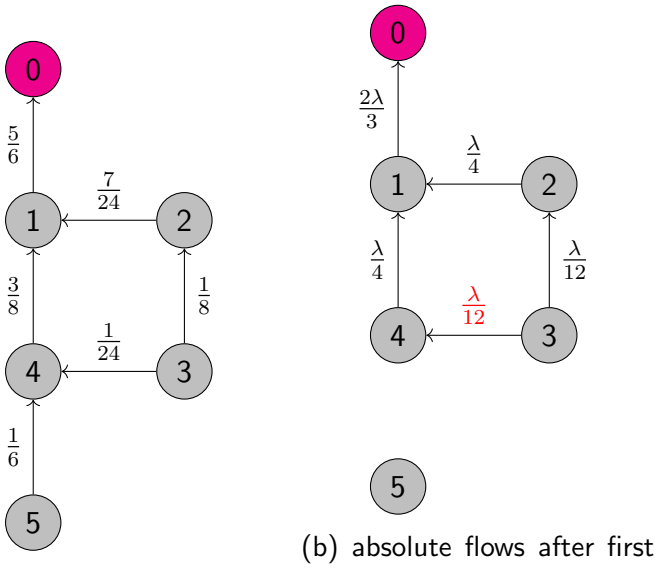
Finally, applying the law of total probability we obtain

$$\mathbb{E}[|S|] = \sum_{\alpha=1}^{\frac{n}{2}} \mathbb{E}[|S| \mid K = \alpha] \mathbb{P}(K = \alpha) = \sum_{\alpha=1}^{\frac{n}{2}} \alpha \frac{2}{n} \mathbb{I}\left\{\alpha > \frac{1}{2\lambda}\right\} = \frac{2}{n} \sum_{\alpha=\lfloor \frac{1}{2\lambda} \rfloor + 1}^{\frac{n}{2}} \alpha.$$

□

3.4 A combination of trees and cycles

From the analyses in the preceding sections, it's evident that within cycle structures, a cascade may lead to one additional edge failure after the initial edge failure. Conversely, in tree structures, a cascade halts after the first edge failure. However, this section introduces a scenario illustrating that if a tree component is connected to a cycle, a disconnection within the tree can still trigger further cascading events.
This observation underscores that in graphs with cycle components, the likelihood of cascade continuation is notably higher. This increased tendency for propagation is not limited to scenarios where an edge failure occurs within a cycle; it extends to situations where an edge failure happens in a different part of the graph. The presence of cycles in the network thus amplifies the potential for cascading events. Next to that, when the cascade results in a tree network, the cascade does not necessarily stop when the initial network is not a tree. We illustrate this with an example.



(a) Emergency edge limits

(b) absolute flows after first edge failure

Figure 3.5: Combination of tree and cycle structures

In Figure 3.5, we observe a unique graph that combines elements of both cycle and tree structures, accompanied by its edge limits. Unlike the more predictable flow behaviours in standalone cycle and tree graphs, this hybrid structure exhibits a more complex pattern of edge limits. This complexity arises primarily due to the distribution of flow from node 3. Instead of an equal distribution, the generation from node 3 is unevenly divided, with $\frac{-\lambda}{8}$ directed towards node 2 and $\frac{-\lambda}{24}$ towards the other path. Note that despite this uneven distribution, the total flow from node 3 remains consistent with its generation of $\frac{-\lambda}{6}$. This flow strategy ensures that no single edge bears excessive flow unless it represents the sole

route. In this case, the edge $(1,4)$ is avoided because at node 4, the tree component merges, resulting in more pressure on the edge $(1,4)$. Consequently, more flow is routed from node 3 along paths that circumvent this edge to prevent overloading. This adjustment results in a relatively low edge limit on edge $(3,4)$, and therefore in a higher probability of this edge breaking in the next step.
In Figure 3.5, following the disruption of the edge $(4,5)$, node 3 reverts to an equal flow division. This change leads to a scenario where $|f_{(3,4)}|$ exceeds $|F^0_{(3,4)}|$. Additionally, due to this disconnection, the demand and generation vectors undergo necessary redistribution. Despite the redistribution, the edge $(4,3)$ will still be the next to break for a certain λ, as we see in the following derivation.
In the described situation, the new demand and generation vectors are

$$d' = \left(1 - \frac{\lambda}{6}, 0, 0, 0, 0\right) \text{ and } g' = \left(\frac{\lambda}{6} + (1-\lambda), \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}\right).$$

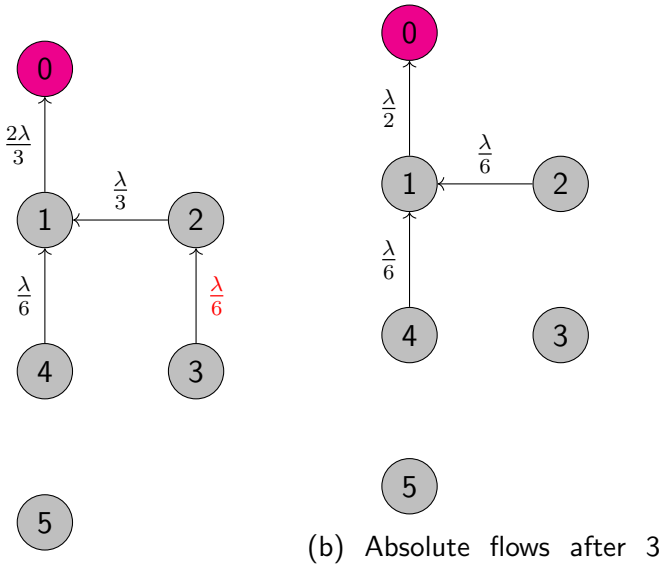
Then,

$$\begin{aligned} g' - d' &= \left(-\lambda + \frac{2\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}\right) \\ &= \left(\frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}, \frac{\lambda}{6}\right) + \left(-\lambda + \frac{\lambda}{6}, 0, 0, 0, 0\right) \end{aligned}$$

The entry in the first column of the PTDF V corresponding to the edge $(3,4)$ is $\frac{1}{10}$. As discussed in Equation 2, the flow can be computed as follows

$$f_{(3,4)} = \frac{1}{10}\left(-\lambda + \frac{\lambda}{6}\right) = -\frac{\lambda}{12}$$

Therefore, if $\lambda > \frac{1}{2}$, then $|f_{(3,4)}| = \frac{\lambda}{12} > \frac{1}{24} = |F^0_{(3,4)}|$ and $(3,4)$ is the next edge to break. Consequently, in this example for $\lambda > \frac{1}{2}$ the second edge to break is $(3,4)$.
Note that, this is not the end of the cascade. Figure 3.6 (a) shows that we end up with a tree after the second edge failure. From Section 3.1, we may have the intuition that the cascade would stop, but this example shows that the intuition is incorrect. The results shown about cycles and trees in the previous section are only valid for the initialization of the cascade. Therefore, obtaining a tree after a few edge failures does not imply that the cascade stops. In this case, the cascade does stop after the last edge failure $(2,3)$ as is depicted in Figure 3.6 (b).
The key observation is that in a graph with both cycle and tree components, an edge failure in the tree part can still trigger cascading effects due to the cycle elements. Still, the presence of more cycles and generally the presence of more connectivity seem to further the propagation of the cascade. Yet, the question is whether the additional edge failures lead to larger and more components being disconnected. We study these effects in terms of the ER graph in the next chapter.



(a) Absolute flows

(b) Absolute flows after 3 edge failures, cascade stops

Figure 3.6: Combination of tree and cycle structures

3.5 Insights

Before we discuss the results, we summarize the insights gained from the previous chapter. Our key findings are for a cascade on a graph $\mathcal{G}$:

1. If the initial graph $\mathcal{G}$ is a tree, a cascade stops after the initial failure,
2. If the initial graph $\mathcal{G}$ is a cycle, a cascade stops after at most two failures,
3. If $\mathcal{G}$ consists of a mix of cycle and tree components, and the initial edge failure is in either a tree or cycle component, a cascade might propagate further.

It is essential to underline that the findings mentioned in points 1 and 2 are applicable exclusively when the original structure of the graph $\mathcal{G}$ begins as a tree or a cycle. Thus, the emergence of a subgraph that is entirely a tree after a series of edge failures or disconnections does not guarantee the end of the cascade. This raises the question of how the cascade behaves on more complex topology structures. Therefore, we consider the ER model in the next chapter.

4 Results for cascade in Erdős–Rényi model

In this chapter, we explore the numerical results of the cascade in the Erdős–Rényi model, denoted by $ER_n(p)$, which is defined by n vertices and edge probability p. The largest connected component generated by $ER_n(p)$ is referred to as $\mathcal{G}$, with edges $\mathcal{E}$ and nodes $\mathcal{N}$.

Recall from Section 2.2.1 that as p increases (for fixed number of nodes n), the expected degree as well as the expected number of cycles grows. Furthermore, in Section 3 we show that a cascade stops immediately in a tree and has at most two failures in a cycle. However, in Section 3.4 we found that when $\mathcal{G}$ consists of trees and cycles, the cascade might continue even if the first disconnection occurs in a tree component. Thus, when $\mathcal{G}$ has a more complex structure with several cycles, the cascade behaviour becomes more complex. Therefore, the more complex structures are numerically analyzed in the rest of this chapter by looking at the Erdős–Rényi model.

4.1 Ratio of edge failures

In Section 2.2.1, we observed an increasing trend in the average degree K of the giant component $\mathcal{G}$ with the rise in probability p. The average degree, defined as $K = \frac{2|\mathcal{E}|}{|\mathcal{N}|}$, illustrates the average number of connections per node within the giant component. This relationship is confirmed by simulation results presented in Figure 4.1, where a clear positive correlation between p and K is evident.

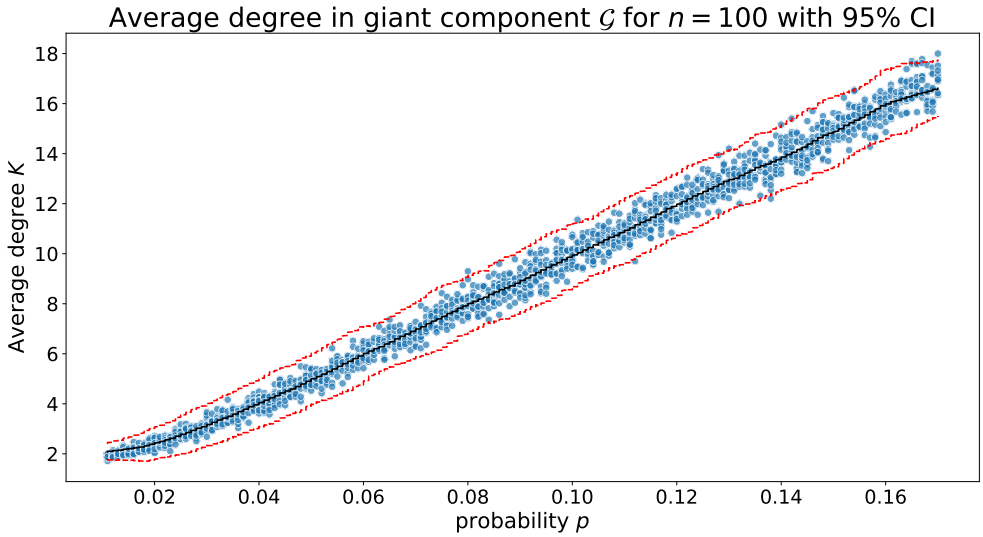


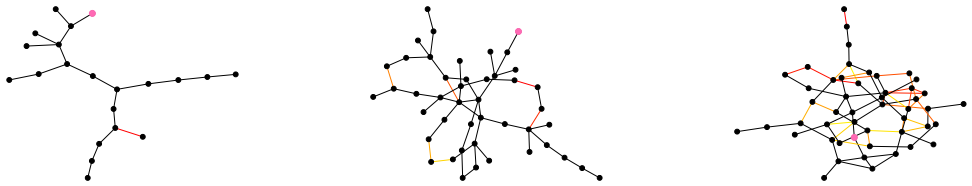
Figure 4.1: Figure on average degree K in giant component with 1600 simulation runs of a cascade with $n = 100$. The black line shows the trend and the red line shows the confidence interval.

With high probability an average degree of at least 2 is needed for cycles to exist in the giant $\mathcal{G}$. Near the percolation threshold $p = \frac{1}{n}$ (which equates to 0.01 for $n = 100$ in Figure 4.1), as discussed in Section 2.2.1, the average degree hovers around 2. Consequently, at this

point, the probability of cycle formation within $\mathcal{G}$ is small. However, as p increases, we observe that K significantly surpasses 2, ensuring the existence of cycles within $\mathcal{G}$.
Building upon the findings from Section 3, which suggest that the occurrence of cycles contributes to a higher amount of edge failures, it stands to reason that as the probability p increases, the tendency for cascades to propagate may be strengthened. This is attributed to the higher occurrence of cycles when p increases. In light of this, we establish a metric that enables us to study this hypothesis.

Definition 4.1. *(Ratio of edge failures) Let L be the number of edges that fail due to having maximum relative edge exceedance at some point during the cascade until the cascade stops in $\mathcal{G}$. The ratio of edge failures is given by $R = \frac{L}{|\mathcal{E}|}$.*

In Figure 4.2 we observe a few graph examples for increasing probability p from left to right. In Figure 4.2a we observe only one edge failure as the initial graph is a tree and in Figure 4.2b multiple edge failures happen in different cycles and there seem to be two failures per cycle. While it might be difficult to observe a pattern in Figure 4.2c note that as p increases the number of cycles increases, therefore Figure 4.2c has the highest value of p but also the highest number of cycles. Due to the increase in cycles, we also notice that the ratio of edge failures R increases as p increases, which validates our hypothesis.



(a) $p = 0.0202$, $|\mathcal{E}| = 20$, $L = 1$, $R = 0.05$ (b) $p = 0.0402$, $|\mathcal{E}| = 51$, $L = 6$, $R = 0.118$ (c) $p = 0.062$, $|\mathcal{E}| = 71$, $L = 29$, $R = 0.408$

Figure 4.2: The figures show a cascade with $\lambda = 0.9$ on the giant component $\mathcal{G}$ for $n = 50$. The colouring from red to yellow shows the first edge failure until the last, respectively. The pink node represents the unit demand node.

To confirm our hypothesis we look at Figure 4.3 and see that indeed there is a clear concave trend. The points that are not participating in this trend are those forming the bottom horizontal line. For these cases, the first edge failure happens in a component at the end of the network, which is connected to the component with a unit demand node via one edge. Therefore, the disconnection of this component does not result in redirection over edges within the component with the unit demand node. Thus, in these cases, there is just a small number of disconnections in components which are on the ends of the network. Consequently, these edge failures do not result in a significant cascade propagation as they

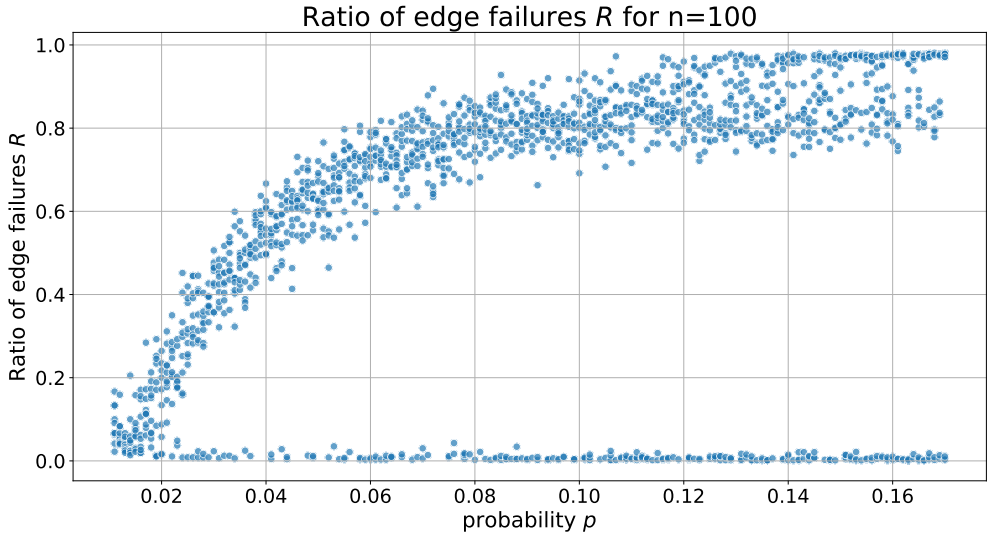


Figure 4.3: Figure on the ratio of edge failures R in giant component with 1600 simulation runs of a cascade with $n = 100$.

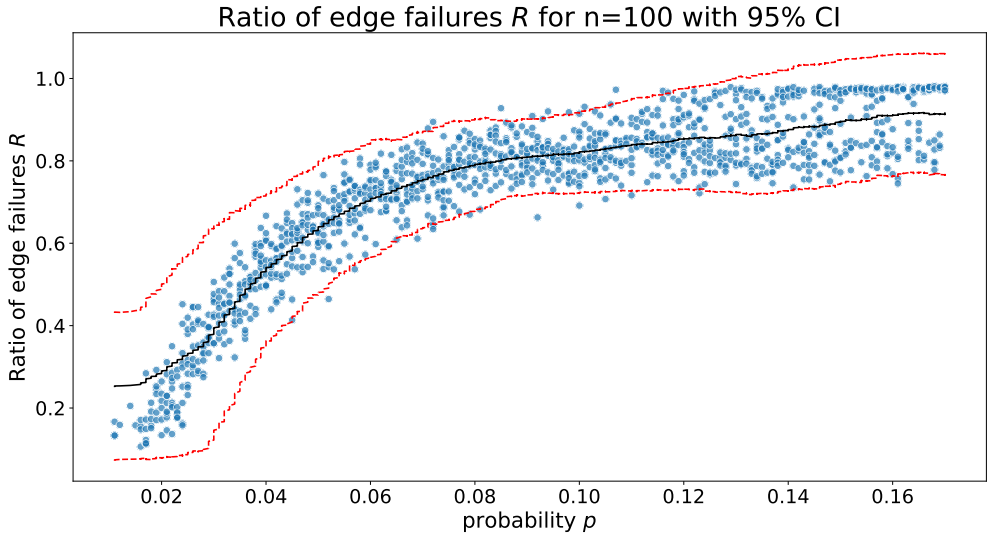


Figure 4.4: Figure on the ratio of edge failures R in giant component, where $R > 0.1$ with 1600 simulation runs of a cascade with $n = 100$. The black line shows the trend and the red line shows the confidence interval.

stop after a small number of edge failures.
To be able to say more about the cases of the points that do have more edge failures, we

take a look at Figure 4.4 which only takes the points on the trend line into consideration. In this figure, we see that the steepest increase is on the interval $0 < p < 0.065$. When $p > 0.065$ the gradient decreases and the ratio of edge failures R becomes more stable. The behaviour for $p > 0.065$ can be explained as follows. Firstly, the number of initial edges increases as p increases. Secondly, as we have more edges we have a higher average degree and more edges will fail. However, as both failures and initial edges increase at approximately the same rate the ratio of edge failures R is not affected. Lastly, in Figure 4.4 as $p > 0.12$ more cases get close to losing almost all their edges resulting in a wider upper confidence interval.
With the established increase in edge failures, it becomes relevant to explore the node connectivity and the resulting disconnections caused by the cascade. Intuitively, one might suspect that an increased ratio of edge failures R would correlate with increased disconnections within the graph. However, it is also crucial to recognize that a higher R coincides with a larger probability p, indicative of a more densely connected graph. Consequently, the likelihood of an edge failure leading to a node disconnection may actually diminish with higher values of p. In the next section, we therefore explore the intricate relationship between cascade propagation and graph disconnections further.

4.2 Size of individual disconnected component

In this section, we explore the individual disconnections that happen during the cascade on $\mathcal{G}$. This, along with our insights for the ratio edge failures, will help to investigate the number of nodes and components that are disconnected at the end of the cascade. In the previous section, we already argued that more edge failures might not result in more disconnections due to p being positively correlated with the number of edge failures, resulting in a more densely connected graph in which edge failures might have less of an effect. Considering this, we first investigate if more edge failures are needed to disconnect a graph constructed with higher probability p.
In Figure 4.5, this conjecture appears to be true: the number of edge failures needed to achieve a disconnection for the first 5 disconnections increases as p increases. Therefore, we conclude that as we have a higher p, a higher average degree and thus more cycles, more edge failures are needed for disconnections to occur.
As stated before, the magnitude of these disconnections is important as it can help to find the total amount of nodes that get disconnected during the cascade, together with the results of the edge failures. Another reason to look into the magnitude of these individual disconnections is their importance when redistributing the demand and generation. Given our understanding that a higher number of edge failures is needed to disconnect a highly connected graph (with a higher probability p), there are two likely scenarios. The first is that the disconnections manifest in such a way as to create large, isolated segments of the graph, i.e. splitting the graph into components of equal size. Conversely, the second scenario is that an increase in edge failures may result in the separation of the graph into

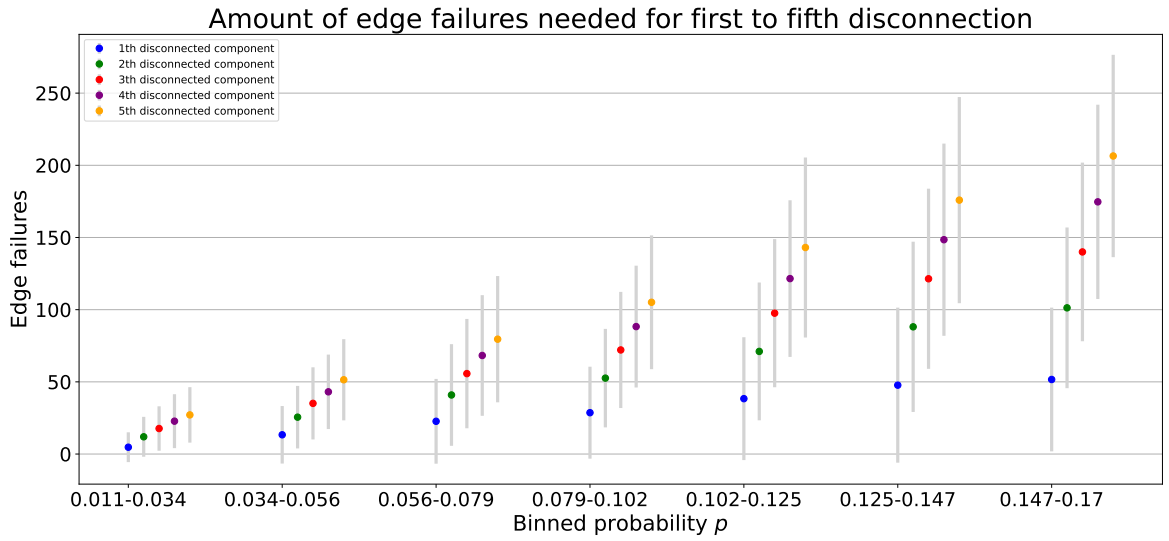


Figure 4.5: Figure on the number of edge failures needed per ith disconnection for a cascade with $\lambda = 0.9$ on an ER model with $n = 100$ for 1600 simulation runs.

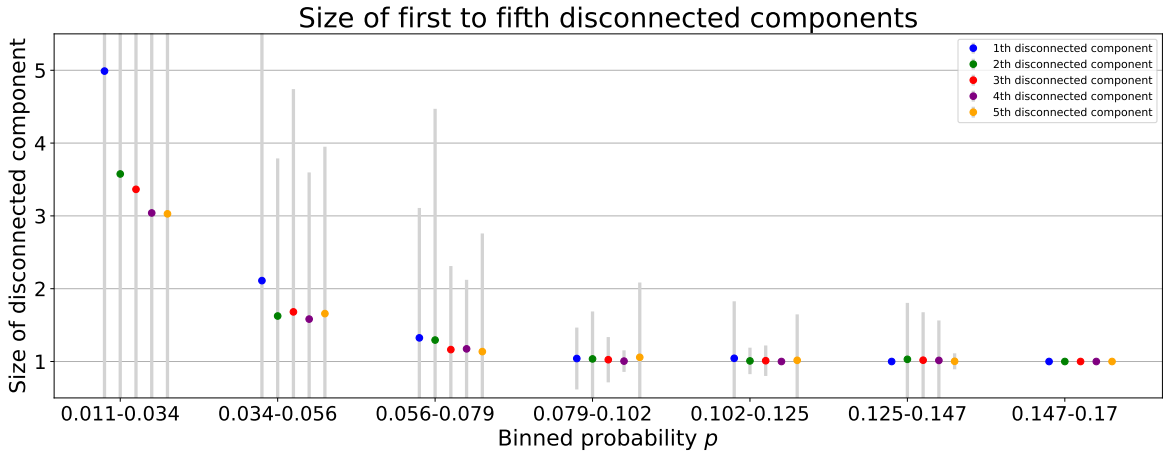
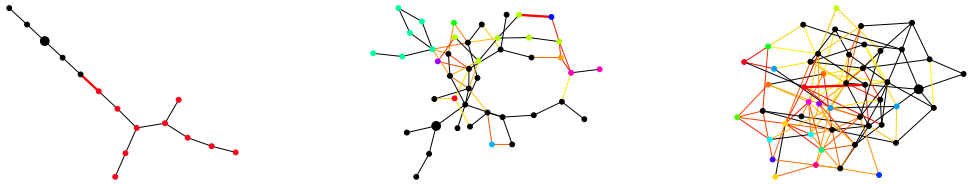


Figure 4.6: Figure on cascade with 1600 simulation runs of a cascade with $n = 100$ and $\lambda = 0.9$.

numerous smaller, disconnected subgraphs, like single nodes or line components.
Figure 4.6 confirms our second hypothesis as the size of the disconnected component decreases as p increases. As we already saw in Figure 4.5, this also results in more edge failures before a first disconnection as failures happen throughout the whole graph and not just at one spot to prevent large disconnections. Therefore, smaller components are more likely to disconnect, as more edge failures decrease the size of a component before it disconnects.



(a) $p = 0.0216$ (b) $p = 0.0412$ (c) $p = 0.104$

Figure 4.7: The figures show a cascade with $\lambda = 0.9$ on the giant component $\mathcal{G}$ for $n = 50$. The colouring from red to yellow shows the first edge failure until the last, respectively. The different coloured nodes show the different components that have disconnected. The black cluster is the component with the enlarged unit demand node.

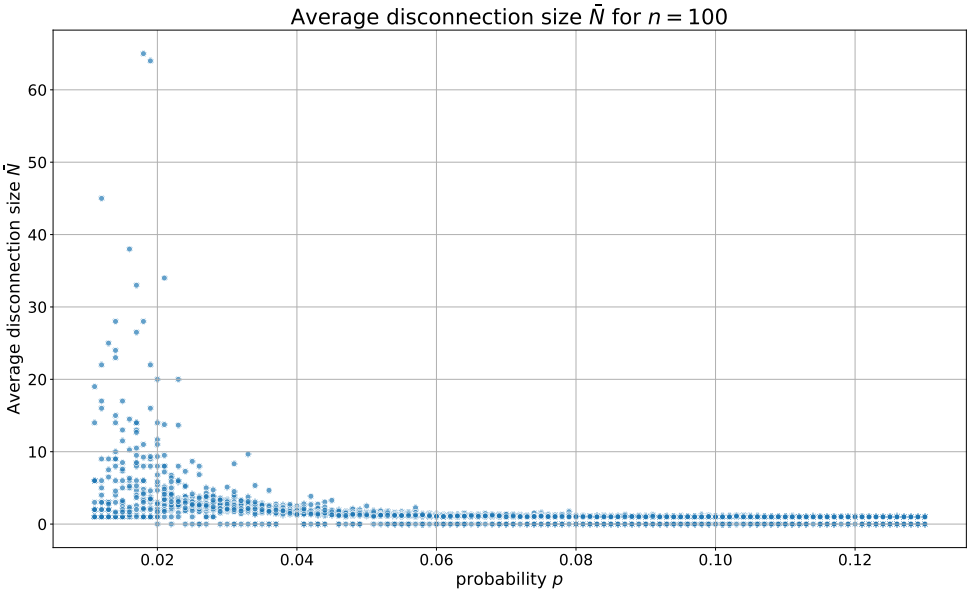


Figure 4.8: Figure on average disconnection size $\bar{N}$ for a cascade with $\lambda = 0.9$ on an ER graph with $n = 100$ for 2400 simulation runs.

Figure 4.7 demonstrates the results found from Figure 4.6 with examples. We see in the graph with the smallest $p = 0.0216$ a tree which results in a big disconnection, illustrated by the red colour. In the next example for a higher $p = 0.0412$ there are still some tree structures on the outside of the big cycle component that are disconnected, leading to smaller disconnection than in the previous one. The biggest disconnection is of size 6 compared to

the disconnection of size 10 in the tree but the components are still of significant size and complexity, i.e. having cycles. The last example is constructed with the highest p value. Here we see that the components that disconnect are either single nodes or line components. For the first 5 disconnections and some examples, the behaviour is clear, a decrease in the size of the components that disconnect for increasing p, followed by convergence towards single nodes and small line components. However, to get a better picture of the full cascade and the disconnection size we take a look at the average disconnection size defined below.

Definition 4.2. *(Average disconnection size) Let N_i be the number of nodes that are disconnected from the demand node during the cascade in $\mathcal{G}$ at i edge failures. The average disconnection size is given by*

$$\bar{N} = \frac{1}{\sum_i \mathbb{1}\{N_i > 0\}} \sum_i N_i.$$

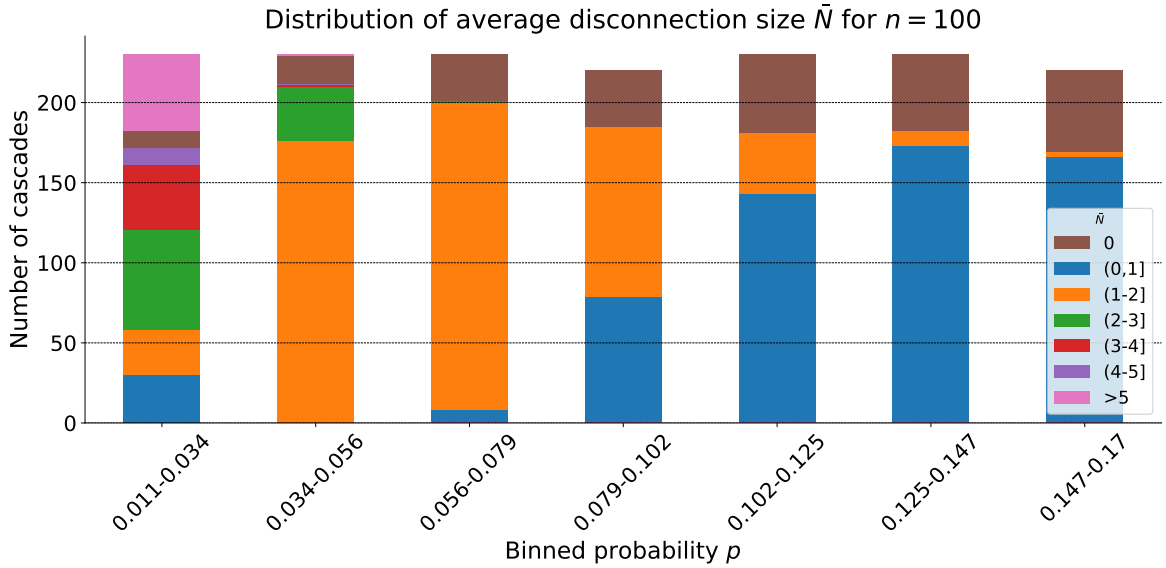


Figure 4.9: The figure shows the number of cascades for which the average disconnection size $\bar{N}$ is within a certain category per bin for edge probability p. The cascade was performed with $\lambda = 0.9$ on an ER graph with $n = 100$ and 1600 simulation runs are shown in total. Note that not all counts are equal per bin, due to the binning.

In Figure 4.8, we see that as p increases the components that disconnect are single nodes or very small components. In addition, note the large variability for the lower p values. This is due to the fact that these values of p are close to the percolation threshold, therefore, there are trees but also cycle graphs represented. The disconnection in trees can be very large when the first disconnection is close to the demand node resulting in a large deviation

here as shown in Figure 4.7a.
Lastly, note that in both Figure 4.8 and Figure 4.6, the interval also mentioned in the study of the ratio of edge failures, $0 < p < 0.065$, is present as we see the most significant decrease in this interval while above this interval the metric stabilizes. Furthermore, in Figure 4.8, we see that the average disconnection gets small, however, the quantity can not be obtained from this figure. Therefore, an additional figure, Figure 4.9 shows the distribution of the average. We see that as p increases the average of the disconnected components in each cascade decreases to the range of 0 to 2. Thus, we can conclude that as p increases the disconnections that occur are mostly nodes or small line components. However, to get an even better idea of the size of the single disconnections, Figure 4.10 shows the distribution of the size for every disconnection that happens in a cascade for a certain value p. We see that as p increases the only disconnections that happen are those of size 1. Therefore, we can conclude that as p gets larger only single nodes disconnect. Now that we have established results for the number of edge failures and the size of individual disconnections, in the next section we take a look at the effect on the total number of nodes that become disconnected after the cascade.

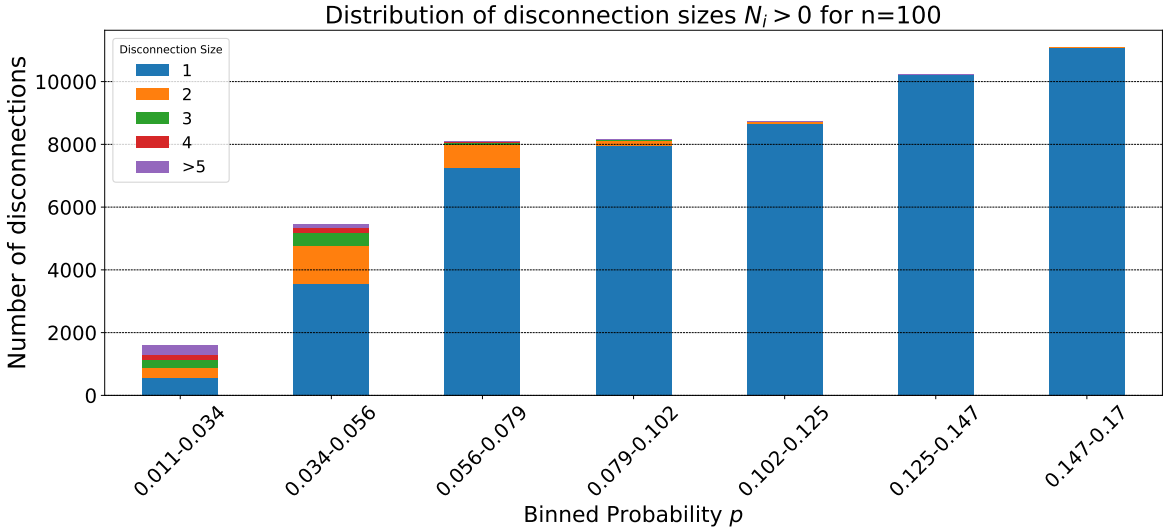


Figure 4.10: The figure shows the number of disconnections for which the size of the disconnection of a certain value, without 0, per bin for edge probability p. The cascade was performed with $\lambda = 0.9$ on an ER graph with $n = 100$ and 1600 simulation runs are shown in total. Note that, on top of the last two bars there are still a small number of nodes in other categories but because of the much larger count of disconnections of size 1.

4.3 Total disconnection

In the previous two sections, both the edge failures and the individual disconnection are studied. In this section, we combine these findings by studying the total disconnection ratio defined as follows.

Definition 4.3. *(Total disconnection ratio) Let N_i be the number of nodes that are disconnected from the demand node during the cascade in $\mathcal{G}$ at i edge failures. The total disconnection ratio is given by*

$$T = \frac{\sum_i N_i}{|\mathcal{N}|}.$$

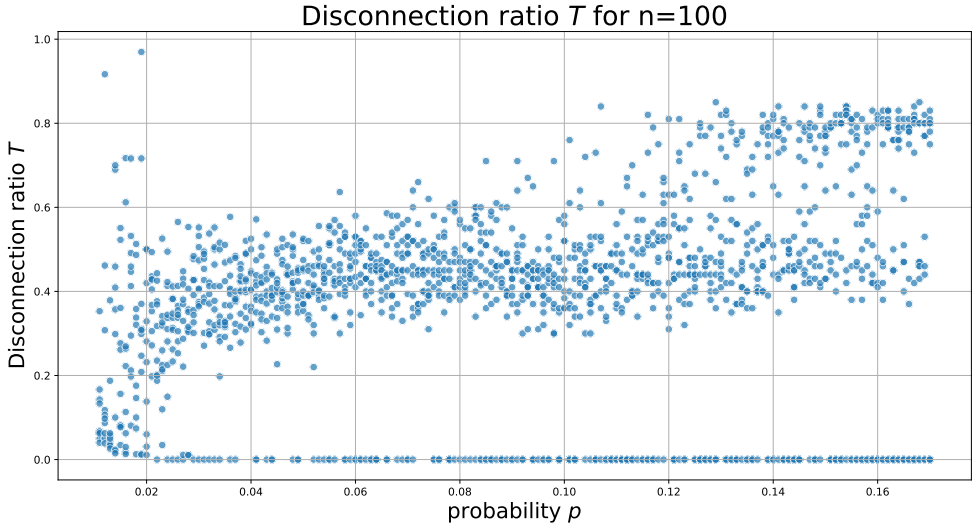


Figure 4.11: Figure disconnection ratio T for a cascade with $\lambda = 0.9$ on the ER model with $n = 100$, for 1600 simulation runs.

We found that for $0 < p < 0.065$, the ratio of edge failure increases, while the sizes of the individual disconnected components decrease. Figure 4.11 shows that for this interval there is an increase in the total ratio of disconnection T. Thus the edge failures have a greater effect and/or steeper increase as p increases compared to the decrease in the size of the disconnection, since there is an increase in the total disconnection of nodes as p increases over this interval.
After this interval we again see the metric becoming more stable. This can be explained by the ratio of edge failures and the individual disconnection sizes staying stable as well. As we have a fixed ratio of edge failures and a fixed size for the components that disconnect we

expect the total disconnection, which is a consequence of these two, to be stable as well.
As we have already seen in Section 4.1 for $p > 0.65$, as p increases more edges fail but we also have more edges to start with due to a higher p. However, as both failures and initial edges increase at approximately the same rate the ratio of edge failures R is not affected. Furthermore, as p is large enough there are only disconnections of size 1, see Definition 4.2, and giant component consist of all, in this case 100, initial nodes. However, as p increases, the connectivity increases, therefore, the increase in edge failures does not result in a higher ratio of disconnection and seems to cluster around $T = 0.45$ and $T = 0.8$ and in Figure 4.11.
The first cluster of $T = 0.45$ could be linked to the nodes in Figure 4.3 that are situated around and slightly above $R = 0.8$. The reasoning for the difference in value for both ratios is explained for the second cluster.
This second cluster at $T = 0.8$ is situated in the top right of Figure 4.11 and can be linked to the cluster that we see in Figure 4.3, which is positioned close to a ratio of edge failures $R = 1$. The link between $R = 1$ and $T = 0.8$ can be derived from the influence of p on the giant component. As p increases the giant component will start to consist of all 100 nodes but the amount of edges increases along with p and could encounter the following situation. Suppose that there is a giant component with 100 nodes and 1000 edges before the cascade and at the end of the cascade, the component that contains the demand node has 20 nodes and edges. In this case, the total disconnection ratio $T = 0.8$ and the ratio of edge failures $R = 0.98$, match the results that we observe in the figures.
Therefore, we see that as $p > 0.65$ the metrics are not influenced anymore as the structure of the graph is not altered in terms of cycles and trees. The number of edges does increase, resulting in more edge failures which is balanced out by the increase in initial edges. Furthermore, the increase in connectivity, due to the increase in edges, balances out the disconnection as more edge failures do not lead to greater disconnections.
To conclude this chapter, we sum up five findings.

1. As p increases on the interval $0 < p < 0.065$ the ratio of edge failures R increases substantially because with increasing p the number of edges that are part of cycles grows resulting in more edge failures. When $p > 0.065$ the gradient of the increase slows down as the rate at which new edges become part of the initial giant component, and the rate at which the number of edges become part of cycles is approximately the same.

2. As p increases more edge failures are needed to result in disconnections.

3. As p increases, the average disconnection size $\bar{N}$ decreases up to a range of zero to one in almost all cases when p gets large.

4. As p gets large, the components that disconnect are either single nodes or small line components.

5. The total disconnection ratio T increases on the same interval as R, after which the gradient of increase slows down and two clusters arise at approximately $T = 0.8$ and $T = 0.45$ which can be linked to the two clusters $R = 1.0$ and $R = 0.8$, respectively. The cluster mentioned can be found in Figure 4.3 and Figure 4.11.

5 Lattice

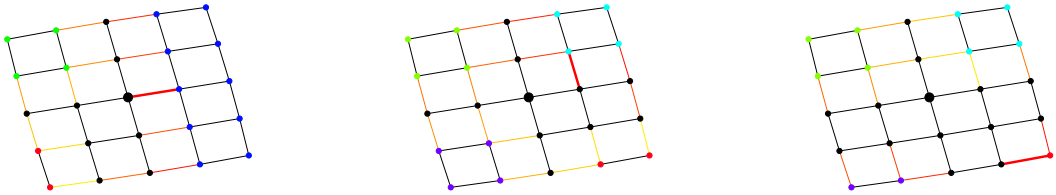
In the preceding chapter, we explored how cascades could be forecasted for certain graph structures like a cycle and a tree and their combinations. When doing cascade analysis on the ER model we were limited to numerical results. Therefore, this chapter shifts focus to a more structured graph type: the lattice. Due to the structure of a lattice, analytical results can become obtainable. To find the analytical results, we start by getting an initial understanding of the behaviour of the cascade on the lattice by looking at examples. Afterwards, numerical results will build upon this knowledge and will be validated with additional examples. Lastly, a first analytical approach is explored.

5.1 Examples of the cascade

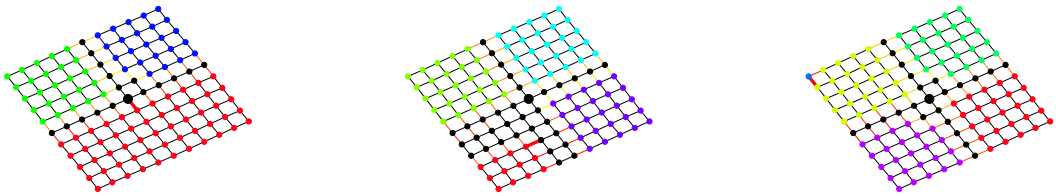
In this section, we consider a square lattice $\mathcal{L}$ as described in Section 2.2.2 and we assume the centre node is the demand node. A reason for placing the unit demand node centrally in the lattice is that as the lattice size grows to infinity, every node has an infinity distance to the border, thus every node could be the centre node. For $\mathcal{L}$ to have a central node, it must be structured as an $n \times n$ lattice where n is an odd number, consisting of n^2 nodes in total.
Due to the particular structure of the lattice, the distance from the demand node to the first edge failure is crucial to the behaviour of the cascade as we see in this section. Therefore we define the following notion of the distance.

Definition 5.1. *(Distance d from demand node to first edge failure.) d is the distance on the shortest path from the first edge failure to the demand node.*

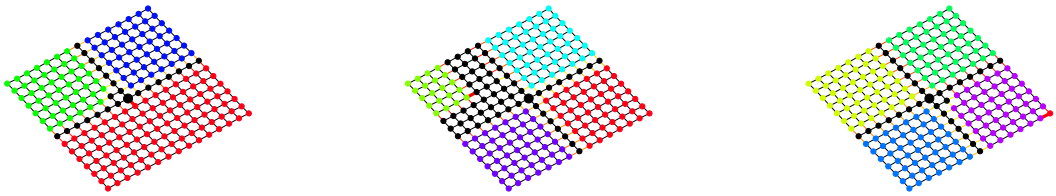
Next, we show the influence of the distance on the cascade progression in lattices using examples in Figure 5.1. This figure shows 9 examples of the cascade on $n \times n$ lattices. Through these examples, we can distinguish between various types of edges, noting differences between those located on the border, in the middle, those that consistently remain connected to the demand node, and those that frequently disconnect.
In the first column, we see the cases in which the first edge failure is at a distance $d = 1$ from the unit demand node. In all three examples for $n = 5$, $n = 11$, and $n = 15$, a T-shape is visible as the edges fail along the same line as the initial edge failure and perpendicular to the line created between the first edge failure and the demand node. In the second column,



(a) n=5, d=1 (b) n=5, d=2 (c) n=5, d=4



(d) n=11, d=1 (e) n=11, d=5 (f) n=11, d=10

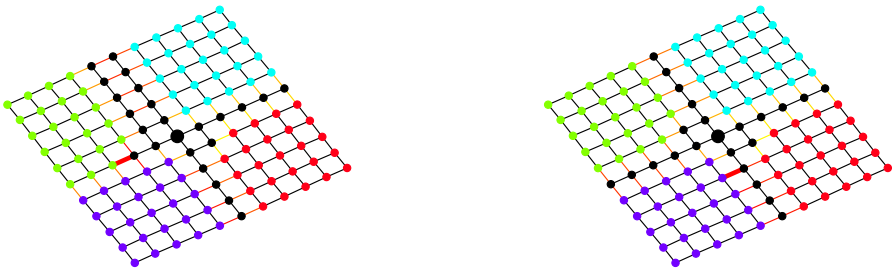


(g) n=15, d=1 (h) n=15, d=7 (i) n=15, d=14

Figure 5.1: This figure shows 9 example cascades on lattices of size $n \times n$ such that the first edge failure is at distance d. The different coloured nodes represent different disconnected components and the coloured lines represent edge failures, from the first one to the last, coloured in a red-to-yellow gradient, respectively. The first edge failure is depicted as a wider red coloured line. The black cluster is the component with the enlarged unit demand node.

we have three examples where the edge failure is somewhere in the middle of one of the quarters of the lattice. In this case, we see three separate quarters and one quarter in which the initial edge failure happened. In this quarter a smaller piece is disconnected from the unit demand node. In the last column, the first edge failure is placed at the corners of the lattice. The cascade then propagates along a cross creating a cross-shape.
What we have seen so far is that the edges on the cross never fail or disconnect in the

cascade unless the first edge failure is on the cross. The latter happens for $d = 1$, which results in all the other edges in that part of the cross failing as well. The question is what happens if an edge farther away from the demand node is the first edge failure, which we present in we present in Figure 5.2a. In this figure, we observe the same behaviour as to



(a) First edge failure on cross (b) First edge failure perpendicular to cross

Figure 5.2: This figure shows 2 example cascades on lattices of size 11 by 11 such that the first edge failure is at distance $d = 3$.

the ones for $d = 1$ in Figure 5.1: the edges between the demand node and the first edge failure on the cross do not fail but the others do fail during the cascade.
In the second and third columns of Figure 5.1 and Figure 5.2b we have examples of the first edge failure happening within a quarter and not on the cross. There is always some disconnection within the quarter. There exist two different cases: either there is a block disconnected but the quarter itself is still connected to the cross or both a block and the quarter are disconnected. When the first edge failure happens more in the centre of the quarter, only the block disconnects and as we get closer to the corner or to the cross the quarter also disconnects.
Lastly, Figure 5.2b shows, that when the first edge failure is one of the edges connected and perpendicular to a cross edge, the cascade results in a cross shape as well.
In the next part, we discuss the insights that can be used for understanding the numerical results given in Section 5.2.

5.1.1 Insights

We can use the insights gained from the examples so far to approximate the following metrics which are used in the numerical analysis in the next section: the ratio of edge failures, the maximum and minimum ratio of the components that disconnect from the demand node and the total disconnection ratio. For the ratio of edge failures R we use Definition 4.1

in which $|\mathcal{E}| = 2(n^2 - n)$ and for the total disconnection ratio T we use Definition 4.3 in which $|\mathcal{N}| = n^2$. We define minimum and maximum disconnection ratio next.

Definition 5.2. *(Minimum and maximum disconnection ratio) Let N_i be the number of nodes that are disconnected from the demand node during the cascade in $\mathcal{G}$ at i edge failures, excluding the instances when no disconnection occurs, i.e. $N_i > 0$ for all i . The maximum M and minimum m disconnection are*

$$M = \max_i \frac{N_i}{n^2} \text{ and } m = \min_i \frac{N_i}{n^2}.$$

Firstly, in each example, the cascade ends when the lattice has a T-shape or a cross. Therefore, the ratio of edge failure R ranges approximately between

$$\left(\frac{1}{2(n^2-n)}\right)\left(\frac{6(n-1)}{2}\right)$$

and

$$\left(\frac{1}{2(n^2-n)}\right)\left(\frac{8(n-1)}{2}\right).$$

These equations arise from the fact that $\frac{2(n-1)}{2}$ disconnections are needed to get one out of four parts of the cross.
Additionally, smaller values for d result in a minimum disconnection of one quarter, which equals to $\left(\frac{(n-1)}{2}\right)^2$ number of nodes, but as $d = (n-1)$, only one node disconnects. Therefore m ranges approximately between

$$\frac{1}{n^2}$$

and

$$\left(\frac{1}{n^2}\right)\left(\left(\frac{(n-1)}{2}\right)^2\right).$$

Furthermore, the maximum disconnection is at least one quarter, which happens in most cases. However, the biggest disconnection happens when $d = 1$ and half the lattice disconnects. Consequently, M has approximately values between

$$\left(\frac{1}{n^2}\right)\left(\left(\frac{(n-1)}{2}\right)^2\right)$$

and

$$\left(\frac{1}{n^2}\right)\left(\frac{n(n-1)}{2}\right).$$

Additionally, in terms of total disconnection, at least 3 quarters disconnect and the biggest total disconnection again occurs for $d = 1$. For $d = 1$ half of the lattice, one quarter and a quarter missing one node disconnect. Thus T ranges approximately between

$$\left(\frac{1}{n^2}\right)\left(3\left(\frac{(n-1)}{2}\right)^2\right).$$

and

$$\left(\frac{1}{n^2}\right)\left(2\left(\frac{(n-1)}{2}\right)^2 + \frac{n(n-1)}{2} - 1\right).$$

In the next section, we look at numerical results for R, m, M and T for a 11×11 square lattice. Therefore, in Table 1 the values for the above-described approximate bounds are given for a 11×11 square lattice.

metric	lower bound	upper bound
R	0.136	0.181
m	0.008	0.207
M	0.207	0.455
T	0.620	0.867

Table 1: Table for the initial approximations of multiple metrics for the cascade on a 11×11 square lattice.

Note that we have not seen all cases and examples, therefore we can expect the actual values to be different from those shown in Table 1. In the next section, we explore the numerical results and check the correctness of the approximations given in Table 1.

5.2 Numerical results

Now that we have an initial idea of the behaviour of the cascade in this lattice, in this section, we look at the following metrics: the ratio of edge failures, the maximum and minimum ratio of the components that disconnect from the demand node and the total disconnection ratio. We discuss these metrics based on distance d from the first edge failure to the demand node for $n = 11$. The reason for differentiating for d is, that as stated in Section 5.1 there is a distinction to be made for the edges that are on the 'cross', the middle edges and the borders and the distance d is a way to identify the different kinds of edges. Therefore, in our results, we expect, that for $d = 1$ and $d = 10$ there will be no confidence intervals, as there is only one kind of edge that can be at that distance, a cross or corner edge respectively. Thus these cascades will always propagate in the same manner. However,

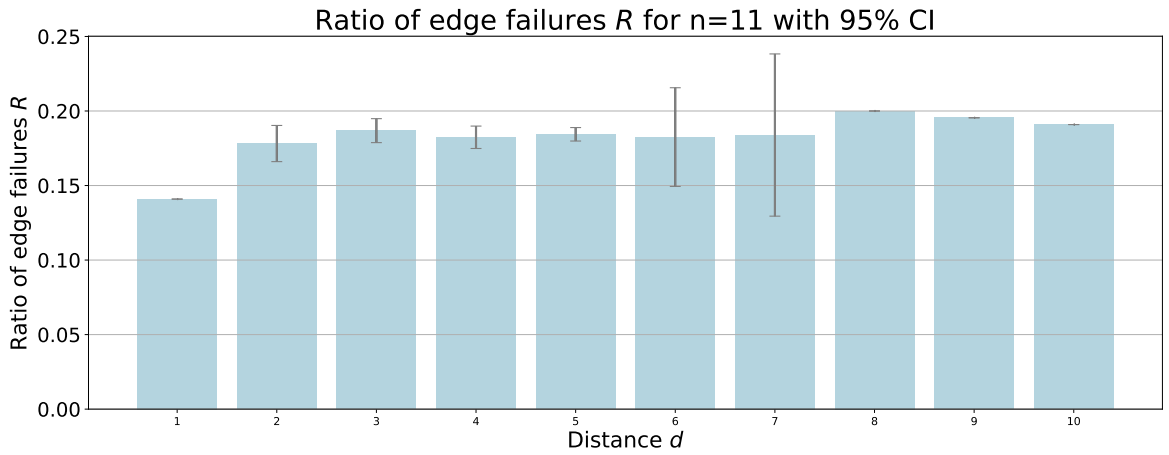
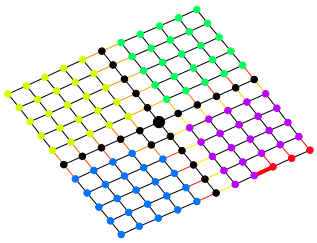


Figure 5.3: Figure on the ratio of edge failures R of a cascade with $\lambda = 0.9$ on a 11×11 square lattice with 2000 runs.

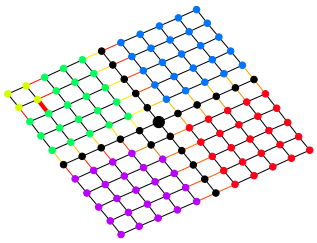
for other distances, such a conclusive statement can not be made yet. With this in mind, we discuss the first metric, the ratio of edge failures.

In Figure 5.3 we see that indeed as hinted in the previous part in terms of edge failures, the cascade has a fixed behaviour when $d = 1$ and $d = 10$ as there is no confidence interval. We might expect that for all the other cases for d, there is a confidence interval because at these distances cross edges, middle edges and border or corner edges can be the first edge failure. However, this is not the case for $d = 8$ and $d = 7$. Besides $d = 1$ and $d = 10$ $d = 8$ and $d = 9$ do not have a confidence interval.

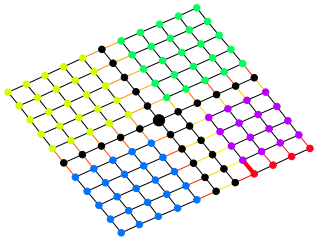
Since the transition seems to happen between $d = 7$ and $d = 8$ we take a look at Figure 5.4 in which examples cascade are displayed on lattices with initial edge failures at distance $d = 7$ and $d = 8$. In the first column, there is very similar behaviour for both $d = 7$ and $d = 8$ as a border edge is the first edge failure, resulting in the full quarter being disconnected and a small block from the quarter getting cut off. However, in the second column, we do see a difference. In both cases, $d = 7$ and $d = 8$, a block within the quarter is disconnected, however, for $d = 8$ also the full quarter is disconnected but for $d = 7$ the quarter stays connected.
This difference occurs because the first edge failure at $d = 8$ results in a disconnection of the 4 nodes in the corner. This symmetric first disconnection close to the corner results in the quarter being disconnected as well, compared to $d = 7$ where the first disconnected block is closer to the cross shape which consists of the stronger edges in the lattice structure. Therefore, when disconnection happens far enough from the strong cross edges, the cross shape is formed as we see in Figure 5.4b. A similar pattern can be seen in the last column of Figure 5.1 and in Figure 5.5 as the first disconnection is in the corner and results in a small disconnection away from the centre cross.



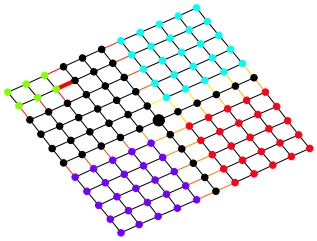
(a) $n = 11$, $d = 8$



(b) $n = 11$, $d = 8$

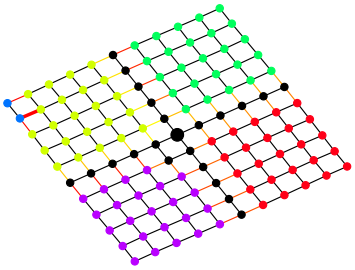


(c) $n = 11$, $d = 7$

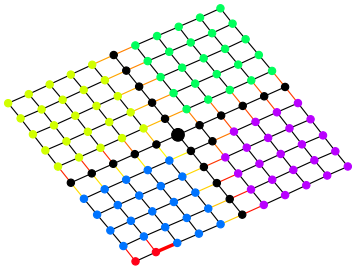


(d) $n = 11$, $d = 7$

Figure 5.4: This figure shows 4 example cascades on lattices of size 11×11 such that the first edge failure is at distance $d = 7$ and $d = 8$.



(a) First edge failure perpendicular to border



(b) First edge failure on border

Figure 5.5: This figure shows 4 example cascades on lattices of size 11 by 11 such that the first edge failure is at distance $d = 9$.

Furthermore, note that as demonstrated in Figure 5.4b for $d = 8$, the example not only illustrates the presence of an approximate cross shape but also highlights initial edge failures within a quarter of the lattice. This observation indicates that the upper bound presented in Table 1 may be underestimated. However, it is important to emphasize that this does not undermine the accuracy of the approximations for the bounds on m, M, and T. Figure 5.4b clearly shows that, despite the edge failures, the number of disconnected nodes in the left quarter remains consistent with these approximations.

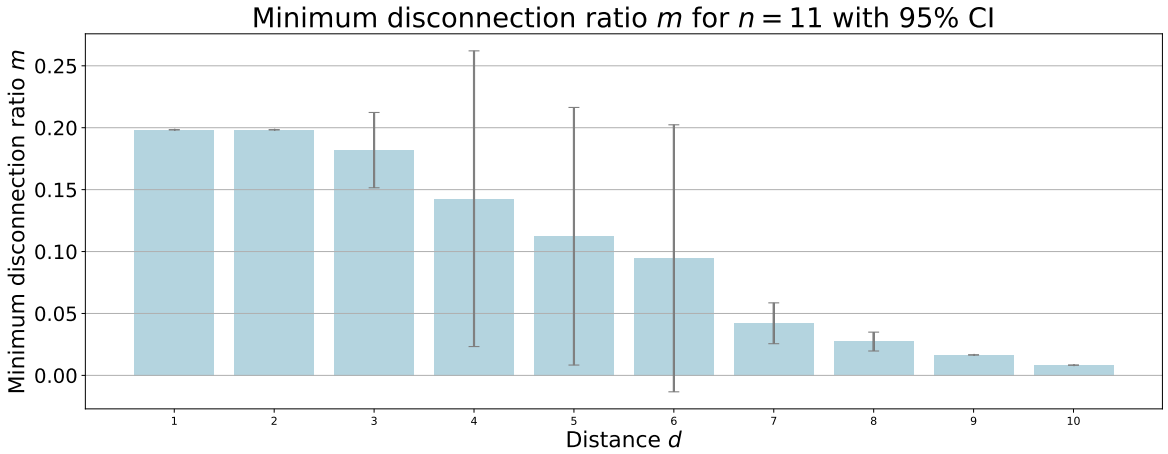
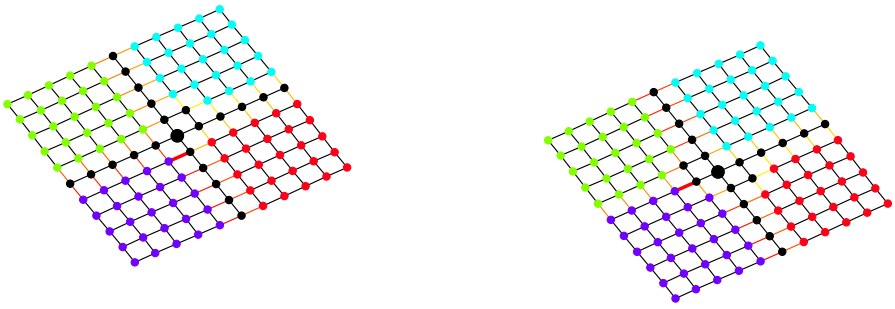


Figure 5.6: Figure on the minimum disconnection ratio m for a cascade with $\lambda = 0.9$ on a 11×11 square lattice with 2000 runs.

We have now seen that the cascade with the first edge failure at $d = 8$ does not have a fixed behaviour. But the two examples displayed in Figure 5.4 coincidentally result in the same amount of edge failures, resulting in the absence of variability in Figure 5.3. However, the disconnection sizes will be different so for these metrics there will be variability.
Next to that, for the first edge failure at distance $d = 9$, Figure 5.5 shows that the two possible disconnections, on the border or perpendicular to the border result in the same cascade. This explains the absence of variance in Figure 5.3. Furthermore, we can also see that in this case, the cascades are the same so we expect the other metrics to also be conclusive without variability for $d = 9$. To verify we now take a look at the minimum and maximum disconnection ratio.
In Figure 5.6, we see everything we might expect up to now. Variability for $d = [3, 8]$, for $d = 9$ we have $m = \frac{2}{11^2}$ and $d = 10$ gives $m = \frac{1}{11^2}$, see Figure 5.5 and Figure 5.1 respectively. Furthermore, for $d = 1$ we have $m = \frac{(n-1)^2-1}{n^2} = \frac{24}{121}$ calculated by looking at Figure 5.1.
Also the bounds in Table 1 are partially verified as for $d = 1$ we have the upper bound $\frac{(n-1)^2-1}{n^2} = \frac{24}{121} = 0.198$ which is a bit off from the approximated upper bound $\frac{(n-1)^2}{n^2} = \frac{24}{121} = 0.207$. Next to that, the lower bound in Table 1 is approximated correctly as this is

shown in Figure 5.6 and the example for $d = 10$ in Figure 5.1 to be $\frac{1}{11^2} = 0.008$.
However, for $d = 2$ we need to look at some more examples to understand these results as at distance $d = 2$ the first edge failure can either be on the cross or an edge perpendicular to the cross. These two options would expectedly lead to different cascades as we already saw in the ratio of edge failures Figure 5.3.



(a) First edge failure perpendicular to the cross

(b) First edge failure on the cross

Figure 5.7: This figure shows 2 example cascades on lattices of size 11×11 such that the first edge failure is at distance $d = 2$.

Figure 5.7 shows the two possible cascades when $d = 2$. We indeed see that the number of edge failures is different and also the way the cascade propagates differs. However, the minimal disconnection in both cases is the same as for $d = 1$ which is $m = \frac{(n-1)^2-1}{n^2} = \frac{24}{121}$.
The difference that does occur, and that is visible in Figure 5.7, is in terms of the maximum disconnection. Therefore, we see variability for $d = 2$ in Figure 5.8 which shows the maximum disconnection ratio.
Furthermore, in Figure 5.8 we see for $d = 1$ that the maximum disconnection ratio is $M = \frac{\frac{n-1}{2}n}{n^2} = \frac{55}{121}$. Lastly, for $d = [3, 10]$ the maximum disconnection is one-quarter of the lattice which equals $M = \frac{(\frac{n-1}{2})^2}{n^2} = \frac{25}{121}$. Note that, $\frac{55}{121} = 0.455$ and $\frac{25}{121} = 0.207$ are exactly the bounds given in Table 1.
Figure 5.9 shows the last metric: the total disconnection ratio. Now that we have analyzed the other metrics, this does not result in any surprises. We see variability for the distances that have multiple cascade outcomes. For the transition distances $d = 2$ and $d = 8$ we see that $d = 2$ does have variability as we can obtain from Figure 5.7. However, $d = 8$ does not have variability as we can see in Figure 5.4, although the structures in which these nodes disconnect are different therefore we did see the variability in Figure 5.6 for $d = 8$. Lastly, as we can see in Figure 5.9 the bounds for T given in Table 1 seem to be good approximations.
All these numerical results are for one specific case of n. However, as we also saw in

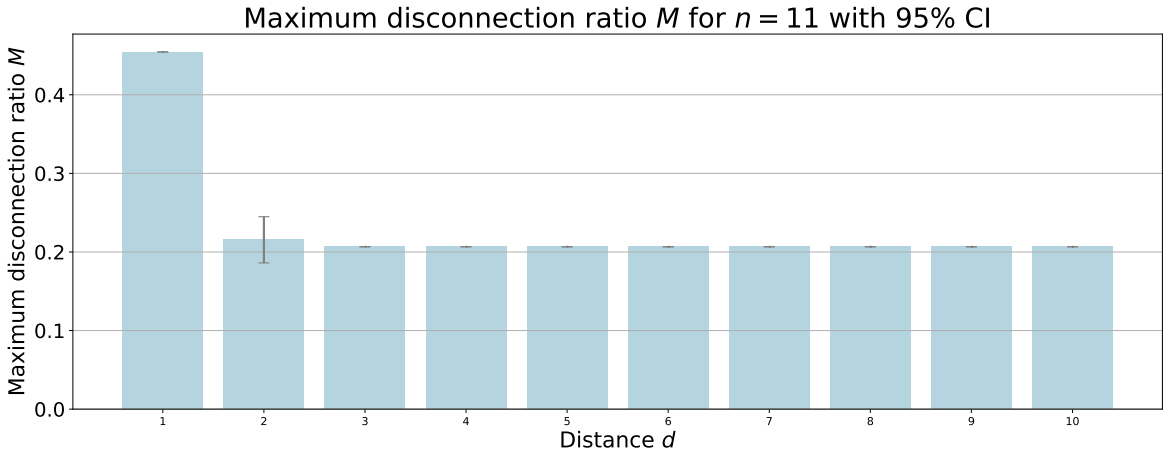


Figure 5.8: Figure on maximum disconnection ratio M of a cascade with $\lambda = 0.9$ on a 11×11 square lattice with 2000 runs.

Figure 5.1, as n increases the behaviour of the cascade does not change significantly and is comparable to that of the studied $n = 11$. Therefore, results from this section can be extended to different values for n.
Now that we have gathered numerical insight, the highly structured lattice invites for an initial analytical approach which is the topic of the next section.

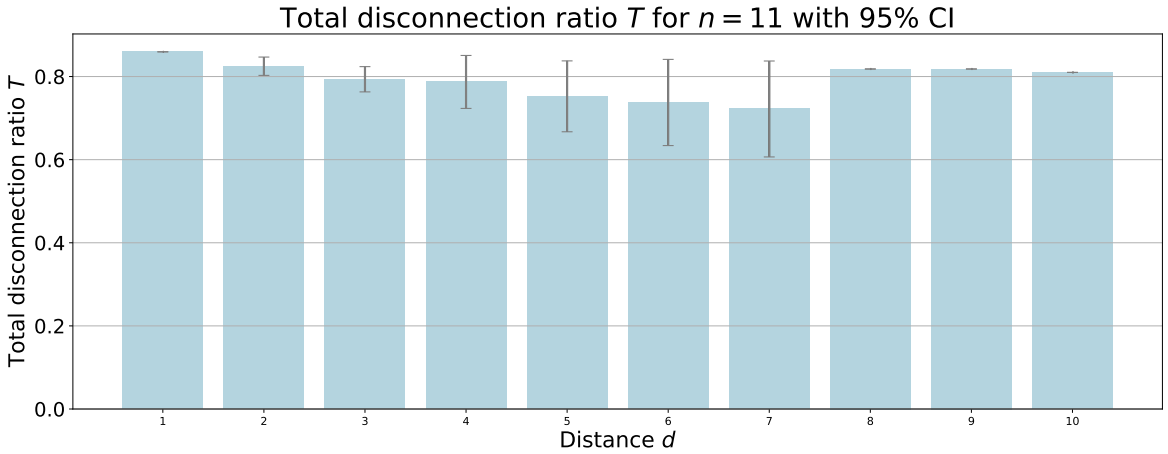


Figure 5.9: Figure on the total disconnection ratio T of a cascade with $\lambda = 0.9$ on a 11×11 square lattice with 2000 runs.

5.3 Emergency edge limits approximation

In this section, we start by analyzing the cascade in an analytical way. Due to time constraints, the analytical work in this thesis is limited to determining the emergency edge limits, which is the first step in understanding the cascade. Due to the lattice being highly structured, we find the emergency edge limits by finding the incidence matrix C, the Laplacian matrix L and finally the PTDF matrix V as described in Section 2.1.
For this purpose we number the lattice starting from the top left node, going left to right and top to bottom. This labelling convention is visualized in the 3 by 3 lattice in Figure 5.10. Furthermore, for any matrix that is used in the rest of the section, we use notation as defined below.

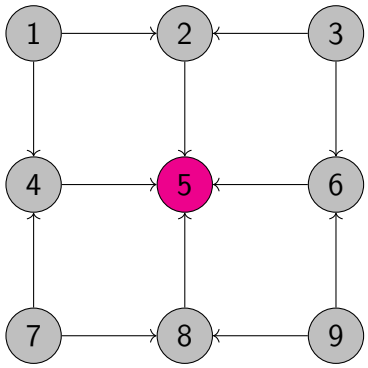


Figure 5.10: Labelling convention of nodes in the lattice with the pink unit demand node in the centre.

Notation 2. *(Matrix for a $n \times n$ lattice) Let A_n be a matrix that describes a matrix A that corresponds to the $n \times n$ lattice.*

In the rest of this section, we denote a particular matrix A that depends on the size of a square $n \times n$ lattice by A_n. Our next steps are finding the structure of the incidence matrix C_n and Laplacian matrix L_n. With the structural findings of the matrix L_n an approximation approach is used to have a closed-form solution to the emergency edge limits.

5.3.1 Indidence matrix C_n

To find the incidence matrix C_n we start by looking at the horizontal edges. Note that these edges are always of the form $(l, l+1)$ or $(l, l-1)$. We can represent these edges on one horizontal line, with the following matrix A_{C_n}, which is a $n \times n-1$ matrix. To describe all the horizontal edges we need n of these blocks A_{C_n} on the diagonal of the upper block of the matrix C_n. The matrix $A_{C_n} \in \mathbb{R}^{(n-1)\times n}$, where the rows correspond to edges and the

columns to nodes, has the form

$$A_{C_n} = \begin{pmatrix} 1 & -1 & & & & & & & \\ & 1 & -1 & & & & & & \\ & & \ddots & \ddots & & & \mathbf{0} & & \\ & & & 1 & -1 & & & & \\ & & & & -1 & 1 & & & \\ & & \mathbf{0} & & & -1 & 1 & & \\ & & & & & & \ddots & \ddots & \\ & & & & & & & -1 & 1 \end{pmatrix}.$$

Now that the horizontal edges are constructed, note that for the vertical edges, there is always a connection between the node l and the node $l+n$ or $l-n$ and l. Therefore we can construct the matrix $B_{C_n} \in \mathbb{R}^{n(n-1)\times n^2}$, where the rows correspond to edges and the columns to nodes, as

$$B_{C_n} = \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 & & & & & \\ & 1 & 0 & \cdots & 0 & -1 & & & \mathbf{0} & \\ & & \ddots & 0 & \cdots & 0 & \ddots & & & \\ & & & 1 & 0 & \cdots & 0 & -1 & & \\ & & & & -1 & 0 & \cdots & 0 & 1 & \\ & & \mathbf{0} & & & \ddots & 0 & \cdots & 0 & \ddots \\ & & & & & & -1 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

such that we have $n-2$ zeros between two non-zero entries on a row and $\frac{n(n-1)}{n}$ rows for which the first non-zero entry is 1 following $\frac{n(n-1)}{n}$ rows for which the first entry is -1. Combining the matrices the incidence matrix for the horizontal lines A_{C_n} and the incidence matrices for the vertical lines B_{C_n} we get the full incidence matrix

$$C_n = \begin{pmatrix} A_{C_n} & & 0 \\ & \ddots & \\ 0 & & A_{C_n} \\ & B_{C_n} & \end{pmatrix}$$

where there are n blocks of A_{C_n}.
Now that we have described the structure of the matrix C_n, the pattern of the blocks in the Laplacian $L_n = C_n^\top C_n$ can be expressed.

5.3.2 Matrix L_n

Now that the construction of the matrix C_n is clear, the matrix $L_n = C_n^\top C_n$ can be given. The matrix L_n is a block matrix, therefore, we define the blocks, after which we explain the

implications.
Firstly, $A_{L_n} \in \mathbb{R}^{n\times n}$ is given

$$A_{L_n} = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 3 & -1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -1 & 3 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}.$$

Furthermore, there is the similar matrix $B_{L_n} \in \mathbb{R}^{n\times n}$,

$$B_{L_n} = \begin{pmatrix} 3 & -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 4 & -1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 3 \end{pmatrix}.$$

Using these blocks and the $n \times n$ negative identity matrix $-I^{n\times n}$, the matrix L_n can be described as

$$L_n = \begin{pmatrix} A_{L_n} & -I^{n\times n} & 0 & 0 & \cdots & 0 & 0 & 0 \\ -I^{n\times n} & B_{L_n} & -I^{n\times n} & 0 & \cdots & 0 & 0 & 0 \\ 0 & -I^{n\times n} & B_{L_n} & -I^{n\times n} & \cdots & 0 & 0 & 0 \\ 0 & 0 & -I^{n\times n} & B_{L_n} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & B_{L_n} & -I^{n\times n} & 0 \\ 0 & 0 & 0 & 0 & \cdots & -I^{n\times n} & B_{L_n} & -I^{n\times n} \\ 0 & 0 & 0 & 0 & \cdots & 0 & -I^{n\times n} & A_{L_n} \end{pmatrix}$$

which has $n-2$ matrices B_L on its diagonal.
Note that the number on the diagonal represents the degree of the node. Therefore, we have four entries of 2 on the diagonal which represent the corners of the lattice. Next to that, there are $2(n-2)+2(n-2) = 4(n-2)$ entries of 3 which represent the borders of the lattice that are not a corner. Lastly, all the other entries on the diagonal are 4's which are all the other nodes.
The matrix L_n displays a notable structure, characterized primarily by the presence of 4's

along its diagonal as n increases. However, this pattern is occasionally interrupted by other numbers, detracting from the uniformity of the diagonal. Due to these disruptions, the matrix L_n is also singular resulting in having to find the more complicated pseudoinverse instead of the inverse. To the best of our knowledge there is no closed-form solution for the pseudoinverse of the matrix L_n. Therefore, in the following section, we approximate the matrix L_n by a matrix that does have a closed-form solution for its inverse. We will explore a matrix similar to L_n but with an exclusively uniform diagonal, consisting solely of 4's. It is important to note that this approach applies only to the inner part of the lattice, excluding the border and corners, to maintain this uniformity. Lastly, note that this matrix will not be singular anymore making it easier to find a closed-form solution for its inverse.

5.3.3 Block Toeplitz matrix

Determining the emergency edge limits poses a challenge due to the non-uniformity present on the diagonal of L_n. However, as suggested in the previous section, we can approximate the matrix L_{n+2} for nodes with a degree of 4 using a specific matrix,

$$T_n = \begin{pmatrix} A_{T_n} & -I^{n\times n} & 0 & \dots & 0 & 0 \\ -I^{n\times n} & A_{T_n} & -I^{n\times n} & \dots & 0 & 0 \\ 0 & -I^{n\times n} & A_{T_n} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & A_{T_n} & -I^{n\times n} \\ 0 & 0 & 0 & \dots & -I^{n\times n} & A_{T_n}, \end{pmatrix}$$

where

$$A_{T_n} = \begin{pmatrix} 4 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 4 & -1 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 4 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 4 & -1 & 0 \\ 0 & 0 & 0 & \cdots & -1 & 4 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 4 \end{pmatrix}.$$

For this Toeplitz block matrix, we can derive a closed-form expression for the inverse by determining the eigenvalue decomposition [16].

$$T_n = S\mathsf{diag}(\sigma)S^{-1}$$

where

$$\sigma = \left\{ 4 - 2\cos\left(\frac{p\pi}{n+1}\right) - 2\cos\left(\frac{q\pi}{n+1}\right) \right\}_{p,q=1}^{n},$$

$$S = \begin{pmatrix} s_{1,1} & s_{1,2} & \dots & s_{1,n} & s_{2,1} & \dots & s_{n,n}, \end{pmatrix}$$

and

$$s_{p,q} = \left(\left(\sin\left(\tfrac{ip\pi}{n+1}\right) \sin\left(\tfrac{jq\pi}{n+1}\right) \right)_{i,j=1}^{n} \right)^{\top}.$$

Note that by making all vectors $s_{p,q}$ unit vectors $\tilde{s}_{p,q} = \frac{s_{p,q}}{\|s_{p,q}\|_2}$ we can define

$$\tilde{S} = \begin{pmatrix} \tilde{s}_{1,1} & \tilde{s}_{1,2} & \dots & \tilde{s}_{1,n} & \tilde{s}_{2,1} & \dots & \tilde{s}_{n,n} \end{pmatrix}$$

and

$$T_n = \tilde{S}\mathsf{diag}(\sigma)\tilde{S}^{\top}.$$

As the matrix T_n is non-singular we can find the inverse

$$(T_n)^{-1} = \tilde{S}\mathsf{diag}\left(\frac{1}{\sigma}\right)\tilde{S}^{\top}.$$

Before we can use this as an approximation we need some more notation.

Notation 3. *(Emergency edge limit for T_n) $F^0_{T_n}$ are the emergency edge limits computed from the matrix T_n. Thus, $F^0_{T_n} = C_n(T_n^{-1})$*

Notation 4. *(Emergency edge limit without the outer band for L_n) $\mathcal{F}^0_{L_n}$ is the subset of edges of $C_n(L_n^+)$ of the edges that are part of the inner $n-2 \times n-2$ lattice.*

As mentioned before, the hypothesis is that $F^0_{T_n}$ is a good approximator for $\mathcal{F}^0_{L_{n+2}}$. To test this we have the definition for the error term as follows

Definition 5.3. *(Approximation error ϵ) Let F and $\mathcal{F}$ be two vector of size n, then the approximation error is given by $\epsilon = \frac{\|F-\mathcal{F}\|_2}{n}$.*

Note that according to these definitions the vectors $F^0_{T_n}$ and $\mathcal{F}^0_{L_{n+2}}$ are of the same length. In Figure 5.11 we therefore plot the error between these two. We see that as n increases the error ϵ decreases, implying that $F^0_{T_n}$ could be a good approximation for $\mathcal{F}^0_{L_{n+2}}$.
Therefore, this approximation could be used to simulate the cascade on larger lattices as less computational time is needed to compute the pseudoinverse of L since we can use the closed-form solution for the approximation inverse. However, verifying this assumption is potentially future work as we discuss in the final next chapter. In this final chapter, we summarize this thesis and discuss possible future challenges and work.

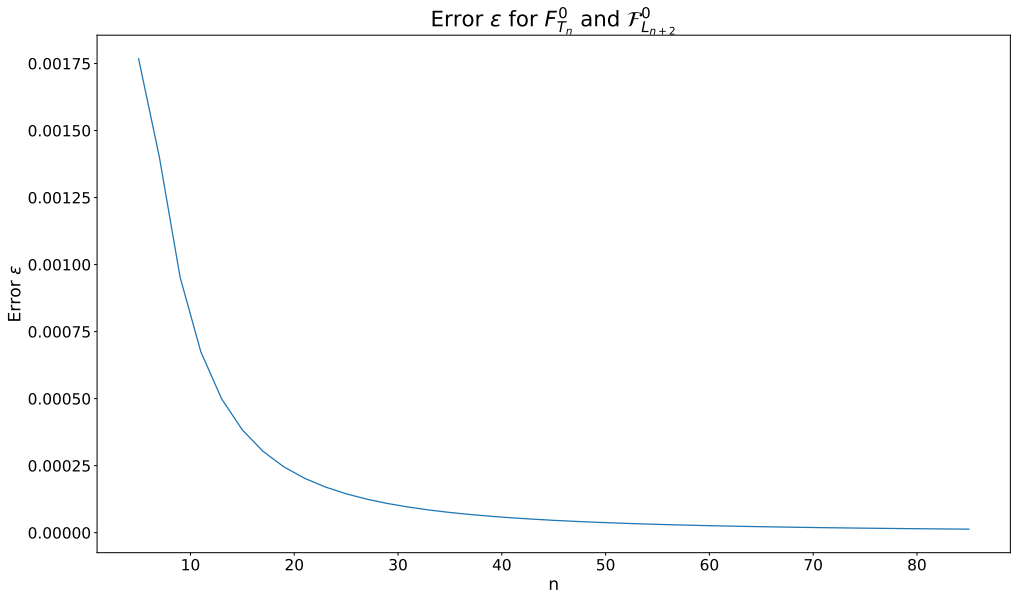


Figure 5.11: Figure on the error ϵ for the approximation of emergency edge limits with a Toeplitz matrix.

6 Conclusion and discussion

The goal of this thesis has been to study the effect of network topology on a particular cascade mechanism. To do so, we focused on trees and cycles in Section 3, the Erdős–Rényi model in Section 4 and lattice topology in Section 5. Next, we summarize the main takeaways from our results, after which we look at the potential future work.

Firstly, our analytical analysis has conclusively demonstrated that in scenarios where the initial graph configuration is a tree, the cascade phenomenon ceases after the initial failure. In contrast, our findings reveal that when the initial graph forms a cycle, the cascade process typically halts after the occurrence of two edge failures. Note that both for cycles and trees this results in a single disconnection that separates the graph into two components. Moreover, our research indicates that in cases involving a hybrid structure of trees and cycles within the graph, the cascade potentially extends further, irrespective of whether the initial edge failure occurs within a cycle or a tree component.

To get an understanding of the cascade in the Erdős–Rényi model that has parameters p, edge probability, and n number of nodes. We looked at the giant component that arises as p is larger than the percolation threshold $\frac{1}{n}$. The insights from Section 3 were useful, as we saw with an increase in p, there is an increase in the probability for cycles in the giant component. As our first result, we found that on the approximate interval $0 < p < 0.065$

the ratio of edge failures R, which is a metric describing the amount of edge failures relative to the initial size of the giant component, increases. This increase occurs because, within this interval, the giant component experiences an increase in the number of cycles. Consequently, the rate at which edge failures occur accelerates, surpassing the rate of the initial number of edges in the giant component, resulting in a higher ratio of edge failures. As p increases beyond this interval, the rate at which edge failures occur begins to align with the rate at which initial edges are added to the giant component, slowing down the gradient of increase. Secondly, although with larger p there seems to be an increase in edge failures on this initial interval, the average disconnection size $\bar{N}$, describing the average for the individual disconnections during the cascade, decreases. This is because, from our results, we have seen that more edge failures are needed to result in disconnections as p increases. When more edge failures happen before disconnection, the component that disconnects is already of a smaller size. Even more so, as p becomes large, the components that disconnect are single nodes or small line components. Additionally, just like the metric for edge failures T, the total disconnection ratio T, which is the number of nodes disconnected from the demand node relative to the initial amount of nodes of the giant component, increases on the interval $0 < p < 0.065$. As we already described how the size of the disconnections becomes smaller for increasing p, there must be more disconnection for T to increase. Additionally, we have seen that both metrics, ratio of edge failures R and total disconnection ratio T increase up to a certain point when the gradient of increase declines. This could imply the existence of certain values to which these metrics converge, however, this could be a topic of further research to explore these thresholds, which is discussed in the next section.

Next to the ER model, the effect of the more structured square lattice on the cascade was also explored. The most notable finding was that the cascade will always result in a shape that approximately resembles a T or a cross in the square lattice when the demand node is placed in the centre. Along the same lines, we have seen that approximately three quarters always disconnect from the demand node.

Lastly, we also had an initial approach to analytical findings for the lattice model, in which the emergency edge limits were approximated reasonably well by a Block Toeplitz matrix. However, due to time constraints, the full implications of this were not studied yet. This could be explored further in research as we discuss in more detail in the next section.

6.1 Potential future work

In this section, we discuss the future work that could be explored based on the results and insights gained from this thesis. Firstly, we take a look at the mathematical contributions and potential next steps. Secondly, we take a look at the models used in this thesis and how other models might be explored in the future.

6.1.1 Mathematical contributions

In the introduction of this thesis in Section 1.2 different graphs for modelling a power grid were mentioned. In this thesis, the network topologies chosen were the ER model and the lattice model. Different numerical results were obtained in Section 4 andSection 5.2, and these results could hold for different (random) graphs, like those mentioned in Section 1.2. Thus, it would be of interest to study other network topologies and see whether there are universal properties that can be derived for certain types of networks.
Another interesting avenue to study could be to explore the numerical results from an analytical perspective. We can look at whether the numerical results can be proven rigorously. A first approach would be to explore the potential converging behaviour that we found for the ER model for both the metrics T and R. It would be interesting to see whether this converging behaviour prevails in other random graphs and whether it can be proven rigorously that a certain limit exists. Additionally, for the lattice model, we could try to rigorously show that certain bounds, like a minimal disconnection of three quarters, exist.
Additionally, our research employed specific metrics for measuring cascade effects, like the ratio of edge failures R, average disconnection $\bar{N}$ and disconnection ratio T. However, additional metrics are a potential avenue for further investigation to improve the understanding of cascading failures in the ER and lattices model. For instance, a metric called the average mutual edge flow change ratio, which quantifies the amount of flow change for each edge, could be explored to get an idea of the amount of flow redistribution [25]. For this metric, a higher value for the average mutual edge flow change ratio would imply a higher probability for the cascade to propagate as more edges might exceed their emergency edge limit. Another interesting metric would be to look at the emergency edge limit of the first edge failure. This is of interest as an edge with a higher emergency edge limit will carry more flow and thus when this edge is removed more flow is redirected to other edges. Furthermore, another avenue of exploration would be the influence of different loads λ on the cascade as this determines what percentage of the emergency edge limits is initially assigned to the network [15].
Lastly, in Section 5.3, we introduced an initial analytical approach to describe the cascading failures on a square lattice by approximating the emergency edge limits. This approximation offers a computationally efficient method for estimating emergency edge limits. Throughout this thesis, we have shown that this approximation serves as a reliable method for determining emergency edge limits. However, future work could involve exploring the precision of this approximation for specific types of edges. For example, it may perform better for edges

close to the demand node compared to those at a greater distance. Another avenue for research could involve applying this approximation to model emergency edge limits within a cascade, to evaluate how closely the resulting cascade mirrors one modeled with actual emergency edge limits.

6.1.2 Different models

In exploring the effect of network topology on cascade behaviour, our study employed a limiting case of the Pareto distribution as discussed in Section 1.1 and Section 2.1.1, recognizing that a more delicate approach is needed to study the model in generality with actual Pareto distributed demands involve the Pareto distribution itself. The choice for the limiting case was driven by the need to simplify the analysis, with the goal that future research could benefit from the results of this thesis. This requires the derivation of results that show the convergence of the cascade sequence to the limiting case with unit demand. If so, the resulting graph after the cascade would be the same from which one can derive relatively directly the blackout size [23].
Furthermore, our analysis utilized a Direct Current (DC) flow model over an Alternating Current (AC) model as discussed in Section 1.2. While the DC model offers a way to analyze a cascade analytically, it abstracts away some of the nuances present in real-world AC systems, like reactive power [17, 24]. This simplification was necessary for the scope of our study, in which we aimed for analytical results, but represents a limitation in fully capturing the electrical properties of power grids. As AC models are useful for simulation, in future work, the results of this thesis can be tested using simulation results on the AC model.
Furthermore, the distinction between high-voltage and low-voltage networks is crucial. While our analysis primarily focuses on the simplifications afforded by the DC model—typically more relevant to high-voltage transmission systems—low-voltage networks often feature tree or radial topologies for effective electricity distribution over shorter distances. These topologies are mentioned in [30] and analyzed in our study, especially in Section 3.1, suggesting that the insights gained could potentially extend to low-voltage system analyses. Future work might include examining how the findings from our DC model-based analysis could inform or be adapted to the specific conditions and operational challenges of low-voltage networks.
These reflections highlight that there is a rich unexplored research area within this topic where this thesis provides a first consideration. By adopting a Pareto distribution for the demand, integrating AC models to capture high-voltage power grids, considering tree structures with respect to a different flow model to capture low to medium-voltage systems, finding potential universal properties for certain types of networks, exploring alternative cascade measurement metrics, and experimenting with the approximation for emergency edge limits in lattices, subsequent studies can build upon our findings to offer deeper insights into the effect of network topology on cascading behaviours.

References

[1] Réka Albert and Albert-László Barabási. "Statistical mechanics of complex networks". In: *Reviews of Modern Physics* 74.1 (Jan. 2002), pp. 47–97. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.74.47.

[2] Béla Bollobás. "A Probabilistic Proof of an Asymptotic Formula for the Number of Labelled Regular Graphs". In: *European Journal of Combinatorics* 1.4 (Dec. 1980), pp. 311–316. ISSN: 01956698. DOI: 10.1016/S0195-6698(80)80030-8.

[3] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman. "Complex dynamics of blackouts in power transmission systems". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 14.3 (Sept. 2004), pp. 643–652. ISSN: 1054-1500. DOI: 10.1063/1.1781391.

[4] Benjamin A. Carreras, David E. Newman, and Ian Dobson. "North American Blackout Time Series Statistics and Implications for Blackout Risk". In: *IEEE Transactions on Power Systems* 31.6 (Nov. 2016), pp. 4406–4414. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2015.2510627.

[5] Benjamin A. Carreras, David E. Newman, and Ian Dobson. "North American Blackout Time Series Statistics and Implications for Blackout Risk". In: *IEEE Transactions on Power Systems* 31.6 (Nov. 2016), pp. 4406–4414. ISSN: 0885-8950. DOI: 10.1109/TPWRS.2015.2510627.

[6] David P. Chassin and Christian Posse. "Evaluating North American electric grid reliability using the Barabási–Albert network model". In: *Physica A: Statistical Mechanics and its Applications* 355.2-4 (Sept. 2005), pp. 667–677. ISSN: 03784371. DOI: 10.1016/j.physa.2005.02.051.

[7] Eduardo Cotilla-Sanchez, Paul D. H. Hines, Clayton Barrows, and Seth Blumsack. "Comparing the Topological and Electrical Structure of the North American Electric Power Infrastructure". In: *IEEE Systems Journal* 6.4 (Dec. 2012), pp. 616–626. ISSN: 1932-8184. DOI: 10.1109/JSYST.2012.2183033.

[8] Priyanka Dey, Rachit Mehra, Faruk Kazi, Sushama Wagh, and Navdeep M. Singh. "Impact of Topology on the Propagation of Cascading Failure in Power Grid". In: *IEEE Transactions on Smart Grid* 7.4 (July 2016), pp. 1970–1978. ISSN: 1949-3053. DOI: 10.1109/TSG.2016.2558465.

[9] I. Dobson and L. Lu. "Voltage collapse precipitated by the immediate change in stability when generator reactive power limits are encountered". In: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 39.9 (1992), pp. 762–766. ISSN: 10577122. DOI: 10.1109/81.250167.

[10] Ian Dobson, Benjamin A. Carreras, Vickie E. Lynch, and David E. Newman. "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 17.2 (June 2007). ISSN: 1054-1500. DOI: 10.1063/1.2737822.

[11] P Erdős and A Rényi. *ON THE EVOLUTION OF RANDOM GRAPHS*. Tech. rep.

[12] Geoffrey Grimmett. "What is Percolation?" In: 1999, pp. 1–31. DOI: 10.1007/978-3-662-03981-6{_}1.

[13] P. Hines, K. Balasubramaniam, and E.C. Sanchez. "Cascading failures in power grids". In: *IEEE Potentials* 28.5 (Sept. 2009), pp. 24–30. ISSN: 0278-6648. DOI: 10.1109/MPOT.2009.933498.

[14] Studiosus Kirchhoff. "Ueber den Durchgang eines elektrischen Stromes durch eine Ebene, insbesondere durch eine kreisförmige". In: *Annalen der Physik* 140.4 (Jan. 1845), pp. 497–514. ISSN: 0003-3804. DOI: 10.1002/andp.18451400402.

[15] Dan Lv, Ali Eslami, and Shuguang Cui. "Load-dependent cascading failures in finite-size erdös-rényi random networks". In: *IEEE Transactions on Network Science and Engineering* 4.2 (Apr. 2017), pp. 129–139. ISSN: 23274697. DOI: 10.1109/TNSE.2017.2685582.

[16] Dr Joseph Maubach. *Scientific Programming (2MMN20): Speeding up Image, Graph, Grid, Sparse Matrix and Database Programs*. Tech. rep.

[17] Dhagash Mehta, Daniel K. Molzahn, and Konstantin Turitsyn. "Recent advances in computational methods for the power flow equations". In: *2016 American Control Conference (ACC)*. IEEE, July 2016, pp. 1753–1765. ISBN: 978-1-4673-8682-1. DOI: 10.1109/ACC.2016.7525170.

[18] David Mikkelson. *Snopes*. Aug. 2003. URL: https://www.snopes.com/fact-check/north-american-blackout/.

[19] Robert Andrews Millikan and Edwin Sherwood Bishop. *Elements of electricity a practical discussion of the fundamental laws and phenomena of electricity and their practical applications in the business and industrial world.* American Technical Society, 1917.

[20] Michael Molloy and Bruce Reed. "A critical point for random graphs with a given degree sequence". In: *Random Structures & Algorithms* 6.2-3 (Mar. 1995), pp. 161–180. ISSN: 1042-9832. DOI: 10.1002/rsa.3240060204.

[21] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. *The Fundamentals of Heavy Tails*. Cambridge University Press, June 2022. ISBN: 9781009053730. DOI: 10.1017/9781009053730.

[22] Tommaso Nesti, Fiona Sloothaak, and Bert Zwart. "Emergence of Scale-Free Blackout Sizes in Power Grids". In: *Physical Review Letters* 125.5 (July 2020). ISSN: 10797114. DOI: 10.1103/PhysRevLett.125.058301.

[23] Tommaso Nesti, Fiona Sloothaak, and Bert Zwart. *Supplemental Material for: Emergence of scale-free blackout sizes in power grids*. Tech. rep.

[24] K. Purchala, L. Meeus, D. Van Dommelen, and R. Belmans. "Usefulness of DC power flow for active power flow analysis". In: *IEEE Power Engineering Society General Meeting, 2005.* IEEE, pp. 2457–2462. ISBN: 0-7803-9157-8. DOI: 10.1109/PES.2005.1489581.

[25] Saleh Soltan, Dorian Mazauric, and Gil Zussman. "Cascading failures in power grids - Analysis and algorithms". In: *e-Energy 2014 - Proceedings of the 5th ACM International Conference on Future Energy Systems*. Association for Computing Machinery, 2014, pp. 195–206. ISBN: 9781450328197. DOI: 10.1145/2602044.2602066.

[26] Ido Tishby, Ofer Biham, Eytan Katzav, and Reimer Kühn. "Revealing the microstructure of the giant component in random graph ensembles". In: *Physical Review E* 97.4 (Apr. 2018), p. 042318. ISSN: 2470-0045. DOI: 10.1103/PhysRevE.97.042318.

[27] U.S.- Canada Power System Outage Task Force. *Final Report on the Blackout in the United States and Canada: Causes and Recommendations*. Tech. rep. 2003.

[28] Duncan J. Watts and Steven H. Strogatz. "Collective dynamics of 'small-world' networks". In: *Nature* 393.6684 (June 1998), pp. 440–442. ISSN: 0028-0836. DOI: 10.1038/30918.

[29] H.H. Yan. "PTDF and TLR from a power marketer's perspective". In: *199 IEEE Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.99CH36364)*. IEEE, July 1999, pp. 156–161. ISBN: 0-7803-5569-5. DOI: 10.1109/PESS.1999.784339.

[30] Haipeng Zhang, Jian Zhao, Xiaoyu Wang, and Yi Xuan. "Low-Voltage Distribution Grid Topology Identification With Latent Tree Model". In: *IEEE Transactions on Smart Grid* 13.3 (May 2022), pp. 2158–2169. ISSN: 1949-3053. DOI: 10.1109/TSG.2022.3146205.