

---

# SCALE-FREE CASCADING FAILURES

## GENERALIZED APPROACH FOR ALL SIMPLE, CONNECTED GRAPHS

---

Agnieszka Janicka<sup>1</sup>, Fiona Sloothak<sup>1</sup>, and Maria Vlasiou<sup>2</sup>

<sup>1</sup>*Eindhoven University of Technology, 5612 AZ Eindhoven, the Netherlands*

<sup>2</sup>*University of Twente, 7522 NB Enschede, the Netherlands*

### ABSTRACT

Cascading failures, wherein the failure of one component triggers subsequent failures in complex interconnected systems, pose a significant risk of disruptions and emerge across various domains. Understanding and mitigating the risk of such failures is crucial to minimize their impact and ensure the resilience of these systems. In multiple applications, the failure processes exhibit scale-free behavior in terms of their total failure sizes. Various models have been developed to explain the origin of this scale-free behavior. A recent study proposed a novel hypothesis, suggesting that scale-free failure sizes might be inherited from scale-free input characteristics in power networks. However, the scope of this study excluded certain network topologies. Here, motivated by power networks, we strengthen this hypothesis by generalizing to a broader range of graph topologies where this behavior is manifested. Our approach yields a universal theorem applicable to all simple, connected graphs, revealing that when a cascade leads to network disconnections, the total failure size exhibits a scale-free tail inherited from the input characteristics. We do so by characterizing cascade sequences of failures in the asymptotic regime.

**Keywords** Cascading failures · Heavy-tailed failures · Scale-free failures · Overload failures · Principle of a single big jump

## 1 Introduction

A cascading failure is a dynamic process that occurs within complex networks composed of interconnected parts, where the failure of an initial component triggers a chain reaction, causing subsequent failures of successive parts [1, 2]. Cascading failures have the potential to propagate throughout the entire network, resulting in significant disruptions and disturbances for its users. Consequently, the study of cascading failures aims to understand their emergence, analyze their behavior, and develop effective strategies to mitigate associated risks. Prominent instances of cascading failures span diverse domains, including but not limited to power grid blackouts [3], the spread of epidemics [4] or viruses in computer systems [5], congestion in transportation systems [6], disruption of supply chains [7], systemic risk in finance [8], and various other domains [9].

In many such cascading networks, it has been observed that failure sizes, which can be the blackout size, the number of individuals affected by the epidemics, or congestion delay, have heavy tails [10, 11, 12]. Specifically, tails are often *scale free*, which means that

$$\mathbb{P}(S > x) \sim Cx^{-\alpha}, \quad (1)$$

where  $C > 0$  and  $\alpha > 0$  are constants, and  $f(x) \sim g(x)$  denotes that  $\frac{f(x)}{g(x)} \rightarrow 1$  as  $x \rightarrow \infty$ . This heavy-tailed nature implies that the occurrence of significantly large failures, which can be considered “catastrophic”, cannot be dismissed as virtually impossible. Hence, understanding cascading failure behavior and devising strategies to prevent large-sized failures becomes imperative. With this motivation in mind, numerous studies have developed cascading failure models aiming to explain the emergence of scale-free behavior [3]. In this paper, we are motivated by applications in power networks. In this setting, the cascading failures lead to blackouts and it is well-established that the distribution of blackout sizes, measured in the amount of lost power, exhibits a scale-free behavior [11].

Studies of cascading failures often focus on the topological aspect, i.e., to understand how the underlying network influences the spread of failures. Large research effort in this field is devoted to the analysis of *structural* failures [13, 14, 15, 16], which refers to the mechanism where the failure spreads locally through direct neighbors. Cascading failures in power networks, which motivate the study in this paper, exhibit an intrinsically different behavior. If a line exhibits a failure, the next line to trip may be disconnected from the original source [3]. This type of mechanism is called *overload* failures and has also received a lot of attention [17]. For instance, in [18] the authors consider a nodal overload model on *scale-free networks* in which some relevant quantity is transported between nodes via the shortest path. In the context of complex networks, the term *scale-free networks* refers to networks with a scale-free degree distribution [19]. Other models with similar characteristics under various performance measures are studied in [5], [20], and [21]. The models in [22, 23, 24, 25] are motivated by power networks, which is why they assume that the load on the network is redistributed in a manner dictated by physical laws. The objective of these studies is to understand the influence of the network’s structure on the vulnerability of the system, with a particular focus on scale-free networks. As such, the findings are induced by the assumed network structure and do not generalize to other network topologies. In this paper, we show that scale-free failure sizes can emerge in other network structures as well.

The study of scale-free failures has expanded beyond the topological aspect, providing other hypotheses for the emergence of scale-free failures. For example, the Sandpile model, proposed in [26], introduces the notion of *Self-Organized Criticality* (SoC) as a potential emergence mechanism. This notion assumes that a system comprising particles naturally evolves towards *criticality* – a stage where even a small disturbance can lead to a radical behavioral change of the system, called a *phase transition*. This transition induces a redistribution of particles, leading to a cascade. For example, in the context of power networks, a *critical* stage occurs when the system operates at its full capacity, and a *phase transition* can be caused by a local disturbance, such as a failure of a line, leading to a cascade of line failures. The authors show that the *phase transition* in self-organized systems leads to scale-free failure sizes. Models inspired by SoC have been applied to study phenomena in physics, biology, computer science or geology [27, 28, 29, 30]. For a comprehensive account of the SoC theory, we refer to [31].

Specifically for power networks, which is the focus of our work, the notion of SoC has been applied to explain scale-free blackout sizes. For example, in [32] the authors present initial evidence for SoC in power networks, using time-series analysis methods. The cascade size in this and other SoC-inspired models exhibits a heavy tail [33], which means that SoC could be a potential explanation for the scale-free nature of blackout sizes. Another notable model in this field is the OPA model [34]. Similar to the Sandpile model, the redistribution of power in the OPA model can lead to line overloads and subsequent failures, creating a cascade. The model assumes two time scales, ‘slow’ and ‘fast’. The ‘slow’ time scale is used to model the continuous increase of power demand over time. This increase drives the system to operate near its full capacity. On the other hand, cascading failures happen in the ‘fast’ time scale. A cascade may be triggered either by random line failures or when a maximum capacity has been reached due to a demand increase. The authors use a standard approach in modeling power networks by using a linear approximation of true power flow, called the *DC power flow* equations [34, 35], which we also apply in this paper.

Other than Self-Organized Criticality, many other mechanisms have been proposed to explain the emergence of scale-free blackouts. An example of such a model is the CASCADE model [36]. It is a probabilistic, overload failure model that can exhibit heavy-tailed failure size, but only in a critical regime. The authors show that the total failure size is given by a quasi-binomial distribution. This distribution can be approximated by the offspring distribution of a branching process, which, in a critical state, is heavy-tailed with parameter  $\alpha = \frac{1}{2}$  as in (1).

While this literature offers many different approaches for the analysis of blackouts in power networks, the emergence of scale-free blackout sizes is still not well understood. The SoC-inspired models exhibit the desired behavior; however, there is not enough statistical evidence showing that real power networks are indeed self-organized [37]. The CASCADE model is highly stylized and the load distribution mechanism is too simplistic to capture the behavior of power networks. Hence, it does not offer a credible explanation of the emergence of scale-free blackouts. The research in complex networks offers explanations for particular graph structures, but these findings do not generalize to other topologies. Thus, the study of cascading failures in power networks continues to be a flourishing area of research.

Our work enriches the literature on the emergence of scale-free cascading failures by strengthening a recent result. Our approach is inspired by a paper by Nesti et al. [38], which presents a novel probabilistic explanation for the emergence of scale-free blackouts. The authors propose a stochastic overload model and, under certain assumptions, derive results for the tail of the blackout distribution. The results suggest that scale-free blackout sizes may be induced by the scale-free demand distribution. In this paper, we adopt the model of [38] and generalize the main result of the paper. By taking an alternative proof approach, we remove some of their assumptions. Thus, we obtain a universal result applicable to all connected and simple graph topologies i.e. graphs with no multi-edges or self-loops, thereby advancing our understanding of cascading failures.

The model in [38] assumes a power network with  $n$  nodes and  $m$  edges. Each node  $i \in \{1, \dots, n\}$  has a specified demand  $d_i$  and generation  $g_i$ . The nodal demand  $d_i$  is a realization of the Pareto-tailed demand random variable  $X_i$  with parameter  $\alpha$ , where  $X_i$ 's are mutually independent. In the context of power networks, this assumption is justified by the well-established observation that city-size distributions are scale free [39, 40, 41]. The model assumes that the demand is exogenous and the generation is planned to meet this demand. In particular, the generation is a solution to a convex direct-current optimal power flow problem (DC-OPF) [42]. While individual nodes may not maintain balance, the total generation and demand in the network are equal, i.e.,  $\sum_{i=1}^n d_i - g_i = 0$ . This is achieved by transporting power between nodes through the edges according to the DC power flow equations. Each edge has a capacity limit that is initially not exceeded. A cascade of failures is triggered by an external disturbance, leading to the failure of an edge. If the failure leads to network disconnection, the demand-generation balance is restored in each connected component by proportionally decreasing whichever resource is in abundance. Consequently, the flow from the failed edge is redistributed. The edge with the largest *relative exceedance* (if any) – the absolute ratio of flow and capacity – fails next and the flow is redistributed again. This process continues until there are no more flows on edges that exceed the edge capacity, marking the end of the cascading failure process.

The main object of analysis in [38] is the total cascade size  $S$ . It is defined as the total demand lost due to balance restoration. Specifically,

$$S := \sum_{i=1}^n \left( d_i - d_i^{(\text{end})} \right),$$

where  $d_i^{(\text{end})}$  is the demand of node  $i$  after the cascade stops. The authors show that, under certain assumptions, the total cascade size inherits the Pareto tail from the demand distribution, i.e., Equation (1) holds whenever network disconnection occurs. In the proof, the authors rely on the principle of a single big jump; they show that a large cascade is predominantly caused by scenarios where one node has a significantly larger demand, compared to other nodes. Specifically, for every  $\varepsilon > 0$ ,

$$\mathbb{P}(S > x) \approx \mathbb{P} \left( S > x; \sum_{\substack{X_i \neq M, \\ i \in \{1, \dots, n\}}} X_i < \varepsilon M \right), \quad (2)$$

with  $M = \max\{X_1, \dots, X_n\}$  and  $x > 0$  large. Furthermore, the authors show that the cascade sequence remains the same for all  $\varepsilon \geq 0$  sufficiently small. As a consequence, to obtain the tail behavior of  $S$ , it is sufficient to understand the total cascade size in the limiting case of  $\varepsilon = 0$ , which is exploited in the proof.

The results in [38] rely on two assumptions, ensuring that the cascade is well-defined. Our first contribution is showing that these assumptions are not merely technical but essential, as not all graphs satisfy them. More specifically, Theorem IV.12 in [38] requires a unique maximizer of the exceedance at every step of the cascade. In Lemma 14, given in Section 5, we derive necessary and sufficient conditions that characterize this uniqueness. Then, based on this insight, in Example 1, we construct a graph for which there are non-unique maximizers during the cascade. This raises the need for an “exceedance tie-breaking rule”, specifying which edge should fail when a non-unique maximizer is encountered. With such a rule in place, the cascade of failures is well-defined for any simple, connected graph.

After describing the model in Section 2, we give in Section 3 our second and main contribution, Theorem 2. The theorem states that for a simple, connected graph  $\mathcal{G}$  with Pareto-tailed demand vector  $d = (d_1, \dots, d_n)$ , if a cascade leads to network disconnection, then the total failure size has a Pareto tail (see Equation (1)). Note that Theorem IV.12 of [38] is a special case of our theorem, which we show in Lemma 18 in Section 6, where we also discuss the impact of the “exceedance tie-breaking rule” on our results. To prove the theorem, in the first step, we rely on the principle of a single big jump, as given in Equation (2), to show that we can limit our analysis to scenarios when one node has a large demand and the sum of other demands is bounded. Specifically, we consider all demands satisfying  $\sum_{i \neq j} d_i < \varepsilon d_j$  for fixed  $\varepsilon > 0$  and  $j \in \{1, \dots, n\}$ . Observe that even for a fixed node  $j$  and  $\varepsilon > 0$ , there are infinitely many demand vectors that satisfy this constraint. This is where our proof strategy diverges from [38], where the authors solely consider the case of  $\varepsilon = 0$ . The next step is to obtain the optimal generation vector  $g^*(d) = (g_1^*, \dots, g_n^*)$ , given as the solution of the DC-OPF problem with input  $d$  (see Equation (5)). To the best of our knowledge, the closed-form solution for  $g^*(d)$  is not known. Hence, in Lemma 4 we characterize the optimal solution  $g^*(d)$ , showing it is a projection onto a facet of a well-defined convex polytope  $\mathcal{F}_d$ , given by the feasible region of the DC-OPF problem. Exploiting this characterization, in Proposition 5 we demonstrate that the demands can be partitioned into a *finite* number of subsets, each corresponding to a projection onto a different facet of  $\mathcal{F}_d$ . Consequently, in Proposition 7, we show that the cascade sequence does not depend on  $\varepsilon$  in the neighborhood of 0, which is a technical, yet crucial result required in the proof of Theorem 2. As a result, all demands in a given subset induce a fixed cascade sequence, which we show in Lemma 8. Finally, we prove that  $S$  exhibits a scale-free tail behavior by conditioning on each of the

finitely many cascading sequences and exploiting the independence with respect to  $\varepsilon$ . All major results are proven in Section 4 and proofs of auxiliary results are given in Appendix A.

## 2 Model and notation

In this section, we describe the model of power networks used in this paper. The basic element of the model is a graph with nodes and edges, where each node has an exogenously given demand. In addition, each node may generate power, which together with the demand, induces power flow through the edges of the graph. Edge flow limits, optimal power generation, and the cascading failure process are determined sequentially in three stages:

1. *Planning stage.* In this stage, we determine *operational* edge flow limits following a basic network design process. First, given nodal demands, we determine nodal *planning* generations that yield the minimal generation costs, assuming that there are no power flow limitations. Then, we choose *planning* edge capacities equal to the *planning* flow, induced by the exogenous demand and planning generation. Finally, we set the *operational* edge flow limits as a fraction  $\lambda \in (0, 1)$  of the *planning* edge capacities, ensuring that the system can withstand some unexpected flow changes.
2. *Operational stage.* This stage describes the standard operation of the system, i.e., before the cascade. First, given the graph and all demands, the optimal nodal generations are determined by minimizing generation costs subject to operational edge flow limits. Given these generations, power is distributed through the edges to meet the demand at each node. By construction, the power flow on the edges does not exceed the operational edge flow limits, hence the system is stable.
3. *Emergency stage.* In this stage, the stability of the system is disrupted by a random edge failure, which may induce a cascade of failures in the system. We compute the cascade sequence iteratively, by redistributing the power after each edge failure and consequently identifying the edge with the biggest strain, i.e. the next edge failure, if any.

Given this broad overview of the mechanics of the model, we now proceed to describe the model in detail.

### Network characteristics

As a part of the network characteristics, we need to describe the network, the demands, and the mechanics of the power flow in the system. Regarding the network, consider a simple, connected, directed graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  with  $n$  nodes and  $m$  edges, represented by the incidence matrix  $C \in \mathbb{R}^{m \times n}$ . In particular,

$$C_{e,i} = \begin{cases} 1 & \text{if edge } e \text{ enters node } i, \\ -1 & \text{if edge } e \text{ leaves node } i, \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, let  $\mathcal{E}_i$  denote the set of edges adjacent to node  $i$ , namely  $\mathcal{E}_i := \{e \in \mathcal{E} : C_{e,i} \neq 0\}$ . We emphasize that the direction of the edges is not a property of the underlying network but a modeling choice. As a result, the orientation can be selected in a convenient manner without loss of generality (see Lemma 1).

Let  $d = (d_1, \dots, d_n) \in \mathbb{R}_{\geq 0}^n$  be the demand vector, where  $d_i$  denote the demand of node  $i \in \mathcal{N}$ . The demand  $d_i$  is a realization of a random variable  $X_i$ , which has a Pareto tail with parameter  $\alpha > 0$ . Specifically,

$$\mathbb{P}(X_i > x) \approx Kx^{-\alpha}, \quad (3)$$

for constant  $K$  and large  $x$ . Moreover, we assume that  $X_i$ 's are mutually independent.

Power flows in the network according to physical laws, following the AC power flow equations [43]. However, due to their non-linear nature, typically approximations of these equations are used in practice. In this work, we use a standard approximation, the DC approximation [35], which linearizes the AC power flow equations. This approximation scheme neglects the edge losses due to the assumption that the resistance on edges is much smaller than their reactance, leading to negligible conductance [44, 45]. Moreover, this approach assumes that the voltage angles are small and the voltage profile is flat.

To describe the flow on each edge using the DC power flow approximation, we first need to introduce the notion of edge susceptance that captures the physical characteristics of edges in the network, and the Power Transfer Distribution Factors (PTDF) matrix  $V$ . To this end, let  $b_e$  denote the susceptance of edge  $e$  and define a susceptance matrix  $B \in \mathbb{R}^{m \times m}$  as  $B := \text{diag}(b_1, \dots, b_m)$ . The matrix  $V$  is uniquely determined by the network and is given by

$V := BC(C^T BC)^+$ , where  $+$  denotes the matrix pseudo-inverse [38]. Given this, the edge flow vector  $f \in \mathbb{R}^m$  on graph  $\mathcal{G}$  with demand  $d$  and generation  $g$  is given by

$$f := V(d - g). \quad (4)$$

In this paper, for simplicity, we assume  $B = I$ , i.e., we assume identical edges.

### DC-OPF formulation

The DC-OPF model is used in both *planning* and *operational* stages in order to determine the operational flow limit vector  $\bar{f} \in \mathbb{R}^m$  and the optimal generation vector  $g^*(d) \in \mathbb{R}_{\geq 0}^n$ . Hence, we first introduce the DC-OPF program [42] with a quadratic objective function given below. Recall that  $d_i$  and  $g_i$  denote the demand and generation of node  $i \in \mathcal{N}$ , respectively.

$$\min_{g \in \mathbb{R}_{\geq 0}^n} \frac{1}{2} \sum_{i=1}^n g_i^2 : \quad (5a)$$

$$\sum_{i=1}^n g_i = \sum_{i=1}^n d_i, \quad (5b)$$

$$-\bar{f} \leq V(d - g) \leq \bar{f}. \quad (5c)$$

Observe that this version of DC-OPF is adjusted in the sense that it does not account for generation limits that are typically given as an additional condition. Note that Condition (5b) ensures the balance between the total generation and demand in the network and Condition (5c) ensures that the flows do not exceed these operational limits. Note that in this formulation, we allow for negative flows; the negative sign indicates that the flow is opposite to the initially chosen direction.

### Planning, operational and emergency stage

We now move on to the description of the stages of the model, beginning with the *planning* stage. It is crucial to note that in order to choose a suitable vector of operational edge limits  $\bar{f}$ , Problem (5) is solved without taking the edge capacity condition into account, i.e. Condition (5c) is disregarded. This results in the optimal planning generation vector denoted by  $g^{(pl)}$  which is unique and equal to  $g^{(pl)} = \bar{d}e$  [38], where  $e = (1, \dots, 1) \in \mathbb{R}^n$  is the all-ones vector and  $\bar{d} := \frac{1}{n} \sum_{i=1}^n d_i$  is the average demand per node. Moreover, the corresponding planning flow vector is given by  $f^{(pl)} = V(d - g^{(pl)})$ . Then, we set the *operational* edge limits to  $\bar{f} = \lambda |f^{(pl)}|$ , where  $\lambda \in (0, 1)$  represents a safety tuning parameter called the loading factor. Notably, since  $\text{Ker}(V) = \langle e \rangle$  [38], we deduce that the planning flow vector is given by  $f^{(pl)} = Vd$  and the operational edge limits become  $\bar{f} = \lambda |Vd|$ . This marks the end of the *planning* stage.

In the next stage of the problem, i.e. the *operational* stage, we determine the operational generation vector  $g^*(d) = (g_1^*, \dots, g_n^*) \in \mathbb{R}_{\geq 0}^n$ . The generation vector  $g^*(d)$  is a function of the demand vector  $d$  and is the unique solution to the DC-OPF problem (5). Hence, in the operational stage, Program (5) is solved again, now with Condition (5c) in place, which yields  $g^*(d)$ . With the exogenous demand  $d$  and generation  $g^*(d)$  the system operates in equilibrium until an external disturbance occurs, leading to an edge failure and initiating the *emergency* stage.

At the start of the *emergency* stage, an edge fails uniformly at random, resulting in a modified graph  $\mathcal{G}^{(2)} = (\mathcal{N}^{(2)}, \mathcal{E}^{(2)})$  with corresponding PTDF matrix  $V^{(2)}$ . This may result in an adapted demand  $d^{(2)}$  and generation  $g^{(2)}$ . Specifically, the demand and generation may be modified if the removal of the failed edge leads to a disconnected graph. This can cause an imbalance of generation and demand in the connected components. In such a case, we restore it by proportionally reducing either demand or generation in each node within each component according to Equation (8) given in the next paragraph with input  $r = 1$ ,  $d^{(1)} = d$ , and  $g^{(1)} = g^*(d)$ . In the modified graph, the flow vector is given by  $f^{(2)} = V^{(2)}(d^{(2)} - g^{(2)})$ .

In order to determine which edge will break next, we define the *emergency* edge limit

$$F := \frac{\lambda^*}{\lambda} \bar{f} \quad (6)$$

for some  $\lambda^* \geq \lambda$ . If  $\lambda^* = 1$ , we can interpret the emergency limit  $F$  as the true capacity limit of an edge, while the operational limit  $\bar{f}$  is an additionally imposed safety feature ensuring that the system does not operate at its full capacity in the operational stage. To determine the next edge to break, we look at the ratio of edge flow and the emergency edge limit. Specifically, if any component of  $f^{(2)}$  exceeds its corresponding component in  $F$ , i.e. the flow

exceeds the emergency edge limit for some edge, the cascade effect continues. The next edge to break is the one with the maximum relative exceedance  $\psi_e^{(2)} := |f_e^{(2)}|/F_e$ , with edge  $e$  determined as

$$\arg \max_{e \in \mathcal{E}^{(2)}} \{\psi_e^{(2)} : \psi_e^{(2)} > 1\}. \quad (7)$$

The subsequent steps of the cascade occur in the same manner. Suppose that in the  $r$ -th step of the cascade,  $r \in \mathbb{N}$ , an edge has just failed. This results in graph  $\mathcal{G}^{(r+1)} = (\mathcal{N}^{(r+1)}, \mathcal{E}^{(r+1)})$  with the PTDF matrix  $V^{(r+1)}$ . The demand vector  $d^{(r+1)}$  and the generation vector  $g^{(r+1)}$  are obtained from  $d^{(r)}$  and  $g^{(r)}$  through the process of balance restoration. Specifically, for a connected component  $\mathcal{C} = (\mathcal{N}_{\mathcal{C}}, \mathcal{E}_{\mathcal{C}})$  of  $\mathcal{G}^{(r+1)}$ , let  $\theta^{(r+1)} = \sum_{i \in \mathcal{N}_{\mathcal{C}}} d_i^{(r)} / \sum_{i \in \mathcal{N}_{\mathcal{C}}} g_i^{(r)}$  denote the ratio of total demand and generation in this component. Then, for  $i \in \mathcal{N}_{\mathcal{C}}$ ,

$$(d_i^{(r+1)}, g_i^{(r+1)}) = \begin{cases} \left( \frac{1}{\theta^{(r+1)}} d_i^{(r)}, g_i^{(r)} \right) & \text{if } \theta^{(r+1)} \geq 1, \\ \left( d_i^{(r)}, \theta^{(r+1)} g_i^{(r)} \right) & \text{if } \theta^{(r+1)} \leq 1. \end{cases} \quad (8)$$

This process ensures that if there is excess of demand in the component, i.e.,  $\theta^{(r+1)} > 1$ , the nodal demands are lowered by the factor  $1/\theta^{(r+1)}$ , restoring the balance of total generation and demand. Otherwise, the nodal generations are lowered in a corresponding manner. Lastly, the flow vector  $f^{(r+1)}$  is given by  $V^{(r+1)}(d^{(r+1)} - g^{(r+1)})$ . If the cascade continues, the next edge to break is the maximizer of the relative exceedance  $\psi_e^{(r+1)} := |f_e^{(r+1)}|/F_e$  with  $e$  determined similarly as in (7).

In the model by Nesti et al. [38], all graphs displaying non-uniqueness of flow exceedance maximizers are excluded from the analysis. Therefore, in order to analyze all simple, connected graphs, we must define how the cascade sequence should proceed when a non-uniqueness of a maximizer is encountered. More specifically, we introduce a fixed ‘‘tie-breaking rule’’ that determines which edge(s) should break next in the event of a non-unique exceedance maximizer. One commonly used tie-breaking rule in cascade models is to simultaneously break all edges that maximize the exceedance during a single cascade step, given that the exceedance is greater than 1. Alternatively, we may choose the next edge to fail based on the labeling of the maximizers, such as always breaking the edge with the smallest label. We refrain from imposing any specific tie-breaking rule, but we do require that the ties are resolved according to some deterministic tie-breaking rule  $T$ , defined as follows.

**Definition 1.** *A deterministic tie-breaking rule  $T$  is a function that takes as input a subset of edges  $\mathcal{M} \subseteq \mathcal{E}$ , representing the set of maximizers of the relative exceedance during a single cascade step. It outputs a non-empty subset  $\overline{\mathcal{M}} \subseteq \mathcal{M}$ , representing the set of edges that break during the current step of the cascade. Formally,  $T : 2^{\mathcal{E}} \rightarrow 2^{\mathcal{E}}$  such that  $T(\mathcal{M}) \neq \emptyset$  is the set of edges selected to fail. Note that if  $|\mathcal{M}| = 1$ , meaning that there is only one maximizer, then  $\overline{\mathcal{M}} = \mathcal{M}$ .*

Notably, this definition of a tie-breaking rule excludes any randomized approach, such as breaking one of the maximizer edges uniformly at random. Nevertheless, it is worth mentioning that the results presented in this paper can be extended in a straightforward manner to accommodate randomized tie-breaking rules as well.

The cascade stops if there are no more edges exceeding the emergency limit. The resulting graph and demand vector are denoted respectively by  $\mathcal{G}^{(\text{end})}$  and  $d^{(\text{end})}$ . The total failure size  $S$  is defined as the total demand lost during the process of balance restoration, i.e.  $S = \sum_{i=1}^n d_i - d_i^{(\text{end})}$ . If the cascade leads to no network disconnection, then  $S = 0$ .

### Additional assumptions and notation

With the model fully specified, we proceed to describe additional notation and assumptions used throughout this paper. Next, let  $X := (X_1, X_2, \dots, X_n)$  be a vector of Pareto-tailed demands with parameter  $\alpha$ . We define  $Y$  as the reordering of  $X$  such that the maximum demand is shifted to the first position; that is,

$$Y = (X_i, X_1, X_2, \dots, X_{i-1}, X_{i+1}, \dots, X_n), \text{ with } i := \arg \max_{i=1, \dots, n} \{X_i\}. \quad (9)$$

This notation ensures that the node with the largest demand always has label 1 in  $Y$ . Furthermore, in our analysis, we typically focus on specific normalized demand realizations  $d(\varepsilon, \gamma)$  given by

$$d(\varepsilon, \gamma) := e_1 + \varepsilon\gamma, \text{ for some } \varepsilon > 0 \text{ and } \gamma \in \Gamma, \quad (10)$$

where  $\Gamma := \{\gamma = (\gamma_1, \dots, \gamma_n) \mid \gamma_1 = 0, \gamma \geq 0, e^T \gamma = 1\}$  and  $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^n$ . This particular choice is motivated by the principle of a single big jump, as explained in the introduction. Namely, this choice assumes that

the first node has a significantly larger demand, compared to other nodes. Note that this is a normalized demand profile, as it ensures that the largest nodal demand is equal to 1. In other words,  $d(\varepsilon, \gamma)$  corresponds to a realization  $y = (y_1, \dots, y_n)$  of  $Y$  scaled by  $y_1$ , where  $\frac{1}{y_1}(0, y_2, \dots, y_n) = \varepsilon\gamma$  for some vector  $\gamma \in \Gamma$ . Here,  $\gamma$  represents the ratios between all but the first component of  $y$ , and  $\varepsilon$  represents the proportion of demand between the node with the largest demand and all other nodes.

For convenience of exposition and calculations, it is suitable to choose a specific orientation of the edges. In what follows, we choose the orientation induced by the following lemma.

**Lemma 1.** *Given a graph  $\mathcal{G}$  and  $\gamma \in \Gamma$ , it is possible to adjust the orientation of edges in the graph so that  $Vd \geq 0$  for all  $\varepsilon \geq 0$  sufficiently small, where  $d = e_1 + \varepsilon\gamma$ .*

*Proof.* Let  $i \in \mathcal{E}$  and let  $d$  be a demand vector as given in (10). Suppose that  $(Ve_1)_i > 0$ . Then,  $(Vd)_i = (Ve_1)_i + \varepsilon(V\gamma)_i$ . For  $\varepsilon < \frac{(Ve_1)_i}{|(V\gamma)_i|}$ , we observe that  $\varepsilon|(V\gamma)_i| < (Ve_1)_i$ . Therefore, for this choice of  $\varepsilon$ , we find that  $(Vd)_i \geq 0$ .

Now, suppose that  $(Ve_1)_i = 0$  and  $(Vd)_i < 0$ . Then, we swap the orientation of edge  $i$ . This operation results in swapping signs in the  $i$ -th row of  $V$ , and it does not affect the other edges. Since  $(Ve_1)_i = 0$ , it remains 0 after swapping, and now  $(Vd)_i > 0$ . Both cases together imply that we can assume  $Ve_1 \geq 0$  and  $Vd \geq 0$ , given  $\varepsilon$  sufficiently small. This completes the proof.  $\square$

Note that we can assume the orientation induced by Lemma 1 without loss of generality because the orientation of edges influences only the sign of the flow, which does not influence the cascade of failures.

In our analysis, the optimal generation  $g^*(d)$  will be characterized as a projection onto the boundary of the feasible region of the DC-OPF problem. To this end, let  $\mathcal{F}_d$  denote the feasible region defined by the Conditions (5b) and (5c) of the DC-OPF problem:

$$\mathcal{F}_d := \{g \in \mathbb{R}_{\geq 0}^n : e^T g = e^T d, \quad -\lambda|Vd| \leq V(d - g) \leq \lambda|Vd|\}.$$

Under the orientation of edges induced by Lemma 1, the inequality constraints of the DC-OPF problem simplify to  $(1 - \lambda)Vd \leq Vg \leq (1 + \lambda)Vd$ . Therefore,  $\mathcal{F}_d$  can be written as

$$\mathcal{F}_d = \{g \in \mathbb{R}_{\geq 0}^n : e^T g = e^T d, \quad (1 - \lambda)Vd \leq Vg \leq (1 + \lambda)Vd\}. \quad (11)$$

Observe that  $\mathcal{F}_d$  is constrained by  $2m + 1$  hyperplanes, which are given by

$$\begin{aligned} H_0 &:= \{g \in \mathbb{R}^n : e^T g = e^T d\}, \\ H_i &:= \{g \in \mathbb{R}^n : (1 - \lambda)(Vd)_i = (Vg)_i\}, \quad i = 1, \dots, m, \\ H_j &:= \{g \in \mathbb{R}^n : (1 + \lambda)(Vd)_{j-m} = (Vg)_{j-m}\}, \quad j = m + 1, \dots, 2m. \end{aligned}$$

Next, let  $\mathcal{I} := 2^{\{0, \dots, 2m\}}$  be the power set of all hyperplane indices and let the hyperplane index set  $I^*(d) \in \mathcal{I}$  be the subset of indices of hyperplanes containing the optimal solution  $g^*(d)$ . Specifically,

$$I^*(d) := \{i \in \{0, \dots, 2m\} : g^*(d) \in H_i\}. \quad (12)$$

We now define the affine subspace  $S_{I^*(d)}$  where the optimal solution  $g^*(d)$  lies, which is given by

$$S_{I^*(d)} := \bigcap_{i \in I^*(d)} H_i.$$

Observe that as we show in the proof of Lemma 10,  $g^*(d)$  is a unique point on the boundary of  $\mathcal{F}_d$  and therefore  $S_{I^*(d)}$  is non-empty. Later on in the proofs of our results, we use an alternative characterization of  $g^*(d)$  through a projection matrix  $A_{I^*(d)}$  defined in Lemma 10, where we show that

$$g^*(d) = A_{I^*(d)}d.$$

Intuitively,  $A_{I^*(d)}$  projects the demand onto the affine space  $S_{I^*(d)}$ . This characterization allows us to exploit linear properties of  $g^*(d)$ . Moreover, in Proposition 7, we show that for all  $\varepsilon > 0$  small,  $I^*(d)$  depends on  $d$  only through  $\gamma$ . In particular, there exists a mapping  $\phi : \Gamma \rightarrow \mathcal{I}$  such that  $I^*(d) = \phi(\gamma)$  for all  $\varepsilon > 0$  small. Note that we use the notation  $I^* \in \mathcal{I}$  when we refer to some element of  $\mathcal{I}$  and not the particular element corresponding to demand  $d$ . Observe that since  $\mathcal{I}$  is a finite set, there are finitely many matrices  $A_{I^*}$  with  $I^* \in \mathcal{I}$ .

Next, we give additional notation used in the analysis of the total cascade size. Let the random variable  $N \in \mathcal{N}$  be the label of the node with the largest demand and let  $L^{(1)} \in \mathcal{E}$  be the random variable denoting the edge that

failed first. Since all nodal demands are i.i.d., each node has the same probability of having the largest demand; i.e.,  $\mathbb{P}(N = i) = \frac{1}{n}$  for all  $i \in \mathcal{N}$ . Moreover, the first edge failure is chosen uniformly at random; hence,  $\mathbb{P}(L^{(1)} = l) = \frac{1}{m}$  for all  $l \in \mathcal{E}$ . Furthermore, let  $Z(i, l, \gamma)$  be the number of nodes disconnected from node  $N = i$  after the completion of the cascade that was initiated by the failure of edge  $L^{(1)} = l$  and initial demand  $d(\varepsilon, \gamma)$ . In our analysis, we show that  $Z(i, l, \gamma)$  is well defined for all  $\varepsilon > 0$  sufficiently small, as in this setting the cascade depends on  $d(\varepsilon, \gamma)$  only through  $\gamma$ . Note that  $Z(i, l, \gamma)$  takes values in  $\mathcal{Z} := \{0, \dots, n-1\}$ . With this, for a given node with largest demand  $i \in \mathcal{N}$ , first edge failure  $l \in \mathcal{E}$ , number of disconnected nodes  $z \in \mathcal{Z}$ , and hyperplane index set  $I^* \in \mathcal{I}$ , we define  $\Gamma(i, l, z, I^*)$  as the subset of  $\gamma \in \Gamma$ , for which the hyperplane index set is equal to  $\phi(\gamma) = I^*$  and the number of disconnected nodes is equal to  $Z(i, l, \gamma) = z$  in the neighborhood of  $\varepsilon = 0$ . Specifically,

$$\Gamma(i, l, z, I^*) = \{\gamma \in \Gamma : \phi(\gamma) = I^*, Z(i, l, \gamma) = z\}.$$

What we show next is that subsets  $\Gamma(i, l, z, I^*)$  partition  $\Gamma$ . We show in Lemma 8 that for a given node  $N = i$  and an edge  $L^{(1)} = l, \gamma$  uniquely defines the number of disconnected nodes  $Z(i, l, \gamma) = z$ . Further, recall that  $\gamma$  uniquely defines the hyperplane index set  $I^*$  through  $\phi$ . As a consequence, for a given  $N = i$  and  $L^{(1)} = l$ , the set  $\Gamma$  has the following partition

$$\Gamma = \bigcup_{z \in \mathcal{Z}, I^* \in \mathcal{I}} \Gamma(i, l, z, I^*).$$

Note that the partition is finite because  $\mathcal{Z}$  and  $\mathcal{I}$  are finite sets.

Next, we divide possible realizations of the random variable  $Y$  (see Equation (9)) into subsets  $Y(i, l, z, I^*)$ . Within each subset, the ratios between components of  $Y$  correspond to the ratios  $\gamma \in \Gamma(i, l, z, I^*)$ . Specifically,

$$Y(i, l, z, I^*) := \{Y \mid Y_3 = \frac{\gamma_3}{\gamma_2} Y_2, \dots, Y_n = \frac{\gamma_n}{\gamma_2} Y_2, \text{ for some } \gamma = (0, \dots, \gamma_n) \in \Gamma(i, l, z, I^*)\}. \quad (13)$$

Observe that the set  $\Gamma$  denotes all possible ratios between components  $i = 2, \dots, n$  of a vector in  $\mathbb{R}_{\geq 0}^n$ . Hence, for a given  $i \in \mathcal{N}$  and  $l \in \mathcal{E}$ , the subsets  $Y(i, l, z, I^*)$  form a finite partition of all possible realizations of  $Y$ . We exploit this partition in the analysis of the total cascade size by conditioning on the finitely many events  $\{Y \in Y(i, l, z, I^*)\}$ .

Lastly, we define the  $\mathcal{O}$  notation. Consider arbitrary functions  $f(x)$  and  $g(x)$ . Then,  $f(x) = \mathcal{O}(g(x))$  as  $x \rightarrow \infty$  means that there exists  $M > 0$  and  $x^* \geq 0$  such that  $|f(x)| \leq M g(x)$  for all  $x \geq x^*$ .

We conclude by summarizing the notation used in this paper in Table 1.

### 3 Main results

In this section, we introduce and prove our main result, Theorem 2. First, in Section 3.1, we state the theorem, provide all results required for its proof, and sketch the steps of the proof, which we fully give in Section 3.2. The proofs of the remaining results, unless stated otherwise, are provided in Section 4.

#### 3.1 Main theorem and roadmap of the proof

Our main result, Theorem 2, shows that in the case of graph disconnection, the total failure size has a Pareto tail with the same parameter as the demand distribution. More specifically, if under demand profile (10) there is a positive probability of network disconnection, then the total failure size  $S$  is Pareto-tailed with parameter  $\alpha$ . Otherwise,  $S$  has a lighter tail.

**Theorem 2** (Tail of the total cascade size). *Let  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  and  $\lambda \in (0, 1)$ . If there exist a node  $i \in \mathcal{N}$ , a first edge failure  $l \in \mathcal{E}$ , a hyperplane index set  $I^* \in \mathcal{I}$ , and a number of nodes  $z \in \mathcal{Z} \setminus \{0\}$  disconnected from node  $i$  such that  $\mathbb{P}(Y \in Y(i, l, z, I^*)) > 0$ , then*

$$\mathbb{P}(S > x) \sim C x^{-\alpha} \text{ as } x \rightarrow \infty,$$

where

$$C := \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \frac{K}{m} \left( \frac{\lambda z}{n} \right)^\alpha \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l)$$

and  $K$  is the normalizing constant of a Pareto-tailed random variable as given in (3). Otherwise,  $\mathbb{P}(S > x) = \mathcal{O}(x^{-2\alpha})$ .

The theorem is a generalization of [38, Theorem IV.12], as it does not require additional assumptions on  $\lambda$  and  $\mathcal{G}$ . In Section 5, we show that these assumptions are not merely technical, as they exclude certain graphs. In particular, in



Variable	Description
$\mathbb{R}_{\geq 0}$	Set of non-negative real numbers, $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ .
$\mathcal{G} = (\mathcal{N}, \mathcal{E})$	Directed graph of the underlying network with node set $\mathcal{N}$ and edge set $\mathcal{E}$ .
$\mathcal{E}_i$	Set of edges adjacent to (either entering or leaving) node $i$ .
$n, m$	Number of nodes and edges in graph $\mathcal{G}$ , respectively, that is, $n =  \mathcal{N} $ and $m =  \mathcal{E} $ .
$C$	Incidence matrix of $\mathcal{G}$ , $C \in \mathbb{R}^{m \times n}$ .
$L$	Laplacian matrix of $\mathcal{G}$ , $L \in \mathbb{R}^{n \times n}$ . Specifically, $L = C^T C$ .
$V$	Power Transfer Distribution Factors (PTDF) matrix of $\mathcal{G}$ , $V \in \mathbb{R}^{m \times n}$ . Specifically, $V = CL^+$ .
$e$	Vector of ones of appropriate size, $e = (1, \dots, 1)^T$ .
$e_i$	$i^{\text{th}}$ unit vector of appropriate size.
$d, g^*(d)$	Initial vectors of nodal demands and generations, respectively, with $d, g^*(d) \in \mathbb{R}_{\geq 0}^n$ .
$\Gamma$	Set of all ratios $\gamma$ , i.e. $\Gamma := \{\gamma = (\gamma_1, \dots, \gamma_n) \mid \gamma_1 = 0, \gamma \geq 0, e^T \gamma = 1\}$ .
$d(\varepsilon, \gamma)$	Particular demand profile $d(\varepsilon, \gamma) := e_1 + \varepsilon \gamma$ , for some $\varepsilon > 0$ and $\gamma \in \Gamma$ .
$\bar{d}$	Average demand per node, $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$ .
$\lambda$	Operational safety parameter, $\lambda \in (0, 1)$ .
$f$	Vector of edge flows on $\mathcal{G}$ with demand $d$ and generation $g^*(d)$ , $f \in \mathbb{R}^m$ . Specifically, $f = V(d - g^*(d))$ .
$\bar{f}$	Vector of operational edge limits on $\mathcal{G}$ , $\bar{f} \in \mathbb{R}^m$ . Specifically, $\bar{f} = \lambda V d$ .
$F$	Vector of emergency edge limits on $\mathcal{G}$ , $F \in \mathbb{R}^m$ . Specifically, $F = \frac{\lambda^*}{\lambda} V d$ for some $\lambda^* \geq \lambda$ .
$\mathcal{F}_d$	Feasible region of the DC-OPF optimization problem; see (5).
$H_j$	$j$ -th hyperplane constraining polytope $\mathcal{F}_d$ .
$I^*(d)$	Hyperplane index set for demand $d$ .
$S_{I^*(d)}$	Face of $\mathcal{F}_d$ containing $g^*(d)$ .
$A_{I^*(d)}$	Matrix characterizing the optimal generation, i.e. $g^*(d) = A_{I^*(d)} d$ .
$\phi(\gamma)$	Hyperplane index set for $\gamma$ around $\varepsilon = 0$ , that is, $\phi(\gamma) = I^*(d)$ for all $\varepsilon > 0$ small.
$\mathcal{G}^{(r+1)}$	Directed graph of the network after $r$ steps of the cascade with node set $\mathcal{N}^{(r+1)}$ and edge set $\mathcal{E}^{(r+1)}$ .
$\mathcal{G}^{(\text{end})}$	Directed graph of the network at the end of the cascade.
$d^{(r+1)}$	Vector of nodal demands after $r$ steps of the cascade, $d^{(r+1)} \in \mathbb{R}_{\geq 0}^n$ , $d^{(1)} = d$ .
$d^{(\text{end})}$	Vector of nodal demands at the end of the cascade, $d^{(\text{end})} \in \mathbb{R}_{\geq 0}^n$ .
$g^{(r+1)}(d)$	Vector of nodal generations after $r$ steps of the cascade, $g^{(r+1)}(d) \in \mathbb{R}_{\geq 0}^n$ , $g^{(1)}(d) = g^*(d)$ .
$V^{(r+1)}$	PTDF matrix of $\mathcal{G}$ after $r$ steps of the cascade, $V^{(r+1)} \in \mathbb{R}^{m \times n}$ .
$f^{(r+1)}$	Vector of edge flows on $\mathcal{G}$ after $r$ steps of the cascade, $f^{(r+1)} \in \mathbb{R}^m$ .
$\psi^{(r+1)}$	Vector of relative edge flow exceedances after $r$ steps of the cascade, $\psi^{(r+1)} \in \mathbb{R}^m$ .
$T$	Deterministic tie-breaking rule; see Definition 1.
$X$	Vector of i.i.d. Pareto distributed random variables with parameter $\alpha > 0$ , $X \in \mathbb{R}_{\geq 0}^n$ .
$Y$	Reordering of $X$ by shifting the maximum demand to the first position; see (9).
$S$	Total failure size.
$N$	Node with the largest demand, $N \in \mathcal{N}$ .
$L^{(1)}$	First edge to fail in the cascade, $L^{(1)} \in \mathcal{L}$ .
$\mathcal{I}$	Power set of $\{0, \dots, 2m\}$ .
$Z(i, l, \gamma)$	Number of nodes disconnected from node $i$ after edge $l$ initiated the cascade with nodal demand $d(\varepsilon, \gamma)$ .
$\mathcal{Z}$	Set of the number of nodes disconnected from the node with the largest demand, i.e., $\mathcal{Z} = \{0, \dots, n-1\}$ .
$\Gamma(i, l, z, I^*)$	Subset of $\gamma \in \Gamma$ , for which $Z(i, l, \gamma) = z$ and $g^*(d(\varepsilon, \gamma)) = A_{I^*} d(\varepsilon, \gamma)$ .
$Y(i, l, z, I^*)$	Subset of realizations of $Y$ having ratios between its components given by some $\gamma \in \Gamma(i, l, z, I^*)$ .

Table 1: Summary of notation.

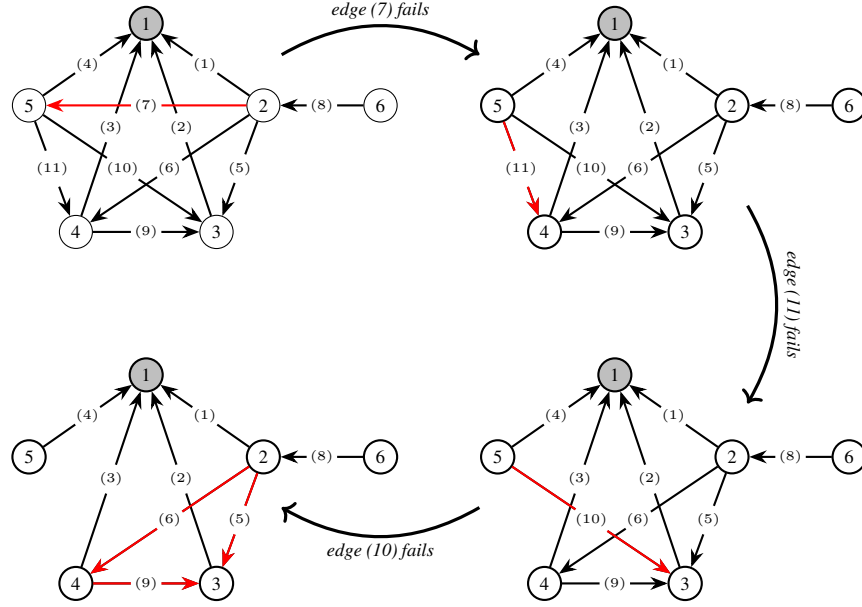


Figure 1: Example of a graph with a positive probability of a non-unique cascade sequence. Specifically, if node 1 has the largest demand, then the failure of edge (7) leads to the subsequent failure of edges (11) and (10). At this stage of the cascade, edges (5), (6), and (9) are the maximizers of the flow exceedance.

Example 1, we demonstrate that the graph in Figure 1 does not meet the assumptions of the theorem in [38] because in the 4-th step of the cascade, three edges maximize the relative flow exceedance. With our theorem, the analysis of the cascade size in this graph becomes feasible.

To be able to prove the theorem, we first take the following steps:

1. Motivated by Lemma 3, we show that it is sufficient to consider demand profiles as given in (10).
2. In Lemma 4, we reformulate the optimal generation in terms of a projection onto the polytope  $\mathcal{F}_d$ .
3. In Proposition 5, we decompose the optimal generation into the matrix equation  $g^*(d) = A_{I^*(d)}d$ .
4. In Lemma 6, we show that for all demand vectors  $d(\varepsilon, \gamma)$  in the neighborhood of  $\varepsilon = 0$ , the first column of the matrix  $A_{I^*(d)}$  is fixed and equal to  $(1 - \lambda)e_1 + \frac{\lambda}{n}e$ . We exploit this property in the proofs of Proposition 7 and Theorem 2.
5. In Proposition 7, we show that  $A_{I^*(d)}$  depends on  $d(\varepsilon, \gamma)$  only through  $\gamma$  in the neighborhood of  $\varepsilon = 0$ .
6. Using the result of Lemma 8, we partition  $\Gamma$  into finitely many subsets, each corresponding to a particular subset  $I^* \in \mathcal{I}$  and a particular cascade sequence in the neighborhood of  $\varepsilon = 0$ . This yields the corresponding partition of the feasible demand space into subsets  $Y(i, l, z, I^*)$ .
7. In Lemma 9, we derive the asymptotic behavior of the probability that  $Y_1 > x$ ,  $\sum_{j=2}^n Y_j \leq \varepsilon Y_1$ , and  $Y \in Y(i, l, z, I^*)$ , given that the node with the largest demand has label  $i$  and edge  $l$  fails first, as  $x \rightarrow \infty$ . This is a crucial ingredient in the proof of the main result.

Steps 1 to 6 carefully partition the feasible demand space into subsets with certain generation and cascade characteristics. Using this insight, we prove Theorem 2 in the following way. We first divide the demand space into two sets, one of which induces only light-tailed cascade sizes. For the remaining set, we apply the previously derived partition to ensure certain cascade characteristics within each subset. Then, we construct upper and lower bounds for the total cascade size  $S$ . Using the result of Step 7, we prove that both bounds are asymptotically equal and yield heavy-tailed behavior. This proof strategy is executed, using the following steps.

8. We condition on the label of the node with the largest demand and the first edge failure.
9. Motivated by Lemma 3, we consider the two cases  $\sum_{j=2}^n Y_j \leq \varepsilon Y_1$  and  $\sum_{j=2}^n Y_j > \varepsilon Y_1$ . For the latter case, the probability of the total cascade size is of order  $\mathcal{O}(x^{-2\alpha})$ .

10. In the former case, we study the total cascade for each event  $\{Y \in Y(i, l, z, I^*)\}$  that occurs with positive probability. For  $Y \in Y(i, l, z, I^*)$  and  $\sum_{j=2}^n Y_j \leq \varepsilon Y_1$ , the optimal generation and the cascade sequence are known due to Proposition 7 and Lemma 8. If none of the events induce a positive total cascade size, we conclude that the tail probability of  $S$  is of order  $\mathcal{O}(x^{-2\alpha})$ , as given in Step 9. Otherwise, we provide an upper and lower bound for the total cascade size  $S$  as functions of  $i, l, z, I^*$ , and  $\varepsilon$ . These bounds scale linearly with the largest demand. Importantly, as  $\varepsilon \downarrow 0$ , both bounds converge to the same limit, which is known due to Lemma 6.
11. We apply the results from Steps 7 and 10 to obtain asymptotic expressions for the upper and lower bound for the probability of the total cascade size  $S$ , showing that both depend on  $\varepsilon$ .
12. We take  $\varepsilon \downarrow 0$  for both bounds derived in Step 11. The limits are equal and yield an asymptotic expression for the probability of the total cascade size with  $Y \in Y(i, l, z, I^*)$  and  $\sum_{j=2}^n Y_j \leq \varepsilon Y_1$ . Summing up over all possible large demand nodes  $i \in \mathcal{N}$ , first edge failures  $l \in \mathcal{E}$ , number of disconnected nodes  $z \in \mathcal{Z}$  and hyperplane index sets  $I^* \in \mathcal{I}$  yields the result of the theorem.

Our proof bears similarities with the proof of Theorem IV.12 in [38]. However, major differences occur in Steps 10 and 11. In Step 10, since we do not impose additional assumptions, the continuity argument used in [38] is not valid. To circumvent this problem, we partition all feasible demands into subsets with properties that allow us to derive the expressions for the total cascade size in each scenario. Furthermore, the proof of [38, Theorem IV.12] relies on existing results for the asymptotic behavior of  $\mathbb{P}(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1)$  as  $x \rightarrow \infty$ . Our proof requires an analogous asymptotic result, derived in Lemma 9, which additionally takes into account the partition of the demand space. We now proceed to describe Steps 1 up to 7 in more detail, provide formal results, and formally prove Theorem 2.

In the first step, we rely on the following lemma, proven in the existing literature. This lemma implies that the dominant scenario for a large total cascade involves one node possessing a significantly greater demand than all other nodes combined. Therefore, we can confine our analysis to the demands specified in (10); all other scenarios can be disregarded in the asymptotic regime.

**Lemma 3** (Lemma IV.1 of [38]). *Suppose  $X_i, i = 1, \dots, n$ , are independent and identically distributed Pareto-tailed demands with parameter  $\alpha > 0$ , i.e.  $\mathbb{P}(X_i > x) \sim Kx^{-\alpha}$  as  $x \rightarrow \infty$ . Let  $S$  be the total cascade failure size. Then, for every  $\varepsilon > 0$ , as  $x \rightarrow \infty$ ,*

$$\mathbb{P}\left(S > x; \sum_{j=1, j \neq i}^n X_j \geq \varepsilon X_i\right) = \mathcal{O}(x^{-2\alpha}),$$

where  $i := \arg \max_{i=1, \dots, n} \{X_i\}$ .

In the second step, we give an alternative formulation of the DC-OPF program (5) that provides a geometrical interpretation of the optimal generation. Specifically,  $g^*(d)$  is a point in  $\mathcal{F}_d$  minimizing the Euclidean distance to  $\bar{d}e$ . In other words,  $g^*(d)$  can be constructed as an orthogonal projection. This is formally stated in the following lemma.

**Lemma 4.** *The vector  $g^*(d)$  is an optimal solution of the problem*

$$\min_{g \in \mathcal{F}_d} \{\|g - \bar{d}e\|_2\}$$

if and only if it is an optimal solution of (5).

In the third step, we leverage the theory of orthogonal projections to establish that for a given  $d$ , there exists a matrix  $A_{I^*(d)}$ , such that  $g^*(d) = A_{I^*(d)}d$ . Furthermore, we characterize the matrix  $A_{I^*(d)}$  in terms of the hyperplane index set  $I^*(d)$ , which defines the face on which  $g^*(d)$  lies. There are finitely many subsets  $I^*(d)$  because the polytope  $\mathcal{F}_d$  has finitely many faces. Hence, the set of solutions  $g^*(d)$  can be fully characterized by using a finite number of matrices. This result is given in the following proposition.

**Proposition 5.** *Let  $d$  be a demand vector satisfying (10) with  $\varepsilon \geq 0$  sufficiently small. Then,*

1. *there exists a matrix  $A_{I^*(d)} \in \mathbb{R}^{n \times n}$  such that*

$$g^*(d) = A_{I^*(d)}d;$$

2. *there exists a finite set of matrices  $\mathcal{A}$  such that  $g^*(d) = A_{I^*(d)}d$  for some  $A_{I^*(d)} \in \mathcal{A}$ .*

In the fourth step, we derive the following lemma, stating that the first column of  $A_{I^*(d)}$  is fixed for all  $\gamma \in \Gamma$  and all  $\varepsilon > 0$  sufficiently small. This column is equal to the optimal generation induced by demand vector  $e_1$ , i.e., the demand vector in the limiting case of  $\varepsilon = 0$ .

**Lemma 6.** *Let  $I^*(d)$  be the hyperplane index set for a demand vector  $d$  satisfying (10) with  $\varepsilon \geq 0$  sufficiently small. Then,*

$$A_{I^*(d)}e_1 = (1 - \lambda)e_1 + \frac{\lambda}{n}e.$$

In the fifth step, using the Karush–Kuhn–Tucker conditions, we show that for  $\varepsilon$  sufficiently small, the matrix  $A_{I^*(d)}$  depends on  $d$  only through  $\gamma$ . The subsequent proposition makes this statement explicit.

**Proposition 7.** *For a demand vector  $d(\varepsilon, \gamma)$  defined in (10) and  $\varepsilon > 0$  small enough, the optimal generation  $g^*(d)$  is given by*

$$g^*(d) = A_\gamma d, \quad (14)$$

for some  $A_\gamma \in \mathbb{R}^{n \times n}$ . More specifically, there exists a mapping  $\phi : \Gamma \rightarrow \mathcal{I}$  such that the hyperplane index set is equal to  $I^*(d) = \phi(\gamma)$  for all  $\varepsilon > 0$  sufficiently small. Hence, Equation (14) holds with  $A_\gamma := A_{\phi(\gamma)}$ .

In the sixth step, we derive the following lemma, which shows that the cascade sequence also depends on  $d$  only through  $\gamma$  in the neighborhood of  $\varepsilon = 0$ .

**Lemma 8.** *Assume that node  $i \in \mathcal{N}$  has the largest demand and edge  $e \in \mathcal{E}$  is the first to fail. Then, for a given  $\gamma \in \Gamma$ , the cascade sequence and the resulting graph  $\mathcal{G}^{(end)}$  are independent of  $\varepsilon$  for all  $\varepsilon > 0$  sufficiently small.*

As a consequence of Lemma 8, given the largest demand node  $N = i$  and the first edge failure  $L^{(1)} = l$ , the number of nodes disconnected from  $i$  is uniquely defined by  $\gamma$  in the neighborhood of  $\varepsilon$ . As a result, we can classify the demands into subsets  $Y(i, l, z, I^*)$ , as defined in (13). Each demand in a subset has the following properties: first, it has its largest component at node  $i$ ; second, it yields generation characterized by the hyperplane index set  $I^*$ ; and last, it results in  $z$  nodes disconnected from node  $i$  after the cascade initiated by edge  $l$ .

In the seventh step, we derive an asymptotic result for  $\mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in Y(i, l, z, I^*)\right)$  as  $x \rightarrow \infty$ . Our result shows that, in the limit, the event  $\{Y \in Y(i, l, z, I^*)\}$  becomes independent of the event  $\{Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1\}$ . Moreover,  $\mathbb{P}(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1) \sim n\mathbb{P}(X_1 > x)$ . This is an extension of the *principle of a single big jump* that states that  $\mathbb{P}(Y_1 > x) \sim n\mathbb{P}(X_1 > x)$  [46]. Our result is given in the following lemma.

**Lemma 9.** *Consider  $X_1, \dots, X_n$  independent and identically distributed Pareto random variables with parameter  $\alpha > 0$ , and the corresponding random variable  $Y$  as defined in (9). Then, for  $\varepsilon > 0$  sufficiently small and all  $i \in \mathcal{N}$ ,  $l \in \mathcal{E}$ ,  $z \in \mathcal{Z}$ , and  $I^* \in \mathcal{I}$ , for which  $\mathbb{P}(Y \in Y(i, l, z, I^*)) > 0$ , we have that*

$$\mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in Y(i, l, z, I^*) | N = i, L^{(1)} = l\right) \sim n\mathbb{P}(X_1 > x)\mathbb{P}(Y \in Y(i, l, z, I^*) | N = i, L^{(1)} = l).$$

Using the above results, in the following section we provide the proof of the main theorem.

### 3.2 Proof of Theorem 2

In this section, we prove the main result, following Steps 8 – 12.

*Proof.* First, following Step 8, we observe that the cascade sequence depends the node  $N$  that has the largest demand as well as on which edge failed first, denoted by  $L^{(1)}$ . Therefore, we condition on these events. Recall that  $\mathbb{P}(N = i) = 1/n$  for all  $i$  and  $\mathbb{P}(L^{(1)} = l) = 1/m$  for all  $l \in \mathcal{E}$ . Thus, we obtain

$$\mathbb{P}(S > x) = \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \frac{1}{nm} \mathbb{P}(S > x | N = i, L^{(1)} = l).$$

Next, let  $\varepsilon$  be sufficiently small so that  $Z(i, l, \gamma)$  is well-defined. This can be done due to Lemma 8, which shows that the cascade sequence is independent of  $\varepsilon$  in this regime. Then, we decompose the tail probability of the conditional total failure size in two cases:

$$\mathbb{P}(S > x | N = i, L^{(1)} = l) = \mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 | N = i, L^{(1)} = l) + \mathbb{P}(S > x, \sum_{j=2}^n Y_j \geq \varepsilon Y_1 | N = i, L^{(1)} = l).$$

From Lemma 3, it follows that  $\mathbb{P}(S > x, \sum_{j=2}^n Y_j \geq \varepsilon Y_1 | N = i, L^{(1)} = l) = \mathcal{O}(x^{-2\alpha})$ . Thus, only the first term requires further consideration. This marks the end of Step 9.

Next, let  $Z(i, l)$  be the expected number of nodes disconnected from node  $i$  after a cascade sequence that was initiated by the failure of edge  $l$  in the neighborhood of  $\varepsilon = 0$ , i.e.

$$Z(i, l) := \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} z \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l).$$

Suppose that for all nodes  $i$ , edges  $l, z \in \mathcal{Z}/\{0\}$ , and  $I^* \in \mathcal{I}$ , we have  $\mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) = 0$ . This implies that for all  $i \in \mathcal{N}$  and  $l \in \mathcal{E}$ ,  $Z(i, l) = 0$ . Consequently, almost surely, no node has been disconnected from node  $i$  for  $\varepsilon > 0$  sufficiently small. Thus, the cascade size in this setting is equal to 0 with probability 1. In particular, for  $x > 0$ ,

$$\mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l) = 0.$$

Altogether, if  $\mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) = 0$  for all  $i \in \mathcal{N}$ ,  $l \in \mathcal{E}$ ,  $z \in \mathcal{Z}/\{0\}$ , and  $I^* \in \mathcal{I}$ , then the cascade size satisfies

$$\mathbb{P}(S > x) = \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \frac{1}{nm} (0 + \mathcal{O}(x^{-2\alpha})) = \mathcal{O}(x^{-2\alpha}),$$

as argued in Step 10. Hence, from now on, we assume that there exist  $i \in \mathcal{N}$ ,  $l \in \mathcal{E}$ ,  $z \in \mathcal{Z}/\{0\}$ , and  $I^* \in \mathcal{I}$  such that  $\mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) > 0$ .

It is known that the solution of the DC-OPF program (5) is scale-invariant, i.e.  $g^*(\beta Y) = \beta g^*(Y)$ , for any demand vector  $Y$  and  $\beta > 0$  [38]. Moreover, we know that the normalized demand vector  $Y/Y_1$  is equal to some demand profile  $d(\varepsilon, \gamma)$ . Hence, due to the scale-invariance of  $g^*(d)$  and Proposition 7, it holds that

$$g^*(Y) = Y_1 g^*(Y/Y_1) = Y_1 g^*(d(\varepsilon, \gamma)) = Y_1 A_{I^*(d)} d(\varepsilon, \gamma) = A_{I^*(d)} Y.$$

This means that the linear form of the optimal generation holds not only for the normalized demand profiles but also more generally for  $Y$ , which we exploit from now on. Notably, the cascade corresponding to demand  $Y$  is the same as the cascade corresponding to  $d(\varepsilon, \gamma)$ . This is due to the linearity of the flow vector  $f$ , see (4), and the emergency edge limit vector  $F$ , see (6), with respect to demand  $d$  and generation  $g$ . This property implies that the relative exceedance vector  $\psi$ , see (7), depends solely on the proportions between the elements of  $d$ , i.e. the vector  $\gamma$ , and not the scale factor  $Y_1$ .

Let,  $A(i, l, z, I^*)$  denote the set of nodes connected to the node with the largest demand after the cascade with demand  $Y \in Y(i, l, z, I^*)$ . Note that due to the previous argument,  $A(i, l, z, I^*)$  is well-defined. Again, using the law of total probability and the fact that we know the solution form of  $g^*(Y)$  given in Proposition 7, we find

$$\begin{aligned} & \mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l) \\ &= \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) \\ &= \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \mathbb{P}\left(\sum_{j \notin A(i, l, z, I^*)} g_j^* - Y_j > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l\right) \\ &= \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \mathbb{P}\left(\sum_{j \notin A(i, l, z, I^*)} e_j^T (A_{I^*} - I) Y > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l\right). \end{aligned}$$

Note that the conditions  $\sum_{j=2}^n Y_j < \varepsilon Y_1$  and  $Y \in Y(i, l, z, I^*)$  together imply that  $Y = (e_1 + \hat{\varepsilon}\gamma)Y_1$  for some  $\hat{\varepsilon} < \varepsilon$  and  $\gamma \in \Gamma(i, l, z, I^*)$ . Let  $j \notin A(i, l, z, I^*)$ , which implies that  $j \neq 1$ . Then, using Lemma 6, we obtain

$$e_j^T (A_{I^*} - I) e_1 = \frac{\lambda}{n} e_j^T e - \lambda e_j^T e_1 = \frac{\lambda}{n}.$$

Furthermore, let  $B(i, l, z, I^*) := \max_{j \notin A(i, l, z, I^*), \gamma \in \Gamma(i, l, z, I^*)} \{|e_j^T (A_{I^*} - I)\gamma|\}$ . Thus, we find the following bounds:

$$z \left( \frac{\lambda}{n} - B(i, l, z, I^*) \varepsilon \right) \leq \sum_{j \notin A(i, l, z, I^*)} e_j^T (A_{I^*} - I) (e_1 + \hat{\varepsilon}\gamma) \leq z \left( \frac{\lambda}{n} + B(i, l, z, I^*) \varepsilon \right).$$

Using the above bounds, we find

$$\begin{aligned} & \mathbb{P}\left(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l\right) \\ & \leq \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \mathbb{P}\left(z \left(\frac{\lambda}{n} + B(i, l, z, I^*)\varepsilon\right) Y_1 > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l\right). \end{aligned}$$

Similarly, we can derive a lower bound:

$$\begin{aligned} & \mathbb{P}\left(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l\right) \\ & \geq \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \mathbb{P}\left(z \left(\frac{\lambda}{n} - B(i, l, z, I^*)\varepsilon\right) Y_1 > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l\right). \end{aligned}$$

We now move on to Step 11, where we derive the asymptotic behavior of the above bounds. Applying the result of Lemma 9 to all summands, for which  $\mathbb{P}(Y \in Y(i, l, z, I^*)) > 0$ , we find that

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{\mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l)}{n\mathbb{P}(X_1 > x)} \\ & = \lim_{x \rightarrow \infty} \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \frac{\mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l)}{n\mathbb{P}(X_1 > x)} \\ & = \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \lim_{x \rightarrow \infty} \frac{\mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l)}{n\mathbb{P}(X_1 > x)} \\ & \leq \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \lim_{x \rightarrow \infty} \frac{\mathbb{P}\left(z \left(\frac{\lambda}{n} + B(i, l, z, I^*)\varepsilon\right) Y_1 > x, \sum_{j=2}^n Y_j < \varepsilon Y_1, Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l\right)}{n\mathbb{P}(X_1 > x)} \\ & = \sum_{I^* \in \mathcal{I}} \sum_{\substack{z \in \mathcal{Z}: \\ \mathbb{P}(Y \in Y(i, l, z, I^*)) > 0}} K\left(z \left(\frac{\lambda}{n} + B(i, l, z, I^*)\varepsilon\right)\right)^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) \\ & = \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} K\left(z \left(\frac{\lambda}{n} + B(i, l, z, I^*)\varepsilon\right)\right)^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l), \end{aligned}$$

for some  $K > 0$ . Similarly, we find a lower bound:

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{\mathbb{P}(S > x, \sum_{j=2}^n Y_j < \varepsilon Y_1 \mid N = i, L^{(1)} = l)}{n\mathbb{P}(X_1 > x)} \\ & \geq \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} K\left(z \left(\frac{\lambda}{n} - B(i, l, z, I^*)\varepsilon\right)\right)^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l). \end{aligned}$$

In the final step, taking the limit  $\varepsilon \rightarrow 0$ , we obtain the equality of both bounds. Hence, we conclude that, as  $x \rightarrow \infty$ ,

$$\begin{aligned} \mathbb{P}(S > x) & \sim \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} \frac{K}{m} \left(\frac{\lambda z}{n}\right)^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) \cdot x^{-\alpha} \\ & = \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \frac{K}{m} \left(\frac{\lambda}{n}\right)^\alpha \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} z^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) \cdot x^{-\alpha}. \end{aligned}$$

Note that since we assume that  $\mathbb{P}(Y \in Y(i, l, z, I^*) \mid N = i, L^{(1)} = l) > 0$  for some  $i, l, I^*$  and  $z > 0$ , the constant in front of  $x^{-\alpha}$  is strictly positive. Thus, the result of the theorem follows.  $\square$

Theorem 2 shows us that the total failure size exhibits a Pareto tail with parameter  $\alpha$ , which is the same parameter that governs the Pareto tail of the demand distribution. This finding is in line with the previous results that show this property for a particular subset of graphs. In fact, Lemma 18 in Section 6 shows that our theorem, when applied to a graph that meets the conditions of Theorem IV.12 of [38], yields the exact same result.

## 4 Proofs and supporting results

In this section, we prove the previously introduced lemmas and propositions. Some of the proofs rely on auxiliary results, which are introduced and proved when appropriate.

### 4.1 Proof of Lemma 4

To prove this lemma, we rewrite the generation cost function (5a), in terms of a squared distance between  $g$  and  $\bar{d}e$ .

*Proof.* Observe that for all  $g \in \mathcal{F}_d$ ,

$$\|g - \bar{d}e\|_2^2 = \sum_{i=1}^n (g_i \bar{d})^2 = \sum_{i=1}^n g_i^2 - 2\bar{d} \sum_{i=1}^n g_i + \sum_{i=1}^n \bar{d}^2 = \sum_{i=1}^n g_i^2 - n\bar{d}^2,$$

where the final equality follows since  $e^T g = e^T d$  for all  $g \in \mathcal{F}_d$ . Since  $\bar{d}$  is a constant, and taking the root is a concave function, we conclude that the minimum of  $\|g - \bar{d}e\|_2$  is attained at the same point as the minimum of  $\frac{1}{2}g^T g$  over the feasible area  $\mathcal{F}_d$ .  $\square$

### 4.2 Proof of Proposition 5

We prove Proposition 5, using insights from Lemma 4. This is done in two steps; first, we state Lemma 10, which provides an explicit formula for  $A_{I^*(d)}$  using projection theory. Based on this result, we then prove the proposition.

**Lemma 10.** *For a demand vector  $d$  satisfying (10), let  $I^*(d)$  be defined as in (12). Consider the following partition of  $I^*(d)$ :*

$$I^*(d) = \{0\} \cup I_1 \cup I_2, \text{ such that } I_1 \subseteq \{1, \dots, m\}, I_2 \subseteq \{m+1, \dots, 2m\},$$

with  $s_1 := |I_1|$ ,  $s_2 := |I_2|$ , and define

$$V_{I_1} := \begin{bmatrix} (V)_{j_1} \\ \vdots \\ (V)_{j_{s_1}} \end{bmatrix}, \quad V_{I_2} := \begin{bmatrix} (V)_{k_1} \\ \vdots \\ (V)_{k_{s_2}} \end{bmatrix}, \quad I_1 = \{j_1, \dots, j_{s_1}\}, \quad I_2 = \{k_1, \dots, k_{s_2}\}.$$

Then, the optimal solution  $g^*(d)$  is given by

$$g^*(d) = A_{I^*(d)} d,$$

where  $A_{I^*(d)} := \frac{1}{n}J + \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \left( \begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix} \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \right)^+ \begin{bmatrix} (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \in \mathbb{R}^{n \times n}$  and  $J$  is the all-ones  $n \times n$  matrix.

The proof of this lemma relies on the theory of orthogonal projections and is provided in Appendix A.1. Next, using Lemma 10, we prove Proposition 5.

*Proof of Proposition 5.* The first statement of the proposition follows directly from Lemma 10. It remains to show the existence of a finite set  $\mathcal{A}$ . Let  $\mathcal{A} := \{A_{I^*(d)} | I^*(d) \subseteq \{0, \dots, 2m\}\}$ . Clearly, for each  $g^*(d)$ , there exists a matrix  $A \in \mathcal{A}$ , namely  $A_{I^*(d)}$ , such that  $g^*(d) = A_{I^*(d)} d$ . Matrix  $A_{I^*(d)}$ , as defined in Lemma 10, depends only on  $I^*(d) \subseteq \{0, \dots, 2m\}$ . Since there are at most  $2^{2m+1}$  subsets of  $\{0, \dots, 2m\}$ , the size of  $\mathcal{A}$  is bounded by  $2^{2m+1}$ . This completes the proof.  $\square$

### 4.3 Proof of Lemma 6

In this section, we prove Lemma 6, which characterizes the first column of matrices  $A_{I^*(d)}$  for certain demand vectors  $d$ . However, the proof requires three auxiliary results, Lemmas 11 – 13, which we introduce below.

In Lemma 11, we show that the number of possible hyperplane index sets  $I^* \in \mathcal{I}$  that lead to the optimal solution can be restricted. This arises from the fact that, for every optimal solution, the flows on edges adjacent to the node with the largest demand are equal to the operational edge limit  $\lambda V d$ . Consequently, the inequality constraints for these edges are always tight. This notion is formalized in the following lemma.

**Lemma 11.** *Let  $d$  be a demand vector as defined in (10) with  $\varepsilon > 0$  sufficiently small. Then, for all edges  $e \in \mathcal{E}_1$ , the following holds:*

$$(V g^*(d))_e = (1 - \lambda)(V d)_e. \quad (15)$$

*In other words, the flow on all edges adjacent to node 1 is equal to the operational limit.*

We prove the lemma by constructing an alternative solution to the DC-OPF problem (5), leading to a contradiction. The proof is provided in Appendix A.2.

Next, we show that constraint  $(1 + \lambda)(Vd)_i, i = 1, \dots, m$  can be tight only if  $(Ve_1)_i = 0$ .

**Lemma 12.** *For all  $\varepsilon > 0$  sufficiently small, if  $(Ve_1)_i > 0$  for  $i \in \{1, \dots, m\}$ , then the constraint  $(1 + \lambda)(Vd)_i$  is not tight. Specifically,*

$$(Vg^*(d))_i < (1 + \lambda)(Vd)_i.$$

The proof of the lemma relies on a continuity argument, for which the details can be found in Appendix A.3. Next, we show that the first row of matrices  $A_{I^*}$  is known for certain choices of hyperplane index set  $I^* \in \mathcal{I}$ .

**Lemma 13.** *Let  $I^* \in \mathcal{I}$  and let  $\mathcal{E}_1$  be the set of edges adjacent to node 1. If  $\mathcal{E}_1 \subseteq I^*$ , then*

$$e_1^T A_{I^*} = (1 - \lambda)e_1^T + \frac{\lambda}{n}e^T.$$

*Proof.* Let  $I^* \in \mathcal{I}$ , with  $\mathcal{E}_1 \subseteq I^*$  and let  $d$  be a demand vector satisfying (10). Let  $f_{\mathcal{E}_1}^{(pl)} := \sum_{i \in \mathcal{E}_1} f_i^{(pl)}$  and  $f_{\mathcal{E}_1}^{(op)} := \sum_{i \in \mathcal{E}_1} f_i$  be the total flow at edges adjacent to node 1 during the *planning* and *operational* phase, respectively. The assumption that  $\mathcal{E}_1 \subseteq I^*$  implies that the flow on edges adjacent to node 1 is equal to the *operational* limit, which is the fraction  $\lambda$  of flows in the *planning* phase. This means that  $f_{\mathcal{E}_1}^{(op)} = \lambda f_{\mathcal{E}_1}^{(pl)}$ . During the *planning* phase, the flow into node 1 is equal to the shortage of demand at node 1, namely

$$f_{\mathcal{E}_1}^{(pl)} = d_1 - g_1^{(pl)} = 1 - \frac{1 + \varepsilon}{n}.$$

Therefore, we conclude that  $f_{\mathcal{E}_1}^{(op)} = \lambda(1 - \frac{1+\varepsilon}{n})$ .

Matrix  $A_{I^*}$  maps  $d$  to some generation vector  $g$  for which the flow into node 1 is fixed. Since the flow into a node is given by a difference between the demand and generation, we know that  $d_1 - g_1 = \lambda(1 - \frac{1+\varepsilon}{n})$ . Hence,

$$g_1 = e_1^T A_{I^*} d = 1 - \lambda \left( 1 - \frac{1 + \varepsilon}{n} \right). \quad (16)$$

Define  $a_1 := e_1^T A_{I^*}$ . Then,

$$e_1^T A_{I^*} d = a_1(e_1 + \varepsilon \gamma) = a_1 e_1 + \varepsilon a_1 \gamma.$$

Then, (16) yields

$$a_1 e_1 + \varepsilon a_1 \gamma = 1 - \lambda + \frac{\lambda}{n} + \frac{\lambda}{n} \varepsilon.$$

Since the above statement must be true for all  $\varepsilon$  and all  $\gamma$ , we find that  $a_1 e_1 = 1 - \lambda + \frac{\lambda}{n}$ . Moreover,  $a_1 \gamma = \frac{\lambda}{n}$  for all  $\gamma$  if and only if  $a_1$  is constant in every component except for the first one, which follows from the fact that  $\gamma_1 = 0$ . This implies that  $a_1 = e_1^T A_{I^*} = (1 - \lambda)e_1 + \frac{\lambda}{n}e$ .  $\square$

We are now ready to prove Lemma 6. In the proof, we show that the first row and the first column of the matrix  $A_{I^*(d)}$  are equal, using Lemmas 11 and 12. This, together with Lemma 13, yields the result.

*Proof.* Let  $d$  be a demand vector as given in (10) for some  $\varepsilon \geq 0$  sufficiently small and let  $I^*(d)$  be the corresponding hyperplane index set. Due to Lemma 10, we have that  $g^*(d) = A_{I^*(d)} d$ , with

$$A_{I^*(d)} = \frac{1}{n}J + \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \left( \begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix} \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \right)^+ \begin{bmatrix} (1 - \lambda)V_{I_1} \\ (1 + \lambda)V_{I_2} \end{bmatrix}. \quad (17)$$

Without loss of generality, assume that  $\begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix}$  has full rank. Otherwise, remove all dependent rows from  $\begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix}$ . This action does not change the underlying subspace or the projection matrix; however, it simplifies the computations. This is because removing all dependent rows of  $\begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix}$  ensures that  $\begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix} \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix}$  is a full-rank matrix. Consequently, its pseudoinverse is equal to its inverse.

Using the above representation of  $A_{I^*(d)}$ , we show that the first row and the first column of  $A_{I^*(d)}$  are equal. Denote  $\left( \begin{bmatrix} V_{I_1} \\ V_{I_2} \end{bmatrix} \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \right)^{-1} =: Q = \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix}$ , with  $Q_1 \in \mathbb{R}^{|I_1| \times |I_1|}$ ,  $Q_2 \in \mathbb{R}^{|I_1| \times |I_2|}$ ,  $Q_3 \in \mathbb{R}^{|I_2| \times |I_1|}$ , and  $Q_4 \in \mathbb{R}^{|I_2| \times |I_2|}$ . Note that  $Q$  is symmetric by construction, hence  $Q_2 = Q_3^T$ ,  $Q_1 = Q_1^T$ , and  $Q_4 = Q_4^T$ .



The first column of  $A_{I^*(d)}$  is given by

$$\begin{aligned} A_{I^*(d)}e_1 &= \frac{1}{n}Je_1 + [V_{I_1}^T \quad V_{I_2}^T] \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix} \begin{bmatrix} (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} e_1 \\ &= \frac{1}{n}e + [V_{I_1}^T \quad V_{I_2}^T] \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix} \begin{bmatrix} (1-\lambda)V_{I_1}e_1 \\ 0 \end{bmatrix} \\ &= \frac{1}{n}e + (1-\lambda)V_{I_1}^T Q_1 V_{I_1} e_1 + (1-\lambda)V_{I_2}^T Q_3 V_{I_1} e_1, \end{aligned}$$

where in the second equality we applied the result of Lemma 12. Similarly, we derive an expression for the first row of  $A_{I^*(d)}$ :

$$\begin{aligned} e_1^T A_{I^*(d)} &= \frac{1}{n}e^T + [e_1^T V_{I_1}^T \quad 0] \begin{bmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{bmatrix} \begin{bmatrix} (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= \frac{1}{n}e^T + [e_1^T V_{I_1}^T Q_1 \quad e_1^T V_{I_1}^T Q_2] \begin{bmatrix} (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= \frac{1}{n}e^T + (1-\lambda)e_1^T V_{I_1}^T Q_1 V_{I_1} + (1+\lambda)e_1^T V_{I_1}^T Q_2 V_{I_2} \\ &= \frac{1}{n}e^T + (1-\lambda)e_1^T V_{I_1}^T Q_1 V_{I_1} + (1+\lambda)e_1^T V_{I_1}^T Q_3^T V_{I_2}. \end{aligned}$$

Hence, we observe that  $(e_1^T A_{I^*(d)})^T = A_{I^*(d)}e_1$  if and only if  $V_{I_2}^T Q_3 V_{I_1} e_1 = 0$ .

To show that  $V_{I_2}^T Q_3 V_{I_1} e_1 = 0$ , we first write an explicit expression for  $Q_3$ . In general, the inverse of an invertible block matrix  $\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix}$  is given by

$$\begin{bmatrix} P_1 & P_2 \\ P_3 & P_4 \end{bmatrix}^{-1} = \begin{bmatrix} (P_1 - P_2 P_4^{-1} P_3)^{-1} & -(P_1 - P_2 P_4^{-1} P_3)^{-1} P_2 P_4^{-1} \\ -(P_4 - P_3 P_1^{-1} P_2)^{-1} P_3 P_1^{-1} & (P_4 - P_3 P_1^{-1} P_2)^{-1} \end{bmatrix}.$$

Hence,  $Q_3 = -Q_4 V_{I_2} V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1}$  and  $-V_{I_2}^T Q_4 V_{I_1} e_1 = V_{I_2}^T Q_4 V_{I_2} V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} e_1$ .

Next, we show that  $V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} e_1 = e_1 - \frac{1}{n}e$ . For this purpose, we apply Lemma 13 to  $\tilde{I} = I^*(d) \setminus \{m+1, \dots, 2m\}$ . Note that the conditions of this lemma are met, namely that  $\mathcal{E}_1 \in \tilde{I}$ , due to Lemma 11. Hence, we obtain that  $e_1^T A_{\tilde{I}} = (1-\lambda)e_1^T + \frac{\lambda}{n}e^T$ , with

$$A_{\tilde{I}} = \frac{1}{n}J + (1-\lambda)V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1}.$$

We observe that  $A_{\tilde{I}}$  is a symmetric matrix. In particular:

$$\begin{aligned} A_{\tilde{I}}^T &= \frac{1}{n}J^T + (1-\lambda) \left( V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} \right)^T \\ &= \frac{1}{n}J + (1-\lambda)V_{I_1}^T \left( (V_{I_1} V_{I_1}^T)^T \right)^{-1} V_{I_1} \\ &= \frac{1}{n}J + (1-\lambda)V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} = A_{\tilde{I}}. \end{aligned}$$

The symmetry implies that  $A_{\tilde{I}}e_1 = (1-\lambda)e_1 + \frac{\lambda}{n}e$ . Hence,

$$V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} e_1 = e_1 + \frac{1}{1-\lambda} \left( \frac{\lambda}{n} - \frac{1}{n} \right) e = e_1 - \frac{1}{n}e.$$

Lemma 12 shows that  $g^*(d)$  lies on hyperplane  $H_j$  for some  $j \in \{m+1, \dots, 2m\}$  if and only if  $(Ve_1)_{j-m} = 0$ . This implies that the first column of  $V_{I_2}$  is equal to 0. In addition, recall that  $V_{I_2}e = 0$ . Hence,

$$V_{I_2} V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} e_1 = V_{I_2} \left( e_1 - \frac{1}{n}e \right) = V_{I_2} e_1 = 0.$$

Therefore, we conclude that

$$V_{I_2}^T Q_3 V_{I_1} e_1 = -V_{I_2}^T Q_4 (V_{I_2} V_{I_1}^T (V_{I_1} V_{I_1}^T)^{-1} V_{I_1} e_1) = 0.$$

Hence,  $A_{I^*(d)}e_1 = e_1^T A_{I^*(d)}$ . Again, due to Lemma 11, we know that  $\mathcal{E}_1 \subseteq I^*(d)$ . Therefore, we can apply Lemma 13 to  $I^*(d)$ , which yields that

$$A_{I^*(d)}e_1 = (1 - \lambda)e_1 + \frac{\lambda}{n}e.$$

□

#### 4.4 Proof of Proposition 7

In this section, we prove Proposition 7, which provides a characterization of the optimal generation  $g^*(d)$ . The proof uses the Karush-Kuhn-Tucker conditions to prove the optimality of the proposed solution of the DC-OPF problem (5).

*Proof.* Let  $\gamma = (0, \gamma_2, \dots, \gamma_n)$ , with  $e^T \gamma = 1$ ,  $\gamma > 0$ . Fix  $\hat{\varepsilon} > 0$  small enough. Using Lemma 10, for  $\hat{d} := e_1 + \hat{\varepsilon}\gamma$ , we have that the optimal solution of (5) is given by  $g^*(\hat{d}) = \hat{A}_{I^*(\hat{d})}\hat{d}$ , with  $\hat{A}_{I^*(\hat{d})} := \hat{A}_{I^*}$ . Note that  $\hat{\varepsilon}$  is fixed, therefore  $\hat{A}_{I^*}$  depends only on  $\gamma$ .

The main idea behind the proof is to use the Karush-Kuhn-Tucker (KKT) conditions associated with  $g^*(\hat{d})$  to establish that  $\hat{A}_{I^*}d$  is the optimal solution of the optimization problem (5) for any  $d(\varepsilon, \gamma)$  with  $\varepsilon < \hat{\varepsilon}$ . The optimality of  $\hat{A}_{I^*}d$  is ensured by finding suitable Lagrange multipliers that satisfy the corresponding KKT conditions.

We begin by proving that strong duality holds for our optimization problem using Slater's condition [47]. Consider a solution  $g = d$ , where each node generates its own demand. This solution satisfies the equality constraint (5b) because  $e^T g = e^T d$ . It remains to show that, for this choice of  $g$ , all inequality constraints (5c) are not tight. Let  $l \in \mathcal{E}$  and suppose that  $(Vd)_l > 0$ . Then,

$$(1 - \lambda)(Vd)_l < (Vd)_l = (Vg)_l < (1 + \lambda)(Vd)_l.$$

Thus, the constraints corresponding to edge  $l$  are not tight. Furthermore, if  $(Vd)_l = 0$ , then  $(1 - \lambda)(Vd)_l = (1 + \lambda)(Vd)_l = 0$ , so the inequality constraints together form an equality constraint. Hence, by Slater's condition, we conclude that strong duality holds.

To prove that  $g^*(d)$  is the optimal solution to the optimization problem with input  $d$ , we need to find suitable Lagrange multipliers  $\mu, \nu$  and  $\delta$  that satisfy the following KKT conditions:

$$g^*(d) + V^T(\nu - \mu) + \delta e = 0, \quad (18)$$

$$\mu \geq 0, \nu \geq 0, \quad (19)$$

$$\mu_i \left( (1 - \lambda)Vd - Vg^*(d) \right)_i = 0, \quad i = 1, \dots, m, \quad (20)$$

$$\nu_i \left( (1 + \lambda)Vd - Vg^*(d) \right)_i = 0, \quad i = 1, \dots, m, \quad (21)$$

$$(1 - \lambda)Vd \leq Vg^*(d) \leq (1 + \lambda)Vd, \quad (22)$$

$$e^T g^*(d) = e^T d. \quad (23)$$

Generation  $g^*(\hat{d})$  is the optimal solution of (5) for input  $\hat{d}$ . Therefore, there exists a set of Lagrange multipliers  $\hat{\mu} \geq 0$ ,  $\hat{\nu} \geq 0$  and  $\hat{\delta}$  that satisfy Conditions (18) – (23) for  $d = \hat{d}$ . In what follows, we use these Lagrange multipliers, to show that  $\hat{A}_{I^*}d$  is the optimal solution of (5) for every  $d = e_1 + \varepsilon\gamma$  with  $\varepsilon = \alpha\hat{\varepsilon}$  and  $\alpha \in (0, 1)$  by demonstrating that  $\hat{A}_{I^*}d$  satisfies all KKT conditions (18) – (23).

Condition (23): Equation (23) applied to  $\hat{d}$  gives us

$$e^T \hat{a}_1 + e^T \hat{A}_{I^*}(\hat{\varepsilon}\gamma) = e^T \hat{A}_{I^*}\hat{d} = e^T g^*(\hat{d}) \stackrel{(23)}{=} e^T \hat{d} = 1 + \hat{\varepsilon},$$

where  $\hat{a}_1 := \hat{A}_{I^*}e_1$ . From Lemma 6, we know that  $\hat{a}_1 = (1 - \lambda)e_1 + \frac{\lambda}{n}e$ . We obtain that  $e^T \hat{a}_1 = 1$ ; hence  $e^T \hat{A}_{I^*}\gamma = 1$ . Consequently,

$$e^T g^*(d) = e^T \hat{A}_{I^*}d := e^T \hat{a}_1 + e^T \hat{A}_{I^*}(\varepsilon\gamma) = 1 + \varepsilon = e^T d.$$

This shows that Condition (23) is satisfied for demand vector  $d$ .

Condition (22): Next, we verify that Condition (22) holds. Equation (22) applied to  $\hat{d}$  gives us

$$(1 - \lambda)V\hat{d} \leq Vg^*(\hat{d}) \leq (1 + \lambda)V\hat{d}.$$

For the expression in the middle, we find

$$Vg^*(\hat{d}) = V\hat{A}_{I^*}(e_1 + \hat{\varepsilon}\gamma) = V(\hat{a}_1 + \hat{\varepsilon}\hat{A}_{I^*}\gamma) = V((1 - \lambda)e_1 + \frac{\lambda}{n}e + \hat{\varepsilon}\hat{A}_{I^*}\gamma) = V((1 - \lambda)e_1 + \hat{\varepsilon}\hat{A}_{I^*}\gamma),$$

where in the last line follows from the property  $Ve = 0$  of the PTDF matrix. Therefore, Equation (22) becomes

$$(1 - \lambda)V(e_1 + \hat{\varepsilon}\gamma) \leq V((1 - \lambda)e_1 + \hat{\varepsilon}\hat{A}_{I^*}\gamma) \leq (1 + \lambda)V(e_1 + \hat{\varepsilon}\gamma).$$

We first focus on the left-hand side inequality, which, after simplification, yields

$$(1 - \lambda)V(\hat{\varepsilon}\gamma) \leq V(\hat{\varepsilon}\hat{A}_{I^*}\gamma).$$

Multiplying both sides by  $\varepsilon/\hat{\varepsilon}$  and adding  $(1 - \lambda)Ve_1$  yields the left inequality of (22) for demand  $d$ . Rewriting the right-hand side inequality gives us

$$-\lambda Ve_1 + \hat{\varepsilon}V\hat{A}_{I^*}\gamma \leq \lambda Ve_1 + (1 + \lambda)\hat{\varepsilon}V\gamma,$$

and hence,

$$\hat{\varepsilon}(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma) \leq 2\lambda Ve_1.$$

If  $(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma)_i \leq 0$  for some  $i = 1, \dots, m$ , then inequality  $\varepsilon(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma)_i \leq 0 \leq 2\lambda(Ve_1)_i$  is satisfied irrespective of the value of  $\varepsilon$ , since  $Ve_1 \geq 0$ . On the other hand, if  $(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma)_i \geq 0$ , then

$$\varepsilon(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma)_i \leq \hat{\varepsilon}(V\hat{A}_{I^*}\gamma - (1 + \lambda)V\gamma)_i \leq 2\lambda(Ve_1)_i.$$

Hence, Condition (22) is satisfied for demand  $d$ .

For Conditions (18) – (21), we choose Lagrange multipliers  $\nu = (1 - \alpha)\tilde{\nu} + \alpha\hat{\nu}$ ,  $\mu = (1 - \alpha)\tilde{\mu} + \alpha\hat{\mu}$  and  $\delta = \alpha\hat{\delta} + (1 - \alpha)\tilde{\delta}$ , where  $\tilde{\nu}$ ,  $\tilde{\mu}$  and  $\tilde{\delta}$  are the Lagrange multipliers corresponding to the optimization problem with  $d = e_1$ , for which the optimal solution is known. In particular,  $\tilde{\delta} := -\frac{1}{n}$ ,  $\tilde{\nu} := 0$  and  $\tilde{\mu} := (1 - \lambda)Ce_1$  [38, Lemma IV.2].

Condition (19): Since  $\tilde{\nu}$ ,  $\tilde{\mu}$ ,  $\hat{\nu}$ ,  $\hat{\mu}$  satisfy (19), by definition,  $\mu$  and  $\nu$  also satisfy Condition (19).

Condition (18): Substituting  $g^*(d) = \hat{a}_1 + \hat{A}_{I^*}(\varepsilon\gamma)$  into Condition (18) yields

$$\hat{a}_1 + \varepsilon\hat{A}_{I^*}\gamma + V^T(\nu - \mu) + \delta e = 0,$$

and hence

$$(1 - \lambda)e_1 + \frac{\lambda}{n}e + \varepsilon\hat{A}_{I^*}\gamma + V^T(\nu - \mu) + \delta e = 0. \quad (24)$$

We know that (24) holds for  $\hat{\mu}$ ,  $\hat{\nu}$ ,  $\hat{\delta}$  and  $\hat{\varepsilon}$ . Moreover, (24) holds for  $\tilde{\mu}$ ,  $\tilde{\nu}$ ,  $\tilde{\delta}$  and  $\tilde{\varepsilon} = 0$ . A weighted sum of these two equalities with weights  $\alpha$  and  $(1 - \alpha)$ , respectively, yields the following equality:

$$\begin{aligned} 0 &= (1 - \lambda)e_1 + \frac{\lambda}{n}e + \varepsilon\hat{A}_{I^*}\gamma + V^T(\alpha\hat{\nu} + (1 - \alpha)\tilde{\nu} - \alpha\hat{\mu} - (1 - \alpha)\tilde{\mu}) + (\alpha\hat{\delta} + (1 - \alpha)\tilde{\delta})e \\ &= (1 - \lambda)e_1 + \frac{\lambda}{n}e + \varepsilon\hat{A}_{I^*}\gamma + V^T(\nu - \mu) + \delta e. \end{aligned} \quad (25)$$

Hence, Condition (18) is satisfied for  $d$ .

Condition (21): To be able to say more about  $\hat{\nu}$ , we investigate Equation (21) for  $\hat{d}$ . It can be rewritten as follows:

$$\hat{\nu}_i \left( (1 + \lambda)Ve_1 + (1 + \lambda)\hat{\varepsilon}V\gamma - (1 - \lambda)Ve_1 - \hat{\varepsilon}V\hat{A}_{I^*}\gamma \right)_i = 0, \quad i = 1, \dots, m.$$

Thus, we have

$$\hat{\nu}_i \left( 2\lambda Ve_1 + (1 + \lambda)\hat{\varepsilon}V\gamma - \hat{\varepsilon}V\hat{A}_{I^*}\gamma \right)_i = 0, \quad i = 1, \dots, m.$$

Suppose that  $(Ve_1)_i > 0$ . Then, since  $\hat{\varepsilon}$  is small enough, Lemma 12 yields  $\left(2\lambda Ve_1 + (1 + \lambda)\hat{\varepsilon}V\gamma - \hat{\varepsilon}V\hat{A}_{I^*}\gamma\right)_i > 0$ , which implies that  $\hat{\nu}_i = 0$ . In addition, we know that  $\tilde{\nu}_i = 0$ , which yields that  $\nu_i = 0$ .

Now, suppose that  $(Ve_1)_i = 0$ . If  $\hat{\nu}_i = 0$ , then again,  $\nu_i = 0$ . On the other hand, if  $\hat{\nu}_i \neq 0$ , then

$$2\lambda(Ve_1)_i + (1 + \lambda)\hat{\varepsilon}(V\gamma)_i - \hat{\varepsilon}(V\hat{A}_{I^*}\gamma)_i = (1 + \lambda)\hat{\varepsilon}(V\gamma)_i - \hat{\varepsilon}(V\hat{A}_{I^*}\gamma)_i = 0.$$

Multiplying the above equation with  $\alpha$  yields

$$2\lambda(Ve_1)_i + (1 + \lambda)\varepsilon(V\gamma)_i - \varepsilon(V\hat{A}_{I^*}\gamma)_i = 0,$$

which means that the  $i$ -th equation of (21) is satisfied. Altogether, this implies that Equation (21) holds for all  $i$ .

**Condition (20):** Clearly, if  $\mu_i = 0$ , then the constraint is satisfied for  $i$ . Therefore, suppose that  $\mu_i > 0$ . This means that either  $\tilde{\mu}_i > 0$  or  $\hat{\mu}_i > 0$  (or both). We know that  $\tilde{\mu} = Ce_1$  with  $C$  the incidence matrix of graph  $\mathcal{G}$ . Therefore,  $\tilde{\mu}_i > 0$  if and only if  $i \in \mathcal{E}_1$ . Then by Lemma 11, Equation (20) holds for such a choice of  $i$ .

It remains to consider cases when  $\hat{\mu}_i > 0$ ,  $i \neq \mathcal{E}_1$ . In this case, Equation (20) for  $\hat{\varepsilon}$  gives us

$$(1 - \lambda)Ve_1 + (1 - \lambda)\hat{\varepsilon}V\gamma = (1 - \lambda)Ve_1 + \hat{\varepsilon}V\hat{A}_{I^*}\gamma.$$

Hence,

$$(1 - \lambda)\hat{\varepsilon}V\gamma = \hat{\varepsilon}V\hat{A}_{I^*}\gamma.$$

Multiplying the above by  $\alpha$  and adding factor  $(1 - \lambda)Ve_1$  gives us

$$(1 - \lambda)Ve_1 + (1 - \lambda)\varepsilon V\gamma = (1 - \lambda)Ve_1 + \varepsilon V\hat{A}_{I^*}\gamma,$$

so (20) holds for this choice of  $i$  as well.

Since all KKT conditions (18)–(23) are satisfied, we conclude that  $g^*(d) = \hat{A}_{I^*}d$  for all  $\varepsilon > 0$  sufficiently small. Lastly, we set  $\phi(\gamma) := I^*(\hat{d})$ , which completes the proof.  $\square$

#### 4.5 Proof of Lemma 8

The proof of Lemma 8 relies on the continuity of the demand and generation with respect to  $\varepsilon > 0$  at every step of the cascade. After showing the continuity, we use Taylor's expansion theory to show that the edges maximizing the relative flow exceedance at every step of the cascade are the same for all  $\varepsilon > 0$  sufficiently small.

*Proof.* Assume that the node with the largest demand has label  $i$ , that edge  $l$  fails first, and fix  $\gamma \in \Gamma$ . To prove the lemma, we recursively show that the cascade sequence is independent of  $\varepsilon$  in the neighborhood of 0.

Consider the graph after the first step of the cascade. Due to Proposition 7, we know that the initial demand vector  $d$  and the optimal generation vector  $g^*(d)$  are affine functions of  $\varepsilon$  for all  $\varepsilon > 0$  small, and hence, they are continuously differentiable. Next, we argue that  $d^{(2)}$  and  $g^{(2)}$  are continuous with respect to  $\varepsilon$ . Consider a connected component  $\mathcal{C} = (\mathcal{N}_{\mathcal{C}}, \mathcal{E}_{\mathcal{C}})$  of  $\mathcal{G}^{(2)}$ . Equation (8) implies that for  $\theta^{(2)} = \sum_{i \in \mathcal{N}_{\mathcal{C}}} d_i / \sum_{i \in \mathcal{N}_{\mathcal{C}}} g_i^*(d)$ :

- (i) if there is excess of demand in  $\mathcal{C}$ , i.e.  $\theta^{(2)} > 1$ , then  $d_i^{(2)} = d_i / \theta^{(2)}$  and  $g_i^{(2)} = g_i^*(d)$  for  $i \in \mathcal{N}_{\mathcal{C}}$ ,
- (ii) if there is excess of generation in  $\mathcal{C}$ , i.e.  $\theta^{(2)} < 1$ , then  $d_i^{(2)} = d_i$  and  $g_i^{(2)} = \theta^{(2)} g_i^*(d)$  for  $i \in \mathcal{N}_{\mathcal{C}}$ ,
- (iii) if there is a balance of generation and demand in  $\mathcal{C}$ , i.e.  $\theta^{(2)} = 1$ , then  $d_i^{(2)} = d_i$  and  $g_i^{(2)} = g_i^*(d)$  for  $i \in \mathcal{N}_{\mathcal{C}}$ .

Due to the linearity of the initial demand and generation with respect to  $\varepsilon$ , it follows that  $\theta^{(2)}$  is continuous in  $\varepsilon$ . Hence, we conclude that  $d^{(2)}$  and  $g^{(2)}$  are also continuous in  $\varepsilon$ .

Next, we argue that the maximizers of the relative exceedance  $\psi^{(2)}$  are independent of  $\varepsilon$ . First, we observe that for any  $e \in \mathcal{E}^{(1)}$ , the relative exceedance  $\psi_e^{(2)}$  is a rational function  $\frac{P(\varepsilon)}{Q(\varepsilon)}$  where both  $P$  and  $Q$  are some polynomials in  $\varepsilon$  of degree  $k = 2$ . This follows from the fact that both flows and emergency flow limits are linear functions of demand and generation and from the above definition of the demand and generation. Given this, we consider two cases:

Case 1: There is only one edge  $e^*$  that maximizes the exceedance at  $\varepsilon = 0$ . Then, due to continuity of rational functions on their domain,  $e^*$  is the exceedance maximizer for all  $\varepsilon > 0$  sufficiently small.

Case 2: There is a set of edges  $\mathcal{E}^* \subset \mathcal{E}$ , with  $|\mathcal{E}^*| \geq 2$  such that each  $e^* \in \mathcal{E}^*$  maximizes the exceedance at  $\varepsilon = 0$ . In this case, we need to compare the speed of growth of *unique* functions  $\psi_{e^*}^{(2)}$  at  $\varepsilon = 0$ . This is possible by comparing the derivatives, since rational functions are infinitely often differentiable. If the first derivative test is inconclusive, namely, there are multiple functions  $\psi_{e^*}^{(2)}$  that maximize  $\psi_{e^*}^{(2)}$  at  $\varepsilon = 0$ , we continue the test with higher order derivatives until the test is conclusive. We emphasise that there could be multiple edges  $e^*$  that have the exact same exceedance for all  $\varepsilon \geq 0$  sufficiently small, therefore, we only compare the derivatives of *unique* functions  $\psi_{e^*}^{(2)}$ . As the result of the derivative test, we obtain some function  $\psi^*$ , which has the fastest growth amongst functions  $\psi_{e^*}^{(2)}$ . Hence, the set of exceedance maximizers is unique for all  $\varepsilon > 0$  sufficiently small and given by  $\{e^* \in \mathcal{E}^* : \psi_{e^*}^{(2)} = \psi^*\}$ .

Since in both cases, the set of exceedance maximizers is independent of  $\varepsilon$ , we conclude that the graph after the second cascade step,  $\mathcal{G}^{(3)}$ , is also independent of  $\varepsilon$  in this setting.

Finally, consider the graph after  $r$  cascade steps for some  $r \in \mathbb{N}$ . We can use the same line of arguments to argue that if  $\mathcal{G}^{(r)}$  is independent of  $\varepsilon$ , then the resulting graph  $\mathcal{G}^{(r+1)}$  is also independent of  $\varepsilon$ . Hence, by induction, it follows that in a sufficiently small neighborhood of  $\varepsilon = 0$ , the cascade sequence and the resulting graph  $\mathcal{G}^{(\text{end})}$  does not depend on  $\varepsilon$ . □

#### 4.6 Proof of Lemma 9

In order to prove the asymptotic convergence of  $\mathbb{P}(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l)$ , we first rewrite the expression in a suitable manner and then construct upper and lower bounds. Consequently, we show that both bounds are equal in the limit of  $x \rightarrow \infty$ .

*Proof.* First, we observe that the distribution of  $Y$  is independent of the first edge failure  $L^{(1)} = l$ . Moreover, because  $X_i$ 's are independent and identically distributed, conditioning on the label of the node with the largest demand  $N = i$  does not change the distribution of  $Y$ . Therefore, the conditioning only influences the set  $\mathcal{Y}(i, l, z, I^*)$ . For simplicity of notation, we write  $\mathcal{Y} = \mathcal{Y}(i, l, z, I^*)$ , as the dependence on  $i, l, z$ , and  $I^*$  is not exploited. In the proof, it is only relevant that the initial set of demands  $Y$  is in *some* well-defined subset  $\mathcal{Y}$ , but the shape of this subset, specified through variables  $i, l, z$ , and  $I^*$  does not play a role. It becomes relevant in the proof of Theorem 2, where we study the result of cascade as a function of the initial demand.

We observe that we can write the event  $\{Y_1 > x\}$  as a union of disjoint events  $\{X_1 > x, X_2 < X_1, \dots, X_n < X_1\} \cup \dots \cup \{X_n > x, X_1 < X_n, \dots, X_{n-1} < X_n\}$ , where we implicitly use the fact that the probability of two continuous random variables being equal is 0. Using this, we write

$$\begin{aligned} & \mathbb{P}(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l) \\ &= \sum_{k=1}^n \mathbb{P}(X_k > x, X_1 \leq X_k, \dots, X_n \leq X_k, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l) \quad (X_k \text{ is the maximum}) \\ &= \sum_{k=1}^n \mathbb{P}(X_1 > x, X_2 \leq X_1, \dots, X_n \leq X_1, \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l) \quad (X_k \text{'s are i.i.d.}) \\ &= n \mathbb{P}(X_1 > x, X_2 \leq X_1, \dots, X_n \leq X_1, \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l). \end{aligned}$$

For  $\varepsilon < 1$  we have that  $\{X_1 \leq X_k, \dots, X_n \leq X_k\} \supset \{\sum_{j=1, j \neq k}^n X_j < \varepsilon X_k\}$  because  $X_j$ 's take positive values. This means that we can omit the requirements  $X_2 \leq X_1, \dots, X_n \leq X_1$  in the expression above without changing the probability of the event in question. In particular, we obtain

$$\mathbb{P}(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l) = n \mathbb{P}(X_1 > x, \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l).$$

Our goal is to derive upper and lower bounds for this probability, converging to the same limit as  $x \rightarrow \infty$ , under appropriate scaling. To this end, we write the event  $\{\sum_{j=2}^n X_j \leq \varepsilon X_1\}$  as a union of two disjoint events, namely  $\{\sum_{j=2}^n X_j \leq \varepsilon x\} \cup \{\varepsilon x < \sum_{j=2}^n X_j \leq \varepsilon X_1\}$ . Then,

$$\begin{aligned} \mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) &= n\mathbb{P}\left(X_1 > x, \sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) \\ &\quad + n\mathbb{P}\left(X_1 > x, \varepsilon x < \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right). \end{aligned} \quad (26)$$

When event  $\{X_1 > x, \sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\}$  is considered,  $X_1$  is the maximum of  $X_1, \dots, X_n$ , and therefore  $Y = X$ . This means that  $\{Y \in \mathcal{Y} | N = i, L^{(1)} = l\}$  does not depend on  $X_1$  in this particular setting. Hence, using the independence of  $X_i$ 's, we obtain

$$\mathbb{P}\left(X_1 > x, \sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) = \mathbb{P}(X_1 > x) \mathbb{P}\left(\sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right). \quad (27)$$

Upper and lower bounds are obtained by approximating the second summand of (26), i.e.,

$$\mathbb{P}\left(X_1 > x, \varepsilon x < \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right). \quad (28)$$

For a lower bound, we approximate (28) with 0, and for an upper bound with  $\mathbb{P}\left(X_1 > x, \varepsilon x < \sum_{j=2}^n X_j\right)$ . Clearly, this is an upper bound as we remove the condition on  $Y$  and  $X_1$ . Then, leveraging the independence of  $X_i$ 's we obtain

$$0 \leq \mathbb{P}\left(X_1 > x, \varepsilon x < \sum_{j=2}^n X_j \leq \varepsilon X_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) \leq \mathbb{P}(X_1 > x) \mathbb{P}\left(\varepsilon x < \sum_{j=2}^n X_j\right). \quad (29)$$

Using Equations (26), (27), and (29), altogether we obtain that

$$\begin{aligned} &n\mathbb{P}(X_1 > x) \mathbb{P}\left(\sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) \\ &\leq \mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) \\ &\leq n\mathbb{P}(X_1 > x) \left( \mathbb{P}\left(\sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) + \mathbb{P}\left(\varepsilon x < \sum_{j=2}^n X_j\right) \right). \end{aligned}$$

We observe that  $\lim_{x \rightarrow \infty} \mathbb{P}\left(\sum_{j=2}^n X_j \leq \varepsilon x, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) = \mathbb{P}(Y \in \mathcal{Y} | N = i, L^{(1)} = l)$ . Note that this probability is positive, by the condition of the lemma. Moreover, as  $X_i$ 's are i.i.d. Pareto-tailed,  $\mathbb{P}(\sum_{j=2}^n X_j > \varepsilon x) = \mathcal{O}(x^{-\alpha}) = o(1)$ . Hence, both the lower and upper bound of  $\mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) / (n\mathbb{P}(X_1 > x))$  converge to  $\mathbb{P}(Y \in \mathcal{Y} | N = i, L^{(1)} = l) > 0$  as  $x \rightarrow \infty$ . In other words,

$$\mathbb{P}\left(Y_1 > x, \sum_{j=2}^n Y_j \leq \varepsilon Y_1, Y \in \mathcal{Y} | N = i, L^{(1)} = l\right) \sim n\mathbb{P}(X_1 > x) \mathbb{P}(Y \in \mathcal{Y} | N = i, L^{(1)} = l).$$

□

## 5 Uniqueness of flow exceedance

In this section, we discuss the uniqueness of flow exceedance at an arbitrary step of the cascade. Our main goal is to show that there exist graphs for which, with positive probability, the maximizers of the flow exceedance are not

always unique. To this end, we first identify a set of conditions that need to be satisfied for two edges to have the same flow exceedance. Then, we present an example of a graph that satisfies the aforementioned conditions and show that there exists a cascade leading to non-unique flow exceedance maximizers. This implies that the assumptions of [38, Theorem IV.12] are not merely technical, but essential, and thus our theorem offers a relevant generalization. In addition, the example demonstrates that a tie-breaking rule, as given in Definition 1, is crucial to ensure that the cascade process is well-defined.

Consider a graph  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  and demand  $d(\varepsilon, \gamma)$  as defined in (10). We assume that  $d$  meets the conditions of Proposition 7, which means that the optimal generation is given by

$$g^*(d) = A_{\phi(\gamma)}d.$$

We consider the setting after  $r$  edge failures. For  $l \in \mathcal{E}^{(r+1)}$  let  $v_l := e_l^T V$  and  $v_l^{(r+1)} := e_l^T V^{(r+1)}$  denote the  $l$ -th row of  $V$  and  $V^{(r+1)}$ , respectively. For simplicity, we assume that  $\mathcal{G}^{(r+1)}$  is still a connected graph. In such a case, there is no demand and generation reduction in the network so  $g^{(r+1)}(d) = g^*(d)$  and  $d^{(r+1)} = d$ . In this setting, we derive the following lemma, which provides necessary and sufficient conditions for equal flow exceedance of two edges for all  $\varepsilon \geq 0$ .

**Lemma 14.** *Suppose that  $\mathcal{G}^{(r+1)}$  is a connected graph. Edges  $j$  and  $k \in \mathcal{E}^{(r+1)}$  have equal flow exceedance for all  $\varepsilon \geq 0$  sufficiently small if and only if the following equations are simultaneously satisfied*

$$\begin{cases} e_1^T Q^{(r+1)} (A_{\phi(\gamma)} - I) e_1 = 0, & (30) \\ e_1^T [Q^{(r+1)} (A_{\phi(\gamma)} - I) + (A_{\phi(\gamma)} - I)^T (Q^{(r+1)})^T] \gamma = 0, & (31) \\ \gamma^T Q^{(r+1)} (A_{\phi(\gamma)} - I) \gamma = 0, & (32) \end{cases}$$

where

$$Q^{(r+1)} := \begin{cases} v_j^T v_k^{(r+1)} - v_k^T v_j^{(r+1)} & \text{if } f_k^{(r+1)}(d) \cdot f_j^{(r+1)}(d) \geq 0, \\ v_j^T v_k^{(r+1)} + v_k^T v_j^{(r+1)} & \text{otherwise,} \end{cases}$$

and  $f^{(r+1)}(d)$  is the flow vector corresponding to demand  $d$  after  $r$  edge failures, i.e.  $f^{(r+1)}(d) := V^{(r+1)}(d^{(r+1)} - g^{(r+1)})$ .

To prove the lemma, we equate the flow exceedances of two arbitrary edges and write it as a polynomial in  $\varepsilon$ . We conclude that this polynomial is equal to 0 for all  $\varepsilon \geq 0$  if and only if all of its coefficients are equal to 0.

*Proof.* Suppose that  $j, k \in \mathcal{E}^{(r+1)}$  have the same relative flow exceedance for all  $\varepsilon \geq 0$  sufficiently small. The relative flow exceedance of an arbitrary edge  $i \in \mathcal{E}^{(r+1)}$  after  $r$  edge failures is given by

$$\psi_i^{(r+1)}(d) = \frac{|f_i^{(r+1)}(d)|}{F_i(d)}.$$

Rewriting the expression for  $\psi_i^{(r+1)}$ , we find

$$\psi_i^{(r+1)}(d) = \frac{|f_i^{(r+1)}(d)|}{F_i(d)} = \frac{|(V^{(r+1)}(g^{(r+1)}(d) - d^{(r+1)}))_i|}{\lambda^*(Vd)_i} = \frac{|v_i^{(r+1)}(A_{\phi(\gamma)} - I)(e_1 + \varepsilon\gamma)|}{\lambda^* v_i(e_1 + \varepsilon\gamma)}.$$

Edges  $j$  and  $k$  have the same exceedance if and only if  $\psi_j^{(r+1)}(d) = \psi_k^{(r+1)}(d)$ .

First, we consider the case when  $f_k^{(r+1)}(d) \cdot f_j^{(r+1)}(d) \geq 0$ . For  $\psi_k^{(r+1)}(d) = \psi_j^{(r+1)}(d)$  to be equal, it must hold that

$$\begin{aligned} & \left( v_k^{(r+1)} A_{\phi(\gamma)} e_1 + v_k^{(r+1)} A_{\phi(\gamma)} \gamma \varepsilon - v_k^{(r+1)} e_1 - v_k^{(r+1)} \gamma \varepsilon \right) \left( v_j (e_1 + \varepsilon\gamma) \right) \\ &= \left( v_j^{(r+1)} A_{\phi(\gamma)} e_1 + v_j^{(r+1)} A_{\phi(\gamma)} \gamma \varepsilon - v_j^{(r+1)} e_1 - v_j^{(r+1)} \gamma \varepsilon \right) \left( v_k (e_1 + \varepsilon\gamma) \right). \end{aligned} \quad (33)$$

We rewrite Equation (33) as a polynomial in  $\varepsilon$  and obtain

$$\begin{aligned} 0 &= e_1^T \left( v_j^T v_k^{(r+1)} - v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) e_1 + e_1^T \left( v_j^T v_k^{(r+1)} - v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) \gamma \varepsilon \\ &+ \gamma^T \left( v_j^T v_k^{(r+1)} - v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) e_1 \varepsilon + \gamma^T \left( v_j^T v_k^{(r+1)} - v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) \gamma \varepsilon^2. \end{aligned}$$

If  $f_k^{(r+1)}(d) \cdot f_j^{(r+1)} \leq 0$ , we use the same approach and obtain the following equation:

$$0 = e_1^T \left( v_j^T v_k^{(r+1)} + v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) e_1 + e_1^T \left( v_j^T v_k^{(r+1)} + v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) \gamma \varepsilon \\ + \gamma^T \left( v_j^T v_k^{(r+1)} + v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) e_1 \varepsilon + \gamma^T \left( v_j^T v_k^{(r+1)} + v_k^T v_j^{(r+1)} \right) (A_{\phi(\gamma)} - I) \gamma \varepsilon^2.$$

Then, using the definition of  $Q^{(r+1)}$ , in both cases we obtain

$$0 = e_1^T Q^{(r+1)} (A_{\phi(\gamma)} - I) e_1 + \left( e_1^T Q^{(r+1)} (A_{\phi(\gamma)} - I) \gamma + \gamma^T Q^{(r+1)} (A_{\phi(\gamma)} - I) e_1 \right) \varepsilon \\ + \gamma^T Q^{(r+1)} (A_{\phi(\gamma)} - I) \gamma \varepsilon^2 \\ = e_1^T Q^{(r+1)} (A_{\phi(\gamma)} - I) e_1 + e_1^T \left[ Q^{(r+1)} (A_{\phi(\gamma)} - I) + (A_{\phi(\gamma)} - I)^T \left( Q^{(r+1)} \right)^T \right] \gamma \varepsilon \\ + \gamma^T Q^{(r+1)} (A_{\phi(\gamma)} - I) \gamma \varepsilon^2.$$

Edges  $j$  and  $k$  have the same relative exceedance for all  $\varepsilon$  small enough if and only if the coefficients at  $\varepsilon^0$ ,  $\varepsilon^1$  and  $\varepsilon^2$  are equal to 0. Employing this argument yields the result.  $\square$

Note that Lemma 14 can be generalized in a relatively straightforward manner to hold in the cases of network disconnections. Under this scenario, it is necessary to account for the possible reduction in demand and generation at the given step of the cascade. This would only require adaptation of matrix  $A_{\phi(\gamma)} - I$  in order to take into account the scaling of the demand and generation, but the structure of the proof remains the same.

The necessary and sufficient conditions in Lemma 14 can be difficult to verify. Therefore, in Corollary 15 we show that a sufficient condition for non-uniqueness is the skew-symmetry of  $Q^{(r+1)}(A_I - I)$ . This is a convenient, easily-verifiable condition and can be used to find graph instances that exhibit non-uniqueness.

**Corollary 15.** *Let  $A \in \mathcal{A}$  and let  $\Gamma(A) = \{\gamma : g^*(d) = Ad \text{ with } d = e_1 + \varepsilon\gamma\}$ . If for edges  $j, k \in \mathcal{E}^{(r+1)}$ ,  $Q^{(r+1)}(A - I)$  is skew-symmetric, i.e.*

$$\left( Q^{(r+1)}(A - I) \right)^T = -Q^{(r+1)}(A - I),$$

*then the relative flow exceedances at edges  $j$  and  $k$  after  $r$  edge failures are equal for all  $\varepsilon \geq 0$  sufficiently small and all  $\gamma \in \Gamma(A)$ , i.e.*

$$\psi_j^{(r+1)}(d) = \psi_k^{(r+1)}(d) \text{ with } d = e_1 + \varepsilon\gamma.$$

The proof of this corollary follows from Lemma 14 and properties of skew-symmetric matrices.

*Proof.* Let  $j, k \in \mathcal{E}^{(r+1)}$ ,  $\gamma \in \Gamma(A)$ , and suppose that  $Q^{(r+1)}(A - I)$  is a skew-symmetric matrix. A well-known result for skew-symmetric matrices shows that  $w^T Q^{(r+1)}(A - I)w = 0$  for all  $w \in \mathbb{R}^n$ . If we choose  $w = e_1$ , then skew-symmetry of  $Q^{(r+1)}(A - I)$  implies that  $e_1^T Q^{(r+1)}(A - I)e_1 = 0$ , and therefore Equation (30) holds. Similarly, for  $w = \gamma$  skew-symmetry implies that  $\gamma^T Q^{(r+1)}(A - I)\gamma = 0$ , hence Equation (32) holds. Finally, since  $Q^{(r+1)}(A - I)$  is skew-symmetric, we have that  $(A - I)^T \left( Q^{(r+1)} \right)^T = \left( Q^{(r+1)}(A - I) \right)^T = -Q^{(r+1)}(A - I)$  and therefore Equation (31) holds as well. Hence, by Lemma 14, we conclude that  $\psi_j^{(r+1)}(d) = \psi_k^{(r+1)}(d)$  with  $d = e_1 + \varepsilon\gamma$ .

Since  $\gamma$  was chosen arbitrarily, we conclude that the relative flow exceedance on edges  $j$  and  $k$  are equal for all  $\gamma \in \Gamma(A)$  and  $\varepsilon \geq 0$  small enough.  $\square$

It turns out that, indeed, there exist graphs that yield a non-unique cascade sequence with positive probability. In what follows, we present and discuss an example of such a graph instance. However, before we can do so, we need to introduce an additional lemma. The lemma shows that for certain choices of  $d$ , matrix  $A_{I^*(d)}$  has a closed-form solution, which we exploit in Example 1.

**Lemma 16.** *If  $Cd \geq 0$  for some demand vector  $d \in \mathbb{R}_{\geq 0}^n$ , then the optimal solution  $g^*(d)$  to problem (5) is given by*

$$g^*(d) := (1 - \lambda)d + \lambda \bar{d}e.$$

*Note that the optimal solution can be also written as the following matrix equation:*

$$g^*(d) = \left( (1 - \lambda)I + \frac{\lambda}{n}J \right) d.$$



The proof can be found in Appendix A.4. It uses similar ideas as used in the proof of Proposition 7. It also bears similarities with the proof of [38, Lemma IV.2.]. With this, we now discuss the graph example.

**Example 1.** Consider the graph on 6 nodes, depicted in Figure 1, where node 1 has the largest demand.

We consider demands  $d = e_1 + \varepsilon\gamma$ , with

$$\gamma \in \tilde{S} := \{\gamma \in \mathbb{R}_{\geq 0}^6 : e^T \gamma = 1, \gamma_1 = 0, \gamma_3 > \gamma_4 \geq \gamma_5 \geq \gamma_2 \geq \gamma_6, 3\gamma_3 > 4\gamma_4 - \gamma_5\}, \quad (34)$$

and we choose the graph orientation such that  $Cd \geq 0$  for  $\varepsilon$  sufficiently small. For these demands,  $g^*(d) = (1 - \lambda)d + \lambda \bar{d}e$ , which follows from Lemma 16. It can be shown that failure of edge (7) initiates cascade sequence (7)  $\rightarrow$  (11)  $\rightarrow$  (10), after which edges (5), (6), and (9) are the maximizers of the exceedance at this point of the cascade.

Let  $V$  denote the original PTDF matrix of the graph and let  $V^{(4)}$  denote the PTDF matrix of the graph after the failure of edges (7), (10), and (11). We find that

$$V = \frac{1}{30} \begin{pmatrix} 7 & -5 & 1 & 1 & 1 & -5 \\ 6 & 0 & -6 & 0 & 0 & 0 \\ 6 & 0 & 0 & -6 & 0 & 0 \\ 6 & 0 & 0 & 0 & -6 & 0 \\ 1 & -5 & 7 & 1 & 1 & -5 \\ 1 & -5 & 1 & 7 & 1 & -5 \\ 1 & -5 & 1 & 1 & 7 & -5 \\ 5 & 5 & 5 & 5 & 5 & -25 \\ 0 & 0 & 6 & -6 & 0 & 0 \\ 0 & 0 & 6 & 0 & -6 & 0 \\ 0 & 0 & 0 & 6 & -6 & 0 \end{pmatrix} \quad \text{and} \quad V^{(4)} = \frac{1}{24} \begin{pmatrix} 6 & -6 & 0 & 0 & 6 & -6 \\ 5 & -1 & -7 & -1 & 5 & -1 \\ 5 & -1 & -1 & -7 & 5 & -1 \\ 4 & 4 & 4 & 4 & -20 & 4 \\ 1 & -5 & 7 & 1 & 1 & -5 \\ 1 & -5 & 1 & 7 & 1 & -5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 4 & 4 & 4 & 4 & -20 \\ 0 & 0 & 6 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Next, we use Corollary 15 to show that after the failure of edges (7), (10), and (11), the exceedance on edges (5), (6), and (9) is equal for all  $\gamma \in \tilde{S}$ .

Let  $v_i = e_i^T V$  and  $v_i^{(4)} = e_i^T V^{(4)}$  for  $i = 1, \dots, 11$ . We observe that  $v_i^{(4)} = \frac{30}{24}v_i = \frac{5}{4}v_i$  for  $i = 5, 6, 9$ . Moreover, we notice that the flow on edges (5), (6), and (9) is positive for all  $\varepsilon$  small enough. Hence, let  $Q^{(4)} = v_i^T v_j^{(4)} - v_j^T v_i^{(4)}$  for  $i \neq j, i, j \in \{5, 6, 9\}$ . Then,

$$Q^{(4)} = v_i^T v_j^{(4)} - v_j^T v_i^{(4)} = \frac{5}{4}v_i^T v_j - \frac{5}{4}v_j^T v_i,$$

$$\left(Q^{(4)}\right)^T = \frac{5}{4}v_j^T v_i - \frac{5}{4}v_i^T v_j = -Q^{(4)}.$$

This implies that  $Q^{(4)}$  is skew-symmetric. To establish non-uniqueness, it suffices to show that  $Q^{(4)}(A_I - I)$  is skew-symmetric, where  $A_I = (1 - \lambda)I + \frac{\lambda}{n}J$ . We find that

$$Q^{(4)}(A_I - I) = -\lambda Q^{(4)} + \frac{\lambda}{n}Q^{(4)}J = -\lambda Q^{(4)},$$

since  $e \in \text{Ker}(Q^{(4)})$ . Finally, as  $Q^{(4)}$  is skew-symmetric, it follows that  $Q^{(4)}(A_I - I) = -\lambda Q^{(4)}$  is also skew-symmetric. This shows the non-uniqueness of flow exceedances in our example.

To summarize, we have shown that if edges (7), (10), and (11) fail due to a cascade sequence, then the exceedance on edges (5), (6), and (9) are equal for all  $\gamma \in \tilde{S}$ . Hence, it remains to show that cascade sequence (7)  $\rightarrow$  (11)  $\rightarrow$  (10), indeed takes place for a certain subset of  $\gamma$ 's and that edges (5), (6), and (9) are the maximizers of the exceedance at this point of the cascade. An interested reader will find the step-by-step computations in Appendix B.  $\diamond$

The example shows that the tie-breaking rule is necessary to fully define the cascade progression in certain graphs. Therefore, in the following section, we discuss the impact of the tie-breaking rule on the total cascade size.

## 6 Impact of the tie-breaking rule

In this section, we discuss how the choice of the tie-breaking rule influences our results. We consider the theoretical impact from the perspective of Theorem 2 as well as simulation results. Finally, we show that in graphs where the tie-breaking rule does not play a role because the exceedance maximizer is unique at every step of the cascade, Theorem 2 and [38, Theorem IV.12.] are equivalent.

Theorem 2 holds true for any tie-breaking rule  $T$  given by Definition 1. Therefore, it is natural to investigate the influence of the choice of a tie-breaking rule on the total cascade size. In the occurrence of non-uniqueness, the tie-breaking rule plays a pivotal role as it determines the cascade sequence, which, in turn, can impact the number of nodes  $z \in \mathcal{Z}$  disconnected from the node with the largest demand. Consequently, the choice of the tie-breaking rule only affects the constant  $C$ . Hence, in general, the scale-free behavior of the total failure size prevails, independent of the choice of the rule. The only exception arises when the choice of the rule influences the network's connectivity after the cascade. In such a case, depending on the tie-breaking rule, one would observe that the tail either exhibits a scale-free behavior with parameter  $\alpha$  or it scales as  $\mathcal{O}(x^{-2\alpha})$ .

Despite expectations coming from theoretical results, suggesting that a tie-breaking rule plays a role in the total cascade size, practical evidence seems to contradict this notion. Through a simulation experiment, we studied all possible connected graph instances up to and including 7 nodes. We found that even though the choice of the tie-breaking rule leads to different cascades, the resulting total failure size remained consistent across all choices of the tie-breaking rule. In other words, the post-cascade connected components were not affected by the tie-breaking rule. We further extended our simulation experiment to randomly chosen graph instances with more than 7 nodes and, again, we found no evidence suggesting that a tie-breaking rule influences the total cascade size in the asymptotic regime. Based on this, we state the following conjecture:

**Conjecture 17.** *The result of Theorem 2 holds independent of the choice of the exceedance tie-breaking rule.*

Based on our simulation results, we believe in the validity of this conjecture. However, proving it remains a challenge for future research. Nonetheless, our experiments provide some insight into the plausible reasons behind this observed independence. We have seen that non-uniqueness typically occurs in locally dense graphs, whereby the failure of a non-unique maximizer does not lead to network disconnection. While in theory, due to flow redistribution any edge could become the next exceedance maximizer, in practice, we observe that it is the remaining maximizers that become additionally loaded and their relative exceedance continues to be the largest compared to all other edges. Thus, these remaining maximizers break in subsequent steps, creating the same connected components at the end of the cascade for each tie-breaking rule. An example of this behavior can be observed for the graph instance presented in Example 1.

Lastly, we consider a scenario in which the choice of a tie-breaking rule does not affect the cascade sequence in the neighborhood of  $\varepsilon = 0$ . In other words, for all  $\varepsilon \geq 0$  sufficiently small, all exceedance maximizers are unique. Note that in this scenario, [38, Theorem IV.12.] holds. Since our main result generalizes this theorem, both theorems should yield the same result. In the following lemma, we show that this is indeed true.

**Lemma 18.** *Let  $\mathcal{G} = (\mathcal{N}, \mathcal{E})$  and  $\lambda \in (0, 1)$  satisfy the following conditions:*

1. *For all edges  $j$ , the ratios between redistributed flows and edge limits*

$$\frac{f_j^{(r)}(e_1)}{F_j(e_1)} = \frac{\lambda |(V^{(r)}e_1)_j|}{|(Ve_1)_j|}$$

*are all different for all  $r \geq 2$  (Assumption IV.4 of [38]).*

2. *For all edges  $j$  and  $r \geq 2$ ,*

$$|f_j^{(r)}(e_1)| - F_j(e_1) \neq 0.$$

*(Assumption IV.7 of [38]).*

*Then, if there exist  $i \in \mathcal{N}$  and  $l \in \mathcal{E}$  for which the cascade results in network disconnection,*

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

*with*

$$C := \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \frac{K}{m} \left( Z(i, l) \cdot \frac{\lambda}{n} \right)^\alpha,$$

*where  $Z(i, l)$  is the number of cities disconnected from node  $i$  after a cascade initiated by edge  $l$  and demand  $d = e_i$ . Otherwise,  $\mathbb{P}(S > x) = \mathcal{O}(x^{-2\alpha})$ .*

*Proof.* The two assumptions of the theorem ensure that for all  $d = e_1 + \varepsilon\gamma$  and  $\varepsilon > 0$  sufficiently small, the cascade sequence is the same and equal to the cascade sequence under demand  $d = e_1$  [38, Proposition IV.9]. This implies that

for all  $\varepsilon$  small enough, the result of the cascade is the same for all  $\gamma \in \Gamma$ . In particular, for all  $\gamma \in \Gamma$ ,  $Z(i, l, \gamma) = Z(i, l)$ . Hence,

$$\begin{aligned} & \sum_{I^* \in \mathcal{I}} \sum_{z \in \mathcal{Z}} z^\alpha \cdot \mathbb{P}(Y \in Y(i, l, z, I^*) | N = i, L^{(1)} = l) \\ &= Z(i, l)^\alpha \sum_{I^* \in \mathcal{I}} \mathbb{P}(Y \in Y(i, l, Z(i, l), I^*) | N = i, L^{(1)} = l) \\ &= Z(i, l)^\alpha. \end{aligned}$$

Thus, the result of Theorem 2 simplifies to

$$\mathbb{P}(S > x) \sim \sum_{i \in \mathcal{N}} \sum_{l \in \mathcal{E}} \frac{K}{m} \left( Z(i, l) \cdot \frac{\lambda}{n} \right)^\alpha x^{-\alpha}$$

in case of network disconnection.

If there exist no  $i$  and  $l$  that lead to network disconnection when  $\sum_{j=2}^n Y_j \leq \varepsilon Y_1$ , then

$$\mathbb{P}(S > x) = \mathbb{P}(S > x, \sum_{j=2}^n Y_j > \varepsilon Y_1) = \mathcal{O}(x^{-2\alpha}).$$

□

To summarize, the tie-breaking rule allows us to extend the existing results on the tail of the total cascade size to all simple, connected graphs. In theory, the choice of the tie-breaking rule can impact the constant  $C$ , describing the tail-scaling of the total size; however, in all of our simulation experiments, the constant  $C$  was unaffected by the choice of the rule.

## 7 Conclusions

Cascading failures in complex networks can lead to significant disruptions and disturbances across diverse domains. This paper addresses the occurrence of heavy-tailed cascading failures in overload networks, building upon recent developments in the field that suggest that heavy-tailed failures may be induced by the scale-free demand distribution. We emphasize that this is one of many paradigms that may lead to scale-free failures in such networks. Depending on the underlying problem other factors, such as criticality, may still be prominent causes of the observed scale-free behavior. To understand the impact of criticality, it would be necessary to study this problem from the perspective of growing networks. This approach has not been pursued so far and remains a subject for future research.

Our contributions include a novel characterization of optimal nodal generation that arises as the optimal solution to the DC-OPF problem, necessary and sufficient conditions for non-uniqueness of exceedance maximizers, and a theorem on the heavy-tailed nature of total failure size for all simple and connected graph topologies.

In our model, failure propagation is governed by edges maximizing the flow exceedance. When the maximizer is not unique, we choose the next edge to break based on some tie-breaking rule. While the tie-breaking rule theoretically influences failure propagation, our simulation experiment suggests its independence from the total cascade size. Formal proof of this independence requires further investigation.

It is crucial to recognize that the fundamental principles behind the scale-free failure size are versatile and can be applied to various cascading failure scenarios. We anticipate that as long as the heavy-tailed mismatch between generation and demand in a cascading failure model is maintained, the failure size will be heavy-tailed. The development of a universal cascading failure model for diverse scenarios beyond the power networks application would be an interesting future research direction. In addition, investigating the impact of network interdependencies on cascade size is crucial. Many real-world systems exhibit intricate interconnectedness, amplifying failures across networks. Integrating multilayer network theory with our model promises deeper insights into cascading failures within these complex systems.

## Acknowledgments

This work is supported by NWO through Gravitation NETWORKS grant no. 024.002.003.

## References

- [1] Sameer Deshkar and Vinayak Adane. Community resilience approach for prioritizing infrastructure development in urban areas. In *Urban Disasters and Resilience in Asia*, pages 245–267. Butterworth-Heinemann, 2016.
- [2] Magdi S. Mahmoud and Yuanqing Xia. *Networked Control Systems*, chapter Cyberphysical Security Methods, pages 389–456. Butterworth-Heinemann, 2019.
- [3] Hengdao Guo, Ciyang Zheng, Herbert Ho-Ching Iu, and Tyrone Fernando. A critical review of cascading failure analysis and modeling of power system. *Renewable and Sustainable Energy Reviews*, 80:9–22, 2017.
- [4] Mahdi Jalili and Matjaž Perc. Information cascades in complex networks. *Journal of Complex Networks*, 5(5):665–693, Jul 2017.
- [5] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4), Apr 2004.
- [6] Li Daqing, Jiang Yinan, Kang Rui, and Shlomo Havlin. Spatial correlation analysis of cascading failures: Congestions and blackouts. *Scientific Reports*, 4(1), Jun 2014.
- [7] Liang Tang, Ke Jing, Jie He, and H. Eugene Stanley. Complex interdependent supply chain networks: Cascading failure and robustness. *Physica A: Statistical Mechanics and its Applications*, 443:58–69, 2016.
- [8] Xuqing Huang, Irena Vodenska, Shlomo Havlin, and H. Eugene Stanley. Cascading failures in bi-partite graphs: Model for systemic risk propagation. *Scientific Reports*, 3(1), Feb 2013.
- [9] Duncan J. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, Apr 2002.
- [10] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.
- [11] Jie Chen, James S. Thorp, Manu Parashar, et al. Analysis of electric power system disturbance data. In *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, volume 1, page 13, 2001.
- [12] Limiao Zhang, Guanwen Zeng, Daqing Li, Hai-Jun Huang, H. Eugene Stanley, and Shlomo Havlin. Scale-free resilience of real traffic jams. *Proceedings of the National Academy of Sciences*, 116(18):8673–8678, Apr 2019.
- [13] Marc Manzano, Eusebi Calle, Jordi Ripoll, Anna M. Fagertun, Victor Torres-Padrosa, Sakshi Pahwa, and Caterina Scoglio. Epidemic and cascading survivability of complex networks. In *2014 6th International Workshop on Reliable Networks Design and Modeling (RNDM)*, pages 187–193, 2014.
- [14] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(6), May 2001.
- [15] Marc Manzano, Eusebi Calle, Jordi Ripoll, Anna Fagertun, and Víctor Torres-Padrosa. Epidemic survivability: Characterizing networks under epidemic-like failure propagation scenarios. In *Proceedings of 9th International Conference on the Design of Reliable Communication Networks, DRCN 2013*, pages 95–102, Jan 2013.
- [16] Mina Youssef, Robert Kooij, and Caterina Scoglio. Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics. *Journal of Computational Science*, 2(3):286–298, 2011. Social Computational Systems.
- [17] Lucas D. Valdez, Louis Shekhtman, Cristian E. La Rocca, Xin Zhang, Sergey V. Buldyrev, Paul A. Trunfio, Lidia A. Braunstein, and Shlomo Havlin. Cascading failures in complex networks. *Journal of Complex Networks*, 8(2), May 2020.
- [18] Adilson E. Motter and Ying-Cheng Lai. Cascade-based attacks on complex networks. *Physical Review E*, 66:065102, Dec 2002.
- [19] Remco van der Hofstad. *Random Graphs and Complex Networks*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2016.
- [20] Vito Latora and Massimo Marchiori. Efficient behavior of small-world networks. *Physical Review Letters*, 87:198701, Oct 2001.
- [21] Ying-Cheng Lai, Adilson E. Motter, and Takashi Nishikawa. *Attacks and Cascades in Complex Networks*, pages 299–310. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.
- [22] Wang Jian-Wei and Rong Li-Li. Effect attack on scale-free networks due to cascading failures. *Chinese Physics Letters*, 25(10):3826–3829, Oct 2008.
- [23] Jian wei Wang, Li li Rong, Liang Zhang, and Zhongzhi Zhang. Attack vulnerability of scale-free networks due to cascading failures. *Physica A: Statistical Mechanics and its Applications*, 387(26):6671–6678, 2008.

- [24] Jian-Wei Wang and Li-Li Rong. A model for cascading failures in scale-free networks with a breakdown probability. *Physica A: Statistical Mechanics and its Applications*, 388(7):1289–1298, 2009.
- [25] Jian-Wei Wang and Li-Li Rong. Cascade-based attack vulnerability on the us power grid. *Safety Science*, 47(10):1332–1336, 2009.
- [26] Per Bak, Chao Tang, and Kurt Wiesenfeld. Self-organized criticality: An explanation of the  $1/f$  noise. *Physical Review Letters*, 59:381–384, Jul 1987.
- [27] Barbara Drossel and Franz Schwabl. Self-organized critical forest-fire model. *Physical Review Letters*, 69:1629–1632, Sep 1992.
- [28] John M. Beggs and Dietmar Plenz. Neuronal avalanches in neocortical circuits. *The Journal of Neuroscience*, 23(35):11167–11177, Dec 2003.
- [29] Zeev Olami, Hans Jacob S. Feder, and Kim Christensen. Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters*, 68:1244–1247, Feb 1992.
- [30] Sergi Valverde and Ricard V. Solé. Self-organized critical traffic in parallel computer networks. *Physica A: Statistical Mechanics and its Applications*, 312(3):636–648, 2002.
- [31] Henrik J. Jensen. *Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems*. Cambridge Lecture Notes in Physics. Cambridge University Press, 1998.
- [32] Benjamin A. Carreras, David E. Newman, Ian Dobson, and A. Bruce Poole. Initial evidence for self-organized criticality in electric power system blackouts. In *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2000.
- [33] Dusko P. Nedic, Ian Dobson, Daniel S. Kirschen, Benjamin A. Carreras, and Vickie E. Lynch. Criticality in a cascading failure blackout model. *International Journal of Electrical Power & Energy Systems*, 28(9):627–633, 2006. Selection of Papers from 15th Power Systems Computation Conference, 2005.
- [34] Benjamin Carreras, Vickie E. Lynch, Ian Dobson, and David E. Newman. Critical points and transitions in an electric power transmission model for cascading failure blackouts. *Chaos*, 12:985–994, Jan 2003.
- [35] Brian Stott, Jorge Jardim, and Ongun Alsac. DC power flow revisited. *IEEE Transactions on Power Systems*, 24(3):1290–1300, 2009.
- [36] Ian Dobson, Benjamin A. Carreras, and David E. Newman. A probabilistic loading-dependent model of cascading failure and possible implications for blackouts. In *Proceedings on th 36th Annual Hawaii International Conference on System Sciences*, 2003.
- [37] Nicholas W. Watkins, Gunnar Pruessner, Sandra C. Chapman, Norma B. Crosby, and Henrik J. Jensen. 25 years of self-organized criticality: Concepts and controversies. *Space Science Reviews*, 198(1–4):3–44, May 2015.
- [38] Tommaso Nesti, Fiona Sloothak, and Bert Zwart. Emergence of scale-free blackout sizes in power grids. *Physical Review Letters*, 125(5), Jul 2020.
- [39] Kenneth T. Rosen and Mitchel Resnick. The size distribution of cities: An examination of the pareto law and primacy. *Journal of Urban Economics*, 8(2):165–186, 1980.
- [40] Martin J. Beckmann. City hierarchies and the distribution of city size. *Economic Development and Cultural Change*, 6(3):243–248, Apr 1958.
- [41] Brian J.L. Berry. City size distributions and economic development. *Economic Development and Cultural Change*, 9(4, Part 1):573–588, Jul 1961.
- [42] Daniel Bienstock. *Electrical Transmission System Cascades and Vulnerability*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.
- [43] Krishnamurthy Dvijotham and Daniel K. Molzahn. Error bounds on the DC power flow approximation: A convex relaxation approach. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 2411–2418, 2016.
- [44] Henrik Ronellenfitsch, Marc Timme, and Dirk Witthaut. A dual method for computing power transfer distribution factors. *IEEE Transactions on Power Systems*, 32(2):1007–1015, 2017.
- [45] Hossein Seifi and Mohammad Sadegh Sepasian. *Electric power system planning: Issues, algorithms and solutions*. Springer, Berlin, Germany, 2011 edition, 2011.
- [46] Jayakrishnan Nair, Adam Wierman, and Bert Zwart. *The Fundamentals of Heavy Tails: Properties, Emergence, and Estimation*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2022.
- [47] Stephen P. Boyd and Lieven Vandenbergh. *Convex optimization*. Cambridge University Press, 2004.

- [48] Carl Meyer. *Matrix Analysis and Applied Linear Algebra*. Society for Industrial and Applied Mathematics, 2000.
- [49] Ján Plesník. Finding the orthogonal projection of a point onto an affine subspace. *Linear Algebra and its Applications*, 422(2):455–470, 2007.
- [50] Branko Grünbaum, Victor Klee, Micha A. Perles, and Geoffrey Colin Shephard. *Convex polytopes*, volume 16. Springer, 1967.
- [51] Petter Tøndel, Tor A. Johansen, and Alberto Bemporad. An algorithm for multi-parametric quadratic programming and explicit MPC solutions. *Automatica*, 39(3):489–497, 2003.

## A Appendix: Remaining proofs

In this section, we provide all remaining proofs of results in this paper.

### A.1 Proof of Lemma 10

The proof of this lemma requires well-known results from the theory of orthogonal projection, which we briefly introduce beforehand. Consider a non-empty affine subspace  $S := \{x \in \mathbb{R}^n : Mx = b\}$ , for some  $M \in \mathbb{R}^{r \times n}$  with rank  $r \leq n$  and  $b \in \mathbb{R}^r$ . An orthogonal projection of point  $z$  onto  $S$  is given by [48, 49]

$$p(z) := \left( I - M^T (MM^T)^{-1} M \right) z + M^T (MM^T)^{-1} b.$$

The projection equation can be generalized to the situations when  $M$  is not a full-rank matrix. In that case, the inverse operator is replaced with the pseudoinverse operator  $+$ . In particular,

$$p(z) := \left( I - M^T (MM^T)^+ M \right) z + M^T (MM^T)^+ b. \quad (35)$$

With this, we are ready to prove the lemma.

*Proof.* We start by showing that  $g^*(d)$  is a unique point on the boundary of  $\mathcal{F}_d$ . The feasible region  $\mathcal{F}_d$  is defined as a finite intersection of hyperplanes, which implies that it is a convex polytope [50]. From the definition of  $\mathcal{F}_d$ , see (11), it follows that the feasible region is closed. Moreover, it is non-empty because  $d \in \mathcal{F}_d$ . Finally,  $\bar{d}e \notin \mathcal{F}_d$ , as it is the solution to the unconstrained DC-OPF problem, and therefore, by construction,  $\bar{d}e$  does not satisfy the operational flow constraints, given in (5c). Specifically, since  $e \in \text{Ker}(V)$ , the constraint  $-\lambda Vd \leq V(d - \bar{d}e) = Vd \leq \lambda Vd$  is violated. Therefore, in light of Lemma 4, we can conclude that the optimal solution  $g^*(d)$  is unique and lies on the boundary of  $\mathcal{F}_d$  [47, Ch. 8].

By the definition given in (12),  $S_{I^*(d)}$  is the affine space on which  $g^*(d)$  lies. Hence, we can use projection theory to derive  $g^*(d)$  as the projection of  $\bar{d}e$  onto  $S_{I^*(d)}$ . Recall that  $S_{I^*(d)} = \bigcap_{i \in I^*(d)} H_i$ , which is equivalently given by

$$S_{I^*(d)} = \{g \in \mathbb{R}^n : M_I g = b_I\}, \text{ with } M_I := \begin{bmatrix} e^T \\ V_{I_1} \\ V_{I_2} \end{bmatrix} \in \mathbb{R}^{(s_1+s_2+1) \times n} \text{ and } b_I := \begin{bmatrix} e^T \\ (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} d \in \mathbb{R}^{s_1+s_2+1}.$$

According to (35), the projection of  $\bar{d}e$  onto  $S_{I^*(d)}$  is given by

$$p_{I^*(d)}(\bar{d}e) := \left( I - M_I^T (M_I M_I^T)^+ M_I \right) \bar{d}e + M_I^T (M_I M_I^T)^+ b_I.$$

Our goal is to show that  $\left( I - M_I^T (M_I M_I^T)^+ M_I \right) \bar{d}e = 0$ , which would imply that  $p_{I^*(d)} = M_I^T (M_I M_I^T)^+ b_I$ . Then, we will rewrite this expression to obtain  $A_{I^*(d)}$ , as given in the lemma.

First, we find that

$$M_I \bar{d}e = \bar{d} \begin{bmatrix} e^T \\ V_{I_1} \\ V_{I_2} \end{bmatrix} e = \bar{d}e^T e e_1 = n \bar{d}e_1$$

because  $\text{Ker}(V) = \langle e \rangle$ . Using this property again, we find that

$$M_I M_I^T = \begin{bmatrix} e^T \\ V_{I_1} \\ V_{I_2} \end{bmatrix} \begin{bmatrix} e & V_{I_1}^T & V_{I_2}^T \end{bmatrix} = \left[ \begin{array}{c|cc} e^T e & e^T [V_{I_1}^T & V_{I_2}^T] \\ \hline [V_{I_1} \\ V_{I_2}] e & [V_{I_1}^T & V_{I_2}^T] \end{array} \right] = \left[ \begin{array}{c|cc} n & & 0 \\ \hline 0 & [V_{I_1}^T & V_{I_2}^T] \end{array} \right].$$

Next, we compute  $(M_I M_I^T)^+ e_1$ . It is a well-known property that

$$\left[ \begin{array}{c|c} D_1 & 0 \\ \hline 0 & D_2 \end{array} \right]^+ = \left[ \begin{array}{c|c} D_1^+ & 0 \\ \hline 0 & D_2^+ \end{array} \right].$$

Therefore, we find

$$(M_I M_I^T)^+ e_1 = \left[ \begin{array}{c|c} \frac{1}{n} & 0 \\ \hline 0 & \left( \left[ \begin{array}{c|c} V_{I_1} & \\ \hline V_{I_2} & \end{array} \right] \left[ \begin{array}{c|c} V_{I_1}^T & V_{I_2}^T \end{array} \right]^+ \right) \end{array} \right]^+ e_1 = \frac{1}{n} e_1.$$

Hence,

$$M_I^T (M_I M_I^T)^+ M_I \bar{d}e = n \bar{d} M_I^T (M_I M_I^T)^+ e_1 = \bar{d} M_I^T e_1 = \bar{d}e.$$

So, altogether, we have found that

$$\left( I - M_I^T (M_I M_I^T)^+ M_I \right) \bar{d}e = 0.$$

Hence,

$$p_{I^*(d)}(\bar{d}e) = M_I^T (M_I M_I^T)^+ \begin{bmatrix} e^T \\ (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} d.$$

Manipulating the above matrix expression yields

$$\begin{aligned} & M_I^T (M_I M_I^T)^+ \begin{bmatrix} e^T \\ (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= \begin{bmatrix} e & V_{I_1}^T & V_{I_2}^T \end{bmatrix} \left[ \begin{array}{c|c} \frac{1}{n} & 0 \\ \hline 0 & \left( \left[ \begin{array}{c|c} V_{I_1} & \\ \hline V_{I_2} & \end{array} \right] \left[ \begin{array}{c|c} V_{I_1}^T & V_{I_2}^T \end{array} \right]^+ \right) \end{array} \right]^+ \begin{bmatrix} e^T \\ (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{n}e & V_{I_1}^T & V_{I_2}^T \end{bmatrix} \left( \left[ \begin{array}{c|c} V_{I_1} & \\ \hline V_{I_2} & \end{array} \right] \left[ \begin{array}{c|c} V_{I_1}^T & V_{I_2}^T \end{array} \right]^+ \right) \begin{bmatrix} e^T \\ (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= \frac{1}{n} J + \begin{bmatrix} V_{I_1}^T & V_{I_2}^T \end{bmatrix} \left( \left[ \begin{array}{c|c} V_{I_1} & \\ \hline V_{I_2} & \end{array} \right] \left[ \begin{array}{c|c} V_{I_1}^T & V_{I_2}^T \end{array} \right]^+ \right) \begin{bmatrix} (1-\lambda)V_{I_1} \\ (1+\lambda)V_{I_2} \end{bmatrix} \\ &= A_{I^*(d)}. \end{aligned}$$

□

## A.2 Proof of Lemma 11

The proof of this lemma relies on a contradiction argument that arises from a construction of an alternative solution to the DC-OPF problem.

*Proof.* Let  $\varepsilon > 0$  be sufficiently small and let  $g^*(d)$  be the optimal solution to the optimization problem (5). In this proof, we argue by contradiction that (15) holds.

For the sake of contradiction, suppose that there exists some  $i \in \mathcal{E}_1$  for which

$$(1-\lambda)(Vd)_i < (Vg^*(d))_i.$$

Let  $x > 0$  denote the total unused capacity on edges in  $\mathcal{E}_1$ , i.e.

$$x := \sum_{i \in \mathcal{E}_1} \lambda(Vd)_i - (V(d - g^*(d)))_i.$$

Recall that the solution to the planning problem is given by  $g^{(pl)}(d) = \frac{1+\varepsilon}{n}e$ . The total flow on edges in  $\mathcal{E}_1$  during the *planning* problem is equal to the unsatisfied demand at node 1, which is  $1 - \frac{1+\varepsilon}{n}$ . Since the *operational* edge limits are set to fraction  $\lambda$  of flows in the *planning* phase, the total *operational* capacity of flow on edges in  $\mathcal{E}_1$  is

equal to  $\lambda(1 - \frac{1+\varepsilon}{n})$ . This means that the generation of node 1 is equal to  $1 - (\lambda(1 - \frac{1+\varepsilon}{n}) - x) = 1 - \lambda + \frac{1+\varepsilon}{n}\lambda + x$ . Because the objective function is a sum of squares of generations at each node, the objective function is minimal if the remaining generation in the network is equally distributed over all other nodes. In particular, each remaining node generates fraction  $\frac{1}{n-1}$  of the remaining generation, which is given by  $(1 - \lambda)\varepsilon + (n - 1)\frac{1+\varepsilon}{n}\lambda - x$ . Therefore, the value of the objective function corresponding to  $g^*(d)$  is at least

$$\frac{1}{2} \sum_{i=1}^n g_i^*(d)^2 \geq \frac{1}{2} (1 - \lambda + \frac{1+\varepsilon}{n}\lambda + x)^2 + \frac{n-1}{2} \left( \frac{1-\lambda}{n-1}\varepsilon + \frac{1+\varepsilon}{n}\lambda - \frac{x}{n-1} \right)^2.$$

Now, consider an alternative solution  $\bar{g}(d) = (1 - \lambda)d + \frac{1+\varepsilon}{n}\lambda e$ . This solution is feasible as  $\sum_{i=1}^n \bar{g}_i(d) = 1 + \varepsilon$  and  $V(d - \bar{g}(d)) = \lambda Vd$ , i.e., the corresponding flow on each edge is equal to the operational capacity limit. We find that the value of the objective function corresponding to this generation is at most

$$\frac{1}{2} \sum_{i=1}^n \bar{g}_i(d)^2 \leq \frac{1}{2} (1 - \lambda + \frac{1+\varepsilon}{n}\lambda)^2 + \frac{n-1}{2} \left( (1 - \lambda)\varepsilon + \frac{1+\varepsilon}{n}\lambda \right)^2.$$

With this, we find that, in the limit  $\varepsilon \downarrow 0$ , the difference in the values of the objective function for  $g^*(d)$  and  $\bar{g}(d)$  is at least

$$\begin{aligned} & \lim_{\varepsilon \downarrow 0} \frac{1}{2} \left( \sum_{i=1}^n g_i^*(d)^2 - \sum_{i=1}^n \bar{g}_i(d)^2 \right) \\ & \geq \frac{1}{2} \left( 1 - \lambda + \frac{\lambda}{n} \right)^2 + \frac{1}{2} x^2 + x \left( 1 - \lambda + \frac{\lambda}{n} \right) + \frac{(n-1)}{2n^2} \lambda^2 \\ & \quad + \frac{x^2}{2(n-1)} - \frac{x}{n} \lambda - \frac{1}{2} \left( 1 - \lambda + \frac{\lambda}{n} \right)^2 - \frac{(n-1)}{2n^2} \lambda^2 \\ & = \frac{1}{2} \left( \frac{n}{n-1} x^2 \right) + (1 - \lambda)x. \end{aligned}$$

We observe that since  $x > 0$ , the limit is positive. Moreover, functions  $\frac{1}{2} \sum_{i=1}^n \bar{g}_i(d)^2$  and  $\frac{1}{2} \sum_{i=1}^n g_i^*(d)^2$  are polynomials in  $\varepsilon$  and therefore continuous. Hence, for all  $\varepsilon$  sufficiently small,  $\frac{1}{2} \sum_{i=1}^n \bar{g}_i(d)^2 < \frac{1}{2} \sum_{i=1}^n g_i^*(d)^2$ . This is a contradiction because we have assumed that  $g^*(d)$  is optimal.  $\square$

### A.3 Proof of Lemma 12

The following lemma is proven using the continuity of the optimal solution of the DC-OPF problem with respect to the demand vector  $d$ .

*Proof.* First, we observe that the inequality is true for the limiting case of  $\varepsilon = 0$ . That is, in this case, the optimal solution is given by  $g^*(e_1) = (1 - \lambda)e_1 + \frac{\lambda}{n}e$  [38]. If  $i$  is such that  $(Ve_1)_i > 0$ , then,

$$(Vg^*(e_1))_i = (1 - \lambda)(Ve_1)_i + \frac{\lambda}{n}(Ve)_i = (1 - \lambda)(Ve_1)_i < (1 + \lambda)(Ve_1)_i.$$

In addition, we know that the solution  $g^*(d)$  of the DC-OPF problem is continuous in  $d$ , as this problem is an example of multi-parametric quadratic programming with a strictly convex objective. For this class of problems, the continuity is a known result [51, Theorem 1]. The continuity implies that there exists some function  $f(\varepsilon, \gamma)$  with  $\lim_{\varepsilon \rightarrow 0} f(\varepsilon, \gamma) = 0$  such that

$$g^*(e_1 + \varepsilon\gamma) = g^*(e_1) + f(\varepsilon, \gamma).$$

Hence, for all  $\varepsilon > 0$  sufficiently small, we have

$$(Vg^*(d))_i = (Vg^*(e_1))_i + (Vf(\varepsilon, \gamma))_i = (1 - \lambda)(Ve_1)_i + (Vf(\varepsilon, \gamma))_i < (1 + \lambda)(Ve_1)_i. \quad \square$$

### A.4 Proof of Lemma 16

The proof of this lemma relies on the KKT conditions and it uses similar ideas as the proof of Proposition 7.



*Proof.* Let  $d \in \mathbb{R}_{\geq 0}^n$  be a demand vector such that  $Cd \geq 0$ . We prove that  $g^*(d) = (1 - \lambda)d + \lambda\bar{d}e$  is the optimal solution by showing that it satisfies the KKT conditions (18) – (23).

Let  $\nu = 0$ ,  $\mu = (1 - \lambda)Cd$  and  $\delta = -\bar{d}$ . This choice of  $\nu$  automatically satisfies (21). Moreover, since  $Cd \geq 0$  and  $\lambda \in (0, 1)$ , we have that  $\mu \geq 0$ . This means that (19) is also satisfied for our choice of  $\mu$ .

Next, we show that  $e^T g^*(d) = e^T d$  (Condition (23)). We find

$$e^T g^*(d) = e^T ((1 - \lambda)d + \lambda\bar{d}e) = (1 - \lambda)n\bar{d} + \lambda n\bar{d} = n\bar{d} = e^T d.$$

Furthermore, we find that both Conditions (22) and (20) hold because

$$Vg^*(d) = (1 - \lambda)Vd + \lambda\bar{d}Ve = (1 - \lambda)Vd.$$

Hence, it remains to show that Condition (18) is satisfied; specifically, that

$$g^*(d) + V^T(\nu - \mu) + \delta e = 0. \quad (36)$$

Using the fact that  $V = CL^+$ ,  $L = C^T C$ ,  $(L^+)^T = L^+$  and  $L^+L = (I - \frac{1}{n}J)$  [38], we find

$$V^T \mu = (1 - \lambda)V^T Cd = (1 - \lambda)(L^+)^T C^T Cd = (1 - \lambda)L^+ Ld = (1 - \lambda)(I - \frac{1}{n}J)d = (1 - \lambda)d - (1 - \lambda)\bar{d}e.$$

We substitute the expressions for  $g^*(d)$ ,  $\mu$ ,  $\nu$  and  $\delta$  into (36). This yields

$$(1 - \lambda)d + \lambda\bar{d}e - (1 - \lambda)d + (1 - \lambda)\bar{d}e - \bar{d}e = 0.$$

This means that Condition (18) holds as well. With this, we conclude that all KKT conditions are satisfied, which proves that  $g^*(d)$  is the optimal solution.  $\square$

## B Appendix: Cascade propagation in Example 1

In this section, we present the step-by-step computations regarding the cascade propagation on the graph example discussed in Section 5. The graph is also visualized in Figure 2.

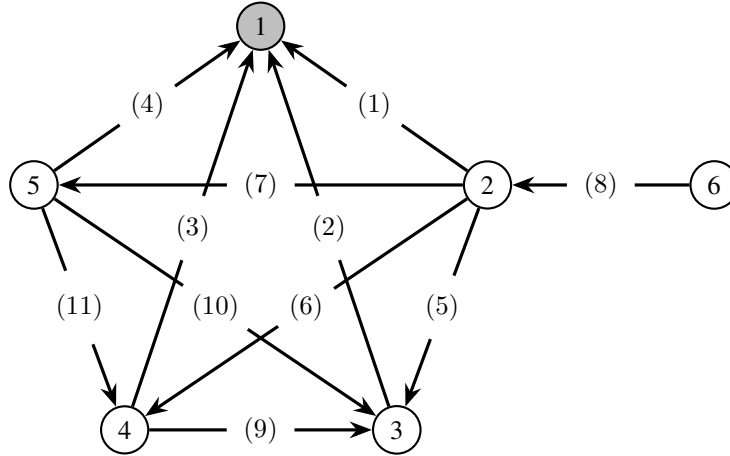


Figure 2: Example of a graph where non-unique relative flow exceedance occurs with positive probability.

Recall that in Section 1, we have shown that after edges (7), (11), and (10) are removed from the graph due to the cascades, the relative flow exceedances on edges (5) (6) and (9) are equal for all  $\gamma \in \tilde{S}$  as defined in (34). Here, we show that the cascade (7)  $\rightarrow$  (11)  $\rightarrow$  (10) happens with positive probability and that edges (5), (6), and (9) are all maximizers of the relative exceedance.

Suppose that the initial failure occurs at edge (7). Then, the corresponding PTDF matrix  $V^{(2)}$  is given by

$$V^{(2)} = \frac{1}{90} \begin{pmatrix} 22 & -20 & 4 & 4 & 10 & -20 \\ 18 & 0 & -18 & 0 & 0 & 0 \\ 18 & 0 & 0 & -18 & 0 & 0 \\ 17 & 5 & -1 & -1 & -25 & 5 \\ 4 & -20 & 22 & 4 & 10 & -20 \\ 4 & -20 & 4 & 22 & 10 & -20 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 15 & 15 & 15 & 15 & 15 & -75 \\ 0 & 0 & 18 & -18 & 0 & 0 \\ -1 & 5 & 17 & -1 & -25 & 5 \\ -1 & 5 & -1 & 17 & -25 & 5 \end{pmatrix}.$$

We observe that rows 2,3,8, and 9 of matrix  $V^{(2)}$  remain unchanged compared to the matrix  $V$ . Therefore, we can immediately conclude that no exceedance occurs on edges (2), (3), (8), and (9). To find the exceedance maximizer, we only need to consider the remaining edges. We obtain the following approximation:

$$\psi_1^{(2)} = \left| \frac{v_1^{(2)}(A_I - I)d}{\lambda^* v_1 d} \right| = \frac{1}{3} \left| \frac{-\lambda(22 + \varepsilon(-20\gamma_2 + 4\gamma_3 + 4\gamma_4 + 10\gamma_5 - 20\gamma_6))}{\lambda^*(7 + \varepsilon(-5\gamma_2 + \gamma_3 + \gamma_4 + \gamma_5 - 5\gamma_5))} \right| \approx \frac{22\lambda}{21\lambda^*},$$

for  $\varepsilon$  small. Note that the approximation is exact for  $\varepsilon = 0$ . Taking the same approach, we find that

$$\psi_4^{(2)} \approx \frac{17\lambda}{18\lambda^*}, \quad \psi_5^{(2)} \approx \frac{4\lambda}{3\lambda^*}, \quad \psi_6^{(2)} \approx \frac{4\lambda}{3\lambda^*}.$$

Finally, we observe that

$$\psi_{10}^{(2)} = \frac{\lambda}{3\lambda^*} \left| \frac{-\frac{1}{\varepsilon} + 5\gamma_2 + 17\gamma_3 - \gamma_4 - 25\gamma_5 + 5\gamma_6}{6\gamma_3 - 6\gamma_5} \right| = \frac{\lambda}{18\lambda^*} \frac{\frac{1}{\varepsilon} - 5\gamma_2 - 17\gamma_3 + \gamma_4 + 25\gamma_5 - 5\gamma_6}{\gamma_3 - \gamma_5},$$

$$\psi_{11}^{(2)} = \frac{\lambda}{3\lambda^*} \left| \frac{-\frac{1}{\varepsilon} + 5\gamma_2 - \gamma_3 + 17\gamma_4 - 25\gamma_5 + 5\gamma_6}{6\gamma_4 - 6\gamma_5} \right| = \frac{\lambda}{18\lambda^*} \frac{\frac{1}{\varepsilon} - 5\gamma_2 + \gamma_3 - 17\gamma_4 + 25\gamma_5 - 5\gamma_6}{\gamma_4 - \gamma_5}.$$

Note that we change the sign of the nominator in these expressions, as for all  $\varepsilon$  small enough, the term  $\frac{1}{\varepsilon}$  dominates the remaining terms and  $\gamma_3 \geq \gamma_4 \geq \gamma_5$  for all  $\gamma \in \tilde{S}$ . With this, we obtain the following approximations:

$$\psi_{10}^{(2)} \approx \frac{\lambda}{18\lambda^* \varepsilon (\gamma_3 - \gamma_5)} \quad \text{and} \quad \psi_{11}^{(2)} \approx \frac{\lambda}{18\lambda^* \varepsilon (\gamma_4 - \gamma_5)}.$$

Based on the above approximations, we can conclude that for all  $\varepsilon$  small enough, either edge (10) or edge (11) has the largest exceedance. Moreover, since  $\gamma_3 > \gamma_4$  it holds that

$$\gamma_3 - \gamma_5 > \gamma_4 - \gamma_5$$

and

$$-17\gamma_3 + \gamma_4 < \gamma_3 - 17\gamma_4.$$

Together, these two conditions lead to  $\psi_{10}^{(1)} < \psi_{11}^{(1)}$ , indicating that edge (11) is the next one to fail in this cascade sequence.

Next, let  $V^{(3)}$  be the PTDF matrix of the graph after the failure of edges (7) and (11). Then,

$$V^{(3)} = \frac{1}{240} \begin{pmatrix} 59 & -55 & 11 & 5 & 35 & -55 \\ 48 & 0 & -48 & 0 & 0 & 0 \\ 49 & -5 & 1 & -65 & 25 & -5 \\ 44 & 20 & -4 & 20 & -100 & 20 \\ 11 & -55 & 59 & 5 & 35 & -55 \\ 10 & -50 & 10 & 70 & 10 & -50 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 40 & 40 & 40 & 40 & 40 & -200 \\ 1 & -5 & 49 & -65 & 25 & -5 \\ -4 & 20 & 44 & 20 & -100 & 20 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Again, we notice that the flow on edges (2) and (8) remains unchanged, resulting in no exceedance occurs on these two edges. For the remaining edges, we take the same approach as before and find the following approximations for the exceedances:

$$\psi_1^{(3)} \approx \frac{59\lambda}{56\lambda^*}, \quad \psi_3^{(3)} \approx \frac{49\lambda}{48\lambda^*}, \quad \psi_4^{(3)} \approx \frac{44\lambda}{48\lambda^*}, \quad \psi_5^{(3)} \approx \frac{11\lambda}{8\lambda^*}, \quad \psi_6^{(3)} \approx \frac{10\lambda}{8\lambda^*},$$

$$\psi_9^{(3)} \approx \frac{\lambda}{48\lambda^*\varepsilon(\gamma_3 - \gamma_4)}, \quad \psi_{10}^{(3)} \approx \frac{4\lambda}{48\lambda^*\varepsilon(\gamma_3 - \gamma_5)}.$$

Clearly, for  $\varepsilon$  small enough, the maximal exceedance is obtained at edge (9) or (10). Since we assume that  $3\gamma_3 > 4\gamma_4 - \gamma_5$ , we obtain that  $4\gamma_3 - 4\gamma_4 > \gamma_3 - \gamma_5$ , from which it directly follows that  $\psi_9^{(3)} < \psi_{10}^{(3)}$ . Hence, the next edge to fail is (10).

To summarize, we have shown that for  $\gamma \in \tilde{S}$ , the edges of the graph will break in the following order:

$$(7) \rightarrow (11) \rightarrow (10).$$

Furthermore, we observe that the  $\tilde{S}$  has a positive Lebesgue measure due to its non-empty interior. For example, consider  $\gamma = (0, \frac{10}{100}, \frac{30}{100}, \frac{28}{100}, \frac{27}{100}, \frac{5}{100})$ , which lies in the interior of  $\tilde{S}$  because  $\gamma_3 > \gamma_4$  and  $3\gamma_3 = \frac{90}{100} > \frac{85}{100} = 4\gamma_4 - \gamma_5$ . Since  $\gamma$  is a realization of  $n-1$  (scaled) i.i.d. Pareto distributions with support equal to  $(0, \infty)$ , the non-empty interior of  $\tilde{S}$  implies that

$$\mathbb{P}(\gamma \in \tilde{S}) > 0.$$

Lastly, we show that after the three edge failures have occurred, the maximizers of the exceedance are edges (5), (6), and (9). By following the same procedure as before, we compute the (approximate) exceedances for each edge and find

$$\psi_1^{(4)} \approx \frac{30\lambda}{28\lambda^*}, \quad \psi_2^{(4)} \approx \frac{25\lambda}{24\lambda^*}, \quad \psi_3^{(4)} \approx \frac{25\lambda}{24\lambda^*}, \quad \psi_4^{(4)} \approx \frac{20\lambda}{24\lambda^*}, \quad \psi_8^{(4)} = \frac{\lambda}{\lambda^*},$$

$$\psi_5^{(4)} = \psi_6^{(4)} = \psi_9^{(4)} = \frac{5\lambda}{4\lambda^*}.$$

We observe that the above fraction is largest for edges (5), (6), and (9). Furthermore, for  $\lambda^* \in [\lambda, \frac{5}{4}\lambda)$ , the exceedance on these edges is larger than 1, implying that these edges should fail next. This proves that the maximizer of the exceedance is non-unique for the provided graph and  $\gamma \in \tilde{S}$ , as we claimed in Example 1.