

# Optimizing the robustness of electrical power systems against cascading failures

Yingrui Zhang\* and Osman Yağan<sup>†</sup>

<sup>1</sup>Department of ECE, Carnegie Mellon University, Pittsburgh, 15213, USA

\*yingruiz@andrew.cmu.edu

<sup>†</sup>oyagan@ece.cmu.edu

## ABSTRACT

Electrical power systems are one of the most important infrastructures that support our society. However, their vulnerabilities have raised great concern recently due to several large-scale blackouts around the world. In this paper, we investigate the robustness of power systems against cascading failures initiated by a *random* attack. This is done under a simple yet useful model based on *global* and *equal* redistribution of load upon failures. We provide a complete understanding of system robustness by i) deriving an expression for the final system size as a function of the size of initial attacks; ii) deriving the critical attack size after which system breaks down completely; iii) showing that complete system breakdown takes place through a first-order (i.e., discontinuous) transition in terms of the attack size; and iv) establishing the optimal load-capacity distribution that maximizes robustness. In particular, we show that robustness is maximized when the difference between the capacity and initial load is the same for all lines; i.e., when all lines have the same redundant space regardless of their initial load. This is in contrast with the intuitive and commonly used setting where capacity of a line is a fixed factor of its initial load.

## 1 Introduction

Electrical power systems are one of the most critical national infrastructures affecting all areas of daily life.<sup>1,2</sup> They also provide crucial support for other national infrastructures such as telecommunications, transportation, water supply systems and emergency services.<sup>3–5</sup> Besides daily life, the destruction of power systems would also weaken or even disable our defense and economic security.<sup>6</sup> Thus, ensuring the robustness of electrical power systems and maintaining the continuous availability of power supply are of utmost priority. Despite its great importance, concerns about the robustness of the power grid have grown recently because of several large-scale outages that took place in different parts of the world. For example, in the 2012 India blackout, 600 million people, nearly a tenth of the world’s population, were left without power.<sup>7</sup> The blackout spread across 22 states in Northern, Eastern, and Northeast India and is the largest power outage in history.<sup>8</sup>

These blackouts often start with natural hazards such as lightning shorting a line or with malicious attacks, and affect only a small portion of the power system initially. But due to the long range nature of electricity, the redistribution of power loads may affect not only geographically co-located lines but also other parts of the system far from the initial affected area. A typical example is the Western Systems Coordinating Council (WSCC) system outage on August 10, 1996,<sup>9</sup> where long range failures have been observed. The large-scale blackouts are often attributed to this initial shock getting escalated due to the intricate dependencies within a power system. For example, when a line is tripped, the flow on all other lines will be updated, and some lines may end up with a total flow (initial plus redistributed after failures) exceeding their capacity. All lines with flows exceeding their capacity will in turn fail and flows on other lines will be updated again, possibly leading to further failures, and so on. This process may continue recursively and lead to a *cascade* of failures, which may potentially breakdown the entire system. For instance, on August 10, 1996, an electrical line sagged in summer heat in Southern Oregon and this initiated a chain reaction that cut power to more than four million people in eleven Western States.<sup>10,11</sup>

Since electrical power systems are among the largest and most complex technological systems ever developed,<sup>12</sup> it is often hard to have a full understanding of their inter- and intra-dependencies and therefore it is hard to predict their behavior under external attacks or random failures. In this work, we aim to shed light on the robustness of power systems using a simple yet useful model. In particular, we assume that when a line fails, its load (i.e., flow) is redistributed *equally* among all other lines. The equal load redistribution model has the ability to capture the *long-range* nature of the Kirchhoff’s law, at least in the mean-field sense, as opposed to the *topological* models where failed load is redistributed only *locally* among neighboring lines.<sup>13,14</sup> This is particularly why this model received recent attention in the context of power systems first in the work by Pahwa et al.<sup>15</sup> and then in Yağan;<sup>16</sup> the model is originally inspired by the *democratic fiber-bundle model*<sup>17</sup> that is used extensively for studying the rupture of fiber-bundles under increasing external force.

Assuming that each of the  $N$  lines in the system is assigned independently an initial load  $L_i$  and a redundant space  $S_i$  – meaning that line capacity equals  $L_i + S_i$  – from a joint distribution  $P_{LS}(x, y) = \mathbb{P}[L \leq x, S \leq y]$ , we study the robustness of this systems against *random* attacks; see Section 2 for the details of our model. In particular, we characterize the *final* fraction  $n_\infty(p)$  of alive (i.e., non-failed) lines at the steady-state, when  $p$ -fraction of the lines are randomly failed. We identify the critical attack size  $p^*$  after which the system breakdowns entirely; i.e.,  $n_\infty(p) = 0$  if  $p > p^*$ . We show that the transition of the system around  $p^*$  is always first-order (i.e., discontinuous). However, depending on the distribution  $P_{LS}$ , this may take place with or without a preceding second-order (i.e., continuous) divergence of  $n_\infty(p)$  from the  $1 - p$  line. Finally, under the constraints that mean load  $\mathbb{E}[L]$  and mean free space  $\mathbb{E}[S]$  are fixed, we show that assigning every line the same free space regardless of its load is *optimal* in the sense of maximizing the robustness  $n_\infty(p)$  for all attack sizes  $p$ . This provably optimal strategy is in sharp contrast with the commonly used<sup>13, 18–22</sup> setting where free-space  $S_i$  is set to be a constant factor of a line’s initial load, e.g.,  $S_i = \alpha L_i$  for all  $i$ . This hints at the possibility that existing power systems are not designed optimally and that their robustness may be significantly improved by reallocating the line capacities (while keeping the total capacity unchanged). Our analytic results are validated via extensive simulations, using both synthetic data for load-capacity values as well as realistic data from IEEE test-case data sets.

We believe that our results provide interesting insights into the dynamics of cascading failures in power systems. In particular, we expect our work to shed some light on the *qualitative* behavior of real-world power systems under random attacks, and help design them in a more robust manner. The results obtained here may have applications in fields other than power systems as well. A particularly interesting application is the study of the traffic jams in roads, where the capacity of a line can be regarded as the traffic flow capacity of a road.<sup>23, 24</sup>

## 2 Model Definitions

**Equal load-redistribution model.** We consider a power system with  $N$  transmission lines  $\mathcal{L}_1, \dots, \mathcal{L}_N$  with initial loads (i.e., power flows)  $L_1, \dots, L_N$ . The *capacity*  $C_i$  of a line  $\mathcal{L}_i$  defines the maximum power flow that it can sustain, and is given by

$$C_i = L_i + S_i, \quad i = 1, \dots, N, \tag{1}$$

where  $S_i$  denotes the *free-space* (or, redundancy) available to line  $\mathcal{L}_i$ . The capacity of a line can be defined as a factor of its initial load, i.e.,

$$C_i = (1 + \alpha_i)L_i \tag{2}$$

with  $\alpha_i > 0$  defining the *tolerance* parameter for line  $\mathcal{L}_i$ . Put differently, the free space  $S_i$  is given in terms of the initial load  $L_i$  as  $S_i = \alpha_i L_i$ ; it is very common<sup>13, 18–20</sup> to use a *fixed* tolerance factor for all lines in the system, i.e., to use  $\alpha_i = \alpha$  for all  $i$ . It is assumed that a line *fails* (i.e., outages) if its load exceeds its capacity at any given time. The key assumption of our model is that when a line fails, the load it was carrying (right before the failure) is redistributed *equally* among all remaining lines.

Throughout we assume that the pairs  $(L_i, S_i)$  are independently and identically distributed with  $P_{LS}(x, y) := \mathbb{P}[L \leq x, S \leq y]$  for each  $i = 1, \dots, N$ . The corresponding (joint) probability density function is given by  $p_{LS}(x, y) = \frac{\partial^2}{\partial x \partial y} P_{LS}(x, y)$ . Throughout, we let  $L_{min}$  and  $S_{min}$  denote the minimum values for load  $L$  and free space  $S$ ; i.e.,

$$L_{min} = \inf\{x : P_L(x) > 0\}$$

$$S_{min} = \inf\{y : P_S(y) > 0\}$$

We assume that  $L_{min}, S_{min} > 0$ . We also assume that the marginal densities  $p_L(x)$  and  $p_S(y)$  are continuous on their support.

The equal load redistribution rule takes its roots from the *democratic* fiber bundle model,<sup>17, 25</sup> where  $N$  parallel fibers with failure thresholds  $C_1, \dots, C_N$  (i.e., capacities) share an applied total force  $F$  *equally*. There, it has been of interest to study the dynamics of recursive failures in the bundle as the applied force  $F$  increases; e.g., see.<sup>26–29</sup> This model has been recently used by Pahwa et al.<sup>30</sup> in the context of power systems, with  $F$  corresponding to the total load shared *equally* by  $N$  power lines. The relevance of the equal load-redistribution model for power systems stems from its ability to capture the *long-range* nature of the Kirchhoff’s law, at least in the mean-field sense, as opposed to the *topological* models where failed load is redistributed only *locally* among neighboring lines.<sup>13, 14</sup> In particular, the equal load redistribution model is expected to be a reasonable assumption under the DC power flow model, which approximates the standard AC power flow model when the phase differences along the branches are small and the bus voltages are fixed;<sup>30</sup> in fact, power flow calculations based on the DC model<sup>31, 32</sup> are known to give accurate results that match the AC model calculations in many cases. Therefore, we expect our work to shed some light on the *qualitative* behavior of real-world power systems under random attacks.

**Problem definition.** Our main goal is to study the robustness of power systems under the equal load redistribution rule. In this work, we assume that failures are initiated by a *random* attack that results with a failure of a  $p$ -fraction of the lines;

of course, all the discussion and accompanying results do hold for the robustness against random *failures* as well. The initial failures lead to redistribution of power flows from the failed lines to *alive* ones (i.e., non-failed lines), so that the load on each alive line becomes equal to its initial load plus its equal share of the total load of the failed lines. This may lead to the failure of some additional lines due to the updated flow exceeding their capacity. This process may continue recursively, generating a *cascade of failures*, with each failure further increasing the load on the alive lines, and may eventually result with the collapse of the entire system. Throughout, we let  $n_\infty(p)$  denote the *final* fraction of alive lines when a  $p$ -fraction of lines is randomly attacked. The *robustness* of a power system will be evaluated by the behavior  $n_\infty(p)$  for all attack sizes  $0 < p < 1$ . Of particular interest is to characterize the *critical* attack size  $p^*$  at which  $n_\infty(p)$  drops to zero.

The problem formulation considered here was introduced by Yağan.<sup>16</sup> This formulation differs from the original democratic fiber-bundle model (and its analog<sup>30</sup> introduced for power systems) in that i) it does not assume that the total load of the system is fixed at  $F$ ; and ii) it allows for power lines to carry different initial loads unlike the democratic fiber bundle model where all lines start with the same initial load. Since power lines in real systems are likely to have different loads at the initial set-up, we believe our formulation is more suitable for studying cascading failures in power systems. In addition,<sup>30</sup> is concerned with failures in the power system that are triggered by increasing the total force (i.e., load) applied. Instead, our formulation allows analyzing the robustness of the system against external attacks or random line failures, which are known to be the source of system-wide blackouts in many interdependent systems.<sup>5,33,34</sup>

A word on notation in use: The random variables (rvs) under consideration are all defined on the same probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . Probabilistic statements are made with respect to this probability measure  $\mathbb{P}$ , and we denote the corresponding expectation operator by  $\mathbb{E}$ . The indicator function of an event  $A$  is denoted by  $\mathbf{1}[A]$ .

## 3 Results

### 3.1 Final system size as a function the attack size, $n_\infty(p)$

Our first main result characterizes the robustness of power systems under any initial load-space distribution  $P_{LS}$  and any attack size  $p$ . Let  $L$  and  $S$  denote generic random variables following the same distribution with initial loads  $L_1, \dots, L_N$ , and free spaces  $S_1, \dots, S_N$ , respectively. Then, with  $x^*$  denoting the smallest solution of

$$h(x) := \mathbb{P}[S > x] (x + \mathbb{E}[L | S > x]) \geq \frac{\mathbb{E}[L]}{1-p} \quad (3)$$

over the range  $x^* \in (0, \infty)$ , the final system size  $n_\infty(p)$  at attack size  $p$  is given by

$$n_\infty(p) = (1-p)\mathbb{P}[S > x^*]. \quad (4)$$

This result, proved in Section 5.1, provides a complete picture about a power system's robustness against random attacks of arbitrary size. In particular, it helps understand the response  $n_\infty(p)$  of the system to attacks of varying magnitude.

For a graphical solution of  $n_\infty(p)$ , one shall plot  $\mathbb{P}[S > x] (x + \mathbb{E}[L | S > x])$  as a function of  $x$  (e.g., see Figure 1a), and draw a horizontal line at the height  $\mathbb{E}[L]/(1-p)$  on the same plot. The leftmost intersection of these two lines gives the operating point  $x^*$ , from which we can compute  $n_\infty(p) = (1-p)\mathbb{P}[S > x^*]$ . When there is no intersection, we set  $x^* = \infty$  and understand that  $n_\infty(p) = 0$ .

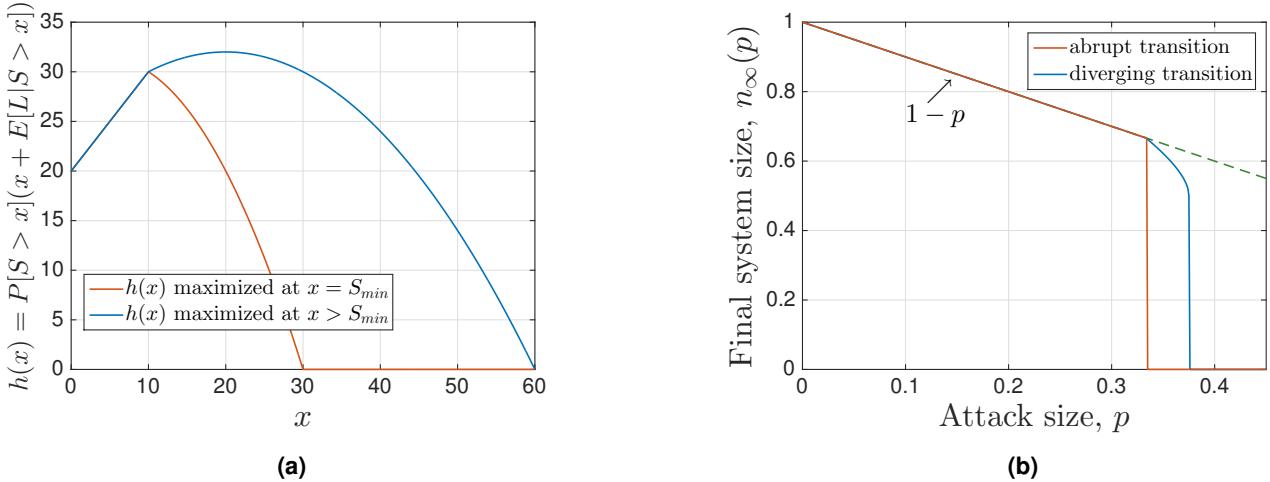
We see from this result that an adversarial attack aimed at a certain part of the electrical power grid may lead to failures in other parts of the system, possibly creating a recursive failure process also known as *cascading failures*. This will often result with a damage in the system much larger than the initial attack size  $p$ . However, in most cases “some” part of the system is expected to continue its functions by undertaking extra load; e.g., with  $n_\infty(p) > 0$ . In such cases, although certain service areas are affected, the power grid remains partially functional. The most severe situations arise when cascading failures continue until the *complete* breakdown of the system where all lines fail; e.g., when  $n_\infty(p) = 0$ . This prompts us to characterize the *critical* attack size  $p^*$ , defined as the largest attack size that the system can sustain.

**The “critical” attack size.** Of particular interest is to derive the *critical* attack size  $p^*$  such that for any attack with size  $p > p^*$ , the system undergoes a complete breakdown leading to  $n_\infty(p) = 0$ ; on the other hand for  $p < p^*$ , we have  $n_\infty(p) > 0$ . More precisely, we define  $p^*$  as

$$p^* = \sup\{p : n_\infty(p) > 0\}.$$

The critical attack size can be derived from the previous results (3)-(4) that characterize  $n_\infty(p)$ . Namely, for any load-free space distribution  $p_{LS}(x, y)$ , the maximum attack size  $p^*$  can be computed from the *global* maximum of the function  $\mathbb{P}[S > x] (x + \mathbb{E}[L | S > x])$ . In particular, we have

$$p^* = 1 - \frac{\mathbb{E}[L]}{\max_x \{\mathbb{P}[S > x] (x + \mathbb{E}[L | S > x])\}}. \quad (5)$$



**Figure 1. Different types of first-order transitions.** We demonstrate the difference between an abrupt first-order transition and a first-order transition with a preceding divergence from the  $1 - p$  line. The lower curves (shown in red) correspond to the case where the load  $L$  and extra space  $S$  are independent and uniformly distributed with  $L_{\min} = S_{\min} = 10$  and  $\mathbb{E}[L] = \mathbb{E}[S] = 20$ . The upper curves (shown in blue) are obtained under the same setting except that we set  $\mathbb{E}[S] = 35$ . We see that the lower curve in Figure 1a reaches its maximum at  $S_{\min} = 10$ , and the corresponding final system size exhibits an abrupt first-order transition as shown in Figure 1b. On the other hand, the upper (i.e., blue) curve in Figure 1a is maximized at  $S = 20 > S_{\min}$ . As expected from our result (e.g., see (8)), the total breakdown of the system takes place after a diverging failure rate is observed.

A proof of this result is given in Section 5.1.

### 3.2 Understanding the “phase transition”: Conditions for abrupt rupture.

It is of significant interest to understand the behavior of the system near the *phase transition*; i.e., when the attack size is very close to but smaller than the critical value  $p^*$ . One main question here is whether  $n_\infty(p)$  decays to zero continuously (i.e., through a second-order transition), or discontinuously (i.e., through a first-order transition). The practical significance of this is that continuous transitions suggest a more stable and predictable system behavior with respect to attacks, whereas with discontinuous transitions system behavior becomes more difficult to predict, for instance, from past data.

Our analysis shows that under the equal-load redistribution model considered here the total breakdown of the system will always be through a first-order (i.e., discontinuous) transition; see Methods for a proof. Namely, we have

$$n_\infty(p^*) > 0, \quad (6)$$

while by definition it holds that  $n_\infty(p^* + \varepsilon) = 0$ , for any  $\varepsilon > 0$  arbitrarily small. This means that regardless of the attack size and the distribution of load and capacity, the transition point where the system has a total breakdown (i.e., where the fraction of alive lines drops to zero) is always discontinuous. These cases are reminiscent of the real-world phenomena of unexpected large-scale system collapses; i.e., cases where seemingly identical attacks/failures leading to entirely different consequences.

Now that we showed that the breakdown of the power system takes place through a first-order transition, an interesting question arises as to whether this first-order rupture at  $p^*$  has any early indicators at smaller attack sizes; e.g., a *diverging* failure rate leading to a non-linear decrease in  $n_\infty(p)$ . Otherwise, an *abrupt* first-order transition is said to take place if the linear decay of  $n_\infty(p)$  (of the form  $1 - p$ ) is followed by a sudden discontinuous jump to zero at  $p^*$ ; i.e., we say that the system exhibits an *abrupt* rupture when it holds that

$$n_\infty(p) = \begin{cases} 1 - p & \text{if } p \leq p^* \\ 0 & \text{if } p > p^* \end{cases} \quad (7)$$

In Figure 1b we demonstrate the distinction between an *abrupt* rupture and a rupture with preceding divergence from the  $1 - p$  line.

We now present our result that reveals the necessary and sufficient condition for an abrupt rupture to take place. We show (in Methods) that the system goes through an abrupt first-order breakdown (e.g., see the below line shown in red in Figure 1b), if and only if the function  $h(x) = \mathbb{P}[S > x](x + \mathbb{E}[L | S > x])$  reaches its maximum at  $x = S_{\min}$ , where  $S_{\min}$  is the minimum value

the extra space  $S$  can take. Namely, an abrupt first-order rupture (*without* a preceding divergence) takes place if and only if

$$\arg \max_{x>0} \{ \mathbb{P}[S > x] (x + \mathbb{E}[L | S > x]) \} = S_{\min}. \quad (8)$$

Otherwise, if  $\arg \max_{x>0} h(x) \neq S_{\min}$ , then a preceding divergence from the  $1 - p$  line will be observed before  $n_\infty(p)$  drops to zero; e.g., see the above line shown in blue in Figure 1b). More precisely, it will hold that  $n_\infty(p) < 1 - p$  for some  $p < p^*$ . A detailed analysis of conditions for these two types of ruptures is presented in Methods.

Figure 1 demonstrates different types of transitions that the system can exhibit in relation to the behavior of  $h(x)$ . In figure 1a, we plot  $h(x)$  in two different cases: the red (i.e., lower) line reaches its maximum at  $S_{\min}$ , while the blue (i.e., upper) line continues to increase after  $S_{\min}$  and reaches its maximum later. Since the function  $h(x)$  represented by the blue line does not satisfy the abrupt rupture condition (8), we see in figure 1b that the corresponding final system size goes through a diverging transition (from the  $1 - p$  line) before entirely breaking down through a first-order transition. On the other hand, we see that  $h(x)$  represented by the red curve reaches its maximum at  $S_{\min}$ . As expected from our results, we see that the corresponding final system size exhibits an abrupt breakdown without any preceding divergence from the  $1 - p$  line.

### 3.3 Achieving optimal robustness

The most important question from a system design perspective is concerned with deriving the *universally optimum* distribution of initial loads  $L_1, \dots, L_N$  and free spaces  $S_1, \dots, S_N$  that leads to *maximum* robustness under the constraints that  $\mathbb{E}[L]$  and  $\mathbb{E}[S]$  are fixed. We believe that the answer to this problem would be very useful in designing real-world power grids with optimum robustness, i.e. with the final system size  $n_\infty(p)$  maximized for any attack size  $p$ . The motivation for the constraints on the mean load  $\mathbb{E}[L]$  and mean free space  $\mathbb{E}[S]$  are as follows. The total load carried by the system is likely to be dictated by system requirements in most real-world cases, which also determines the average load per line. In addition, the total capacity (or, total free space) available to the system is likely to be bounded due to the *costs* associated with using high-capacity lines.

Our results concerning this important problem are presented next. First, we focus on maximizing the critical attack size  $p^*$ . We show in Methods that the critical attack size always satisfies

$$p^* \leq \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mathbb{E}[L]} = \frac{\mathbb{E}[S]}{\mathbb{E}[C]} \quad (9)$$

Namely, regardless of the distribution  $p_{LS}$  that generates load-capacity pairs, the system will always go into a complete breakdown if more than  $\mathbb{E}[S]/\mathbb{E}[C]$ -fraction of lines are attacked; i.e., the system can never sustain a random attack of size exceeding the ratio of mean free space to mean capacity. Next, we show that this critical attack size is in fact attainable *under any load distribution* by a *Dirac* delta distribution for the free-spaces, i.e., by giving every line the same free space. More precisely, let  $p_{dirac}^*$  denote the critical attack size when  $p_{LS}(x,y) = p_L(x)\delta(y - \mathbb{E}[S])$ , where the distribution  $p_L(x)$  of the initial loads  $L_1, \dots, L_N$  is arbitrary. We show in Methods that

$$p_{dirac}^* \geq \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mathbb{E}[L]}.$$

Combined with (9) this shows that assigning every line the same free space (regardless of the initial loads) maximizes the largest attack that the system can sustain.

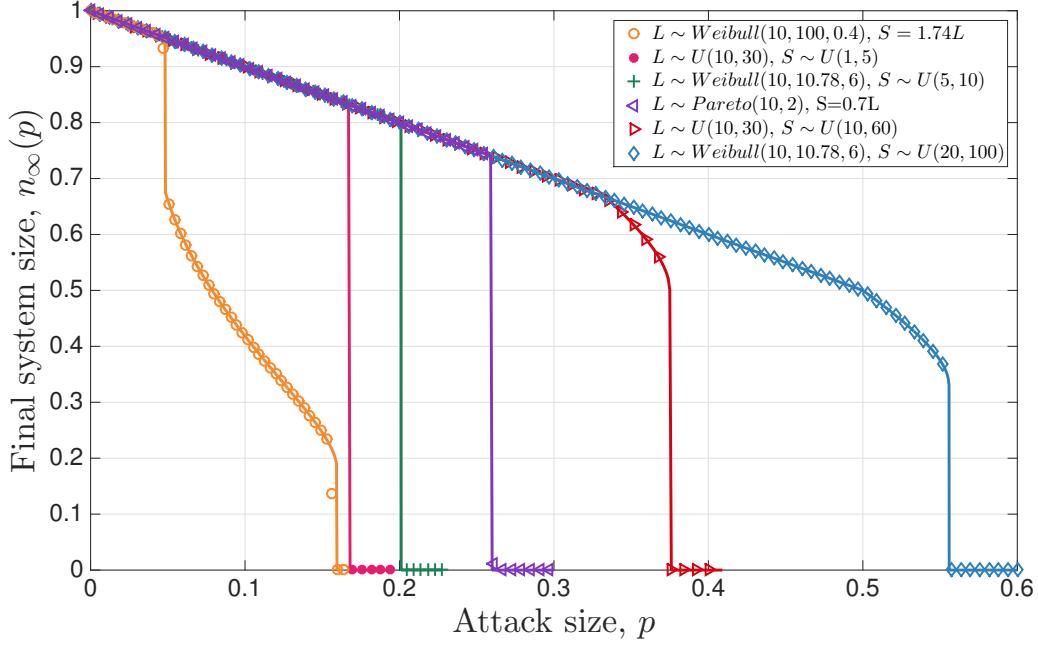
More can be said regarding the optimality of equal free-space allocation. Let  $p_{optimal}^*$  denote the maximum critical attack size as established above, i.e.,  $p_{optimal}^* = \mathbb{E}[S]/(\mathbb{E}[S] + \mathbb{E}[L])$ . In view of the fact that we always have  $n_\infty(p) \leq 1 - p$ , the next result firmly establishes that using the Dirac delta distribution for free space optimizes the robustness of the system uniformly for any attack size  $p$ . In particular, if  $p_{LS}(x,y) = p_L(x)\delta(y - \mathbb{E}[S])$ , then the corresponding final system size  $n_{\infty,dirac}(p)$  satisfies

$$n_{\infty,dirac}(p) = \begin{cases} 1 - p & \text{for } p < p_{optimal}^* \\ 0 & \text{for } p \geq p_{optimal}^* \end{cases} \quad (10)$$

Namely, the distribution  $p_{LS}(x,y) = p_L(x)\delta(y - \mathbb{E}[S])$  maximizes the final system size  $n_\infty(p)$  uniformly for all  $p$ .

This result shows that as far as the *random* attacks are concerned, the system's robustness can be maximized under the constraints of fixed  $\mathbb{E}[L]$  and fixed  $\mathbb{E}[S]$  (and hence fixed  $\mathbb{E}[C]$ ), by giving each line an equal free space  $\mathbb{E}[S]$ , **irrespective of how the initial loads are distributed**. Put differently, the robustness will be maximized by choosing a line's capacity  $C_i$  through  $C_i = L_i + \mathbb{E}[S]$  no matter what its load  $L_i$  is. In view of (2), this then leads to a tolerance factor

$$\alpha_i = \mathbb{E}[S]/L_i \quad (11)$$



**Figure 2.** Final system size under different load-free space distributions. Analytic results (obtained from (3) and (4)) are represented by lines, whereas simulation results (averaged over 200 independent runs) are represented by symbols. We see that in each case theoretical results match the simulation results very well.

meaning that the optimal robustness is achieved when lines with larger initial loads are given smaller tolerance factors. This result is rather counter-intuitive because one might think that lines with large initial loads shall receive extra protection (in the form of larger free space or tolerance factor) given the potentially detrimental effect of their failure to the overall system. Our result prove this intuition incorrect and show that robustness is maximized if all lines share the fixed total free-space equally. In fact, numerical results presented in Section 3.4 show that the standard choice of setting free-space to be a constant factor of initial load (i.e., using the same tolerance factor for all lines) may lead to significantly worse robustness than the optimal choice given at (11).

A possible explanation to this counter-intuitive result is as follows. When all lines have the same extra space, we ensure that the system never goes through a *cascade* of failures. In other words, when  $p$ -fraction of the lines are attacked, we will have either  $n_\infty(p) = 1 - p$  or  $n_\infty(p) = 0$  depending on whether or not, respectively, the total load of failed lines divided by  $1 - p$  is less than the common free space  $S$ . In addition, if the attack size is large enough that total load of failed lines, i.e.,  $p\mathbb{E}[L]$ , is larger than the total free space  $(1 - p)\mathbb{E}[S]$  available in the rest of the system, then regardless of the distribution  $p_{LS}(x, y)$ , the system will collapse. Collectively, these explain why assigning equal free-space to all lines ensures that system will go through an abrupt rupture, but only at the optimal critical attack size  $p_{optimal}^*$ .

The optimality results presented here shed light on the recent findings by Yağan<sup>16</sup> who investigated the model considered here in the special case where  $S_i = \alpha L_i$  for all lines; i.e., the case where  $p_{LS}$  is degenerate with  $p_{LS}(x, y) = p_L(x)\delta(y - x\alpha)$ . There, they found that the optimal robustness is achieved (i.e.,  $n_\infty(p)$  is maximized uniformly across all  $p$ ), if the *load*  $L$  follows a Dirac delta distribution; i.e., the system is most robust when each line carries the same initial load. On the other hand, our results show that the distribution of load has very little to do with optimizing the robustness, and in fact robustness is maximized under *any* initial load distribution if the free-space is distributed equally. This shows that Yağan's result of the optimality of equal load distribution is merely a coincidence. It only arises under the assumption that a line's free space is a constant factor of its load, so that equal allocation of initial loads is equivalent to that of free-space.

### 3.4 Numerical results

We now confirm our theoretical findings via numerical simulations, using both synthetic and real-world data. We focus on the former case first and consider various commonly known distributions for the load and free-space variables.

**Synthetic data.** Throughout, we consider three commonly used families of distributions: i) Uniform, ii) Pareto, and iii) Weibull. The corresponding probability density functions are defined below for a generic random variable  $L$ .

- Uniform Distribution:  $L \sim U(L_{\min}, L_{\max})$ . The density is given by

$$p_L(x) = \frac{1}{L_{\max} - L_{\min}} \cdot \mathbf{1}[L_{\min} \leq x \leq L_{\max}]$$

- Pareto Distribution:  $L \sim \text{Pareto}(L_{\min}, b)$ . With  $L_{\min} > 0$  and  $b > 0$ , the density is given by

$$p_L(x) = L_{\min}^b b x^{-b-1} \mathbf{1}[x \geq L_{\min}].$$

To ensure that  $\mathbb{E}[L] = bL_{\min}/(b-1)$  is finite, we also enforce  $b > 1$ . Distributions belonging to the Pareto family are also known as a *power-law* distributions and have been extensively used in many fields including power systems.

- Weibull Distribution:  $L \sim \text{Weibull}(L_{\min}, \lambda, k)$ . With  $\lambda, k, L_{\min} > 0$ , the density is given by

$$p_L(x) = \frac{k}{\lambda} \left( \frac{x - L_{\min}}{\lambda} \right)^{k-1} e^{-\left(\frac{x-L_{\min}}{\lambda}\right)^k} \mathbf{1}[x \geq L_{\min}].$$

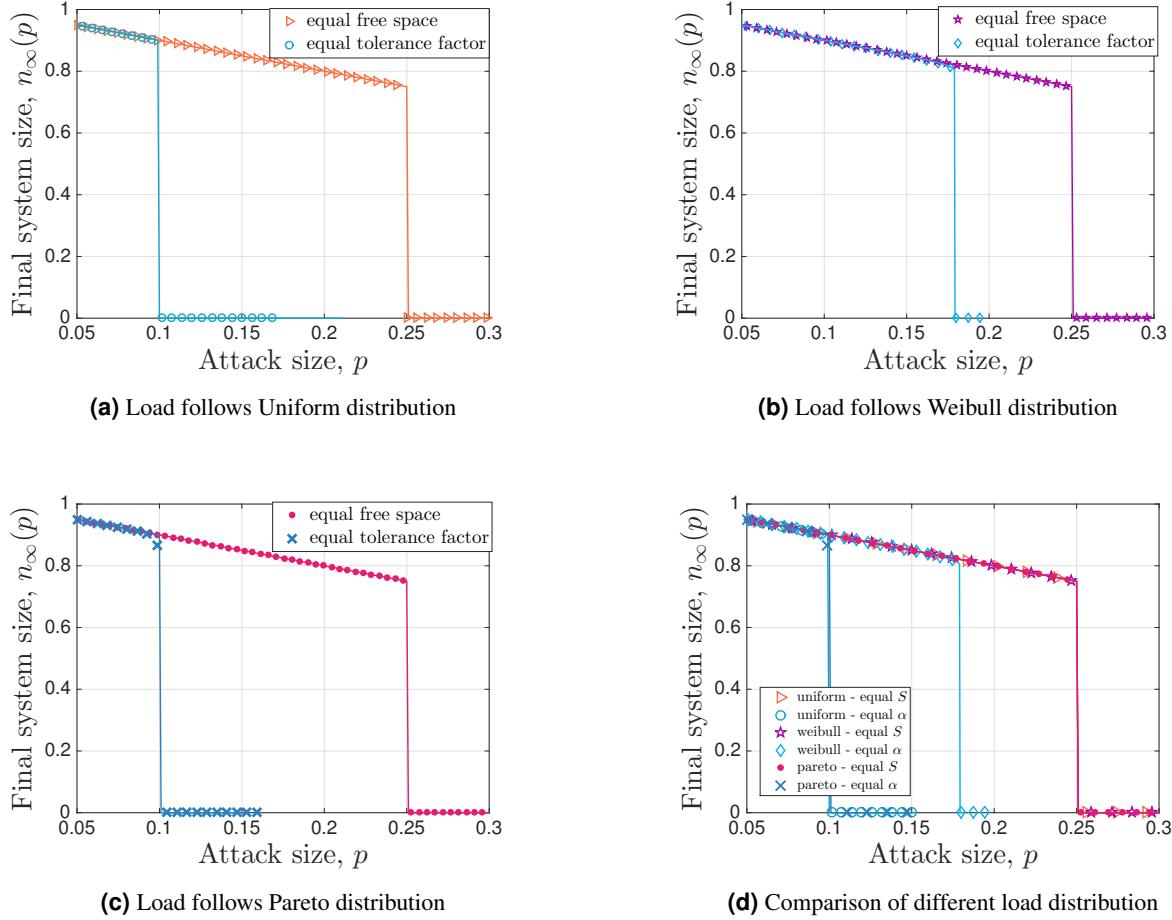
The case  $k = 1$  corresponds to the exponential distribution, and  $k = 2$  corresponds to Rayleigh distribution. The mean load is given by  $\mathbb{E}[L] = L_{\min} + \lambda \Gamma(1+1/k)$ , where  $\Gamma(\cdot)$  is the gamma-function given by  $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ .

First, we confirm our results presented in Sections 3.1 and 3.2 concerning the response of the system to attacks of varying size; i.e. concerning the final system size  $n_\infty(p)$  under different load-extra space distributions including its transition behavior around the critical attack size  $p^*$ . In all simulations, we fix the number of lines at  $N = 10^6$ , and for each set of parameters being considered (e.g., the distribution  $p_{LS}(x, y)$  and attack size  $p$ ) we run 200 independent experiments. The results are shown in Figure 2 where symbols represent the *empirical* value of the final system size  $n_\infty(p)$  (obtained by averaging over 200 independent runs for each data point), and lines represent the analytical results computed from (3) and (4). We see that theoretical results match the simulations very well in all cases.

The specific distributions used in Figure 2 are as follows: From left to right, we have i)  $L$  is Weibull with  $L_{\min} = 10, \lambda = 100, k = 0.4$  and  $S = \alpha L$  with  $\alpha = 1.74$ ; ii)  $L$  is Uniform over [10,30] and  $S$  is Uniform over [1,5]; iii)  $L$  is Weibull with  $L_{\min} = 10, \lambda = 10.78, k = 6$  and  $S$  is Uniform over [5,10]; iv)  $L$  is Pareto with  $L_{\min} = 10, b = 2$ , and  $S = \alpha L$  with  $\alpha = 0.7$ ; v)  $L$  is Uniform over [10,30] and  $S$  is Uniform over [10,60]; and vi)  $L$  is Weibull with  $L_{\min} = 10, \lambda = 10.78, k = 6$  and  $S$  is Uniform over [20,100]. Thus, the plots in Figure 2 demonstrate the effect of the load-free space distribution on the robustness of the resulting power system. We see that both the *family* that the distribution belongs to (e.g., Uniform, Weibull, or Pareto) as well as the specific parameters of the family affect the behavior of  $n_\infty(p)$ . For instance, the curves representing the two cases where  $L$  and  $S$  follow a Uniform distribution demonstrate that both *abrupt* ruptures and ruptures with a preceding divergence are possible in this setting, depending on the parameters  $L_{\min}, L_{\max}, S_{\min}$  and  $S_{\max}$ . In cases where the load follows a Pareto distribution and  $S = \alpha L$ , only abrupt ruptures are possible as shown in.<sup>16</sup> Finally, we see that the Weibull distribution gives rise to a richer set of possibilities for the transition of  $n_\infty(p)$ . Namely, we see that not only we can observe an abrupt rupture, or a rupture with preceding divergence (i.e., a second-order transition followed by a first-order breakdown), it is also possible that  $n_\infty(p)$  goes through a first-order transition (that does not breakdown the system) followed by a second-order transition that is followed by an ultimate first-order breakdown; see the behavior of the orange circled line in Figure 2. We remark that these cases occur when  $h(x)$  has a local maximum at  $x = S_{\min}$ , while its global maximum occurs at a later point  $x > S_{\min}$ ; see<sup>16</sup> for a more detailed discussion of this matter.

In our second set of simulations we seek to verify the results presented in Section 3.3, namely the optimality of assigning the same free space to all lines (regardless of how initial loads are distributed) in terms of maximizing the robustness. In the process, we also seek to compare the robustness achieved under equal free-space distribution versus the commonly used strategy of setting  $S_i = \alpha L_i$  for each line. We note that the latter setting with a universal tolerance factor  $\alpha$  is commonly used in relevant research literature<sup>13,18–20</sup> as well as in industrial applications;<sup>21,22</sup> therein, the term  $(1+\alpha)$  is sometimes referred to as the *Factor of Safety*. The results are depicted in Figure 3 where lines represent the analytical results given in Section 3.3 and symbols are obtained by averaging over 200 independent experiments with  $N = 10^6$  lines. In all cases we fix the mean load at  $\mathbb{E}[L] = 30$  and mean free-space at  $\mathbb{E}[S] = 10$ . With load distributed as Uniform (Figure 3a), Weibull (Figure 3b), or Pareto (Figure 3c), we either let  $S_i = 10$  for all lines, or use  $S_i = \alpha L_i$  with  $\alpha = \mathbb{E}[S]/\mathbb{E}[L] = 1/3$ , the latter choice making sure that the mean free-space is the same in all plots.

We see in all cases that there is an almost perfect agreement between theory and simulations. We also confirm that regardless of how initial load is distributed, the system achieves uniformly optimal robustness (i.e., maximum  $n_\infty(p)$  for all  $p$ ) as long as the free-space is distributed equally; e.g., see Figure 3d that combines all plots in Figures 3a–3c. In other words, we confirm that (10) holds with the critical attack size  $p^*$  matching the optimal value  $p_{optimal}^* = \mathbb{E}[S]/\mathbb{E}[C] = 0.25$ . Finally, by comparing the

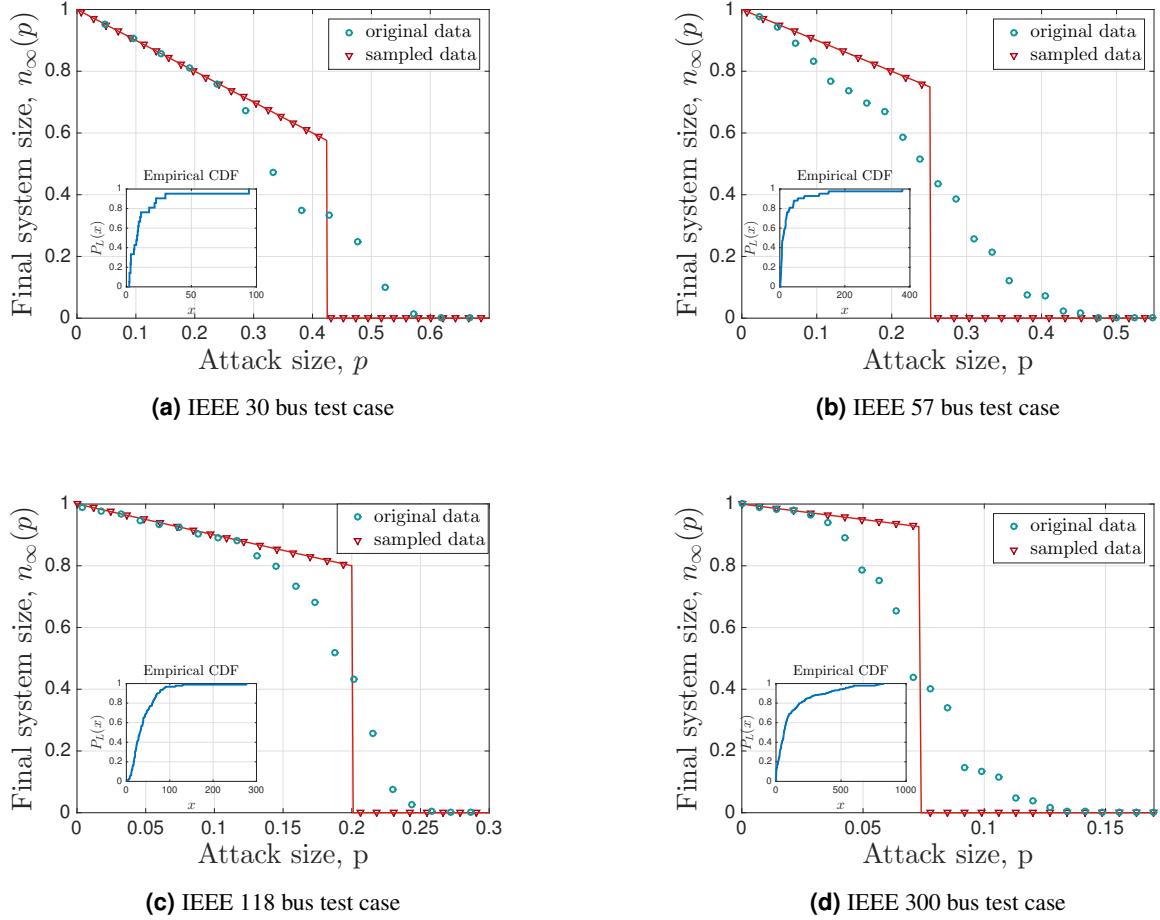


**Figure 3. Final system size under equal free space vs. equal tolerance factor.** In all cases, we set  $L_{min} = 10$ ,  $\mathbb{E}[L] = 30$ , and  $\mathbb{E}[S] = 10$ . When load follows Weibull distribution we let  $k = 6$  and set  $\lambda = 20/\Gamma(1 + 1/k)$  so that  $\mathbb{E}[L] = 30$ . In each of the three cases, we either let  $S \sim \delta(\mathbb{E}[S])$  meaning that all lines have the same free space, or we set  $S_i = \alpha L_i$  with  $\alpha = \mathbb{E}[L]/\mathbb{E}[S] = 1/3$  so that the mean free space still equals 10. We see that analysis (represented by lines) match the simulations (shown in symbols) very well and that robustness is indeed optimized by equal free-space allocation regardless of how initial load is distributed. We also see that system is significantly more robust under equal free space allocation as compared to the case of the equal tolerance factor.

robustness curves under equal free-space and equal tolerance factor, we see the dramatic impact of free-space distribution on the robustness achieved. To give an example, we see from Figure 3d that regardless of how initial load is distributed, the system can be made robust against random attacks that fail up to 25% of the lines; as already discussed this is achieved by distributing the total free-space equally among all lines. However, if the standard approach of setting the free-space proportional to the initial load is followed, the system robustness can be considerably worse with attacks targeting as low as 10% of the lines being able to breakdown the system.

**Real world data.** Thus far, our analytical results are tested only on synthetic data; i.e., simulations are run when load-free space variables  $\{L_i, S_i\}_{i=1}^N$  are generated *randomly* from commonly known distributions. To get a better idea of the real-world implications of our work, we also run simulations on power flow data from the IEEE power system test cases;<sup>35</sup> the IEEE test-cases are widely regarded as realistic test-beds and used commonly in the literature. Here, we consider four power flow test cases corresponding to the 30-bus, 57-bus, 118-bus, and 300-bus systems. For each test case, we take the load values directly from the data-set.<sup>35</sup> Since the data-set does not contain the line capacities, we allocate all lines an equal free-space,  $S = 10$ ; clearly, most of the discussion here would follow with different free-space distributions.

Figure 4 presents the results from the IEEE data set simulations, where blue circles represent the final system size  $n_\infty(p)$  under original load data from each test case; each data point is obtained by averaging the result of 200 independent random attack experiments. As we compare these circles with our analytical results (represented by solid red lines) we see that the



**Figure 4. Simulation results on IEEE test cases.** The initial load values are taken directly from the corresponding IEEE test-case data-sheet,<sup>35</sup> and each line is given an equal free space of  $\mathbb{E}[S] = 10$ . The empirical distribution of load is shown in the Inset of each figure, and the mean load values are given by 13.54, 29.95, 39.95, and 125.02 for the 30-bus system, 57-bus system, 118-bus system, and 300-bus system, respectively. The blue circles represent the simulation results for the final system size  $n_\infty(p)$ . The theoretical results (shown in lines) capture the overall tendency of  $n_\infty(p)$  but fail to predict the numerical results well, especially around the critical attack size. We see that this is merely a finite-size effect as we sample  $N = 10^5$  load values from the empirical distribution and repeat the same experiment. The results are shown in red triangles and are in perfect agreement with the analysis.

overall tendency of  $n_\infty(p)$  is in accords with the theoretical analysis. However, the agreement of theory and simulations is significantly worse than that observed in Figures 2 and 3. This is because our mean field analysis relies on the number of lines  $N$  being *large*, while the IEEE test case data represent very small systems; e.g., the underlying systems have 30, 57, 118, and 300 lines in Figures 4a-4d, respectively. In order to verify that the mismatch is due to the small system size (rather than the load distribution being different from commonly known ones), we re-sample  $10^5$  load values from the *empirical* load distribution obtained from the data-set in each case; the Inset in each figure shows the corresponding empirical distribution  $P_L(x)$ . The simulation results with these  $N = 10^5$  load values are shown in Figure 4 with red triangles. This time with the number of lines increased, we obtain a perfect match between analysis and simulations. This confirms our analysis under realistic load distributions as well. We also see that although analytical results fail to match the system robustness perfectly when  $N$  is very small, they still capture the overall tendency of the robustness curves pretty well. In fact, they can be useful in predicting attack sizes that will lead to a *significant* damage to the system; e.g., in all cases we see that the analytically predicted critical attack size  $p^*$ , ranging from 0.42 in Figure 4a to 0.07 in Figure 4d, leads to the failure of more than 50 % of all lines in the real system.

## 4 Discussion

Our results provide a complete picture of the robustness of power systems against random attacks under the equal load-redistribution model. Namely, with initial load  $L_i$  and extra space  $S_i$  of each line being independently and identically distributed with  $p_{LS}(x, y)$ , our analysis explains how the final system size  $n_\infty(p)$  will behave under attacks with varying size  $p$ . We also demonstrate different types behavior that  $n_\infty(p)$  can exhibit near and around the *critical* attack size  $p^*$ , i.e., the point after which  $n_\infty(p) = 0$  and the system breaks down completely. We show that the final breakdown of the system is always first-order (i.e., discontinuous) but depending on  $p_{LS}(x, y)$ , this may i) take place abruptly meaning that  $n_\infty(p)$  follows the  $1 - p$  line until its sudden jump to zero; or ii) be preceded by a second-order (i.e., continuous) divergence from the  $1 - p$  line. We also demonstrate the possibility of richer behavior where  $n_\infty(p)$  drops to zero through a first-order, second-order, and then a first-order transition. The discontinuity of the final system size at  $p^*$  makes it very difficult to predict system behavior (in response to attacks) from previous data. In fact, this is reminiscent of the real-world phenomena of unexpected large-scale system collapses; i.e., cases where seemingly identical attacks/failures lead to entirely different consequences. On the other hand, the cases that exhibit a preceding second-order transition are less severe, since the deviation from the  $1 - p$  line may be taken as an early warning that the current attack size is close to  $p^*$  and that the system is not likely to sustain attacks much larger than this.

From a design perspective, it is desirable to maximize the robustness of the electrical power system under certain constraints. In our analysis, we address this problem and derive the optimal load-free space distribution  $p_{LS}(x, y)$  that maximizes the final system size  $n_\infty(p)$  uniformly for all attack sizes  $p$ . Namely, we show that under the constraints that  $\mathbb{E}[L]$  and  $\mathbb{E}[S]$  are fixed, robustness is maximized by allocating the same free space to all lines and distributing the initial loads arbitrarily; i.e. the distribution  $p_{LS}(x, y) = p_L(x)\delta(y - \mathbb{E}[S])$  maximizes robustness for arbitrary  $p_L(x)$ . We show that this optimal distribution leads to significantly better robustness than the commonly used strategy of assigning a universal tolerance factor  $\alpha$ , i.e., using  $p_{LS}(x, y) = p_L(x)\delta(y - \alpha x)$ .

Our theoretical results are verified via extensive simulations using both synthetic data and real world data. We show that our results are in perfect agreement with numerical simulations when the system size  $N$  is large; in most cases it suffices to have  $N = 10^4$  to  $N = 10^5$ . However, we see from our simulations with the IEEE test-cases that when  $N$  is very small (we considered  $N = 30$ ,  $N = 57$ ,  $N = 118$ , and  $N = 300$ ), our theory fails to yield the same prediction accuracy. Nevertheless, we see that our results capture the overall tendency of  $n_\infty(p)$  pretty well, and thus can serve as a useful predictor of the critical attack size.

An important direction for future work would be to relax the simplifications and assumptions used here for modeling the failures in an electrical power system. For example, the equal redistribution rule is used here to capture the long-range effect of the Kirchhoff Law, i.e., that the failure of a line may impact the system *globally*. Future work may consider different types of *global* load redistribution rules that are not based on *equal* redistribution; e.g., load may be redistributed randomly or according to some other rules. Hybrid approaches where a fraction of the load is redistributed only locally to the neighboring lines, while the rest being redistributed globally might be considered. Another interesting direction for future work would be to consider the case of *targeted* attacks, rather than random attacks studied here. A good starting point in that direction would be to study possible attack strategies that a capable adversary might use; e.g., given  $L_1, \dots, L_N$  and  $S_1, \dots, S_N$ , which  $k$  lines should an adversary attack in order to minimize the final system size  $n_\infty$ ? A preliminary analysis of this problem can be found in<sup>36</sup> with partial results indicating that optimal attack strategies may be computationally expensive to derive – i.e., that the problem is NP-Hard.

## 5 Methods

### 5.1 Understanding the cascade dynamics

Our proofs are based on a mean-field analysis of the cascading failure dynamics under the equal redistribution model; see Model Definitions for details. Assume that failures take place in discrete time steps  $t = 0, 1, \dots$ , and are initiated at time  $t = 0$  by the random failure of a  $p$ -fraction of the lines. For each  $t = 0, 1, \dots$ , let  $f_t$  denote the fraction of lines that have *failed* until stage  $t$ . The number of links that are still alive at time  $t$  is then given by  $N_t = N(1 - f_t)$ ; e.g.,  $f_0 = 0$  and  $N_0 = N(1 - p)$ . Also, we find it useful to denote by  $Q_t$  the total *extra* load per *alive* line at (the end of) stage  $t$ . In other words,  $Q_t$  is given by the total load of all  $f_t N$  failed lines until this stage divided by  $(1 - f_t)N$ ; e.g.,  $Q_0 = p\mathbb{E}[L]/(1 - p)$  since the initial attack is *random*.

Our main goal is to derive the final system size  $n_\infty(p)$  as a function of the attack size  $p$ . With the above definitions in place, we clearly have  $n_\infty(p) = 1 - f_\infty$ . Thus, the derivation of  $n_\infty(p)$  passes through an understanding of the behavior of  $f_t$  as  $t \rightarrow \infty$ . Here, we will achieve this by first deriving recursive relations for  $f_t$ ,  $Q_t$ , and  $N_t$  for each  $t = 0, 1, \dots$ , and then analyzing the steady-state behavior of the recursions. This method has already proven successful by Yağan,<sup>16</sup> who studied the same problem in a special case where

$$S_i = \alpha L_i, \quad i = 1, \dots, N$$

with  $\alpha > 0$  defining the *universal* tolerance factor. Put differently, the work<sup>16</sup> considers the special case where

$$p_{LS}(x, y) = p_L(x) \delta(y - x\alpha)$$

for arbitrary distribution of initial load  $p_L(x)$ . Here, we start our discussion from the recursions derived by Yağan [16, Eqn. 6] for this special case. Namely, with  $f_0 = p$ ,  $Q_0 = p\mathbb{E}[L]/(1-p)$  (and  $Q_{-1} = 0$ ), they showed for each  $t = 1, 2, \dots$  that

$$f_{t+1} = 1 - (1 - f_t)\mathbb{P}[\alpha L > Q_t \mid \alpha L > Q_{t-1}] \quad (12)$$

$$Q_{t+1} = \frac{p\mathbb{E}[L] + (1-p)\mathbb{E}[L \cdot \mathbf{1}[\alpha L \leq Q_t]]}{(1-p)\mathbb{P}[\alpha L > Q_t]} \quad (13)$$

$$N_{t+1} = (1 - f_{t+1})N \quad (14)$$

An inspection of their derivation reveals that these recursive relations do hold in the general case as well with  $\alpha L$  replaced by the random variable  $S$ . Namely, with no constraints imposed on the distribution  $p_{LS}(x, y)$  (other than those stated in Model Definitions), we have

$$f_{t+1} = 1 - (1 - f_t)\mathbb{P}[S > Q_t \mid S > Q_{t-1}] \quad (15)$$

$$Q_{t+1} = \frac{p\mathbb{E}[L] + (1-p)\mathbb{E}[L \cdot \mathbf{1}[S \leq Q_t]]}{(1-p)\mathbb{P}[S > Q_t]} \quad (16)$$

$$N_{t+1} = (1 - f_{t+1})N \quad (17)$$

These equations can also be validated intuitively. First of all, given that  $Q_t$  defines the extra load per alive line at the end of stage  $t$ , we know that for a line to fail exactly at stage  $t+1$ , it must have a free space smaller than  $Q_t$  but larger than  $Q_{t-1}$ , with the latter condition ensuring that the line does not fail at any previous stages. So, the fraction of lines that fail at stage  $t+1$  among those that survive stage  $t$  is intuitively given by  $\mathbb{P}[S \leq Q_t \mid S > Q_t]$ . Rewriting (15), we get

$$\frac{\# \text{ of lines that survive stage } t \text{ but fail at } t+1}{\# \text{ of lines that survive stage } t} = \frac{f_{t+1} - f_t}{1 - f_t} = 1 - \frac{1 - f_{t+1}}{1 - f_t} = 1 - \mathbb{P}[S > Q_t \mid S > Q_{t-1}] = \mathbb{P}[S \leq Q_t \mid S > Q_{t-1}]$$

confirming this intuitive argument. In fact, it is clear that for a line to survive stage  $t+1$ , it must i) survive the initial attack (which happens with probability  $1-p$ ), and ii) have a free space  $S > Q_t$ . Given the independence of the initial attack from other variables, we thus have  $1 - f_{t+1} = (1-p)\mathbb{P}[S > Q_t]$ . This explains the denominator of (16) since  $Q_{t+1}$  gives the additional load *per alive* line at this stage. The nominator of (16) should then give the mean total load of the lines that have failed until (and including) stage  $t+1$ , normalized with  $N$ ; the normalization is required given that the denominator term is also normalized with  $N$ . To calculate the total load of the failed lines at this stage, first we note that  $p$  fraction of the lines fail randomly as a result of the initial attack, giving the term  $p\mathbb{E}[L]$  in the nominator of (16). In addition, among the remaining  $1-p$  fraction of the lines, those with free space satisfying  $S \leq Q_t$  will fail, leading to the second term in the nominator of (16).

Returning to the recursions (12)-(14), we see that cascading failures will stop and a steady-state will be reached when  $f_{t+2} = f_{t+1}$ . From (15), we see that this occurs if

$$\mathbb{P}[S > Q_{t+1} \mid S > Q_t] = 1,$$

or, equivalently if

$$\mathbb{P}\left[S > \frac{p\mathbb{E}[L] + (1-p)\mathbb{E}[L \cdot \mathbf{1}[S \leq Q_t]]}{(1-p)\mathbb{P}[S > Q_t]} \mid S > Q_t\right] = 1, \quad (18)$$

upon using (16). In order to simplify this further, we let  $x := Q_t$ , and realize that

$$p\mathbb{E}[L] + (1-p)\mathbb{E}[L \cdot \mathbf{1}[S \leq x]] = p\mathbb{E}[L] + (1-p)\mathbb{E}[L \cdot (1 - \mathbf{1}[S > x])] = \mathbb{E}[L] - (1-p)\mathbb{E}[L \cdot \mathbf{1}[S > x]].$$

With these in place, the condition for cascades to stop (18) gives

$$\mathbb{P}\left[S > \frac{\mathbb{E}[L] - (1-p)\mathbb{E}[L \cdot \mathbf{1}[S > x]]}{(1-p)\mathbb{P}[S > x]} \mid S > x\right] = 1 \quad (\text{Cascade stop condition}) \quad (19)$$

It is now clear how to obtain the fraction  $n_\infty(p)$  of power lines that are still alive at the end of the cascading failures. First, we shall find the smallest solution  $x^*$  of (19) that gives the equilibrium value  $Q_\infty$  at which cascades will stop. Then the final-fraction  $n_\infty$  of alive lines is given by

$$n_\infty(p) = 1 - f_\infty = (1 - p)\mathbb{P}[S > x^*] \quad (20)$$

The last relation follows from the fact that  $1 - f_{t+1} = (1 - p)\mathbb{P}[S > Q_t]$  for each  $t = 0, 1, \dots$ . This can be established in the following manner. Applying (15) repeatedly, we get

$$\begin{aligned} 1 - f_{t+1} &= (1 - f_t)\mathbb{P}[S > Q_t \mid S > Q_{t-1}] \\ 1 - f_t &= (1 - f_{t-1})\mathbb{P}[S > Q_{t-1} \mid S > Q_{t-2}] \\ &\vdots \\ 1 - f_1 &= (1 - f_0)\mathbb{P}[S > Q_0], \end{aligned}$$

which gives

$$1 - f_{t+1} = (1 - f_0) \prod_{\ell=0}^t \mathbb{P}[S > Q_\ell \mid S > Q_{\ell-1}],$$

where  $Q_{-1} = 0$  as before. Since  $Q_t$  is monotone increasing in  $t$ , i.e.,  $Q_{t+1} \geq Q_t$  for all  $t$ , we get

$$1 - f_{t+1} = (1 - f_0) \frac{\mathbb{P}[S > Q_t]}{\mathbb{P}[S > Q_{t-1}]} \cdot \frac{\mathbb{P}[S > Q_{t-1}]}{\mathbb{P}[S > Q_{t-2}]} \cdots \frac{\mathbb{P}[S > Q_1]}{\mathbb{P}[S > Q_0]} \cdot \mathbb{P}[S > Q_0] = (1 - p)\mathbb{P}[S > Q_t] \quad (21)$$

as we recall that  $f_0 = p$ .

We now seek to simplify the cascade stop condition given at (19). For notational convenience, let

$$g(x) := \frac{E[L] - (1 - p)E[L \cdot \mathbb{1}[S > x]]}{(1 - p)\mathbb{P}[S > x]}. \quad (22)$$

Then, (19) becomes

$$\mathbb{P}[S > g(x) \mid S > x] = 1, \quad (23)$$

which holds if either one of the following is satisfied:

1.  $x \geq g(x)$ ; or,
2.  $x < g(x)$  and  $\mathbb{P}[S > g(x) \mid S > x] = 1$ .

The next result (proved in the Supplementary Material) shows that it suffices to consider only the first case for the purposes of our discussion.

**Claim 1** *Let  $x^*$  be the smallest solution of (23), and  $x^{**}$  be the smallest solution of  $x \geq g(x)$ . Then, we have*

$$\mathbb{P}[S > x^*] = \mathbb{P}[S > x^{**}]$$

The proof of this important technical result is given in the Supplementary File.

Rewriting the inequality  $x \geq g(x)$  and using Claim 1, we now establish the first main result of the paper given at (3) and (4). Namely, with  $x^*$  denoting the smallest solution of

$$\mathbb{P}[S > x](x + \mathbb{E}[L \mid S > x]) \geq \frac{\mathbb{E}[L]}{1 - p}, \quad x \in (0, \infty) \quad (24)$$

the final system size  $n_\infty(p)$  is given by  $(1 - p)\mathbb{P}[S > x^*]$ . If (24) has no solution, we set  $x^* = \infty$  leading to  $n_\infty(p) = 0$ . ■

**Critical attack size.** Characterizing the critical attack size  $p^*$  is now a simple matter from the discussion above. Recall that critical attack size is defined via

$$p^* = \sup\{p : n_\infty(p) > 0\}.$$

Since  $\mathbb{E}[L]/(1-p) > 0$  always holds, we know that whenever there exists an  $x^* < \infty$  satisfying (24), we must have that  $\mathbb{P}[S > x^*] > 0$  as otherwise the function  $g(x^*) = \mathbb{P}[S > x^*](x^* + \mathbb{E}[L | S > x^*])$  would be zero conflicting with (24). Therefore, we have  $n_\infty(p) > 0$  for any  $p$  for which (24) has a solution; and we get  $n_\infty(p) = 0$  only if (24) has no solution. With these in mind, it is clear that  $p^*$  will be given by the supremum of  $p$  values for which (24) has a solution. Equivalently,  $p^*$  is given by the value of  $p$  for which  $\mathbb{E}[L]/(1-p)$  equals to the *global* maximum of  $g(x)$ . More precisely, we have

$$p^* = 1 - \frac{\mathbb{E}[L]}{\max_x \{\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])\}} \quad (25)$$

This establishes (5). ■

## 5.2 Order of phase transition and condition for abrupt break-downs

**Order of transition.** We now establish the fact that the final breakdown of the system will always be through a *first-order* (i.e., discontinuous) transition. This amounts to establishing (6), namely that  $n_\infty(p^*) > 0$ ; this then implies a discontinuity in  $n_\infty(p)$  at  $p = p^*$ , since by definition of the critical attack size we have  $n_\infty(p^* + \varepsilon) = 0$  for any  $\varepsilon > 0$ .

We now establish  $n_\infty(p^*) > 0$ . From (25), we see that when  $p = p^*$  the cascade stopping condition (24) will be satisfied by  $x_*$  that maximizes  $\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])$ . In other words, we have

$$n_\infty(p^*) = (1 - p^*)\mathbb{P}[S > x_*] \quad (26)$$

where

$$x_* = \arg \max_x \{\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])\}$$

We argue that  $\mathbb{P}[S > x_*] > 0$  by contradiction. Suppose that  $\mathbb{P}[S > x_*] = 0$ . With  $x_*$  denoting the global maximizer of  $\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])$ , this would imply that

$$\mathbb{P}[S > x](x + \mathbb{E}[L | S > x]) \leq \mathbb{P}[S > x_*](x_* + \mathbb{E}[L | S > x_*]) = 0, \quad \text{for all } x \in (0, \infty). \quad (27)$$

However, with  $0 \leq x < S_{\min}$ , we have  $\mathbb{P}[S > x](x + \mathbb{E}[L | S > x]) = x + \mathbb{E}[L] > 0$  contradicting with (27). Therefore, we must have  $\mathbb{P}[S > x_*] > 0$ , and the desired conclusion  $n_\infty(p^*) > 0$  immediately follows. ■

**Condition for an *abrupt* first-order transition.** An abrupt rupture is said to take place if the linear decay of  $n_\infty(p)$  (in the form of  $1 - p$ ) is followed by a sudden discontinuous jump to zero at  $p^*$ ; i.e., when it holds that

$$n_\infty(p) = \begin{cases} 1 - p & \text{if } p \leq p^* \\ 0 & \text{if } p > p^* \end{cases} \quad (28)$$

We now show that this occurs if and only if the function  $h(x) = \mathbb{P}[S > x](x + \mathbb{E}[L | S > x])$  takes its global maximum at  $x_* = S_{\min}$ . First of all,  $n_\infty(p^*) = 1 - p^*$  implies that  $\mathbb{P}[S > x_*] = 1$  in view of (26). This immediately gives  $x_* \leq S_{\min}$ . For the reverse direction, observe that  $h(x)$  is linearly increasing on the range  $0 \leq x \leq S_{\min}$ . In fact, we have

$$h(x) = x + \mathbb{E}[L], \quad 0 \leq x \leq S_{\min}. \quad (29)$$

Therefore,  $h(x)$  reaches at least a local maximum at  $x = S_{\min}$ , implying that  $x_* \geq S_{\min}$ . Collecting, we establish  $x_* = S_{\min}$  as a *necessary* condition for an abrupt rupture (i.e., (28)) to take place. Next we show that this is also a sufficient condition. If  $x_* = S_{\min}$  meaning that  $h(x)$  is maximized at  $S_{\min}$ , then the minimum solution for the inequality  $h(x) \geq \mathbb{E}[L]/(1-p)$ , when exists, will appear at  $x^*$  that satisfies  $x^* \leq S_{\min}$ . This implies that the final system size  $n_\infty(p) = (1-p)\mathbb{P}[S > x^*]$  is either  $1-p$  (when a solution to the inequality exists), or zero (when no solution to  $h(x) \geq \mathbb{E}[L]/(1-p)$  exists). In other words, when  $x_* = S_{\min}$  the final system size will be of the form (28), i.e., an abrupt rupture will take place. Combining, we conclude that (28) takes place if and only if

$$x_* = \arg \max_{x>0} \{\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])\} = S_{\min}. \quad (\text{necessary and sufficient condition for abrupt rupture}) \quad (30)$$

In the Supplementary File, we explore this issue further and provide a *necessary* (but not sufficient) condition for (30) to hold. This leverages the fact that for the maximum of  $h(x)$  to take place at  $x = S_{\min}$ , the derivative of  $h(x)$  must change its sign to negative at this point; this would ensure a *local* maximum of  $g(x)$  take place at  $S_{\min}$ . The resulting necessary condition, given here for convenience, is

$$p_S(S_{\min})(S_{\min} + \mathbb{E}[L | S = S_{\min}]) > 1. \quad (\text{necessary condition for abrupt rupture})$$

### 5.3 Optimal $P_{LS}$ distribution that maximize robustness

We now seek to find the optimal load-free space distribution  $P_{LS}(x,y)$  that maximizes the robustness of the power system, when  $\mathbb{E}[L]$  and  $\mathbb{E}[S]$  are fixed. First, we focus on maximizing the critical attack size  $p^*$ . Recall (25) and observe that

$$\begin{aligned} h(x) &= \mathbb{P}[S > x](x + \mathbb{E}[L | S > x]) \\ &= x\mathbb{P}[S > x] + \mathbb{E}[L \cdot \mathbf{1}[S > x]] \\ &\leq \mathbb{E}[S] + \mathbb{E}[L \cdot \mathbf{1}[S > x]] \\ &\leq \mathbb{E}[S] + \mathbb{E}[L] \end{aligned} \quad (31)$$

where we use the Markov inequality at (31), i.e. that  $\mathbb{P}[S > x] \leq \mathbb{E}[S]/x$ . Reporting this into (25) we get

$$p^* = 1 - \frac{\mathbb{E}[L]}{\max_x \{h(x)\}} \leq 1 - \frac{\mathbb{E}[L]}{\mathbb{E}[S] + \mathbb{E}[L]} = \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mathbb{E}[L]} = \frac{\mathbb{E}[S]}{\mathbb{E}[C]}. \quad (32)$$

This means that regardless of our choice of the joint distribution  $P_{LS}(x,y)$  the critical attack size can never be larger than  $\frac{\mathbb{E}[S]}{\mathbb{E}[C]}$ .

Next, we show that this upper bound is in fact achievable by a *Dirac*-delta distribution of free space  $S$ . Assume that

$$P_{LS}(x,y) = P_L(x) \mathbf{1}[y \leq \mathbb{E}[S]] \quad (33)$$

where the load distribution  $P_L(x)$  is arbitrary; this is equivalent to having  $p_{LS}(x,y) = p_L(x)\delta(y - \mathbb{E}[S])$ . Let  $p_{dirac}^*$  denote the corresponding critical attack size. With  $x = \mathbb{E}[S]^-$ , we have  $\mathbb{P}[S > x] = 1$ , and  $\mathbb{E}[L \cdot \mathbf{1}[S > x]] = \mathbb{E}[L]$ , so that

$$\lim_{x \uparrow \mathbb{E}[S]} h(x) = \lim_{x \uparrow \mathbb{E}[S]} (\mathbb{P}[S > x](x + \mathbb{E}[L | S > x])) = \mathbb{E}[S] + \mathbb{E}[L]$$

Thus, we have  $\max_x \{h(x)\} \geq \mathbb{E}[S] + \mathbb{E}[L]$ , which immediately gives

$$p_{dirac}^* = 1 - \frac{\mathbb{E}[L]}{\max_x \{h(x)\}} \geq 1 - \frac{\mathbb{E}[L]}{\mathbb{E}[S] + \mathbb{E}[L]} = \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mathbb{E}[L]}. \quad (34)$$

Since the lower bound (32) holds for any distribution, we conclude that

$$p_{dirac}^* = \frac{\mathbb{E}[S]}{\mathbb{E}[S] + \mathbb{E}[L]} := p_{optimal}^*. \quad (35)$$

This shows that a degenerate distribution on the extra space  $S$  leads to optimal (i.e., maximum) critical attack size  $p_{optimal}^*$ . ■

We now show that a Dirac-delta distribution of free space not only maximizes  $p^*$  but it maximizes the final system size  $n_\infty(p)$  uniformly across all attack sizes. It is clear that  $n_\infty(p) \leq 1 - p$  for any distribution  $P_{LS}(x,y)$  and any attack size  $p$ . Thus, our claim will be established if we show that the Dirac-delta distribution (33) leads to an *abrupt* rupture and thus the resulting final system size is in the form given at (28). More precisely, we will get the desired result

$$n_{\infty, dirac}(p) = \begin{cases} 1 - p, & \text{if } p \leq p_{optimal}^* \\ 0, & \text{if } p > p_{optimal}^* \end{cases} \quad (36)$$

upon establishing the abrupt rupture condition (30), i.e., that  $h(x)$  takes its maximum at  $x = S_{\min}$ . This follows immediately upon realizing that we have

$$h(x) = \begin{cases} x + \mathbb{E}[L], & \text{for } 0 \leq x < \mathbb{E}[S] \\ 0, & \text{for } x \geq \mathbb{E}[S] \end{cases}$$

under the distribution (33). ■

## References

1. Rinaldi, S. M., Peerenboom, J. P. & Kelly, T. K. Identifying, understanding, and analyzing critical infrastructure interdependencies. *Control Systems, IEEE* **21**, 11–25 (2001).
2. Brummitt, C. D., D’Souza, R. M. & Leicht, E. Suppressing cascades of load in interdependent networks. *Proceedings of the National Academy of Sciences* **109**, E680–E689 (2012).
3. O’Rourke, T. D. Critical infrastructure, interdependencies, and resilience. *BRIDGE-WASHINGTON-NATIONAL ACADEMY OF ENGINEERING-* **37**, 22 (2007).
4. Dobson, I., Carreras, B. A., Lynch, V. E. & Newman, D. E. Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **17**, 026103 (2007).
5. Yağan, O., Qian, D., Zhang, J. & Cochran, D. Optimal allocation of interconnecting links in cyber-physical systems: Interdependence, cascading failures, and robustness. *Parallel and Distributed Systems, IEEE Transactions on* **23**, 1708–1720 (2012).
6. Foundations, C. Protecting america’s infrastructures: The report of the president’s commission on critical infrastructure protection (1997).
7. Romero, J. J. Blackouts illuminate india’s power problems. *Spectrum, IEEE* **49**, 11–12 (2012).
8. Tang, Y., Bu, G. & Yi, J. Analysis and lessons of the blackout in indian power grid on july 30 and 31, 2012. In *Zhongguo Dianji Gongcheng Xuebao(Proceedings of the Chinese Society of Electrical Engineering)*, vol. 32, 167–174 (Chinese Society for Electrical Engineering, 2012).
9. Kosterev, D. N., Taylor, C. W., Mittelstadt, W. et al. Model validation for the august 10, 1996 wscc system outage. *Power Systems, IEEE Transactions on* **14**, 967–979 (1999).
10. Carreras, B., Newman, D., Dobson, I. & Poole, A. Evidence for soc in electric power system blackouts. In *hicss*, 2016 (IEEE, 2001).
11. Sachtjen, M., Carreras, B. & Lynch, V. Disturbances in a power transmission system. *Physical Review E* **61**, 4877 (2000).
12. Anghel, M., Werley, K., Motter, A. E. et al. Stochastic model for power grid dynamics. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, 113–113 (IEEE, 2007).
13. Crucitti, P., Latora, V. & Marchiori, M. Model for cascading failures in complex networks. *Phys. Rev. E* **69**, 045104 (2004). URL <http://link.aps.org/doi/10.1103/PhysRevE.69.045104>.
14. Wang, J.-W. & Rong, L.-L. Cascade-based attack vulnerability on the {US} power grid. *Safety Science* **47**, 1332 – 1336 (2009). URL <http://www.sciencedirect.com/science/article/pii/S0925753509000174>.
15. Pahwa, S., Hodges, A., Scoglio, C. & Wood, S. Topological analysis of the power grid and mitigation strategies against cascading failures. In *Systems Conference, 2010 4th Annual IEEE*, 272–276 (IEEE, 2010).
16. Yağan, O. Robustness of power systems under a democratic-fiber-bundle-like model. *Phys. Rev. E* **91**, 062811 (2015). URL <http://link.aps.org/doi/10.1103/PhysRevE.91.062811>.
17. Daniels, H. The statistical theory of the strength of bundles of threads. i. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 183, 405–435 (The Royal Society, 1945).
18. Motter, A. E. & Lai, Y.-C. Cascade-based attacks on complex networks. *Phys. Rev. E* **66**, 065102 (2002). URL <http://link.aps.org/doi/10.1103/PhysRevE.66.065102>.
19. Wang, W.-X. & Chen, G. Universal robustness characteristic of weighted networks against cascading failure. *Phys. Rev. E* **77**, 026101 (2008). URL <http://link.aps.org/doi/10.1103/PhysRevE.77.026101>.
20. Mirzasoleiman, B., Babaei, M., Jalili, M. & Safari, M. Cascaded failures in weighted networks. *Physical Review E* **84**, 046114 (2011).
21. Bernstein, A., Bienstock, D., Hay, D., Uzunoglu, M. & Zussman, G. Power grid vulnerability to geographically correlated failures—analysis and control implications. In *INFOCOM, 2014 Proceedings IEEE*, 2634–2642 (IEEE, 2014).
22. Kinney, R., Crucitti, P., Albert, R. & Latora, V. Modeling cascading failures in the north american power grid. *The European Physical Journal B-Condensed Matter and Complex Systems* **46**, 101–107 (2005).
23. Su, Z. et al. Robustness of interrelated traffic networks to cascading failures. *Scientific reports* **4** (2014).
24. Scala, A., Lucentini, P. G. D. S., Caldarelli, G. & D’Agostino, G. Cascades in interdependent flow networks. *Physica D: Nonlinear Phenomena* (2015).

25. Andersen, J. V., Sornette, D. & Leung, K.-t. Tricritical behavior in rupture induced by disorder. *Physical Review Letters* **78**, 2140 (1997).
26. da Silveira, R. Comment on “tricritical behavior in rupture induced by disorder”. *Physical review letters* **80**, 3157 (1998).
27. Sornette, D., Leung, K.-T. & Andersen, J. Conditions for abrupt failure in the democratic fiber bundle model. *arXiv preprint cond-mat/9712313* (1997).
28. Pradhan, S. & Chakrabarti, B. K. Failure properties of fiber bundle models. *International Journal of Modern Physics B* **17**, 5565–5581 (2003).
29. Roy, C., Kundu, S. & Manna, S. Fiber bundle model with highly disordered breaking thresholds. *Physical Review E* **91**, 032103 (2015).
30. Pahwa, S., Scoglio, C. & Scala, A. Abruptness of cascade failures in power grids. *Scientific reports* **4** (2014).
31. Overbye, T. J., Cheng, X. & Sun, Y. A comparison of the ac and dc power flow models for Imp calculations. In *Proceedings, 37th Hawaii International Conference on System Sciences* (2004).
32. Stott, B., Jardim, J. & Alsaac, O. Dc power flow revisited. *Power Systems, IEEE Transactions on* **24**, 1290–1300 (2009).
33. Rosato, V. *et al.* Modelling interdependent infrastructures using interacting dynamical models. *International Journal of Critical Infrastructures* **4**, 63–79 (2008). URL <http://dx.doi.org/10.1504/IJCIS.2008.016092>.
34. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).
35. Christie, R. University of washington power systems test case archive (last accessed on 25 nov 2013)(1999). URL <http://www.ee.washington.edu/research/pstca> .
36. Chatziafratis, E., Zhang, Y. & Yağan, O. On the robustness of power systems: optimal load-capacity distributions and hardness of attacking. In *Information Theory and Applications Workshop (ITA), 2016* (2016). URL [http://ita.ucsd.edu/workshop/16/files/paper/paper\\_2357.pdf](http://ita.ucsd.edu/workshop/16/files/paper/paper_2357.pdf).

## Acknowledgements

This research was supported in part by the National Science Foundation through grant CCF #1422165, and in part by the Department of Electrical and Computer Engineering at Carnegie Mellon University. O. Yağan also acknowledges the Berkman Faculty Development Grant from Carnegie Mellon.

## Author contributions statement

Y. Z. and O. Y. have contributed equally to all parts of the paper.

## Additional information

**Competing financial interests.** The authors declare no competing financial interests.

## APPENDIX

### A. A proof of Claim 1

Let  $x^*$  denote the smallest solution of  $x \geq g(x)$  and  $x^{**}$  denote the smallest solution of (23). Since  $x \geq g(x)$  automatically gives (23), we always have  $x^{**} \leq x^*$ . Here, our goal is to show that

$$\mathbb{P}[S > x^*] = \mathbb{P}[S > x^{**}] \quad (\text{A.1})$$

If  $x^*$  is also the smallest solution of (23), then the claim follows immediately by virtue of the fact that  $x^* = x^{**}$ . Now, assume that there is a solution  $x^{**} < x^*$  of (23). Then it must hold that  $x^{**} < g(x^{**})$ . From (23), this yields

$$1 = \mathbb{P}[S > g(x^{**}) \mid S > x^{**}] = \frac{\mathbb{P}[S > g(x^{**}), S > x^{**}]}{\mathbb{P}[S > x^{**}]} = \frac{\mathbb{P}[S > g(x^{**})]}{\mathbb{P}[S > x^{**}]}$$

Thus, we have

$$\mathbb{P}[S > x^{**}] = \mathbb{P}[S > g(x^{**})]. \quad (\text{A.2})$$

Key to the proof of Claim 1 is the observation that the function  $g(x)$  given at (22) is monotone decreasing in  $\mathbb{P}[S > x]$ . Put differently, (A.2) implies that  $g(x^{**}) = g(g(x^{**}))$ , meaning that  $x = g(x^{**})$  is a solution of  $x \geq g(x)$ . Since  $x^*$  is defined to be the smallest of all such solutions, this gives

$$x^* \leq g(x^{**}). \quad (\text{A.3})$$

On the other hand, the continuity assumption on the distribution of  $S$  implies the continuity of  $g(x)$ . Hence, we have  $x^* = g(x^*)$ . Recalling also that  $g(x)$  is monotone increasing in  $x$  and that  $x^{**} < x^*$ , we get

$$g(x^{**}) \leq g(x^*) = x^*. \quad (\text{A.4})$$

Combining (A.3) and (A.4), we conclude that

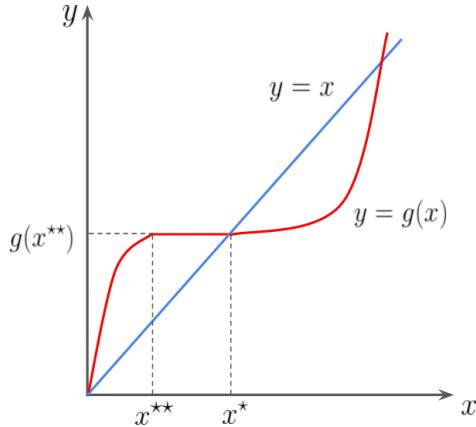
$$x^* = g(x^{**}). \quad (\text{A.5})$$

Graphically, this means that the curve  $y = g(x)$  is constant (and above the line  $y = x$ ) on the range  $[x^{**}, x^*]$  and it intersects with the line  $y = x$  at  $x^*$  (see figure 5 for an illustration).

Combining (A.2) and (A.5), we get

$$\mathbb{P}[S > x^{**}] = \mathbb{P}[S > g(x^{**})] = \mathbb{P}[S > x^*]$$

which establishes the claim (A.1). ■



**Figure 5. Illustration of function relations in Claim 1.** We define  $x^*$  as the first point where curves  $y = x$  and  $y = g(x)$  intersect. We show that if there exists  $x^{**} < x^*$  satisfying (23), then  $g(x)$  must be constant over  $x^{**} \leq x \leq x^*$ , yielding (A.5).

## B. A necessary condition for abrupt rupture

Recall the condition (30) for an abrupt rupture to take place, namely the need for  $h(x)$  to be maximized at  $x = S_{\min}$ . To explore this issue further, we now compute the derivative of  $h(x)$ . We have

$$\begin{aligned}
 & \frac{d}{dx} (\mathbb{P}[S > x] (x + \mathbb{E}[L | S > x])) \\
 &= \frac{d}{dx} (x \mathbb{P}[S > x] + \mathbb{E}[L \cdot \mathbb{1}[S > x]]) \\
 &= \mathbb{P}[S > x] + x(-p_S(x)) + \frac{d}{dx} \left( \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \ell \cdot \mathbb{1}[s > x] p_{LS}(\ell, s) ds d\ell \right) \\
 &= \mathbb{P}[S > x] - x p_S(x) + \frac{d}{dx} \left( \int_{-\infty}^{\infty} \int_x^{\infty} \ell \cdot p_{LS}(\ell, s) ds d\ell \right) \\
 &= \mathbb{P}[S > x] - x p_S(x) + \frac{d}{dx} \left( \int_x^{\infty} \int_{-\infty}^{\infty} \ell \cdot p_{LS}(\ell, s) d\ell ds \right) \\
 &= \mathbb{P}[S > x] - x p_S(x) - \int_{-\infty}^{\infty} \ell \cdot p_{LS}(\ell, x) d\ell \\
 &= \mathbb{P}[S > x] - x p_S(x) - \mathbb{E}[L | S = x] p_S(x).
 \end{aligned} \tag{A.6}$$

For  $x < S_{\min}$ , we have  $\mathbb{P}[S > x] = 1$  and  $\mathbb{E}[L | S > x] = \mathbb{E}[L]$ , so that  $\frac{d}{dx} h(x) = 1$ ; this is already evident from (29). Then, for an abrupt rupture to take place, (A.6) should turn negative at  $x = S_{\min}$ ; i.e., it must hold that

$$p_S(S_{\min})(S_{\min} + \mathbb{E}[L | S = S_{\min}]) > 1.$$

■