2018 IEEE Conference on Decision and Control (CDC)
Miami Beach, FL, USA, Dec. 17-19, 2018

Modeling and Analysis of Cascading Failures in Interdependent Cyber-Physical Systems

Yingrui Zhang and Osman Yağan

Abstract— **Integrated cyber-physical systems (CPSs), such as the smart grid, are becoming the underpinning technology for major industries. A major concern regarding such systems are the seemingly unexpected large scale failures, which are often attributed to a small initial shock getting escalated due to intricate dependencies within and across the individual counterparts of the system. In this paper, we develop a novel interdependent system model to capture this phenomenon, also known as cascading failures. Our framework consists of two networks that have inherently different characteristics governing their *intra*-dependency: i) a *cyber*-network where a node is deemed to be functional as long as it belongs to the largest connected (i.e., giant) component; and ii) a *physical* network where nodes are given an initial *flow* and a *capacity*, and failure of a node results with redistribution of its flow to the remaining nodes, upon which further failures might take place due to *overloading*. Furthermore, it is assumed that these two networks are *inter-dependent*. For simplicity, we consider a one-to-one interdependency model where every node in the cyber-network is dependent upon and supports a single node in the physical network, and vice versa. We provide a thorough analysis of the dynamics of cascading failures in this interdependent system initiated with a random attack. The system robustness is quantified as the *surviving* fraction of nodes at the end of cascading failures, and is derived in terms of all network parameters involved. Analytic results are supported through an extensive numerical study. Among other things, these results demonstrate the ability of our model to capture the *unexpected* nature of large-scale failures, and provide insights on improving system robustness.**

I. Introduction

Today's worldwide network infrastructure consists of a web of interacting cyber-networks and physical systems. Integrated cyber-physical systems (CPSs) are increasingly becoming the underpinning technology for major industries. The smart grid is an archetypal example of a CPS where the power grid and the communication network are coupled together; the grid depends on the communication network for its control, and the communication network depends on the grid for power. While this coupling brings unprecedented improvements, it has been observed [1] that such interdependent systems tend to be fragile against failures, natural hazards, and attacks. For instance, in the event of an attack or random failures in such an interdependent system, the failures in one of the networks can cause failures of the dependent nodes in the other network and vice versa. This process may continue in a recursive manner, triggering a cascade of failures that can potentially collapse the entire system. In fact, the cascading effect of even a partial Internet blackout could disrupt major national infrastructures involving Internet services, power grids, and financial markets [2]. For example, the electrical blackout that affected much of Italy on 28 September 2003 had started with the shutdown of a power station, which led to failures in the Internet communication network, which in turn caused the breakdown of more stations [3] .

While important, traditional studies in network science fall short in characterizing the robustness of interdependent networks since the focus has mainly been on single networks. Despite some recent research activity studying interdependent networks [2], [4]–[6], very few consider engineering aspects of inter-dependent networks and very little is known as to how such systems can be designed to have maximum robustness under certain design constraints; see [7]–[9] for rare exceptions. The current literature is also lacking interdependent system models that capture fundamental differences between *physical* and *cyber* networks, and enable studying robustness of systems that integrate networks with inherently different behavior. There is thus a need to develop new approaches for modeling and analyzing cascading failures in interdependent cyber-physical systems.

In this paper, we develop a model that will help understand how failures would propagate in an interdependent system that constitutes physical *and* cyber networks. We provide a thorough analysis of the dynamics of cascading failures initiated by a *random* attack. The system robustness, defined as the *steady-state* fraction of nodes that survive the cascade, is characterized in terms of all network parameters involved. We also leverage our main result to investigate how the robustness can be improved by adjusting various parameters defining the interdependent system; e.g., load/capacity values in the physical network and the degree distribution of the cyber-network. Our results also reveal an intricate connection between the robustness of each constituent network and the robustness of the interdependent system formed by them; see Section IV for details.

The rest of the paper is organized as follows. In Section II, we present our interdependent system model in details. In Section III, we present the main result of the paper, which allows computing the fraction of surviving nodes at each step of cascading failures initiated by a random attack. In Section IV, we present numerical results demonstrating the accuracy of our analysis in the finite node regime. The paper is concluded in Section V with several suggestions for future work.

This research was supported by National Science Foundation through grant CCF #1422165.

Y. Zhang and O. Yağan are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. Email: yingruiz@andrew.cmu.edu, oyagan@ece.cmu.edu

II. System Model

A. Intra-dependency vs. Inter-dependency

Our modeling framework is motivated with the inherent dependencies that exist in many real-world systems including CPSs. Namely, we will characterize how component failures propagate and cascade, both within the cyber or the physical parts of the system (due to "intra-dependency"), as well as across them due to "inter-dependency". The actual meaning of "failure" is expected to be domain-dependent and can vary from a component being physically damaged to a node's inability to carry out its tasks. For ease of exposition, we consider two sub-systems, say A and B.

Assume that network A consists of nodes $\{a_1, \ldots, a_N\}$ and network B consists of nodes $\{b_1, \ldots, b_N\}$. For illustration purposes, we can think of network A as the power network consisting of generators and substations (i.e., the physical network), and network B as the control and communication network consisting of control centers and routers (i.e., the cyber network). This is a classical example of an interdependent CPS, with the power stations sending data to and receiving control signals from routers, and routers receiving power from substations. Modeling the dependencies within and between networks A and B amounts to answering three questions. First, we have to decide on the set of rules governing how failures would propagate within each network, leading to a characterization of the *intra*-dependencies. For example, we should identify how the failure of a power node a_i affects other substations and generators in the power network A. Similarly, we should identify how the failure of a communication node b_j affects other nodes in B. Finally, we must characterize the *inter*-dependence of the two systems, and how such interdependence may lead to propagation of failures across them. Namely, we must have a set of rules that specify how the failure of a power station a_i impacts the nodes $\{b_1, \ldots, b_N\}$ in the communication network and vice versa. Once these modeling questions are answered, the propagation of failures in an interdependent system can be studied. Without loss of generality, assume that the failures are initiated in network A due to random failures or attacks. To get a better idea about the role of intra- and inter-dependencies in the cascade of failures, consider an *asynchronous* failure update model, where the effect of intra-dependencies and inter-dependencies are considered in two separate batches, following one another. The asynchronous failure update assumption eases the implementation and analysis of the model, and can be shown to yield the same steady-state network structures with a synchronous failure update model; just note that failure propagation process is monotone and that (according to our assumption) nodes can not heal once failed.

B. The Model

Despite the vast literature on interdependent networks [2], [7], [10], there has been little (if any) attempt to characterize the robustness of interdependent systems where the constituent networks have different intra-dependency behaviors.

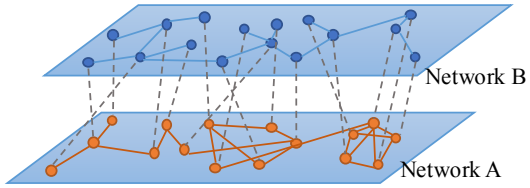

Fig. 1. *System model illustration for the cyber-physical systems, where network A can be the physical grid, and network B can be the communication network that sends control signals. The interdependence across the two networks are realized through random one-to-one support links shown by dashed lines.*

In the case of CPS, it would be expected that the cyber and physical counterparts obey inherently different rules governing how failures would propagate internally in each network. To this end, we study in this paper an interdependent system model that consists of two networks with different characteristics governing their *intra*-dependency: i) a *cyber*-network where a node is deemed to be functional as long as it belongs to the largest connected (i.e., giant) component; and ii) a *physical* network where nodes are given an initial *flow* and a *capacity*, and failure of a node results with redistribution of its flow to the remaining nodes, upon which further failures might take place due to *overloading* (i.e., the flow of a node exceeding its capacity). To the best of our knowledge, this is the first work in the literature that studies interdependence between networks with fundamentally different intra-dependency; most existing works are focused on the interdependency between two physical networks [11], or two cyber-networks [2].

For simplicity, the interdependence across the two networks is assumed to be one-to-one; i.e., every node in the cyber-network is dependent upon and supports a single node in the physical network, and vice versa; see Figure 1. More precisely, we assume that for each $i = 1, \ldots, N$, nodes a_i and b_i are dependent on each other meaning that if one fails, the other will fail as well. Although simplistic, the one-to-one interdependence model is considered to be a good starting point and has already provided useful insights in similar settings [2]; more complicated interdependence models shall be considered in future work including regular allocation strategy, i.e., each node in A is connected to k nodes in B and vice versa, or a more general case where some nodes do not have interdependent links and can function even without any support from the other network.

Intra-dependency in Network A. Let network A represent a flow network on nodes $a_1, \ldots, a_N$. Each node a_i is given an initial load (e.g., power flow) $L_1, \ldots, L_N$. The *capacity* C_i of node a_i defines the maximum flow that it can sustain, and is given by

$$C_i = L_i + S_i, \qquad i = 1, \ldots, N, \tag{1}$$

where S_i denotes the *free-space* (or, redundancy) available to node a_i. It is assumed that a node *fails* (i.e., outages) if its load exceeds its capacity at any given time. The key assumption of our intra-dependency model for network

A is that when a node fails, the load it was carrying is redistributed *equally* among all remaining nodes. This leads to an increase in load carried by all remaining nodes, which in turn may lead to further failures of overloaded nodes, and so on, potentially leading to a cascade of failures.

The equal load redistribution rule takes its roots from the *democratic* fiber bundle model [12], [13], and has been recently used by Pahwa et al. [14] in the context of power systems; see also [8], [15]. 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 [16]. Throughout we assume that the load and free-space 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 $i = 1, \ldots, N$. The corresponding (joint) probability density function is given by $p_{LS}(x, y) = \frac{\partial^2}{\partial x \partial y} P_{LS}(x, y)$. In order to avoid trivial cases, we assume that $S_i > 0$ and $L_i > 0$ with probability one for each a_i. Finally, we assume that the marginal densities $p_L(x)$ and $p_S(y)$ are continuous on their support.

Intra-dependency in Network B. Let network B represent a cyber (e.g., communication) network consisting of nodes $b_1, \ldots, b_N$. In this network, we assume that a node keeps functioning as long as it belongs to the largest (i.e., *giant*) connected component of the network. If a node loses its connection to the giant core of the network, then it is assumed to have failed and can no longer carry out its functions. This percolation-based failure rule, though not suitable for *physical* systems carrying a flow, can be regarded as a reasonable model for *cyber*-networks (e.g., sensor networks) where connectivity to a giant core would be crucial for a node's capability to deliver its tasks. Robustness of networks under the giant-component based failure model has been extensively analyzed in the case of *single* networks [17]–[19]. The focus has recently been shifted towards *interdependent* networks with the work of Buldyrev et al. [2], where robustness of two interdependent networks, both operating under the giant-component based intra-dependence rule, was studied. Their model, and most works that follow, are unable to capture the true nature of a cyber-physical network, where the cyber-network and the physical-network should obey a different set of rules determining their intra-dependencies.

We define the structure of the network B through its *degree distribution*, namely the probabilities $\{d_k, k = 0, 1, \ldots\}$ that an arbitrary node in B has degree k; clearly, we need to have $\sum_{k=0}^{\infty} d_k = 1$. In particular, each node $b_1, \ldots, b_N$ is assigned a degree drawn from the distribution $\{d_k\}_{k=0}^{\infty}$ independently from any other node. Once the degree sequence, $\text{degree}(b_1), \ldots, \text{degree}(b_N)$, of the network is determined, network B is constructed by selecting *uniformly at random* a graph among all graphs on N nodes with the given degree sequence; see [19], [20] for details of such constructions. This class of networks is known in the literature as the *configuration model* or *random graphs with arbitrary degree distribution*. Degree distribution is often regarded as the core property defining a graph, and random networks with arbitrary degree distributions are extensively used as a starting point in the literature on robustness of complex networks.

Interdependent System Model. With the intra-dependency models of both networks specified, we adopt a one-to-one inter-dependency model across networks A and B; i.e., nodes a_i and b_i depend on each other for each $i = 1, \ldots, N$. With these in mind, we are interested in understanding the dynamics of cascading failures in this interdependent system, where failures are initiated by removing a $(1 - p)$-fraction of nodes, selected *randomly*, from network A . As explained before, we assume an asynchronous cascade model, where intra-propagation and inter-propagation of failures are considered in a sequential manner. At any stage $t = 1, 2, \ldots$ of the cascade process, a node a_i in network A will still be functioning if and only if (i) its current flow at time t is less than its capacity; *and* (ii) its counterpart b_i in network B is still functioning (which is equivalent to b_i being contained in the largest connected subgraph of B). Similarly, a node b_j in network B survives cascade step t if and only if i) it belongs to largest connected component of B at time t; and (ii) its counterpart a_j in network A is still functioning (which is equivalent to a_j carrying a flow at time t that is less than its capacity).

Since the cascade process is monotone, a steady-state will eventually be reached, possibly after all nodes have failed. Let $\mathcal{N}_{\text{surviving}} \subset \{1, \ldots, N\}$ be the set of node id's that are still functioning at the steady state. In other words, the surviving interdependent system will consist of nodes $\{a_i : i \in \mathcal{N}_{\text{surviving}}\}$ where each a_i has more capacity than its flow and $\{b_i : i \in \mathcal{N}_{\text{surviving}}\}$ that constitutes a connected subgraph of network B. The primary goal of this paper is to derive the *mean* fraction of nodes that survive the cascades as a function of the initial attack size $1-p$, in the asymptotic limit of large network size N. More precisely, we would like to characterize $S(p)$ defined as

$$S(p) := \lim_{N \to \infty} \frac{\mathbb{E}\left[|\mathcal{N}_{\text{surviving}}(p)|\right]}{N}$$

In what follows, we present our main result that allows computing $S(p)$ under any degree distribution $\{d_k\}_{k=0}^{\infty}$ of the cyber-network B, any load-(free-space) distribution $P_{LS}(x, y)$ of the physical network A, and under any attack size $0 \leq p \leq 1$. This is followed in Section IV by a numerical study that demonstrates the accuracy of our analysis even with finite N, and presents insights on how the robustness of an interdependent CPS can be improved by careful allocation of available resources (e.g., node capacities and degrees).

III. Main Result

Our main result is presented next. The approach is based on recursively deriving the *mean* fraction of surviving nodes from both networks at each stage $t = 1, 2, \ldots$ of the cascade process. The cascade process starts at time $t = 0$ with a random attack that kills $1 - p$ fraction of the nodes from network A. As mentioned earlier, we assume

an asynchronous cascading failure model where at stages $t = 1, 3, \ldots$ we consider the failures in network A and in stages $t = 2, 4, \ldots$ we consider the failures in network B. In this manner, we keep track of the subset of vertices $A_1 \supset A_3 \supset \ldots \supset A_{2i+1}$ and $B_2 \supset B_4 \supset \ldots \supset B_{2i}$ that represent the functioning (i.e., surviving) nodes at the corresponding stage of the cascade. We let f_{A_i} denote the *relative* size of the surviving set of nodes from network A at stage i, i.e.,

$$f_{A_i} = \frac{|A_i|}{N}, \quad i = 1, 3, 5, \ldots$$

We define f_{B_i} similarly as

$$f_{B_i} = \frac{|B_i|}{N}, \quad i = 2, 4, 6, \ldots$$

Our main result, presented next, shows how these quantities can be computed in a recursive manner.

Theorem 3.1: Consider an interdependent system as described in Section II, where the load and free-space values of nodes $a_1, \ldots, a_N$ are drawn independently from the distribution p_{LS}, and network B is generated according to the configuration model with degree distribution $\{d_k\}_{k=0}^{\infty}$; i.e., we have $\mathbb{P}[\text{degree of node } b_i = k] = d_k$ for each $k = 0, 1, \ldots$ and $i = 1, \ldots, N$. Let mean degree be denoted by $\langle d \rangle$, i.e., let $\langle d \rangle = \sum_{k=0}^{\infty} k d_k$. With $f_{B_0} = p_{B_0} = p$, $f_{A_{-1}} = 1$, and $Q_{-1} = 0$, the relative size of the surviving parts of network A and B at each stage of the cascade, initiated by a random attack on $1 - p$ fraction of the nodes, can be computed recursively as follows for each $i = 0, 1, \ldots$

$$p_{A_{2i+1}} = \frac{f_{B_{2i}}}{f_{A_{2i-1}}} \tag{2}$$

$$Q_{2i+1} = Q_{2i-1} + \min\left\{x \in (0, \infty] : \frac{\mathbb{P}[S > Q_{2i-1} + x]}{\mathbb{P}[S > Q_{2i-1}]}(x + Q_{2i-1} + \mathbb{E}[L \mid S > x + Q_{2i-1}]) \geq \frac{Q_{2i-1} + \mathbb{E}[L \mid S > Q_{2i-1}]}{p_{A_{2i+1}}}\right\} \tag{3}$$

$$f_{A_{2i+1}} = f_{A_{2i-1}} \cdot p_{A_{2i+1}} \cdot \mathbb{P}[S > Q_{2i+1} \mid S > Q_{2i-1}] \tag{4}$$

$$p_{B_{2i+2}} = p_{B_{2i}} \frac{f_{A_{2i+1}}}{f_{B_{2i}}} \tag{5}$$

$$u_{2i+2} = \max\left\{u \in [0, 1] : u = 1 - \sum_{k=0}^{\infty} \frac{k d_k}{\langle d \rangle}(1 - u \cdot p_{B_{2i+2}})^{k-1}\right\} \tag{6}$$

$$f_{B_{2i+2}} = p_{B_{2i+2}}\left(1 - \sum_{k=0}^{\infty} d_k \left(1 - u_{2i+2} \cdot p_{B_{2i+2}}\right)^k\right) \tag{7}$$

The notation used in Theorem 3.1 is summarized in Table I. In these iterations, it is assumed that if at any stage i, it happens to be the case that no $x < \infty$ satisfies the inequality at (3), we set $Q_{2i+1} = \infty$. It is then understood that the entire network A (and thus B) have failed, and we get $f_{A_{2i+1}} = f_{B_{2i+2}} = 0$. Similarly, it can be seen that the equality in (6) always holds with $u = 0$. Thus, if at any stage i, there is no $u > 0$ satisfying the equality in (6), we will get $u_{2i+2} = 0$ leading to $f_{B_{2i+2}} = 0$; i.e., the entire network B (and thus A) will have collapsed.

A_i	set of surviving nodes in network A at $i = 1, 3, 5, \ldots$
B_i	set of surviving nodes in network B at $i = 2, 4, 6, \ldots$
$f_{A_{2i+1}}$	fraction $\lvert A_{2i+1}\rvert/N$ of surviving nodes in A at $2i+1$
$f_{B_{2i+2}}$	fraction $\lvert B_{2i+2}\rvert/N$ of surviving nodes in B at $2i+2$
Q_{2i+1}	extra load per surviving node in A at $2i+1$
$p_{A_{2i+1}}$	prob. of a node in A_{2i-1} surviving *inter*-failures at $2i$
$1 - p_{B_{2i+2}}$	*equivalent* prob. of random attack to B that gives B_{2i+2}
u_{2i+2}	auxiliary variable used in computing $f_{B_{2i+2}}$

TABLE I

KEY NOTATION IN THE ANALYSIS OF CASCADING FAILURES

As mentioned before, our goal is to obtain the *final* system size, i.e., the relative size of the surviving nodes at the steady-state. In view of the one-to-one interdependence model, the surviving size of the networks A and B will be the same at the steady-state. Thus, we conclude that $S(p) = \lim_{i\to\infty} f_{A_i} = \lim_{i\to\infty} f_{B_i}$.

Next, we provide an outline of the proof. In [8], we already analyzed the cascade dynamics and derived the final system size in a single flow carrying network (similar to network A in our analysis), when $1 - p$ fraction of its nodes are randomly removed; the result enables computing the final system size in terms of the initial attack size $1 - p$, as well as the load and free space distribution $P_{LS}(x, y)$. The results established in [8] are incorporated in the recursions above through expression (3) that allows us to calculate, in a recursive manner, the extra load that each of the surviving nodes at a particular stage will be carrying in addition to their initial load.

According to the failure propagation model described at the beginning of this section, at odd stages failures from

network B can propagate to network A, causing a fraction of nodes to be removed. Given that the intra-failure dynamics of network B is completely independent from network A, the impact of the failures in B to network A will be equivalent to a *random* attack launched on A. In addition, at each odd stage $t = 2i - 1$, $i = 1, 2, \cdots$, we can treat the remaining part of network A as a new physical network A_{2i-1}, with the appropriately updated size and load-'free-space' distribution. Thus, the random removal of nodes caused by failures in network B (through the one-to-one interdependency links) from last cascade stage can be viewed as a new random attack to A_{2i-1} that keeps only $p_{A_{2i+1}}$ fraction of its nodes alive. Then following a similar approach, we can compute the size of network A at the next stage $2i + 1$, i.e., $f_{A_{2i+1}}$. An important observation is the need to update the load and free-space distributions for each new network A_{2i+1} to incorporate the facts that the surviving nodes in A_{2i+1} are added with Q_{2i-1} amount of extra load, and at the same time the free-space of each surviving node must be at least Q_{2i-1}. We can show that the changes of the distribution can be represented by the initial load and free-space distribution with Q_{2i+1} representing the extra load in each stage. In other words, each time failures propagate between the two networks, network A will shrink to a group of nodes that have a higher free space and that are now carrying more load. The fractional size of this surviving subset of nodes at each time stage can be computed via the equivalent attack size $p_{A_{2i+1}}$ (caused by failures in network B propagated via the one-to-one dependent links), extra load Q_{2i+1} and the load free-space distribution $P_{LS}(x, y)$; see (2)-(4).

Following the same approach, in network B we treat each new failure that comes from network A as a new random attack (or failure) on the existing network B_{2i+2}. For a node in network B to function, it must belong to the giant component, so actually the functioning network B_{2i+2} at time stage $t = 2i + 2, i = 0, 1, 2, \cdots$ is the giant component after the random attack propagated from network A. A key insight here is that the sequential process of applying a first random attack on the cyber-network, then computing the giant component, and then applying a second random attack and then computing the giant component is *equivalent* to (in terms of the fractional size of the set of nodes that survives) the process where the second random attack is applied directly after the first one without computing the giant component [2]. This way, the result of a series of random attack/giant-component calculation processes can be emulated by a single random attack/giant-component calculation, with an appropriately calculated *equivalent* random attack size. In our calculations, this *equivalent* attack size for stage $2i + 2$ is represented by $1 - p_{B_{2i+2}}$ and can be computed recursively as given in (5). This formula is based on treating all *new* failures propagated from network A in the following time stage as the new random attack size launched on B, which is then used to update the equivalent attack size $1 - p_{B_{2i+2}}$ that will be used to emulate the entire cascade sequence up until that stage. Then, the size of network B_{2i+2}, namely the size of the giant component after randomly removing $(1 - p_{B_{2i+2}})$-fraction of nodes, can be computed using the technique of generating functions [2], [18], [19], [21], [22]. The formulas that give the network size $f_{B_{2i+2}}$ at $i = 0, 1, \ldots$ are presented at (6) and (7).

Once we know how to compute the surviving network sizes $f_{A_{2i+1}}$ and $f_{B_{2i+2}}$ at each stage, the propagation of failures between the two networks is seen to be governed via (2) and (5) that reveal how the key quantities p_{A2i+1} and $p_{B_{2i+2}}$ used in computing $f_{A_{2i+1}}$ and $f_{B_{2i+2}}$, respectively, need to be updated based on the result of the last cascade stage. Collecting, a thorough analysis that reveals a full understanding of the system behavior and robustness during the failure process is presented in equations (2)-(7).

IV. Numerical Results

In this section, we confirm our analytic results through numerical simulations under a wide range of parameter choices, with a particular focus on checking the accuracy of the results when the network size N is finite. For physical networks carrying a certain flow (i.e., network A in our analysis), we consider different combinations of probability distributions for the load and free-space variables. Throughout, we consider three commonly used families of distributions: i) Uniform, ii) Pareto, and iii) Weibull. These distributions are chosen here because they cover a wide range of commonly used and representative cases. In particular, uniform distribution provides an intuitive baseline. Distributions belonging to the Pareto family are also known as a *power-law* distributions and have been observed in many real-world networks including the Internet, the citation network, as well as power systems [23]. Weibull distribution is widely used in engineering problems involving reliability and survival analysis, and contains several classical distributions as special cases; e.g., Exponential, Rayleigh, and Dirac-delta.

As explained in Section II-B, the cyber-network where a node is only functional when it belongs to the giant component (i.e., network B in our analysis) is generated according to the configuration model with degree distribution $\{d_k\}_{k=0}^{\infty}$. In the simulations, we consider two representative cases given below:

- Erdős-Rényi (ER) network model [24]. This corresponds to having the degree distribution d_k follow a Binomial distribution, i.e., $d_k \sim \text{Binomial}(N-1; \frac{\langle d \rangle}{N-1})$; as before $\langle d \rangle$ gives the mean node degree.
- The scale-free (SF) network model [17]. We consider the case where the degree distribution $\{d_k\}_{k=0}^{\infty}$ is a *power-law* with *exponential cut-off*, which was observed [25] in many real networks including the Internet; i.e., we have

$$d_k = \begin{cases} 0 & \text{if } k = 0 \\ \frac{1}{\text{Li}_\gamma(e^{-1/\Gamma})} k^{-\gamma} e^{-k/\Gamma} & \text{if } k = 1, 2, \cdots, \end{cases} \tag{8}$$

where γ is the power exponent, Γ is the cut-off point, and $\text{Li}_m(z) := \sum_{k=1}^{k=\infty} z^k k^{-m}$ is the normalizing constant.

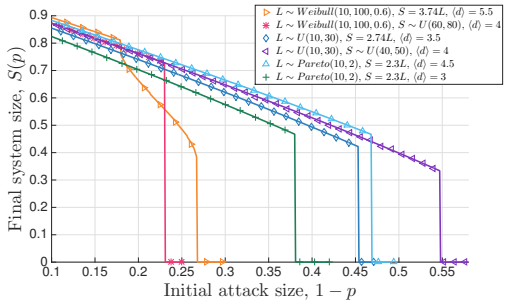

Fig. 2. *Final system size under different network settings, including different load-free space distributions in the physical network and different mean degree in the ER cyber network. Analytic results are represented by lines, whereas simulation results are represented by symbols (averaged over 100 independent runs). We see that in each case theoretical results match the simulation results very well.*

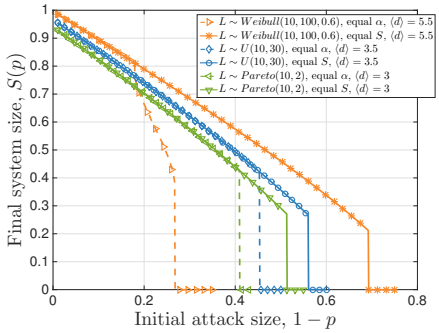

Fig. 3. *Final system size under equal free space (solid lines with symbols) or equal tolerance factor (dashed lines with symbols) when network B is a ER graph with fixed mean degree. The symbols are empirical results over 100 independent runs on network size $N = 10^5$, and lines (dashed or solid) represent analytic results. We can see in all cases equal free space greatly improves system robustness by allowing the system to sustain a larger initial attack size and still not collapsing.*

We remind that although we restrict our attention to these special cases in the simulations, our analysis applies under more general degree distributions as well.

A. Fiber Network Coupled with ER Network

The Erdős-Rényi graph is one of the most basic and widely used network models and often serve as a starting point in simulations. In our study, we start with N nodes, and connect each pair of vertices with an edge with probability $\langle d \rangle/(N-1)$ independently from each other. When N is large, this is equivalent to generating the network via the configuration model using a *Poisson* degree distribution with mean $\langle d \rangle$.

First, we confirm our main result presented in Sec. III concerning the final system size $S(p)$, i.e., the mean fraction of surviving nodes at the end of cascading failures initiated by a random attack that removes $1-p$ fraction of nodes in network A. In all simulations, the number of nodes in both networks is $N = 10^5$, and for each set of parameters considered, we run 100 independent experiments. The results are shown in Figure 2 where symbols represent the *empirical* value of the final system size $S(p)$, and lines represent the analytic results computed via (2)–(7). We see that theoretical results match the simulations very well in all cases. This suggests that although asymptotic in nature, our main result can still be helpful when the network size N is finite.

The plots in Figure 2 show how different load-free space distributions in network A as well as the mean degree in network B affect the system behavior. For example, with the mean degree of network B being $\langle d \rangle = 4$, when load in network A follows Weibull distribution (magenta asterisk) , the final system size drops to zero at a point where the attack size is around 0.23, meaning that any random attack that kills more than 23% of the nodes will destroy the entire system. On the other hand, if the load and free space follows Uniform distribution (purple triangle), the system is quite robust and can sustain initial attack sizes up to 0.55 without collapsing. Similarly, when we fix the distribution in network A, we can see the effect of mean degree in the cyber network on system robustness: when initial load in physical network follows Pareto$(10, 2)$ distribution, and free space is given by $S = \alpha L$ with $\alpha = 2.3$, we see that increasing mean degree of network B from $\langle d \rangle = 3$ (green cross) to $\langle d \rangle = 4.5$ (light blue triangle) leads to a substantial increase on the final system size at *all* attack sizes; i.e., the interdependent CPS becomes more robust. This is intuitive since higher $\langle d \rangle$ values lead to a cyber-network B with higher levels of connectivity enabling the entire CPS to sustain larger attacks while maintaining a larger fraction of nodes in its giant component. An interesting observation from Figure 2 is that in all cases, the final drop of the system size to zero takes place through a *first-order* (i.e., discontinuous) transition[1], making it difficult to predict system behavior from previous data. In fact, this abrupt failure behavior is 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.

From a design perspective, it is of interest to understand how the robustness of the interdependent system can be improved or even maximized under certain constrains. To gain insights on this, we fix the mean degree in network B (the cyber network), and explore the effect of the allocation (i.e., distribution) of node capacities in the physical network. A key determining factor of system robustness is expected to be the free-space distribution as it specifies the extra load a node can receive from the failed ones before it fails due to overloading. The vast majority of the literature and most real world applications employ a *linear* free-space allocation scheme where the free-space assigned to a node is set to be a fixed proportion of its initial load; i.e., $S = \alpha L$ where α is the *tolerance factor* and is usually a fixed value [26], [27] used for the entire network. We already showed in [8] that in a single flow-carrying network, allocating every node exactly the same free-space leads to a higher robustness than the commonly used setting of equal tolerance factor. In fact, in the single network case, the robustness is shown to be *maximized* when all nodes receive the same free space. Our numerical simulations, presented in Figure 3, shows that the above conclusion still applies in interdependent networks. Namely, assigning every node the same free space provides

[1]The nomenclature concerning the order of transitions is adopted from the studies on phase transition in Physics; simply put, first (resp. second) order transitions are associated with *discontinuous* (resp. continuous) variations.

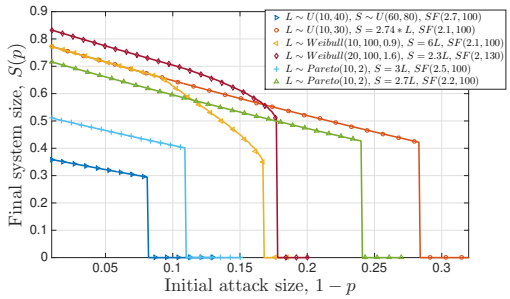

Fig. 4. *Final system size under different network settings, including different load-free space distributions in the physical network and different exponent in the scale-free cyber network. Analytic results are represented by lines, whereas simulation results are represented by symbols (averaged over 100 independent runs). We see that in each case theoretical results match the simulation results very well.*

a much better overall system robustness as compared to the widely used setting of equal tolerance factor (i.e., linear free-space allocation). To provide an overall evaluation of the system robustness, we define the critical attack size $1 - p^\star$ as the minimum attack size that breaks down the whole system. Thus, the larger $1 - p^\star$ is, the more robust will the system be since it can sustain larger attacks. In Figure 3, the comparison between the equal free-space and equal tolerance factor allocations are made with the mean free space $\mathbb{E}[S]$ being fixed (i.e., the total free space in the network is constrained). We see that compared to the equal tolerance factor scheme, the equal free-space allocation enables the system to sustain much larger attacks. In fact, in the case of Weibull distribution, the robustness is almost 2.5 times higher in the case of equal free space as compared to the case with equal tolerance factor; i.e., $1 - p^\star = 0.7$ vs. $1 - p^\star = 0.28$.

B. Fiber Network Coupled with SF Network

Although the ER graph constitutes a simple and useful network model, networks in most real-world applications might have significantly different structure and robustness behavior against attacks. For instance, scale-free networks (SF model) were shown [28] to exhibit fundamentally different robustness behavior with ER networks; the former is very robust against random attacks but fragile against *targeted* attacks, while the situation is exactly the opposite for the latter. In order to better understand the impact of the topology of the cyber-network on the overall robustness of an interdependent CPS, we consider in this section the case where the cyber network (network B) has a power-law degree distribution with exponential cutoff. In addition to being observed in many real-world networks including the Internet, power-law distributions with exponential cut-off also ensure that all moments of the node degree are *finite*, which helps certain convergences take place faster (i.e., with smaller N).

In Figure 4, we verify our analytic results when network B (cyber network) has a degree distribution in the form of a power-law with exponential cut-off; this is denoted by $SF(\gamma, \Gamma)$, where γ is the power exponent and Γ is the cut-off parameter given at (8). In all cases, we fix the number

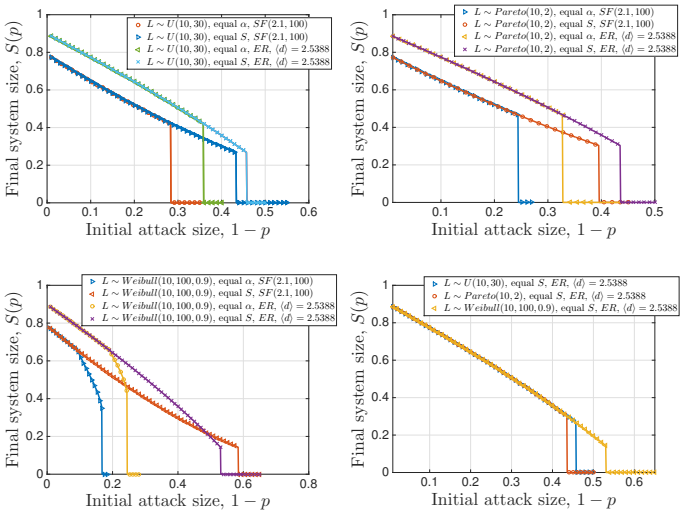

Fig. 5. *Comparison of final system size when equal tolerance factor (equal α) and equal free space (equal S) schemes are used. The mean value of free space is kept the same, as well as the mean degree in SF and ER networks. In all cases, equal S outperform the widely used equal α scheme. The effect of topology in the cyber network is not unitary: in some cases ER leads to better robustness, while in other cases SF is better, contradicting the results [17] concerning the robustness of single networks.*

of nodes in both networks to be $N = 10^6$, and consider several different load-'free space' distributions in network A and different (γ, Γ) values for network B (while noting that in real-world networks, it is often observed that $2 < \gamma < 3$). The simulation results are obtained by averaging over 100 independent experiments for each data point and it is seen that they are in very good agreement with the analytic results.

Next, we seek to obtain an overall understanding of how the free-space allocation in the physical network together with the topology of the cyber network (ER vs. SF model) affect the system robustness. With the discussion from Section IV-A in mind, we consider the widely used equal tolerance factor allocation (equal α), where the free-space S is a fixed factor α of the load on a node (i.e., $S = \alpha * L$) and the equal free-space allocation scheme (equal S) that was shown [8] to be optimal in a single physical network. For fairness, all comparisons are made under the same initial load distribution, and with the mean free-space in network A and the mean node degree in network B being fixed. The results are shown in Figure 5. We can see that no matter how the initial load is distributed, i.e., whether it's Uniform, Pareto or Weibull, and despite of the structure of the cyber network being SF network or ER network, equal free-space allocation can greatly improve system robustness as compared to the equal tolerance factor allocation. We observe that when all nodes in network A are given the same free-space, the overall interdependent CPS can sustain a much larger initial attack size without collapsing; i.e., it has a much larger critical attack size. For example in the case of Weibull distributed load with SF network, the system can only take around 16.8% of initial attack size when using equal α, but can sustain a initial attack that removes 58% of the nodes when equal S is used, making the system about 3 times more robust in terms of the critical attack size.

We also see in Figure 5 that the topology of the cyber network affects the robustness of the interdependent CPS in an intricate way, with some cases showing the exact opposite of what would have been expected from the results on single networks. In particular, SF networks are known [17] to be more robust than ER networks against random attacks. This is often attributed to the fact that SF networks typically have a few nodes with very high degrees and the network will likely contain a large connected component unless these high-degree nodes are removed (which is unlikely to happen if the attack is *random*); this dependence on a few nodes is exactly what makes SF networks very fragile against a *targeted* attack. In the case of the interdependent CPS model, we see that the comparison of the overall robustness between the cases where they cyber network is SF or ER is a much more complicated matter. In fact, depending on the load-'free-space' distribution in the physical network, the cyber network being SF does *not* always lead to a better robustness than the case with ER. For example, in the upper two plots in Figure 5 where the initial load is Uniform and Pareto, respectively, the cases with the ER network leads to a better robustness than that with SF. In the bottom left picture where the initial load in the physical networks is Weibull, the situation is even more intricate. With equal α, the case where the cyber-network is ER leads to a better robustness, while SF network performs better (in terms of the critical attack size) under the equal-S allocation. This shows that an integrated CPS can not be designed in the most robust way by considering the physical and cyber counterparts separately. Instead, a holistic design approach is needed where the robustness of the CPS as a whole is considered.

V. Conclusion

We studied the robustness of an interdependent system against cascading failures initiated by a random attack. This is done through a novel model where the constituent networks exhibit inherently different intra-dependency characteristics. In particular, inspired by many applications of interdependent CPSs, our model consists of a flow network where failure leads to flow redistribution and possible further failures due to *overloading*, and a cyber-network where nodes need to be a part of the giant component to be functional. Through the dynamic relations of cascading failures, we derive the mean fraction of nodes that ultimately survive the cascade as a function of the initial attack size. We confirm our analysis and derive useful insights concerning the robustness of interdependent CPSs through extensive numerical simulations. There are many open directions for future work; one can consider k inter-links instead of one-to-one interdependent links. One might also consider a heterogeneous allocation of inter-links and study the optimal allocation scheme under certain constraints [7]. It would also be interesting the consider more complicated flow redistribution models based on network topology.

References

[1] A. Vespignani, "Complex networks: The fragility of interdependency," *Nature*, vol. 464, pp. 984–985, 2010.

[2] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, p. 1025, 2010.

[3] V. Rosato, L. Issacharoff, F. Tiriticco, S. Meloni, S. Porcellinis, and R. Setola, "Modelling interdependent infrastructures using interacting dynamical models," *International Journal of Critical Infrastructures*, no. 1, pp. 63–79, 01.

[4] S. V. Buldyrev, N. W. Shere, and G. A. Cwilich, "Interdependent networks with identical degrees of mutually dependent nodes," *Phys. Rev. E*, vol. 83, no. 016112, 2011.

[5] R. Parshani, S. V. Buldyrev, and S. Havlin, "Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition," *Phys. Rev. Lett.*, vol. 105.

[6] X. Huang, J. Gao, S. Buldyrev, S. Havlin, and H. E. Stanley, "Robustness of interdependent networks under targeted attack," *Phys. Rev. E*, vol. 83, no. 6, 2011.

[7] O. Yağan, D. Qian, J. Zhang, and D. Cochran, "Optimal Allocation of Interconnecting Links in Cyber-Physical Systems: Interdependence, Cascading Failures and Robustness," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1708–1720, 2012.

[8] Y. Zhang and O. Yağan, "Optimizing the robustness of electrical power systems against cascading failures," *Scientific reports*, vol. 6, 2016.

[9] S. Chattopadhyay, H. Dai, D. Y. Eun, and S. Hosseinalipour, "Designing optimal interlink patterns to maximize robustness of interdependent networks against cascading failures," *IEEE Transactions on Communications*, vol. 65, no. 9, pp. 3847–3862, Sept 2017.

[10] Y. Zhang, A. Arenas, and O. Yağan, "Cascading failures in interdependent systems under a flow redistribution model," *Phys. Rev. E*, vol. 97, p. 022307, Feb 2018.

[11] A. Scala, P. G. D. S. Lucentini, G. Caldarelli, and G. DAgostino, "Cascades in interdependent flow networks," *Physica D: Nonlinear Phenomena*, vol. 323, pp. 35–39, 2016.

[12] J. V. Andersen, D. Sornette, and K.-t. Leung, "Tricritical behavior in rupture induced by disorder," *Physical Review Letters*, vol. 78, no. 11, p. 2140, 1997.

[13] H. Daniels, "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, no. 995. The Royal Society, 1945, pp. 405–435.

[14] S. Pahwa, C. Scoglio, and A. Scala, "Abruptness of cascade failures in power grids," *Scientific reports*, vol. 4, 2014.

[15] O. Yağan, "Robustness of power systems under a democratic-fiber-bundle-like model," *Phys. Rev. E*, vol. 91, p. 062811, Jun 2015.

[16] "Cascade-based attack vulnerability on the {US} power grid," *Safety Science*, vol. 47, no. 10, pp. 1332 – 1336, 2009.

[17] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[18] M. E. Newman, "Spread of epidemic disease on networks," *Physical review E*, vol. 66, no. 1, p. 016128, 2002.

[19] M. E. Newman, S. H. Strogatz, and D. J. Watts, "Random graphs with arbitrary degree distributions and their applications," *Physical review E*, vol. 64, no. 2, p. 026118, 2001.

[20] B. Bollobás, *Random Graphs*. Cambridge (UK): Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2001.

[21] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, "Networks formed from interdependent networks," *Nature physics*, vol. 8, no. 1, p. 40, 2012.

[22] J. Shao, S. V. Buldyrev, S. Havlin, and H. E. Stanley, "Cascade of failures in coupled network systems with multiple support-dependence relations," *Physical Review E*, vol. 83, no. 3, p. 036116, 2011.

[23] S. Pahwa, A. Hodges, C. Scoglio, and S. Wood, "Topological analysis of the power grid and mitigation strategies against cascading failures," in *Systems Conference, 2010 4th Annual IEEE*. IEEE, 2010, pp. 272–276.

[24] P. Erdös and A. Rényi, "On random graphs, i," *Publicationes Mathematicae (Debrecen)*, vol. 6, pp. 290–297, 1959.

[25] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Rev.*, no. 4, pp. 661–703, Nov.

[26] A. E. Motter and Y.-C. Lai, "Cascade-based attacks on complex networks," *Phys. Rev. E*, vol. 66, p. 065102, Dec 2002.

[27] P. Crucitti, V. Latora, and M. Marchiori, "Model for cascading failures in complex networks," *Phys. Rev. E*, vol. 69, p. 045104, Apr 2004.

[28] R. Albert, H. Jeong, and A.-L. Barabási, "Internet: Diameter of the world-wide web," *nature*, vol. 401, no. 6749, p. 130, 1999.