

Suppressing cascades of load in interdependent networks

Charles D. Brummitt^{a,b,1}, Raissa M. D'Souza^{b,c,d,e}, and E. A. Leicht^f

^aDepartment of Mathematics, ^bComplexity Sciences Center, ^cDepartment of Mechanical and Aerospace Engineering, and ^dDepartment of Computer Science, University of California, Davis, CA 95616; ^eSanta Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501; and ^fComplex Agent-Based Dynamic Networks Complexity Centre, Saïd Business School, University of Oxford, Park End, Oxford OX1 1HP, United Kingdom

AUTHOR SUMMARY

Modern society depends on increasingly interdependent systems that are prone to widespread failure. For example, transportation, communication, power grids, and other infrastructures support one another and the world's interconnected economies. Barrages of incidents large and small—downed power lines, grounded aircrafts, natural disasters, and the like—cause avalanches of repercussions that cascade within and among these systems (1). Although interdependence confers benefits, its effect on the risk of failure at the individual system level and at the collective level remains poorly understood. Here we analyze how the interconnectivity (interdependence) between networks affects the sizes of their cascades of load

shedding. For networks derived from interdependent power grids, we show that interdependence can have a stable equilibrium. An isolated network suppresses its large cascades by connecting to other networks, but too many interconnections exacerbate its largest cascades—and those of the whole system. We develop techniques to estimate this optimal amount of interconnectivity, and we examine how differences in network capacity and load affect this equilibrium. Our framework advances the current mathematical tools for analyzing dynamical processes on interdependent (or modular) networks, and it improves our understanding of systemic risk in coupled networks.

In the basic process that we consider, a system contains many elements, each of which sheds its load onto neighboring elements whenever it reaches its capacity. This is captured by the classic sandpile model of Bak–Tang–Wiesenfeld, a paradigm for the power law statistics of cascades in many disciplines, from neuronal avalanches to financial instabilities to electrical blackouts (2). In a basic formulation on a graph of nodes and edges, each node has a capacity for holding grains of sand (interpreted here as load or stress). Grains of sand are dropped randomly on nodes, and whenever a node has more sand than its capacity, it topples and sheds all its sand onto its neighbors, which may in turn have too much sand and topple, and so on. Thus, dropping a grain of sand can cause an avalanche (cascade) of toppling. These avalanches, like blackouts in power grids (3), occur in sizes

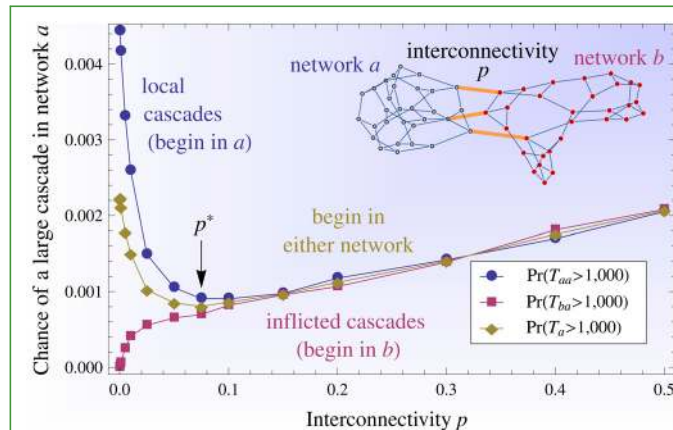


Fig. P1. The chance that a network *a* coupled to another network *b* suffers a cascade spanning more than half the network (gold curve) has a stable minimum at a critical amount of interconnectivity p^* . Networks seeking to mitigate their largest cascades would prefer to build or demolish interconnections to operate at this critical point p^* . The blue (red) curve is the chance that a cascade that begins in *a* (*b*) topples at least 1,000 nodes in *a*. Increasing interconnectivity only exacerbates the cascades inflicted from *b* to *a* (red), but interestingly, it initially suppresses the local cascades in *a*. (From simulations on coupled random three-regular graphs; the inset depicts a small example with 30 nodes per network and $p = 0.1$).

characterized by a power law distribution: They are often small but occasionally enormous.

The Bak–Tang–Wiesenfeld model was originally formulated on a lattice. Given the relevance of networked systems, the dynamics have recently been studied on isolated networks, but not yet on interdependent (or modular) networks. Here we study it on two networks with sparse connections between them. The two networks model either two different types of infrastructures or two modules within one infrastructure, and the interconnections between them model their interdependence. We explicitly study networks extracted from two interdependent power grids in the southeastern United

States and an idealization of them that is more amenable to mathematical study. In this idealization, each node is connected to a node in the other network with probability p (Fig. P1, *Inset*).

Our main result is that interdependence can have a stable equilibrium (Fig. P1). Some interconnectivity is beneficial to an individual network because the other network acts as a reservoir for extra load. The gold curve of Fig. P1 shows that the chance of a large cascade in a network can be reduced by 70% by increasing the interconnectivity p from 0.0005 to 0.075. Too much interdependence, however, becomes detrimental for two reasons. First, new interconnections open pathways for the neighboring network to inflict additional load. Second, each interconnection augments the system's capacity, making available more load that fuels even larger cascades in each network. As a result, the chance of a large

Author contributions: C.D.B., R.M.D., and E.A.L. designed research; C.D.B., R.M.D., and E.A.L. performed research; C.D.B., R.M.D., and E.A.L. contributed new reagents/analytic tools; C.D.B. and R.M.D. analyzed data; and C.D.B., R.M.D., and E.A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: cbrummitt@math.ucdavis.edu.

See full research article on page E680 of www.pnas.org.

Cite this Author Summary as: PNAS 10.1073/pnas.1110586109.

cascade in an individual network eventually increases with interconnectivity p , so p^* is a stable minimum.

This second factor—that new interconnections increase a network's capacity for load—has global consequences. With more load available, larger cascades in the system as a whole become possible. Therefore, networks that interconnect to one another to mitigate their own cascades (Fig. P1) may unwittingly cause larger global cascades in the whole system. This consequence of increased capacity is a warning for the interconnections under construction among, for example, different power grids to accommodate long-distance trade and renewable sources of energy (4).

The results in Fig. P1 show that networks suppressing their largest cascades would seek the optimal level of interconnectivity p^* . However, as shown in the main article, building interconnections to reach p^* increases the occurrence of small cascades. Conversely, networks can suppress their smallest cascades the most by seeking isolation, $p = 0$. But suppressing their smallest cascades exacerbates their largest ones (left side of Fig. P1), just as extinguishing small forest fires can incite large ones, and engineering power grids to suppress small blackouts can increase the risk of large ones (3).

Finally, we determine how asymmetry among networks affects the level of interconnectivity that each network considers optimal. For instance, two interconnected power grids may differ in capacity, load, redundancies, demand, susceptibility to line outages, and age of infrastructure. We

capture these differences with a parameter that controls the rates at which cascades begin in either network. We show that, in any asymmetric situation, the equilibrium will be frustrated, with only one of the grids able to achieve its optimal level of interconnectivity.

Determining how interdependence affects the functioning of networks is critical to understanding the infrastructures so vital to modern society. Whereas others have recently shown that interdependence can lead to alarmingly catastrophic cascades of failed connectivity (5), here we show that interdependence also provides benefits, and these benefits can balance the detriments at stable equilibria. We expect that this work will stimulate calculations of critical points in interconnectivity among networks subjected to other dynamics. As critical infrastructures such as power grids, transportation, communication, banks, and markets become increasingly interdependent, resolving the risks of large cascades and the incentives that shape them becomes ever more important.

1. Little RG (2002) Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. *J Urban Technol* 9:109–123.
2. Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. *Phys Rev A* 38:364–374.
3. Dobson I, Carreras BA, Lynch VE, Newman DE (2007) Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos* 17:026103.
4. Joyce C (April 28, 2009) Building power lines creates a web of problems. (National Public Radio), <http://www.npr.org/templates/story/story.php?storyId=103537250>.
5. Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464:1025–1028.

Suppressing cascades of load in interdependent networks

Charles D. Brummitt^{a,b,1}, Raissa M. D'Souza^{b,c,d,e}, and E. A. Leicht^f

^aDepartment of Mathematics, ^bComplexity Sciences Center, ^cDepartment of Mechanical and Aerospace Engineering, and ^dDepartment of Computer Science, University of California, Davis, CA 95616; ^eSanta Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501; and ^fComplex Agent-Based Dynamic Networks Complexity Centre, Saïd Business School, University of Oxford, Park End, Oxford OX1 1HP, United Kingdom

Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved December 5, 2011 (received for review July 5, 2011)

Understanding how interdependence among systems affects cascading behaviors is increasingly important across many fields of science and engineering. Inspired by cascades of load shedding in coupled electric grids and other infrastructure, we study the Bak–Tang–Wiesenfeld sandpile model on modular random graphs and on graphs based on actual, interdependent power grids. Starting from two isolated networks, adding some connectivity between them is beneficial, for it suppresses the largest cascades in each system. Too much interconnectivity, however, becomes detrimental for two reasons. First, interconnections open pathways for neighboring networks to inflict large cascades. Second, as in real infrastructure, new interconnections increase capacity and total possible load, which fuels even larger cascades. Using a multitype branching process and simulations we show these effects and estimate the optimal level of interconnectivity that balances their trade-offs. Such equilibria could allow, for example, power grid owners to minimize the largest cascades in their grid. We also show that asymmetric capacity among interdependent networks affects the optimal connectivity that each prefers and may lead to an arms race for greater capacity. Our multitype branching process framework provides building blocks for better prediction of cascading processes on modular random graphs and on multitype networks in general.

vulnerability of infrastructure | complex networks

Networks that constitute our critical infrastructure increasingly depend on one another, which enables cascades of load, stress, and failures (1–6). The water network, for instance, turns turbines and cools nuclear reactors in the electrical grid, which powers the transportation and communications networks that underpin increasingly interdependent global economies. Barrages of disturbances at different scales—volcanic eruptions (7), satellite malfunctions (4), earthquakes, tsunamis, wars (8)—trigger cascades of load shedding in interdependent transportation, communication, and financial systems. Interdependence can also increase within a particular infrastructure, with consequences for cascading failures. The electrical grid of the United States, for example, consists of over 3,200 independent grids with distinct ownership—some private, others public—and unique patterns of connectivity, capacities, and redundancies (9). To accommodate rising demand for electricity, long-distance trade of energy (10), and new types of power sources (11), interconnections among grids bear ever more load (12), and many new high-capacity transmission lines are planned to interconnect grids in the United States and in Europe (13). Fig. 1 shows the interconnections planned to transport wind power (11). Though necessary, these interconnections affect systemic risk in ways not well understood, such as in the power grid, where the modular structure affects its large cascades. For example, the August 14, 2003 blackout, the largest in North American history, spread from a grid in Ohio to one in Michigan, then to grids in Ontario and New York before overwhelming the northeast (10, 14).

Researchers have begun to model cascades of load and failure within individual power grids using probabilistic models (15), linearized electric power dynamics (16, 17), and game theory (18).

The first models of interdependent grids use simplified topologies and global coupling to find that interconnections affect critical points of cascades (19), which suggests that they may affect the power law distributions of blackout size (12, 15). Models with interconnections among distinct infrastructure have focused on the spread of topological failures, in which nodes are recursively removed (20–22), and not on the dynamical processes occurring on these networks. These models find that interdependence causes alarmingly catastrophic cascades of failed connectivity (20–22). Yet as we show here, interdependence also provides benefits, and these benefits can balance the dangers at stable critical points.

Here we develop a simple, dynamical model of load shedding on sparsely interconnected networks. We study Bak–Tang–Wiesenfeld sandpile dynamics (23, 24) on networks derived from real, interdependent power grids and on sparsely coupled, random regular graphs that approximate the real topologies. Sandpile dynamics are paradigms for the cascades of load, self-organized criticality, and power law distributions of event sizes that pervade disciplines, from neuronal avalanches (25–27) to cascades among banks (28) to earthquakes (29), landslides (30), forest fires (31, 32), solar flares (33, 34), and electrical blackouts (15). Sandpile cascades have been extensively studied on isolated networks (35–41). On interdependent (or modular) networks, more basic dynamical processes have been studied (42–45), but sandpile dynamics have not.

We use a multitype branching process approximation and simulations to derive at a heuristic level how interdependence affects cascades of load. Isolated networks can mitigate their largest cascades by building interconnections to other networks, because those networks provide reservoirs to absorb excess load. Build too many interconnections, however, and the largest cascades in an individual network increase in frequency for two reasons: Neighboring networks inflict load more easily, and each added interconnection augments the system's overall capacity and load. These stabilizing and destabilizing effects balance at a critical amount of interconnectivity, which we analyze for synthetic networks that approximate interdependent power grids. As a result of the additional load introduced by interconnections, the collection of networks, viewed as one system, suffers larger global cascades—a warning for the increasing interdependence among electrical grids (Fig. 1), financial sectors, and other infrastructure (11–13). Finally, we study the effects of capacity and load imbalance. Networks with larger total capacity inflict larger avalanches on smaller capacity networks, which suggests an arms race for greater capacity. The techniques developed here advance

Author contributions: C.D.B., R.M.D., and E.A.L. designed research; C.D.B., R.M.D., and E.A.L. performed research; C.D.B. and E.A.L. contributed new reagents/analytic tools; C.D.B. and R.M.D. analyzed data; and C.D.B., R.M.D., and E.A.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: cbrummitt@math.ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1110586109/-DCSupplemental.

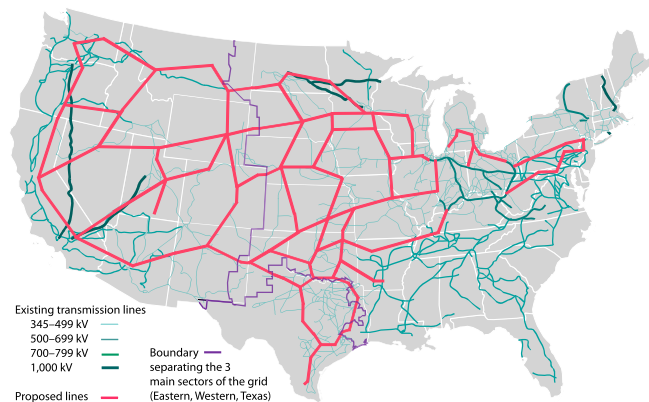


Fig. 1. The power grid of the continental United States, illustrating the three main regions or “interconnects”—Western, Eastern, and Texas—and new lines (in red) proposed by American Electric Power to transport wind power (11, 13).

the theoretical machinery for dynamical processes on multitype networks as well as our heuristic understanding of how interdependence and incentives affect large cascades of load in infrastructure.

Theoretical Formulation

Sandpile Dynamics. Introduced by Bak, et al. in 1987 and 1988 (23, 24), the sandpile model is a well-studied, stylized model of cascades that exhibits self-organized criticality, power laws, and universality classes, which has spawned numerous related models with applications in many disciplines (e.g., refs. 30, 31, 33, 34, and 46). In a basic formulation on an arbitrary graph of nodes and edges, one drops grains of sand (or “load”) uniformly at random on the nodes, each of which has an innate threshold (or capacity). Whenever the load on a node exceeds its threshold, that node topples, meaning that it sheds (or moves) its sand to its neighbors. These neighbors may in turn become unstable and topple, which causes some of their neighbors to topple, and so on. In this way, dropping a grain of sand on the network can cause a cascade of load throughout the system—often small but occasionally large. The cascade finishes once no node’s load exceeds its capacity, whereupon another grain of sand is dropped, and the process repeats. Probability measures of the size, area, and duration of avalanches typically follow power laws asymptotically in the limit of many avalanches (47).

Many classic versions of the sandpile model (23, 46, 47) connect the nodes in a finite, two-dimensional lattice and assign all nodes threshold four, so that a toppled node sheds one sand grain to each of its four neighbors. The lattice has open boundaries, so that sand shed off the boundary is lost, which prevents inundation of sand. A few variants of the model on lattices can be solved exactly if the shedding rules have Abelian symmetry (47).

More recently, sandpile models have been studied on isolated networks, including Erdős-Rényi graphs (35, 36), scale-free graphs (39–41), and graphs generated by the Watts–Strogatz model on one- (37) and two-dimensional (38) lattices. A natural choice for the capacities of nodes—which we use here—is their degree, so that toppled nodes shed one grain to each neighbor (39, 48). Other choices include identical (39), uniformly distributed from zero to the degree k (40), and $k^{1-\eta}$ for some $0 \leq \eta < 1$ (40, 41), but all such variants must choose ways to randomly shed to a fraction of neighbors. Shedding one grain to each neighbor is simpler and exhibits the richest behavior (39). A natural analog of open boundaries on finite lattices is to delete grains of sand independently with a small probability f as they are shed. We choose the dissipation rate of sand f so that the largest cascades topple almost the entire network.

The mean-field solution of sandpile cascades is characterized by an avalanche size distribution that asymptotically obeys a

power law with exponent $-3/2$ and is quite robust to network structure. (For example, on scale-free random graphs, sandpile cascades deviate from mean-field behavior only if the degree distribution has a sufficiently heavy tail, with power law exponent $2 < \gamma < 3$; ref. 39.) Nevertheless, sparse connections among interdependent networks divert and direct sandpile cascades in interesting, relevant ways.

Topologies of Interacting Networks. Here we focus on interdependent power grids and idealized models of them. We obtained topological data on two interdependent power grids—which we label c and d —from the US Federal Energy Regulation Commission (FERC).^{*} (All data shown here are sanitized to omit sensitive information.) Owned by different utilities but connected to one another in the southeastern United States,[†] power grids c and d have similar size (439 and 504 buses) but rather different average internal degrees (2.40 and 2.91, respectively). The grids are sparsely interconnected by just eight edges, making the average external node degrees 0.021 and 0.018, respectively. More information on c and d is in Table S1. As in other studies, we find that these grids have narrow degree distributions (49–51) and are sparsely interconnected to one another (12).

To construct idealized versions of the real grids, consider two networks labeled a and b . Because of the narrow degree distribution of the real grids, we let network a be a random z_a -regular graph (where each node has degree z_a) and network b be a random z_b -regular graph. These two are then sparsely interconnected as defined below. To define this system of coupled networks more formally, we adopt the multitype network formalism of refs. 45 and 52. Each network a, b has its own degree distribution, $p_a(k_{aa}, k_{ab})$ and $p_b(k_{ba}, k_{bb})$, where, for example, $p_a(k_{aa}, k_{ab})$ is the fraction of nodes in network a with k_{aa} neighbors in a and k_{ab} in b . We generate realizations of multitype networks with these degree distributions using a simple generalization of the configuration model: All nodes repeatedly and synchronously draw degree vectors (k_{oa}, k_{ob}) from their degree distribution p_o (where $o \in \{a, b\}$), until the totals of the internal degrees k_{aa}, k_{bb} are both even numbers and the totals of the external degrees k_{ab}, k_{ba} are equal.[‡]

We interconnect the random z_a, z_b -regular graphs by Bernoulli-distributed coupling: each node receives one external “edge stub” with probability p and none with probability $1 - p$. Hence the degree distributions are $p_a(z_a, 1) = p, p_a(z_a, 0) = 1 - p$, and $p_b(1, z_b) = p, p_b(0, z_b) = 1 - p$. We denote this class of interacting networks by the shorthand $R(z_a)-B(p)-R(z_b)$; we illustrate a small example of $R(3)-B(0.1)-R(4)$ in Fig. 2.

Measures of Avalanche Size. We are most interested in the avalanche size distributions $s_a(t_a, t_b), s_b(t_a, t_b)$, where, for example, $s_a(t_a, t_b)$ is the chance that an avalanche begun in network a (indicated by the subscript on s_a) causes t_a, t_b many topplings in networks a, b , respectively. These distributions count the first toppling event, and they are defined asymptotically in that s_a, s_b are frequencies of avalanche sizes after the networks have undergone many cascades. To study s_a and s_b , we simulate sandpile avalanches and approximate them using a multitype branching process.

^{*}We scraped the network topologies of two “areas” of the southeastern region of the US power grid from simulation output provided by the FERC via the Critical Energy Infrastructure Information, <http://www.ferc.gov/legal/ceii-foia/ceii.asp>

[†]These two power grids are also connected to other grids, which we ignore in this paper but could model with more types of nodes.

[‡]A standard practice in the literature regarding admissible joint degree distributions is that the average interdegrees $\langle k_{ab} \rangle, \langle k_{ba} \rangle$ of the degree distributions must be equal (45, 53). This practice is more restrictive than necessary: Most any joint degree distributions can be used, as long as conditioning on valid degree sequences (even internal and equal external total degrees) leaves some probability. We derive the effective degree distributions, which are convolutions of the input ones, in Lemma 1 in the Appendix.

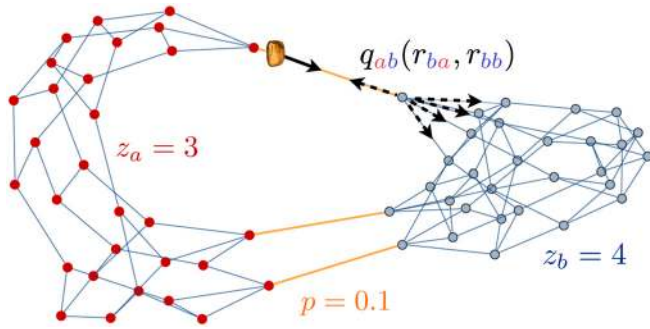


Fig. 2. A random three- and four-regular graph connected by Bernoulli-distributed coupling with interconnectivity parameter $p = 0.1$ [$R(3)-B(0.1)-R(4)$]. We also illustrate the shedding branch distribution $q_{ab}(r_{ba}, r_{bb})$, the chance that an ab shedding causes r_{ba}, r_{bb} many ba, bb sheddings at the next time step. Note these random graphs are small and become tree-like when they are large ($\gtrsim 1,000$ nodes).

Multitype Branching Process Approximation. In these next two sections, we present an overview of our mathematical formulation, with details left to *Materials and Methods*. We develop a branching process approximation that elucidates how sandpile cascades spread in interconnected networks, advances theoretical tools for cascades on multitype networks, justifies using this model as an idealization of real infrastructure like power grids, and establishes an open and relevant mathematical challenge. However, readers more interested in the applications of the model may wish to skip to *Results*.

Sandpile cascades on networks can be approximated by a branching process provided that the network is locally tree-like (i.e., has few short cycles), so that branches of a nascent cascade grow approximately independently. The interacting networks $R(z_a)-B(p)-R(z_b)$ are tree-like provided they are sparse and large enough (with at least several hundred nodes) because the edges are wired uniformly at random. Power grids are approximately tree-like: The clustering coefficient of power grids c and d , for example, is $C \approx 0.05$, an order of magnitude larger than an Erdős-Rényi random graph with equally many nodes and edges, but still quite small. Although tree-based approximations of other dynamical processes work surprisingly well even on non-tree-like graphs, power grid topologies were found to be among the most difficult to predict with tree-based theories (54). Here we find that analytic, tree-based approximations of sandpile dynamics agree well with simulations even on the real power grid topologies (Fig. 3 and Figs. S1 and S2).

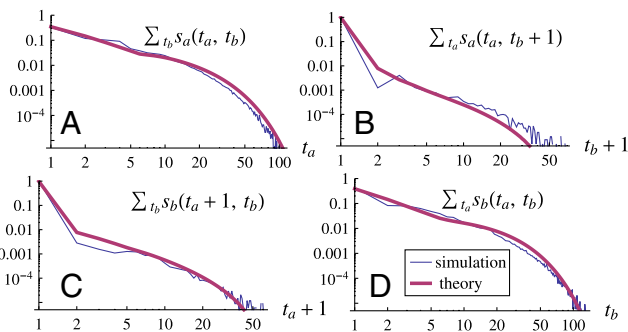


Fig. 3. The multitype branching process (red curves) approximates simulations of sandpile cascades on the power grids c, d (blue curves) surprisingly well, given that power grids are among the most difficult network topologies on which to predict other dynamics (54). We plot the four marginalized avalanche size distributions (using the labels a, b for grids c, d). For example, Fig. 3A, plotting $\sum_{t_b} s_a(t_a, t_b)$ versus t_a , is the chance an avalanche begun in a topples t_a nodes in a . Vertical axes' labels are the expressions in the plots. Here dissipation $f = 0.1$.

Cascades on interacting networks require a multitype branching process, in which a tree grows according to probability distributions of the number of events of various types generated from seed events. We consider two basic event types, a and b topplings—i.e., toppling events in networks a and in b . These simplify the underlying branching process of sheddings, or grains of sand shed from one network to another, of which there are four types: $aa, ab, ba,$ and bb sheddings. (Note that there is no distinction between topplings and sheddings on one, isolated network, because sand can only be shed from, say, a to a .)

A key property of sandpile dynamics on networks, which enables the branching process calculations, is that in simulations the amount of sand on a node is asymptotically uniformly distributed from zero to one less its degree (i.e., there is no typical amount of sand on a node) (48, 55). Hence the chance that a grain arriving at a node with degree k topples it equals the chance that the node had $k - 1$ grains of sand, which is $1/k$. So sandpile cascades are approximated by what we call $1/k$ percolation: The cascade spreads from node u to node v with probability inversely proportional to the degree of v . That this probability decays inversely with degree suggests a direct interpretation for infrastructure: Important nodes have k times more connectivity than unimportant (degree 1) nodes, so they are k times less likely to fail (they are presumably reinforced by engineers). But when important nodes do fail, they cause k times more repercussions (shedded grains of sand). We found some evidence for this inverse dependence on degree in the power flowing through buses (nodes) in power grids: Each additional degree correlates with an additional 124 mVA of apparent power flowing through it ($R^2 = 0.30$; see Fig. S3).

The details of the branching process analysis extend the standard techniques as presented in *Materials and Methods*. We give here only the crux of the derivation. Suppose a grain of sand is shed from network $o \in \{a, b\}$ to network $d \in \{a, b\}$ (origin network, o ; destination network, d). What is the chance that this grain shed from o to d (an od shedding) causes r_{da} and r_{db} many grains to be shed from network d to a and from d to b , respectively, at the next time step? This probability distribution, denoted $q_{od}(r_{da}, r_{db})$, is the branch (or children) distribution of the branching process for sheddings. Fig. 2 illustrates q_{ab} as an example. Neglecting degree-degree correlations [the subject of so-called $P(k, k')$ theory; ref. 54], a grain shed from network o to d arrives at an edge stub chosen uniformly at random, so it arrives at a node with degree $P_d(r_{da}, r_{db})$ with probability proportional to r_{do} because that node has r_{do} many edges pointing to network o . Using this fact and the chance of toppling found above ($1/\text{total degree}$), we approximate asymptotically that

$$q_{od}(r_{da}, r_{db}) = \frac{r_{do} P_d(r_{da}, r_{db})}{\langle k_{do} \rangle} \frac{1}{r_{da} + r_{db}} \quad [1]$$

for $r_{da} + r_{db} > 0$, where $\langle k_{do} \rangle$ is the expected number of edges from d to o , $\sum_{k_{da}, k_{db}} k_{do} P_d(k_{da}, k_{db})$. To normalize q_{od} , set

$$q_{od}(0, 0) := 1 - \sum_{r_{da} + r_{db} > 0} q_{od}(r_{da}, r_{db}), \quad [2]$$

which is the probability that the destination node does not topple (i.e., that it has fewer grains than one less its total degree).

Note that for an individual, isolated network the analogous branch distribution $q(k)$ simplifies considerably: In the equivalent of Eq. 1 there is a cancellation of k in the numerator of with $1/k$ on the right (39–41). Thus the expected number of children events $\langle q(k) \rangle = \sum_k k \frac{k P(k)}{\langle k \rangle} \frac{1}{k} = 1$. Each seed event gives rise to one child on average, which then gives rise to one child on average, etc., which is called a “critical” branching process. (If less than one child on average, the branching process dies out; if more

than one it may continue indefinitely.) The branching process approximations of sandpile cascades on the interacting networks studied here—coupled random regular graphs a, b and power grids c, d —are also critical, because the principle eigenvalue of the matrix of first moments of the branch distributions is one.

The branching process of sheddings is high dimensional, with four types aa, ab, ba, bb recording origin and destination networks. Transforming the shedding branch distributions q_{od} to the toppling branch distributions u_a, u_b is easy; the key is that a node topples if and only if it sheds at least one grain of sand (for details, see *Materials and Methods*). Transforming to toppling distributions also halves the dimensions of the branching process of topplings, simplifying calculations.

Self-Consistency Equations. We analyze implicit equations for the avalanche size distributions s_a, s_b using generating functions (56). Denote the generating functions associated to the toppling branch distributions u_a, u_b and the avalanche size distributions s_a, s_b by capital letters \mathcal{U} and \mathcal{S} ; for example,

$$\mathcal{U}_a(\tau_a, \tau_b) =: \sum_{t_a, t_b=0}^{\infty} u_a(t_a, t_b) \tau_a^{t_a} \tau_b^{t_b} \quad \text{for } \tau_{da}, \tau_{db} \in \mathbb{C}.$$

The theory of multitype branching processes (57) implies the self-consistency equations

$$\mathcal{S}_a = \tau_a \mathcal{U}_a(\mathcal{S}_a, \mathcal{S}_b), \quad \mathcal{S}_b = \tau_b \mathcal{U}_b(\mathcal{S}_a, \mathcal{S}_b), \quad [3]$$

where each \mathcal{S} is evaluated at (τ_a, τ_b) . In other words, the left-hand equation in [3] says that, to obtain the distribution of the sizes of cascades begun in a , the cascade begins with an a toppling (hence the τ_a out front), which causes at the next time step a number of a and b topplings distributed according to \mathcal{U}_a , and these topplings in turn cause numbers of a and b topplings distributed according to \mathcal{S}_a and \mathcal{S}_b .

We wish to solve Eq. 3 for \mathcal{S}_a and \mathcal{S}_b , because their coefficients are the avalanche size distributions s_a, s_b of interest. In practice, however, these implicit equations are transcendental and difficult to invert. Instead, we solve Eq. 3 with computer algebra systems using three methods—iteration, Cauchy integral formula, and multidimensional Lagrange inversion (58)—to compute exactly hundreds of coefficients; for details, see *Materials and Methods*. Fig. 3 shows good agreement between simulation of sandpile cascades on power grids c, d and the branching process approximation (obtained by iterating Eq. 3 seven times starting from $\mathcal{S}_a = \mathcal{S}_b = 1$, with branch distributions calculated from the empirical degree distributions of c, d). For more details on the agreement, including how the degrees of nodes with external links account for the characteristic “blips” in the avalanche size distributions of the power grids, see Figs. S1 and S2.

These numerical methods are computationally feasible for the probabilities of the smallest avalanches, but we are most interested in the probabilities of the largest avalanches—that is, in the asymptotic behaviors of $s_a(t_a, t_b), s_b(t_a, t_b)$ as $t_a, t_b \rightarrow \infty$. Unfortunately the technique used for an isolated network—an expansion at a singularity of \mathcal{U} —fails for sandpile cascades on Bernoulli-coupled random regular graphs and on the power grids, because their generating functions have singularities at infinity and none in the finite plane (see *Materials and Methods*). Generalizing these asymptotic techniques to “multitype cascades” with singularities at infinity poses an outstanding mathematical challenge. Nevertheless, three tactics—simulations, computer calculations of coefficients of $\mathcal{S}_a, \mathcal{S}_b$, and analytical calculations of the first moments of the branch and avalanche size distributions—suffice to obtain interesting conclusions about the effect of interdependence on critical cascades of load, as discussed next.

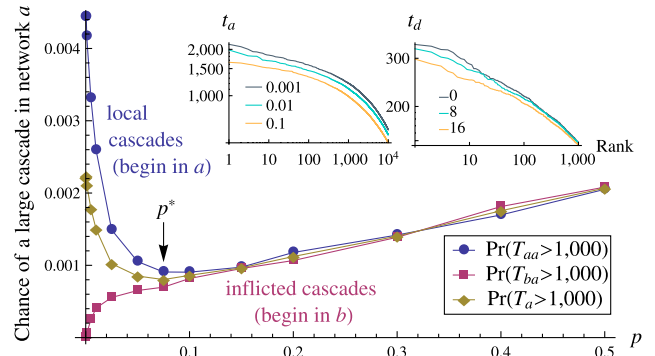


Fig. 4. Interconnectivity is locally stabilizing, but only up to a critical point. The main plot, the results of simulations on $R(3)$ - $B(p)$ - $R(3)$ (2×10^6 grains; $f = 0.01$; 2×10^3 nodes/network), shows that large local cascades decrease and then increase with p , whereas large inflicted cascades only become more likely. Their average (gold curve) has a stable minimum at $p^* \approx 0.075 \pm 0.01$. This curve and its critical point, p^* , are stable for cutoffs \mathcal{C} in the regime $400 \leq \mathcal{C} \leq 1,500$. (Insets) Rank-size plots on log-log scales of the largest cascades in network a (Left) for $p = 10^{-3}, 10^{-2}, 10^{-1}$ and in power grid d connected to c by 0, 8, or 16 edges.

Results

Locally Stabilizing Effect of Interconnections. We first answer this question, would an isolated network suppress its largest cascades of load by connecting to another network? For coupled random regular graphs $R(z_a)$ - $B(p)$ - $R(z_b)$, yes; increasing interconnectivity p suppresses an individual network’s largest cascades, but only up to a critical point p^* (Fig. 4).

First we introduce notation. For a cascade that begins in network a , the random variables T_{aa}, T_{ab} are the sizes of the “local” and “inflicted” cascades: The number of topplings in a and in b , respectively. For example, a cascade that begins in a and that topples 10 a nodes and 5 b nodes corresponds to $T_{aa} = 10, T_{ab} = 5$. We denote T_a to be the random variable for the size of a cascade in network a , without distinguishing where the cascade begins. (We define T_{ba}, T_{bb}, T_b analogously.) Dropping sand uniformly at random on two networks of equal size means that avalanches begin with equal probability in either network, so $\Pr(T_a = t_a) = \sum_{t_b} (s_a(t_a, t_b) + s_b(t_a, t_b)) / 2$.

In Fig. 4 we plot the probability of observing a large avalanche in a (that topples at least half of all its nodes) as a function of interconnectivity p , as measured in numerical simulations on the $R(3)$ - $B(p)$ - $R(3)$ topology. We distinguish between those avalanches that begin in a (blue local cascades), begin in b (red inflicted cascades), or in either network (gold). With increasing interconnectivity p , large inflicted cascades from b to a (red curve) increase in frequency due in large part to the greater ease of avalanches traversing the interconnections between networks. More interesting is that increasing interconnectivity suppresses large local cascades (blue curve) for small p , but amplifies them for large p . The 80% drop in $\Pr(T_{aa} > 1,000)$ and 70% drop in $\Pr(T_a > 1,000)$ from $p = 0.001$ to $p^* \approx 0.075 \pm 0.01$ are the locally stabilizing effects of coupling networks. The left inset in Fig. 4 is the rank-size plot showing the sizes of the largest avalanches and their decrease with initially increasing p , and the same holds for simulations on the power grids c and d (right inset).[§] The curve $\Pr(T_a > \mathcal{C})$ and the location of its critical point p^* in Fig. 4 is robust to changing the cutoff $\mathcal{C} \in [400, 1,500]$. Thus a network such as a seeking to minimize its largest cascades would seek interconnectivity that minimizes $\Pr(T_a > \mathcal{C})$, which

[§]To adjust interconnectivity between the real power grids c and d , we (i) delete the eight interconnections so that they are isolated, (ii) leave the eight original interconnections, and (iii) add eight additional interconnections in a way that mirrors the empirical degree distribution.

we estimate to be $p^* \approx 0.075 \pm 0.01$ for $R(3)-B(p)-R(3)$ with 2×10^3 nodes per network.

This central result appears to be generic: changing the system size, internal degrees, or type of degree distribution (so long as it remains narrow) may slightly change p^* but not the qualitative shape of Fig. 4 (see Figs. S4 and S5). Furthermore, this effect of optimal connectivity is unique to interconnected networks: adding edges to a single, isolated network does not reduce the chance of a large cascade (and hence produce a minimum like in Fig. 4). [This result for an isolated network is expected because $s(t) \sim t^{-3/2}$ for all narrow degree distributions; ref. 39.] Whereas adding links within a network can only increase its vulnerability to large cascades, adding links to another network can reduce it. Note that we cannot derive analytical results for Fig. 4 because the standard techniques for single networks fail for multitype generating functions with singularities at infinity, and inverting Eq. 3 numerically is practical only for exactly computing the probabilities of small cascades ($T_a < 50$) and not large ones ($T_a > 10^3$) (see *Materials and Methods*).

Intuitively, adding connections between networks diverts load, and that diverted load tends to be absorbed by the neighboring network rather than amplified and returned, because most cascades in isolated networks are small. One way to see the diversion of load is in the first moments of the toppling branch distributions u_a, u_b . For $R(z_a)-B(p)-R(z_b)$, the average numbers of topplings at the next time step in the same and neighboring networks, respectively, decrease and increase with the interconnectivity p as $\langle u_a \rangle_a = 1 - p/(1 + z_a), \langle u_a \rangle_b = p/(1 + z_b)$, where, for example, $\langle u_a \rangle_a \equiv \sum_{t_a, t_b} t_a u_a(t_a, t_b)$.

However, introducing too many interconnections is detrimental, as shown in Fig. 4. Interconnections let diverted load more easily return and with catastrophic effect. In addition, each interconnection augments the networks' capacity and hence average load, so that large avalanches increase in frequency in individual networks and in the collection of networks, as discussed next.

Globally Destabilizing Effect of Interconnections. Adding interconnections amplifies large global cascades. That is, the largest avalanches in the collection of networks—viewed as one system with just one type of node—increase in size with interconnectivity. Here we are interested in the total avalanche size distribution $s(t)$, the chance of t topplings overall in a cascade. Fig. 5 shows the extension of the right-hand tail of $s(t)$ in simulations on $R(3)-B(p)-R(3)$ with increasing interconnectivity p . The rank-size plot inset shows more explicitly that the largest avalanches increase with p . Similar results occur in simulations on power grids c, d (see *SI Text*).

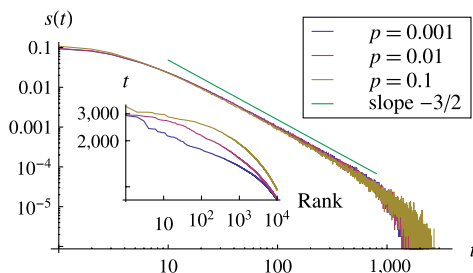


Fig. 5. Increasing the interconnectivity p between two random three-regular graphs extends the tail of the total avalanche size distribution $s(t)$, which does not distinguish whether toppled nodes are in network a or b . The inset shows a rank-size plot on log-log scales of the number of topplings t in the largest 10^4 avalanches (with 2×10^6 grains of sand dropped), showing that adding more edges between random three-regular graphs enlarges the largest global cascades by an amount on the order of the additional number of interconnections. As expected theoretically (39), when a and b nodes are viewed as one network, $s(t) \sim t^{-3/2}$ for large t (green line).

What amplifies the global cascades most significantly is the increase in total capacity (and hence average load available for cascades) and not the increased interdependence between the networks. (Recall that capacities of nodes are their degrees, so introducing new edges between networks increments those nodes' degrees and capacities.) To see this effect on coupled random regular graphs, we perform the following rewiring experiment. Beginning with two isolated random regular graphs, each node changes one of its internal edge stubs to be external with probability p . The degree distributions become, for example, $p_a(z_a - 1, 1) = p, p_a(z_a, 0) = 1 - p$, which we call “correlated-Bernoulli coupling” because the internal and external degrees are not independent. Fig. S6 shows that the largest global avalanches are not significantly enlarged with increasing “rewired interconnectivity” p for random three-regular graphs with such coupling. Furthermore, the enlargement of the largest cascades observed in the rank-size plot in the inset of Fig. 5 is on the order of the extra average load resulting from the additional interconnectivity (and the same holds for simulations on the power grids).

The amplification of global avalanches, though relatively small, is relevant for infrastructure: Additional capacity and demand often accompany—and even motivate—the construction of additional interconnections (11, 13). Furthermore, in reality it is more common to augment old infrastructure with new interconnections as in Fig. 1 (Bernoulli coupling) rather than to delete an existing internal connection and wire it to span across networks (correlated-Bernoulli coupling). Thus, building new interconnections augments the entire system's capacity, and hence average load, and hence largest cascades.

Interconnectivity That Mitigates Cascades of Different Sizes. Fig. 4 shows that networks seeking to mitigate their large avalanches seek optimal interconnectivity p . Networks mitigating their small or intermediate cascades would seek different optimal interconnectivity p , as shown in Fig. 6 [the results of simulations on $R(3)-B(p)-R(3)$]. Fig. 6A shows that the probability of a small cascade in network a ($1 \leq T_a \leq 51$) increases monotonically with p , so that networks mitigating the smallest cascades seek isolation, $p = 0$. By contrast, the chance of a cascade of intermediate size, $100 \leq T_a \leq 150$ (Fig. 6B), has a local maximum at $p^* \approx 0.05$, so networks mitigating intermediate cascades would demolish all interconnections ($p = 0$) or build as many as possible ($p = 1$). By contrast, the largest cascades $400 \leq T_a \leq 1,500$ (Fig. 6C and D) occur with minimal probability at $p^* \approx 0.075 \pm 0.01$. For more

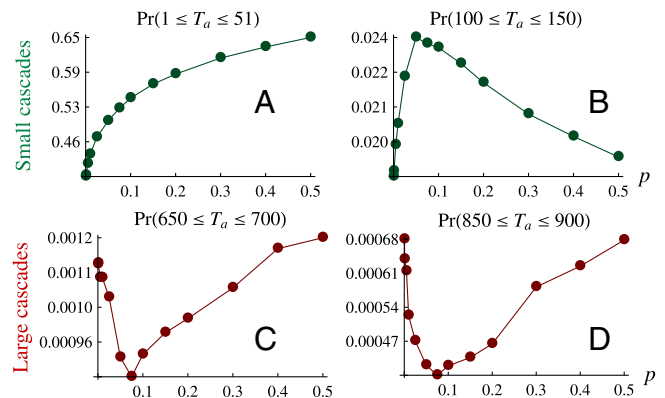


Fig. 6. (A) Networks mitigating the smallest cascades $1 \leq T_a \leq 51$ seek isolation $p = 0$, whereas (B) networks suppressing intermediate cascades $100 \leq T_a \leq 150$ seek isolation $p = 0$ or strong coupling $p = 1$, depending on the initial interconnectivity p in relation to the unstable critical point $p^* \approx 0.05$. But networks like power grids mitigating large cascades (C and D) would seek interconnectivity at the stable equilibrium $p^* \approx 0.075 \pm 0.01$. The bottom figures and the location of p^* are robust to changes in the window $\ell \leq T_a \leq \ell + 50$ for all $400 \leq \ell \leq 1,500$.

plots showing the change in concavity and the robustness of the stable critical point p^* for large cascades, see Fig. S7.

Other models of cascades in power grids conclude that upgrading and repairing the system to mitigate the smallest blackouts may increase the risk of the largest blackouts (15). Similarly, extinguishing small forest fires, a common policy in the 20th century, increases forest vegetation and thus the risk of large forest fires—a phenomenon known as the “Yellowstone effect” (32). The results here augment these conclusions to include interconnectivity. Networks building interconnections in order to suppress their large cascades cause larger global cascades (by an amount on the order of the additional capacity). Networks suppressing their small or intermediate cascades may seek isolation ($p = 0$), which amplifies their large cascades, or (for intermediate cascades) strong coupling ($p = 1$), which amplifies both large local and global cascades.

Capacity Disparity. Two networks that are interdependent are rarely identical, unlike the $R(3)$ - $B(p)$ - $R(3)$ topologies studied thus far, so we determine the effect of capacity disparity on cascades. Because the capacities of nodes are their degrees, we study $R(z_a)$ - $B(p)$ - $R(z_b)$ with different internal degrees, $z_a \neq z_b$. We find that interdependence is more catastrophic for smaller capacity networks, in that they suffer comparatively larger inflicted cascades. They still prefer some interconnectivity, but less than the higher capacity network.

Using the theoretical branching process approximation, we compute how much larger inflicted cascades are from high- to low-capacity networks. Differentiating Eq. 3 with respect to τ_a , τ_b and setting $\tau_a = \tau_b = 1$ yields four linear equations for the first moments of the avalanche size distributions s_a, s_b in terms of the first moments of the branch distributions u_a, u_b . For $R(z_a)$ - $B(p)$ - $R(z_b)$, the four first moments of s_a, s_b are all infinite, as expected, because in isolation these networks’ avalanche size distributions are power laws with exponent $-3/2$ (39–41). Nevertheless, we can compare the rates at which the average inflicted cascade sizes diverge by computing their ratio

$$\frac{\langle s_a \rangle_b}{\langle s_b \rangle_a} = \frac{1 + z_a}{1 + z_b}, \quad [4]$$

where, e.g., $\langle s_a \rangle_a \equiv \sum_{t_a, t_b} t_a s_a(t_a, t_b)$ (see *SI Text* for the derivation). Thus $z_b > z_a$ implies that the inflicted cascades from b to a are larger on average than those from a to b . Fig. S8 shows qualitative agreement with simulations.

As a result of Eq. 4, the low-capacity network prefers less interconnectivity than the high-capacity network. In simulations of $R(3)$ - $B(p)$ - $R(4)$, for instance, low-capacity network a prefers $p_a^* \approx 0.05$, whereas high-capacity b prefers $p_b^* \approx 0.3$. For systems like power grids seeking to mitigate their cascades of load, these results suggest an arms race for greater capacity to fortify against cascades inflicted by neighboring networks.

Incentives and Equilibria in Power Grids. Because different networks have unique susceptibilities to cascades (due to capacity disparity, for example), equilibria among real networks are more nuanced than on identical random graphs. Next we explore how the level of interconnectivity and load disparity affect sandpile cascades on the power grids c and d . [Although sandpile dynamics do not obey Ohm’s and Kirchoff’s laws nor the flow of load from sources to sinks, as in physical power flow models (e.g., refs. 16 and 17), they do closely resemble some engineers’ models of blackouts, and blackout data show evidence of criticality and power laws; ref. 15.] To interpret results, we suppose that the owners of the power grids c, d are rational, in that they wish to mitigate their largest cascades but care little about cascades overwhelming neighboring grids.

To capture different amounts of demand, numbers of redundancies, ages of infrastructure, susceptibility to sagging power lines (16, 17), and other factors that affect the rate at which cascades of load shedding and failures begin in each network, we introduce a load disparity parameter r as follows. Each node in c is r times more likely than a node in d to receive a new grain of sand. Increasing r intensifies the load on grid c , the rate at which cascades begin there, and the sizes of the largest inflicted cascades from c to d . The larger r is, the more volatile power grid c becomes.

Given the spatial structure of the power grid networks, there is no principled way to add arbitrarily many interconnections between them. However, three different levels of interconnectivity are natural: (i) Delete the eight interconnections so that c and d are isolated, (ii) leave the eight original interconnections, and (iii) add eight additional interconnections in a way that mirrors the empirical degree distribution.

Fig. 7 shows that there are two ways to amplify the largest inflicted cascades in d . The first is to increase the number of interconnections (compare red to green-dashed curve in the main figure). The second is to increase r (compare distance from green to gold curve in the main figure and in the inset). At a critical r^* , the largest inflicted cascades in d that begin in c (red and green-dashed curves) are equally large as the largest local cascades in d that begin in d (blue and gold curves). For 16 interconnections, we estimate $r^* \approx 15$ (Fig. 7, *Inset*), and the inflicted cascades are larger and smaller, respectively, for $r = 10, 20$ (Fig. S9)—indicating that inflicted cascades begin to dominate local cascades at $r^* \approx 15$. The actual load disparity between power grids c and d is $r \approx 0.7$, which we estimate by computing the average power incident per node in simulation output from FERC.[§] (There are, however, interdependent power grids in the southeastern United States with $r > 15$.) Because $r = 0.7$, the load is greater on grid d , so grid d prefers more interconnections and c prefers fewer than if r were one. Consequently, any equilibrium between the two grids is frustrated (or semistable): Only one grid can achieve its desired interconnectivity.

Discussion

We have presented a comprehensive study of sandpile cascades on interacting networks to obtain a deeper understanding of cascades of load on interdependent systems, showing both the benefits and dangers of interdependence. We combine a mathematical framework for multitype networks (45, 52) with models of

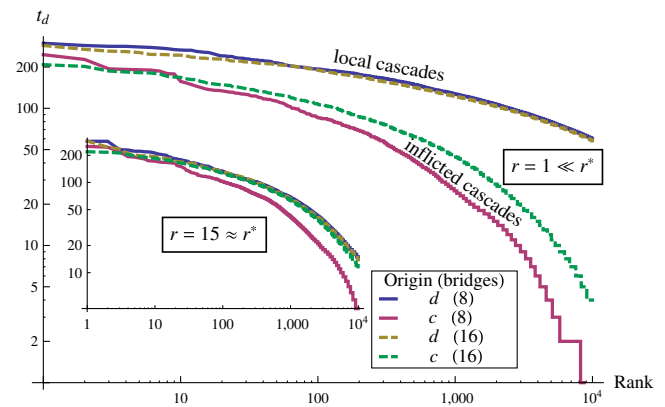


Fig. 7. The critical load disparity at which inflicted cascades in d become equally large as its local cascades is $r^* \approx 15$ (for 16 interconnections, “bridges”). (For eight interconnections, $r^* > 20$; see Fig. S9). Here we show a rank-size plot in log-log scales of the largest 10^4 avalanches in power grid d , distinguishing whether they begin in c (inflicted cascades) or in d (local cascades), for 8 and 16 interconnections (solid, dashed curves), for $r = 1$ (*Main*), $r = 15$ (*Inset*), in simulations with dissipation of sand $f = 0.05$, 10^6 grains dropped after 10^5 transients.

sandpiles on isolated networks (35–41) to derive a multitype branching process approximation of cascades of load between simple interacting networks and between real power grids. We show that some interdependence is beneficial to a network, for it mitigates its largest avalanches by diverting load to neighboring networks. But too much interconnectivity opens new avenues for inflicted load; in addition, each new interconnection adds capacity that fuels even larger cascades. The benefits and detriments in mitigating large avalanches balance at a critical amount of interconnectivity, which is a stable equilibrium for coupled networks with similar capacities. For coupled networks with asymmetric capacity, the equilibria will be frustrated, in that the networks prefer different optimal levels of interconnectivity.

We also show that tuning interconnectivity to suppress cascades of a certain range of sizes amplifies the occurrence of cascades in other ranges. Thus a network mitigating its small avalanches amplifies its large ones (Figs. 4 and 6), and networks suppressing their own large avalanches amplify large ones in the whole system (Fig. 5). Similarly it has been found that mitigating small electrical blackouts and small forest fires appears to increase the risk of large ones (15, 32). Furthermore, the amplification of global cascades due to the increase in capacity (Fig. 5) is a warning for new interconnections among power grids (11, 13) (Fig. 1), financial sectors, and other infrastructure.

These findings suggest economic and game-theoretic implications for infrastructure. (Note that here we consider the strategic interactions of entire networks rather than of individual nodes in one network, which is more standard, e.g., ref. 59.) For example, a power grid owner wishing to mitigate the largest cascades in her grid would desire some interconnections to other grids, but not too many. However, what benefits individual grids can harm society: Grids building interconnections amplify global cascades in the entire system. More detailed models that combine results like these with economic and physical considerations of electrical grids and with costs of building connections may provide more realistic estimates of optimal levels of interconnectivity. Our framework—which models a dynamical process on stable, underlying network topologies—could also be combined with models of topological failure in interdependent networks (20–22). Those studies conclude that systemic risk of connectivity failure increases monotonically with interdependence (“dependency links”). Whether suppressing cascades of load or of connectivity failures is more important might suggest whether to interconnect some ($p^* > 0$) (Fig. 4) or none ($p = 0$) (20–22), respectively. Our results are consistent with recent work showing that an intermediate amount of connectivity minimizes risk of systemic default in credit networks (60), in contrast to refs. 20, 21 and the more traditional view that risk increases monotonically with connectivity in credit networks (e.g., ref. 61).

This work also advances our mathematical understanding of dynamical processes on multitype networks. Because networks with one type of node and edge are impoverished views of reality, researchers have begun to study dynamical processes on multitype networks, such as on modular graphs (42–45). Here we derive a branching process approximation of sandpile cascades on multitype networks starting from the degree distributions, and we discuss open problems in solving for the asymptotic behavior of the generating functions’ coefficients, which elude current techniques for isolated networks. We expect that the computational techniques used here to solve multidimensional generating function equations, such as multidimensional Lagrange inversion (58), will find other uses in percolation and cascades in multitype networks. Finally, in the *Appendix* we derive the effective degree distributions in multitype networks, which expands the admissible degree distributions that others have considered. The machinery we develop considers just two interacting networks, a and b , or equivalently one network with two types of nodes. However, this machinery extends to finitely many types, which may be useful for

distinguishing types of nodes—such as buses, transformers, and generators in electrical grids—or for capturing geographic information in a low-dimensional way without storing explicit locations—such as buses in the interiors of power grids and along boundaries between them.

Here we have focused on mitigating large avalanches in modular networks, but other applications may prefer to amplify large cascades, such as the adoption of products in modular social networks (62) or propagating response-driven surveys across bottlenecks between social groups (63). Cascades in social networks like these may require networks with triangles or other subgraphs added (53, 64, 65); inverting the resulting multidimensional generating function equations for dynamics on these networks would require similar multitype techniques as developed here.

We expect that this work will stimulate calculations of critical points in interconnectivity among networks subjected to other dynamics, such as linearized power flow equations in electrical grids (16, 17) and other domain-specific models. As critical infrastructures such as power grids, transportation, communication, banks, and markets become increasingly interdependent, resolving the risks of large cascades and the incentives that shape them becomes ever more important.

Materials and Methods

Power Grid Topologies. To understand coupling between multiple grids, we requested data from the US Federal Energy Regulation Commission.⁵ Using output of power simulations on many connected areas (distinct grids owned by different utilities) in the southeastern United States, we chose grid d by selecting the grid with the highest average internal degree and the grid, c , to which it had the most interconnections (eight). Grids c , d have 439 and 504 buses (nodes) and average internal degrees 2.40 and 2.91. For our purposes here, the only important details are the narrow degree distributions, the low clustering coefficients, and the number of interconnections between the grids. Other details about the grids are in the *S1 Text*.

Toppling Branch Distributions. To reduce the number of types in the branching process, we count the number of toppling events in each network rather than the number of grains of sand shed from one network to another. Here we derive the toppling branch distributions u_a, u_b from the shedding branch distributions q_{od} . [For instance, $u_a(t_a, t_b)$ is the chance that a toppled node in a causes t_a, t_b many nodes in a, b to topple at the next time step.] Note that a node topples if and only if it sheds at least one grain of sand. Thus a grain traveling from a network o to network d topples its destination node with probability $1 - q_{od}(0, 0)$. Denoting $(0, 0)$ by $\mathbf{0}$ and the binomial distribution by $\beta_k^n(p) \equiv \binom{n}{k} p^k (1-p)^{n-k}$, we have

$$u_a(t_a, t_b) = \sum_{k_a=t_a}^{\infty} \sum_{k_b=t_b}^{\infty} p_a(k_a, k_b) \beta_{k_a}^{t_a} [1 - q_{aa}(\mathbf{0})] \beta_{k_b}^{t_b} [1 - q_{ab}(\mathbf{0})],$$

because the node must have at least t_a many a neighbors and at least t_b many b neighbors, only t_a, t_b of which topple (which are binomially distributed). The expression for u_b is analogous. As an example, the probability generating function of u_a for Bernoulli-coupled random regular graphs $R(z_a)\text{-}B(p)\text{-}R(z_b)$ is

$$\mathcal{U}_a(\tau_a, \tau_b) = \frac{[p - p\tau_a + (z_a + 1)(\tau_a + z_a - 1)]^{z_a} [1 + p(\tau_b - 1) + z_b]}{(z_a + 1)^{z_a} z_a^{z_a} (z_b + 1)}.$$

Three Methods for Numerically Solving the Multidimensional Self-Consistency Equations of a Multitype Branching Process.

The most naïve method to solve the self-consistency equations of a multitype branching process (such as Eq. 3) is to use a computer algebra system like Mathematica or Maple to iterate [3] symbolically starting from $s_a = s_b = 1$, expand the result, and collect coefficients. To obtain the coefficient $s_a(t_a, t_b)$ exactly, it suffices to iterate [3] at least $t_a + t_b + 1$ times. What is more, just a handful of iterations partially computes the tails (coefficients of terms with high powers in s_a, s_b), but the amount of missing probability mass in the tails is undetermined. This method was used in Fig. 3.

A second method is to symbolically iterate Eq. 3 at least $t_a + t_b + 1$ times and to use Cauchy's integral formula

$$s_a(t_a, t_b) = \frac{1}{(2\pi i)^2} \iint_D \frac{\mathcal{S}_a(\tau_a, \tau_b)}{\tau_a^{t_a+1} \tau_b^{t_b+1}} d\tau_a d\tau_b, \quad [5]$$

where $D \subset \mathbb{C}^2$ is a Cartesian product of circular contours centered at the origin, each of radius r smaller than the modulus of the pole of \mathcal{S}_a closest to the origin (58). Then calculate one coefficient at a time using

$$s_a(t_a, t_b) = \frac{1}{4\pi^2} \int_0^{2\pi} \int_0^{2\pi} \frac{\mathcal{S}_a(re^{i\theta}, re^{i\psi})}{r^{t_a+t_b} e^{i(t_a\theta+t_b\psi)}} d\theta d\psi.$$

A third method is Lagrange inversion, generalized to multiple dimensions by Good in 1960 (58), a result that has seen little use in the networks literature. We state the theorem in the language of the two-type branching process considered here, with the notation $\vec{\tau} \equiv (\tau_a, \tau_b)$, $\vec{\delta} \equiv (\delta_a, \delta_b)$, and $\vec{\mathcal{U}} \equiv (\mathcal{U}_a, \mathcal{U}_b)$. For the slightly more general result that holds for arbitrary, finite, initial population, see ref. 58.

Theorem 1. [Good 1960] If $\mathcal{U}_a(\vec{0}) > 0$ and $\mathcal{U}_b(\vec{0}) > 0$ (i.e., every type has a positive probability of being barren), then for $x \in \{0, 1\}$,

$$\mathcal{S}_a(\vec{\tau})^x \mathcal{S}_b(\vec{\tau})^{1-x} = \sum_{m_a, m_b=0}^{\infty} \frac{\tau_a^{m_a} \tau_b^{m_b}}{m_a! m_b!} \left[\frac{\partial^{m_a+m_b}}{\partial \kappa_a^{m_a} \partial \kappa_b^{m_b}} \left\{ \kappa_a^x \kappa_b^{1-x} \times \mathcal{U}_a(\vec{\kappa})^{m_a} \mathcal{U}_b(\vec{\kappa})^{m_b} \left\| \delta_{\mu}^{\nu} - \frac{\kappa_{\mu}}{\mathcal{U}_{\mu}} \frac{\partial \mathcal{U}_{\mu}}{\partial \kappa_{\mu}} \right\| \right\} \right]_{\vec{\kappa}=\vec{0}}, \quad [6]$$

where μ, ν run over the types $\{a, b\}$, δ_{μ}^{ν} is the Kronecker delta, and $\|\cdot\|$ is the determinant.

Taking $x = 0$ gives \mathcal{S}_b , whereas taking $x = 1$ gives \mathcal{S}_a . In practice, hundreds of terms of $\mathcal{S}_a, \mathcal{S}_b$ can be computed exactly by truncating the sum in [6] and using computer algebra systems (see Fig. S10 for more details).

Solving for the Coefficients Asymptotically (and Why Standard Techniques Fail). To determine the asymptotic behaviors of $s_a(t_a, t_b), s_b(t_a, t_b)$ as $t_a, t_b \rightarrow \infty$, the trick is to solve for the inverses of \mathcal{S}_a and \mathcal{S}_b and to expand those inverses \mathcal{S}_a^{-1} and \mathcal{S}_b^{-1} at the singularities of \mathcal{S}_a and \mathcal{S}_b . This technique works for isolated networks (39–41, 66), but \mathcal{S}_a and \mathcal{S}_b have no finite singularities in \mathbb{C}^2 for Bernoulli-coupled random regular graphs nor for power grids c, d .

We demonstrate this failure of standard asymptotic techniques on the networks $R(z_a)\text{-}B(p)\text{-}R(z_b)$. Let $\vec{\omega} := \vec{\delta}(\vec{\tau})$. Assuming an inverse $\vec{\delta}^{-1}$ of $\vec{\delta}$ exists, using [3] we have

$$\vec{\tau} = \vec{\delta}^{-1}(\vec{\omega}) = \frac{\vec{\omega}}{\vec{\mathcal{U}}(\vec{\omega})} = \left(\frac{\omega_a}{\mathcal{U}_a(\vec{\omega})}, \frac{\omega_b}{\mathcal{U}_b(\vec{\omega})} \right). \quad [7]$$

The generating functions $\mathcal{S}_a, \mathcal{S}_b$ have singularities at $\vec{\tau}^*, \vec{\mu}^*$, respectively, if $\vec{D}\mathcal{S}_a^{-1}(\vec{\tau}^*) = \vec{D}\mathcal{S}_b^{-1}(\vec{\mu}^*) = (0, 0)$, where the operator $\vec{D} = (\partial/\partial\omega_a, \partial/\partial\omega_b)$ is a vector of partial derivatives. Differentiating Eq. 7 and equating the numerators to zero gives

$$\mathcal{U}_a(\vec{\tau}^*) - \tau_a^* \frac{\partial \mathcal{U}_a}{\partial \tau_a}(\vec{\tau}^*) = 0, \quad -\tau_a^* \frac{\partial \mathcal{U}_a}{\partial \tau_b}(\vec{\tau}^*) = 0, \quad [8a]$$

$$\mathcal{U}_b(\vec{\mu}^*) - \mu_b^* \frac{\partial \mathcal{U}_b}{\partial \tau_b}(\vec{\mu}^*) = 0, \quad -\mu_b^* \frac{\partial \mathcal{U}_b}{\partial \tau_a}(\vec{\mu}^*) = 0. \quad [8b]$$

The only solution to Eq. 8 is

$$\vec{\tau}_1^* = \left(\frac{z_a^2 + p - 1}{p - z_a - 1}, \frac{p - z_b - 1}{p} \right)_{p \rightarrow 0} \rightarrow (1 - z_a, \infty), \quad [9a]$$

$$\vec{\mu}_1^* = \left(\frac{p - z_a - 1}{p}, \frac{z_b^2 + p - 1}{p - z_b - 1} \right)_{p \rightarrow 0} \rightarrow (\infty, 1 - z_b). \quad [9b]$$

This solution [9] does not recover the singularities $\vec{\tau}^* = (1, c_1), \vec{\mu}^* = (c_2, 1)$ (where c_1, c_2 are arbitrary) that we should obtain when the networks are isolated ($p = 0$) (39). Moreover, although Eqs. 8 vanish at the solutions [9], the derivatives of [7] do not vanish at these solutions [9], because $\vec{D}\mathcal{S}_a^{-1}(\vec{\tau}_1^*) = (\infty, 0), \vec{D}\mathcal{S}_b^{-1}(\vec{\mu}_1^*) = (0, \infty)$. Thus we must discard solutions [9].

Solving only the left-hand equations in [8] yields singularities that do recover the correct singularities when $p = 0$:

$$\vec{\tau}_2^* = \left(\frac{z_a^2 + p - 1}{(1 - z_a)(p - 1 - z_a)}, \bar{\tau}_b \right)_{p \rightarrow 0} \rightarrow (1, \bar{\tau}_b), \quad [10a]$$

$$\vec{\mu}_2^* = \left(\bar{\mu}_a, \frac{z_b^2 + p - 1}{(1 - z_b)(p - 1 - z_b)} \right)_{p \rightarrow 0} \rightarrow (\bar{\mu}_a, 1), \quad [10b]$$

where $\bar{\tau}_b, \bar{\mu}_a$ are arbitrary constants satisfying

$$\bar{\tau}_b \neq \frac{p - z_b - 1}{p}, \quad \bar{\mu}_a \neq \frac{p - z_a - 1}{p}, \quad [11]$$

so that the derivatives of $\vec{\delta}^{-1}(\vec{\omega})$ [8] evaluated at [10] are finite. However, the derivatives on the right-hand sides of Eq. 8 (evaluated at the solutions [10]) are

$$\frac{\partial \mathcal{S}_a^{-1}}{\partial \omega_b}(\tau_2^*) = \left(0, \frac{p(z_b + 1)(z_a^2 + p - 1)(\frac{z_a^2 - 1}{z_a^2 + p - 1})^{z_a}}{(1 - z_a)(1 - p + z_a)[p(\omega_b - 1) + z_b + 1]^2} \right)$$

and similarly for $\partial \mathcal{S}_b^{-1}/\partial \omega_a(\mu_2^*)$ (interchange a and b). This derivative vanishes if and only if $p = 0$ or $z_a = 1$ or $\omega_b = \infty$. Hence we cannot find finite singularities of $\mathcal{S}_a, \mathcal{S}_b$; these generating functions are entire functions with singularities only at $\tau_b = \mu_a = \infty, \tau_a \neq \frac{z_a^2 + p - 1}{p - z_a - 1}, \mu_b \neq \frac{z_b^2 + p - 1}{p - z_b - 1}$. Hence we have no singularities at which to do an asymptotic expansion for the coefficients s_a, s_b , as one can for isolated networks (39–41, 66).

Other techniques exist for asymptotically approximating the coefficients of generating functions, depending on the type of the singularity (ref. 56, sect. 5). Hayman's method (ref. 56, sect. 5.4) works for generating functions with no singularities in the finite plane (i.e., entire functions), like the $\mathcal{S}_a, \mathcal{S}_b$ considered here. However, the theorem requires a closed form expression for the generating function, which we cannot obtain from the self-consistency equations for the synthetic and real interacting networks of interest. Developing techniques for asymptotically approximating the coefficients of multi-dimensional generating functions with singularities at infinity, like those studied here, poses an important challenge for future studies of dynamical processes on multitype networks.

Appendix: Effective Degree Distributions in Multitype Networks

Using the configuration model to generate multitype networks—including bipartite and multipartite graphs, graphs with arbitrary distributions of subgraphs, and the modular graphs considered here—requires matching edge stubs within and among types. For example, for the interacting networks considered here, the number of edge stubs pointing from network a to network b must equal the number from b to a . A standard practice in the literature (45, 53), which is more restrictive than needed, is to require that the averages of the interdegrees over the degree distributions agree (e.g., $\langle k_{ab} \rangle = \langle k_{ba} \rangle$ for two networks a, b of equal size). In fact, most any degree distributions can be used, as long as conditioning on matching edge stubs among the types of nodes leaves some probability. Requiring that the degree distributions satisfy, for example, $\langle k_{ab} \rangle = \langle k_{ba} \rangle$, merely tips the scales in favor of valid degree sequences.

Here we derive the effective degree distribution in multitype networks generated with the configuration model. The idea is simple: Because nodes draw degrees independently, the probability distribution of the total number of edge stubs from, say, network a to b is given by a convolution of the degree distribution. We state it for the two interacting networks considered here, but

it can be easily generalized to, say, role distributions (53), which require matching edge stubs with ratios different from one.

Suppose networks a, b have N_a, N_b many nodes, respectively. Let $\vec{K}_{ab}, \vec{K}_{ba}$ be the random variables for the sequences of “inter-degrees” k_{ab}, k_{ba} of the nodes in networks a, b , respectively, drawn from the input degree distributions $p_a(k_{aa}, k_{ab}), p_b(k_{ba}, k_{bb})$. Suppose, for simplicity, that the internal and external degrees are independent [$p_a(k_{aa}, k_{ab}) = p_{aa}(k_{aa})p_{ab}(k_{ab}), p_b(k_{ba}, k_{bb}) = p_{ba}(k_{ba})p_{bb}(k_{bb})$], although this is not essential. We denote $\Sigma k \equiv \sum_{i=1}^n k_i$ for $k \in \mathbb{Z}^n$

Lemma 1. *With the assumptions in the previous paragraph, the effective interdegree distribution for network a is not the input one, $\Pr(\vec{K}_{ab} = k) = \prod_{i=1}^{N_a} p_{ab}(k_i)$, but rather the conditional probability distribution*

$$\Pr(\vec{K}_{ab} = \vec{k} | \Sigma \vec{K}_{ba} = \Sigma \vec{K}_{ab}) = \Pr(\vec{K}_{ab} = \vec{k}) \frac{p_{ba}^*(\Sigma \vec{k})}{\sum_{\ell} p_{ab}^*(\ell) p_{ba}^*(\ell)},$$

where $p_{ab}^*(\cdot), p_{ba}^*(\cdot)$ are p_{ab}, p_{ba} convolved N_a, N_b times, respectively.

Proof: Using the independence of \vec{K}_{ab} and \vec{K}_{ba} , we have

$$\Pr(\vec{K}_{ab} = \vec{k} | \Sigma \vec{K}_{ba} = \Sigma \vec{K}_{ab}) = \frac{\Pr(\vec{K}_{ab} = \vec{k}) \Pr(\Sigma \vec{K}_{ba} = \Sigma \vec{k})}{\Pr(\Sigma \vec{K}_{ba} = \Sigma \vec{K}_{ab})}.$$

In the numerator, we recognize $\Pr(\Sigma \vec{K}_{ba} = \Sigma \vec{k})$ to be p_{ba} convolved N_b times, evaluated at $\Sigma \vec{k}$. In the denominator, use independence, recognize convolutions, and sum on $\ell \geq 0$.

Lemma 1 shows that the effective interdegree distribution of a is the input degree distribution p_{ab} reduced by an amount given by a fraction of convolutions. This reduction in probability governs how many invalid degree sequences must be generated before generating a valid one. For systems sizes on the order of 10^4 nodes, as considered here, generating degree sequences until producing a valid one is quite feasible, as it takes merely seconds. However, for millions of nodes or more, it is better to generate degree sequences $\vec{K}_{ab}, \vec{K}_{ba}$ once, and then repeatedly choose a node uniformly at random to redraw its degree from its degree distribution until the degree sequences are valid. However, this method does not escape the effect in Lemma 1, which is often subtle but can be substantial if the supports of the convolutions of the degree distributions have little overlap. The interdegree distributions used here, Bernoulli and correlated-Bernoulli with identical expected total interdegree, have “much overlap,” so the effective interdegree distribution is approximately the input one, and the correction factor in Lemma 1 can be neglected.

ACKNOWLEDGMENTS. We thank the members of the Statistical and Applied Mathematical Sciences Institute (SAMSI) Dynamics on Networks working group for useful discussions. C.D.B. was partially supported by National Science Foundation Vertical Integration of Research and Education in the Mathematical Sciences DMS0636297, a Graduate Assistance in Areas of National Need Fellowship, the SAMSI, and the Department of Defense through the National Defense Science and Engineering Graduate Fellowship Program. We also gratefully acknowledge support from the following agencies: the Defense Threat Reduction Agency, Basic Research Award HDTRA1-10-1-0088, the Army Research Laboratory, Cooperative Agreement W911NF-09-2-0053, and the National Academies Keck Futures Initiative, Grant CS05.

- Little RG (2002) Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. *J Urban Technol* 9:109–123.
- Rinaldi SM (2004) Modeling and simulating critical infrastructures and their interdependencies. *38th Hawaii International Conference on System Sciences* (IEEE Computer Society, Washington, DC), pp 1–8.
- Pederson P, Dudenhoefler D, Hartley S, Permann M (2006) Critical infrastructure interdependency modeling: A survey of U.S. and international research. (Idaho National Laboratory, Idaho Falls, ID) Report INL/EXT-06-11464.
- Panzieri S, Setola R (2008) Failures propagation in critical interdependent infrastructures. *Int J Model Identif Control* 3:69–78.
- Amin M (2000) National infrastructure as complex interactive networks. *Automation, Control and Complexity: An Integrated Approach*, eds T Samad and J Weyrauch (Wiley, New York), pp 263–286.
- Grubestic TH, Murray A (2006) Vital nodes, interconnected infrastructures and the geographies of network survivability. *Ann Assoc Am Geogr* 96:64–83.
- Hanlon M (2010) How we could all be victims of the volcano... and why we must hope for rain to get rid of the ash. *Daily Mail*, <http://www.dailymail.co.uk/news/article-1267111/>.
- Cederman LE (2003) Modeling the size of wars: From billiard balls to sandpiles. *Am Pol Sci Rev* 97:135–150.
- US Energy Information Administration (2007) Electric power industry overview, <http://www.eia.doe.gov/electricity/page/prim2/toc2.html>.
- Lerner E (2003) What's wrong with the electric grid? *Ind Phys* 9:8–13.
- Joyce C (April 28, 2009) Building power lines creates a web of problems. (National Public Radio), <http://www.npr.org/templates/story/story.php?storyId=103537250>.
- Amin M, Stringer J (2008) The electric power grid: Today and tomorrow. *Mater Res Soc Bull* 33:399–407.
- Cass S (2011) Joining the dots. *Technol Rev* 114:70.
- Natural Environment Research Council (2004) How and why the blackout began in Ohio. *Final Report on the August 14th Blackout in the United States and Canada* (US-Canada Power System Outage Task Force), pp 45–72.
- Dobson I, Carreras BA, Lynch VE, Newman DE (2007) Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos* 17:026103.
- Pepyne DL (2007) Topology and cascading line outages in power grids. *J Syst Sci Syst Eng* 16:202–221.
- Hines P, Cotilla-Sanchez E, Blumsack S (2010) Do topological models provide good information about electricity infrastructure vulnerability? *Chaos* 20:033122.
- Anghel M, Werley KA, Motter AE (2007) Stochastic model for power grid dynamics. *40th Hawaii International Conference on System Sciences* (IEEE Computer Society, Washington, DC), pp 1–11.
- Newman DE, et al. (2005) Risk assessment in complex interacting infrastructure systems. *38th Hawaii International Conference on System Sciences* (IEEE Computer Society, Washington, DC), 1–10.
- Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464:1025–1028.
- Parshani R, Buldyrev S, Havlin S (2010) Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys Rev Lett* 105:048701.
- Gao J, Buldyrev S, Havlin S, Stanley HE (2011) Robustness of a Network of Networks. *Phys Rev Lett* 107:195701.
- Bak P, Tang C, Wiesenfeld K (1987) Self-organized criticality: An explanation of 1/f noise. *Phys Rev Lett* 59:381–384.
- Bak P, Tang C, Wiesenfeld K (1988) Self-organized criticality. *Phys Rev A* 38:364–374.
- Beggs JM, Plenz D (2003) Neuronal avalanches in neocortical circuits. *J Neurosci* 23:11167–11177.
- Juanico DE, Monterola C (2007) Background activity drives criticality of neuronal avalanches. *J Phys A* 40:9297–9309.
- Ribeiro T, Copelli M, Caixeta F, Belchior H (2010) Spike avalanches exhibit universal dynamics across the sleep-wake cycle. *PLoS One* 5:e14129.
- Haldane AG, May RM (2011) Systemic risk in banking ecosystems. *Nature* 469:351–355.
- Saichev A, Sornette D (2004) Anomalous power law distribution of total lifetimes of branching processes: Application to earthquake aftershock sequences. *Phys Rev E Stat Nonlin Soft Matter Phys* 70:046123.
- Hergarten S (2003) Landslides, sandpiles, and self-organized criticality. *Nat Hazard Earth Sys* 3:505–514.
- Sinha-Ray P, Jensen HJ (2000) Forest-fire models as a bridge between different paradigms in self-organized criticality. *Phys Rev E Stat Nonlin Soft Matter Phys* 62:3215–3218.
- Malamud BD, Morein G, Turcotte DL (1998) Forest fires: An example of self-organized critical behavior. *Science* 281:1840–1842.
- Lu ET, Hamilton RJ (1991) Avalanches and the distribution of solar flares. *Astrophys J* 380:L89–L92.
- Paczuski M, Boettcher S, Baiasi M (2005) Interoccurrence times in the Bak-Tang-Wiesenfeld sandpile model: A comparison with the observed statistics of solar flares. *Phys Rev Lett* 95:181102.
- Bonabeau E (1995) Sandpile dynamics on random graphs. *J Phys Soc Jpn* 64:327–328.
- Lise S, Paczuski M (2002) Nonconservative earthquake model of self-organized criticality on a random graph. *Phys Rev Lett* 88:228301.
- Lahtinen J, Kertész J, Kaski K (2005) Sandpiles on Watts-Strogatz type small-worlds. *Physica A* 349:535–547.
- de Arcangelis J, Herrmann HJ (2002) Self-organized criticality on small world networks. *Physica A* 308:545–549.
- Goh KI, Lee DS, Kahng B, Kim D (2003) Sandpile on scale-free networks. *Phys Rev Lett* 91:148701.
- Lee DS, Goh KI, Kahng B, Kim D (2004) Sandpile avalanche dynamics on scale-free networks. *Physica A* 338:84–91.
- Lee E, Goh KI, Kahng B, Kim D (2005) Robustness of the avalanche dynamics in data-packet transport on scale-free networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 71:056108.
- Vaquez A (2006) Spreading dynamics on heterogeneous populations: Multitype network approach. *Phys Rev E Stat Nonlin Soft Matter Phys* 74:066114.

Supporting Information

Brummitt et al. 10.1073/pnas.1110586109

SI Text

Robustness of Optimal Interconnectivity. How p^* depends on system size, connectivity, and type of degree distribution. Our central result is the minimum $p^* > 0$ in the chance of a large cascade in a network when a fraction p of its nodes are connected to nodes in another network. Important questions are how p^* depends on system size, behaves in the thermodynamic limit (infinite system size), and depends on the internal degrees and on the degree distribution. Here we show that the qualitative form of $\Pr(T_a > \mathcal{C})$ in Fig. 4 appears to be generic.

Doubling the system size and keeping the cutoff \mathcal{C} and the dissipation f fixed (at $\mathcal{C} =$ half the number nodes in a and at $f = 0.01$) does not significantly change Fig. 4, because the dissipation f limits the chance of large cascades. But doubling the system size, doubling the cutoff \mathcal{C} , and halving the dissipation f (so that the largest cascades topple nearly the whole system) slightly decreases p^* . Fig. S4 shows similar results as Fig. 4 but for a system half the size (1,000 nodes per network, $f = 0.02$, cutoff $\mathcal{C} = 500$). In the thermodynamic limit of infinite system size N , we expect p^* to stay bounded away from zero, because what appears to be the determinant is the ratio of edges between the networks (pN) and within them ($z_a N$).

Increasing the internal degrees of nodes [say, to coupled random four-regular graphs $R(4)-B(p)-R(4)$] increases their capacity and hence ability to withstand inflicted cascades, so the minimum in $\Pr(T_a > \mathcal{C})$ is therefore larger, $p^* \approx 0.2$ (Fig. S5), compared to $p^* \approx 0.12$ for random three-regular graphs with equally many nodes (Fig. S4). The networks also have a wider range of optimal p , as expected, given that pN_a external edges are less significant than the $z_a N_a$ internal edges when $z_a = 4$ (Fig. S5) compared to $z_a = 3$ (Fig. 4 and Fig. S4). Nonetheless, the chance of a large cascade eventually increases with p . We also note that power grids more closely resemble random three-regular than four-regular graphs (see Table S1).

Next we determined the effect of introducing some degree heterogeneity by adjusting interconnectivity p between two Erdős-Rényi random graphs. Similar results were obtained as for random regular graphs, indicating that some degree heterogeneity does not affect these results. (We did not test heavy-tailed degree distributions because these rarely occur in the physical infrastructure networks of interest, namely, power grids.)

Finally, we tested whether adding edges at random to a random regular graph produces an optimal connectivity that minimizes the chance of a large cascade. Specifically, each node in a three-random regular graph gains an extra, fourth edge stub independently with probability p [i.e., degree distribution $P(k)$ is $P(3) = 1 - p$, $P(4) = p$]. We find in simulations that $\Pr(T > 500)$ for a system of 10^3 nodes is largely independent of the mean degree z . In particular, the chance of a large cascade does not drop, as it does when initially adding edges between two networks (Fig. 4 and Figs. S4 and S5), which agrees with theoretical results for isolated networks (1): the avalanche size distribution $s(t) \sim t^{-3/2}$ for large t . We conclude that adding links between two networks affects vulnerability to large cascades in a different way than adding links within isolated networks.

Unstable p^* for small cascades, stable p^* for large cascades. Networks seeking to mitigate cascades of small, intermediate, or large sizes would seek different interconnectivity, as shown in the plots of $\Pr(\ell \leq T_a \leq \ell + 50)$ as a function of p in Fig. 6. In Fig. S7, we show more plots to show the change in concavity at intermedi-

ate cascade size and the robustness of the location of the stable minimum $p^* \approx 0.075$ for $R(3)-B(p)-R(3)$.

Increasing capacity fuels larger system-wide cascades. Introducing new interconnections between networks slightly enlarges the largest global cascades. We ask to what extent is this due to the direction of links (pointing internally or externally) and due to the additional capacity for holding sand endowed by the new links (because capacities of nodes are their degrees). To isolate these two effects, we run simulations on two random regular graphs interconnected by “correlated-Bernoulli” coupling: Each node changes an internal edge stub to an external edge stub with probability p .

The resulting total avalanche size distribution, Fig. S6, does not significantly change for different values of interconnectivity p . Thus, what causes the slight enlargement of the largest global avalanches is the slight increase in total capacity of the network, not the direction of links. Moreover, the enlargement of the largest global cascades is consistent with the increase in capacity. For the simulations on random three-regular graphs in Fig. S6, the n th largest avalanche contains on average 1.4 ± 2.6 (mean \pm standard deviation) more topplings with $p = 0.1$ compared to $p = 0.005$, which is an insignificant difference. By contrast, in an analogous simulation on $R(3)-B(0.1)-R(3)$, in which we introduce extra links with probability p , the n th largest avalanche contains 204.8 ± 92.5 more topplings than $R(3)-B(0.005)-R(3)$, a significant difference for networks with 10^3 nodes. Furthermore, this difference is on the order of the additional capacity of $2 \times 10^3 \times (0.1 - 0.005) = 190$ among 2×10^3 nodes with $B(0.1)$ compared to $B(0.005)$ coupling.

We find similar results when introducing eight additional interconnections between power grids c and d . The n th largest global cascades in simulations are larger by an amount on the order of the additional capacity of the two networks. Thus the additional capacity due to the new interconnections, not the direction of links, largely explains the amplification of system-wide cascades.

Bounding the critical load disparity r^* for power grids c, d . In the main paper, we showed that, for 16 interconnections between power grids c and d , the critical load disparity is $r^* \approx 15$. That is, when sand is dropped 15 times more frequently on c nodes than on d nodes, the largest inflicted cascades from c to d approximately equal in size the largest local cascades begun in d . (Recall that by “inflicted cascade from c to d ” we mean the number of topplings in d in an avalanche begun in c , and by “local cascade in d ” we mean the number of topplings in d in an avalanche begun in d .) If d builds more interconnections (or increases the load disparity r), the largest inflicted cascades become larger than the largest local cascades. On the other hand, delete interconnections (or decrease the load disparity r), and d could mitigate its largest local cascades more than the enlargement of the largest inflicted cascades. Hence $r^* = 15$ and 16 interconnections represent a “modularity equilibrium” in local and inflicted cascades.

Here we use simulations to approximately bound the critical load disparity $10 \lesssim r^* \lesssim 20$ for 16 interconnections between c and d . As shown in Fig. S9 for $r = 10, 20$, the largest inflicted cascades (green dashed) are smaller and larger, respectively, than the largest local cascades (gold dashed), indicating that $10 \lesssim r^* \lesssim 20$ for 16 interconnections. For eight interconnections (solid blue and red curves), the critical load disparity r^* is evidently larger than 20, as the largest inflicted cascades are smaller

than the largest local cascades for $r = 20$ (red curve below blue curve).

Power Grid Topologies. Power grids c and d . The readily available datasets on power grids—Institute of Electrical and Electronic Engineers test cases (2), Western States (3), and New York Independent System Operator (4)—did not appear to contain multiple electrical grids connected to one another. As a result, we requested data on power grid connectivity from the US Federal Energy Regulation Commission via the Critical Energy Infrastructure Information program (5). We focused on the southeastern region of the United States, for which we had the output files of power simulations of various grids. The southeastern region consists of areas, distinct grids owned by different utility companies. The largest areas contain thousands of buses, an electrical grid term for the connection points that link generators, transmission lines, and transformers. We ignore wind turbines, for they do not belong to specific areas. Among the areas with at least 100 nodes, we chose grids c and d by first selecting the grid with the highest average internal degree (2.91), and then selecting the area, c , to which it had the most interconnections (eight of them). Grid d consists of one giant component with 504 nodes, whereas grid c consists of one giant component of size 439, and we ignore 14 additional nodes.

Two of the statistics of power grids c and d are important to our study: the narrow degree distribution and small clustering coefficient. Some details are in Table S1. The average clustering coefficient $\langle C \rangle$, the fraction of possible triangles that exist, begins to measure how locally tree-like power grids are. We find that $\langle C \rangle \approx 0.05$ is low, yet an order of magnitude larger than the clustering coefficient of a random three-regular graph with as many nodes as grid d . The average shortest path length among pairs of nodes, $\langle \ell \rangle$, which has been identified as a likely source of difficulty for predicting cascades on networks (6), is rather large in the power grids ($\langle \ell \rangle \approx 10$) because of their two-dimensional, spatial, nearly planar structure. By contrast, random three-regular graphs have significantly smaller diameter $\langle \ell \rangle \approx 7.09 \pm 0.03$.

Correlation between degree and load. In the sandpile model studied here, we choose the capacity of every node to be its degree, so that toppled nodes unambiguously shed one grain to each neighbor. To examine the reasonableness of this assumption for real infrastructure, we sum the apparent power on the edges incident to each bus in power grids c and d , and we plot in Fig. S3 a density histogram of apparent power versus degree of the nodes. Most buses have low degree and low apparent power, and some buses have high degree but low power, whereas others bear high power among few connections. But the general trend is positive: Each additional degree correlates with 123 mVA of additional apparent power, though the correlation is tenuous ($R^2 = 0.30$). This correlation suggests that using the sandpile model with capacity equal to degree in order to study cascades of load in power grids—and by extension other infrastructure—is not unreasonable for the purpose of obtaining heuristic understanding.

Branching Process Approximation. Comparing theory and experiment. We choose the dissipation of sand f (the chance that a grain of sand is deleted as it is shed from one node to another) so that the largest cascades in simulations topple almost all the network. (As a rule of thumb, take $f = 20/N$, where N is the total number of nodes.) Decreasing or increasing f extends or shortens the tail of the avalanche size distribution, respectively. Thus we tune f to utilize the entire system and to achieve a power law over as many scales as possible, because we see power law behavior in, e.g., blackouts in power grids (7).

In Fig. 3 of the main paper, we compare simulations of sandpile cascades with theoretical predictions for dissipation rate of sand $f = 0.1$. This somewhat large dissipation rate mitigates the

largest cascades. In Fig. S1, we show results for the same branching process prediction (seven iterations of the self-consistency equations, started from $\mathcal{S}_a = \mathcal{S}_b = 1$) with two independent simulations with dissipation rate of sand $f = 0.05$. Halving the dissipation rate noticeably extends the tails of the marginalized avalanche size distributions—i.e., the largest avalanches become larger. Iterating the self-consistency equations an eighth time would more accurately compute the probability mass in the tails, but is beyond our computer resources. Showing the results of two independent simulations (blue and green curves of Fig. S1) illustrates the variation in the cascade size distributions. In particular, variation between simulations is apparent only in the tails of the distributions because the largest avalanches are so rare.

Even though the behavior of sandpile cascades on single networks is rather robust to network structure (1, 8–12), we see sensitivity to the sparse connections among modules, as demonstrated in the main paper. What is more, the particular degrees of nodes with external links profoundly affect the inflicted avalanche size distributions, $\sum_{t_a} s_a(t_a, t_b)$ and $\sum_{t_b} s_b(t_a, t_b)$. Notice the characteristic “blips” (deviations from straight lines) that occur in the blue and green curves (the simulations) in the inflicted avalanche size distributions in the upper-right and lower-left plots in Fig. S1. For instance, there is a particularly high probability that an avalanche begun in d topples 8, 9, or 10 nodes in c . We suspect this enhanced probability is because of the outlier node in network c that has one external link and 10 internal links—far more internal links than all other c nodes with an external link. Consequently, a cascade in d that leaks across to c via this node topples with high probability around 8, 9, or 10 nodes in c because of this node’s high internal degree. Similarly, the blip at size three of the distribution of inflicted cascade size from c to d results from the average internal degree 2.5 of d nodes with an external link. It is surprising, given the indifference of sandpile cascades to network structure, that the particular internal degrees of nodes with an external link greatly affect the marginalized avalanche size distributions.

Distance between the avalanche size distributions and the product of their marginals. When we compare the branching process predictions with simulations, such as in Fig. 3 and Fig. S1, we reduce the dimensions of the joint $s_a(t_a, t_b)$ by marginalizing in order to plot one-dimensional curves. This marginalization raises the question: How far are the products of the marginals, such as $(\sum_{t_a} s_a(t_a, t_b))(\sum_{t_b} s_b(t_a, t_b))$, from the joint distributions? Using the first 11^2 coefficients of $s_a(t_a, t_b)$ (for $0 \leq t_a, t_b \leq 11$) computed using Lagrange inversion (Theorem 1), we computed the distance between $s_a(t_a, t_b)$ and the product of its marginals—in the two-norm, one-norm, and Kullback–Leibler divergence—as a function of Bernoulli coupling p between two random three-regular graphs. With increasing coupling p , the cascade sizes t_a, t_b become increasingly correlated, and the joint distribution s_a grows increasingly distant from the product of its marginals. This increasing distance suggests that comparing marginals, as in Fig. 3 and Fig. S1, is an insufficient test of the branching process, and thus we next subject the branching process to its most stringent test.

Fine-grained comparison of branching process and simulation. Because marginalizing s_a, s_b , as in Fig. 3 and Fig. S1, may obscure deviation between theory and the branching process, we next compare the joint distribution $s_a(t_a, t_b)$ to simulation. In Fig. S2, we plot $\Pr(T_a = 10, T_b = x)$ for $0 \leq x \leq 30$ in the simulation (blue) and branching process prediction (red), for interconnectivity $p = 0.005, 0.01, 0.1$. We compute the branching process prediction by calculating $s_a(t_a, t_b), s_b(t_a, t_b)$ for $t_a = 10$ and $0 \leq t_b \leq 30$ using Lagrange inversion (Theorem 1). Because sand is dropped on nodes uniformly at random and each network has 10^3 nodes, the avalanches begin in a or in b with equal probability, so that

$Pr(T_a = 10, T_b = x) = (s_a(10x) + s_b(10x))/2$. Although the branching process underpredicts theory for larger avalanches (especially with greater interconnectivity), we see qualitatively good agreement.

Capacity disparity. Cascades inflicted from large capacity networks to small capacity networks are larger than those from small to large capacity networks. Here we derive the heuristic formula for this effect (Eq. 4) using the multitype branching process, and we show qualitative agreement with simulation.

Differentiating the self-consistency Eqs. 3 with respect to τ_a and τ_b and setting $\tau_a = \tau_b = 1$ yields four equations for the first moments of the avalanche size distributions s_a, s_b :

$$\langle s_a \rangle_a = \frac{\langle u_b \rangle_b - 1}{\gamma}, \quad \langle s_b \rangle_b = \frac{\langle u_a \rangle_a - 1}{\gamma}, \quad \text{[S1a]}$$

$$\langle s_a \rangle_b = \frac{-\langle u_a \rangle_b}{\gamma}, \quad \langle s_b \rangle_a = \frac{-\langle u_b \rangle_a}{\gamma}, \quad \text{[S1b]}$$

with common denominator

$$\gamma = -1 + \langle u_a \rangle_b \langle u_b \rangle_a - \langle u_a \rangle_a (\langle u_b \rangle_b - 1) + \langle u_b \rangle_b. \quad \text{[S2]}$$

For the first moments of the branch distributions u_a, u_b for $R(z_a)-B(p)-R(z_b)$,

$$\langle u_a \rangle_a = 1 - \frac{P}{1 + z_a}, \quad \langle u_a \rangle_b = \frac{P}{1 + z_b} \quad \text{[S3]}$$

(and vice versa for u_b , interchanging a with b), the denominator [S2] is zero, and Eqs. S1 are all infinite. This result is not a surprise, because in isolation the networks' cascade size distributions are asymptotically power laws with exponent $-3/2$, which have infinite first moments. However, one can compare the rates of divergence of the mean inflicted cascade sizes, Eq. S1b, by computing their ratio, $\langle s_a \rangle_b / \langle s_b \rangle_a = (1 + z_a) / (1 + z_b)$. No other ratios appear to contain useful information (nor dependence on p with which to estimate the critical amount of interconnectivity p^*):

$$\frac{\langle s_a \rangle_a}{\langle s_b \rangle_a} = \frac{\langle s_a \rangle_a}{\langle s_b \rangle_b} = \frac{1 + z_a}{1 + z_b}, \quad \frac{\langle s_a \rangle_b}{\langle s_b \rangle_b} = 1.$$

In all simulations on $R(z_a)-B(p)-R(z_b)$ with $z_a \neq z_b$, the heuristic formula (Eq. 4) works qualitatively, in that inflicted cascades are always larger from the larger capacity network (i.e., $z_a < z_b \Rightarrow$

$\langle s_a \rangle_b < \langle s_b \rangle_a$). But finding quantitative agreement with (Eq. 4) is difficult because the avalanche size distributions are power laws with such heavy tails that their variances diverge. Fig. S8 is an example: For $R(3)-B(0.005)-R(3)$, $\langle s_a \rangle_b = 0.086 \pm 0.019$, and $\langle s_b \rangle_a = 1.32 \pm 0.023$ (mean \pm standard error), which yields $\langle s_a \rangle_b / \langle s_b \rangle_a = 0.65 \pm 0.019$, whereas theory predicts $(1 + z_a) / (1 + z_b) = 4/5$ (20% error). Nevertheless, Eq. 4 provides a simple, useful heuristic for the dangers of small capacity networks connecting to large capacity networks.

Computational challenges in showing the locally stabilizing effect.

Although the locally stabilizing effect of interconnections is apparent in simulations, demonstrating it analytically poses mathematical and computational challenges. Here we use multidimensional Lagrange inversion (Theorem 1 in *Materials and Methods*) to solve for the probabilities of the smallest avalanches, $s_a(t_a, t_b)$ for $0 \leq t_a, t_b \leq 10$. Next we plot in Fig. S10A the marginalized avalanche size distribution $\sum_{t_b=0}^{10} s_a(t_a, t_b)$.

Although it appears in Fig. S10A that the largest avalanches become less likely with increasing interconnectivity p , in accordance with the results in simulations, this figure has a caveat: We have only computed up to avalanche size at most 10 in each network. As a result, the right-hand tail of Fig. S10A lacks probability mass compared to the actual marginalized avalanche size distribution $\sum_{t_b=0}^{\infty} s_a(t_a, t_b)$, because, for example, there is a significant chance of an avalanche of size $t_a = 10$ and $t_b = 11, 12$, or 13 . This figure could be improved by computing $s_a(t_a, t_b)$ along "vertical strips" $0 \leq t_b \leq t_b^{\max}$, where $t_b^{\max} \gg 1$, for just a handful of values t_a . However, computing the large coefficients of \mathcal{S}_a requires differentiating expressions raised to large powers, so the largest coefficients are the most difficult to compute.

The Cauchy formula, Eq. 5 of *Materials and Methods*, uses integration rather than differentiation, which can be more numerically stable (13). But computing the largest coefficients $s_a(t_a, t_b)$ with this method requires integrating an increasingly large expression, namely, the result of iterating the self-consistency equations at least $t_a + t_b + 1$ times, starting from $\mathcal{S}_a = \mathcal{S}_b = 1$. This approach, too, takes a long amount of computation time: Computing $s_a(t_a, t_b)$ for $0 \leq t_a, t_b \leq 20$ on a typical laptop computer would take on the order of a week.

Nonetheless, even the smallest 11^2 coefficients of s_a and s_b are useful. Fig. S10B, for example, shows the joint avalanche size distribution $s_a(t_a, t_b)$ for two random regular graphs with Bernoulli (p)-distributed coupling, for $p = 0, 0.1, 1$. As the interconnectivity p increases, the cascades begun in a increasingly "smear out" among a and b , and cascades are more frequently large in both networks at once.

- Goh KI, Lee DS, Kahng B, Kim D (2003) Sandpile on scale-free networks. *Phys Rev Lett* 91:14701.
- University of Washington Electrical Engineering (1999) Power Systems Test Case Archive (Univ Washington, Seattle). Available at <http://www.ee.washington.edu/research/pstca/>. Accessed October 22, 2010.
- Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. *Nature* 393:440–442.
- Wang Z, Scaglione A, Thomas RJ (2010) Generating statistically correct random topologies for testing smart grid communication and control networks. *IEEE Trans Smart Grid* 1:28–39.
- Critical Energy Infrastructure Information (CEII), US Federal Energy Regulatory Commission (FERC), (Received January 5, 2011). Available at <http://www.ferc.gov/legal/ceii-foia/ceii.asp>.
- Melnik S, Hackett A, Porter M, Mucha P, Gleeson J (2011) The unreasonable effectiveness of tree-based theory for networks with clustering. *Phys Rev E Stat Nonlin Soft Matter Phys* 83:036112.
- Dobson I, Carreras BA, Lynch VE, Newman DE (2007) Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization. *Chaos* 17:026103.
- Lee DS, Goh KI, Kahng B, Kim D (2004) Sandpile avalanche dynamics on scale-free networks. *Physica A* 338:84–91.
- Lee E, Goh KI, Kahng B, Kim D (2005) Robustness of the avalanche dynamics in data-packet transport on scale-free networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 71:056108.
- Bonabeau E (1995) Sandpile dynamics on random graphs. *J Phys Soc Jpn* 64:327–328.
- Lahtinen J, Kertész J, Kaski K (2005) Sandpiles on Watts-Strogatz type small-worlds. *Physica A* 349:535–547.
- de Arcangelis J, Herrmann HJ (2002) Self-organized criticality on small world networks. *Physica A* 308:545–549.
- Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys Rev E Stat Nonlin Soft Matter Phys* 64:026118.

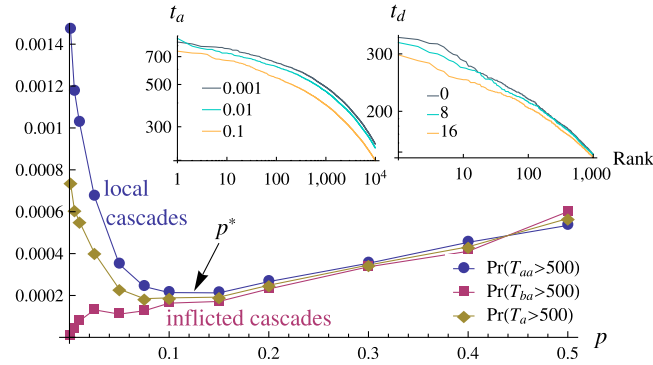


Fig. 54. For half the system size ($N_a = 1,000$), half the cutoff ($\mathcal{C} = 500$), and double the dissipation ($f = 0.02$) of the system in Fig. 4, the chance of a large cascade in a is qualitatively similar, and $p^* \approx 0.12 \pm 0.02$ is slightly larger (in simulations with 2×10^6 grains dropped after 10^5 transients). This gold curve and its critical point is stable to cutoffs \mathcal{C} different from 500 ($200 \leq \mathcal{C} \leq 800$). (Insets) Rank-size plots on log-log scales of the largest cascades in network a (Left) for $p = 10^{-3}, 10^{-2}, 10^{-1}$, and in power grid d (Right) connected to c by 0, 8, or 16 edges.

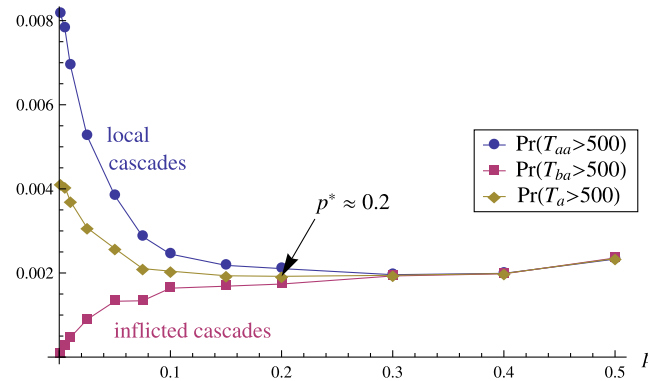


Fig. 55. Increasing the internal degrees of both networks a, b from three to four [i.e., $R(4)-B(p)-R(4)$] slightly increases the minimum $p^* \approx 0.2$ of the chance of a large cascade in network a . (Here, 10^3 nodes/network; $f = 0.02$; 10^6 grains dropped after 10^5 transients.)

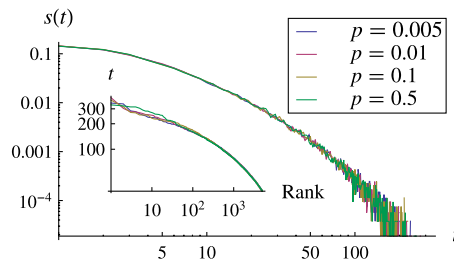


Fig. 56. When nodes change exactly one of their internal edge stubs to be an external edge stub with probability p [regular(z_a)-correlated Bernoulli(p)-regular(z_b)] rather than receiving an external edge stub with probability p [regular(z_a)-Bernoulli(p)-regular(z_b)], global cascades are not significantly amplified. This difference indicates that it is the small increase in capacity (due to the additional edges and the fact that capacities of nodes are their degrees) rather than the direction of edges (internal versus external) that causes the increase in global cascades with increased Bernoulli-distributed coupling between networks. The main plot is the total avalanche size distribution $s(t)$ in a simulation with 10^6 grains of sand dropped (after 10^5 grains dropped without collecting statistics, begun from initial amounts of sand chosen uniformly at random from zero to one less a node's degree) on two networks with 10^3 nodes each, with dissipation of sand $f = 0.05$ and $z_a = z_b = 3$. (Inset) Rank-size plot of largest 5×10^3 cascades, which are nearly indistinguishable.

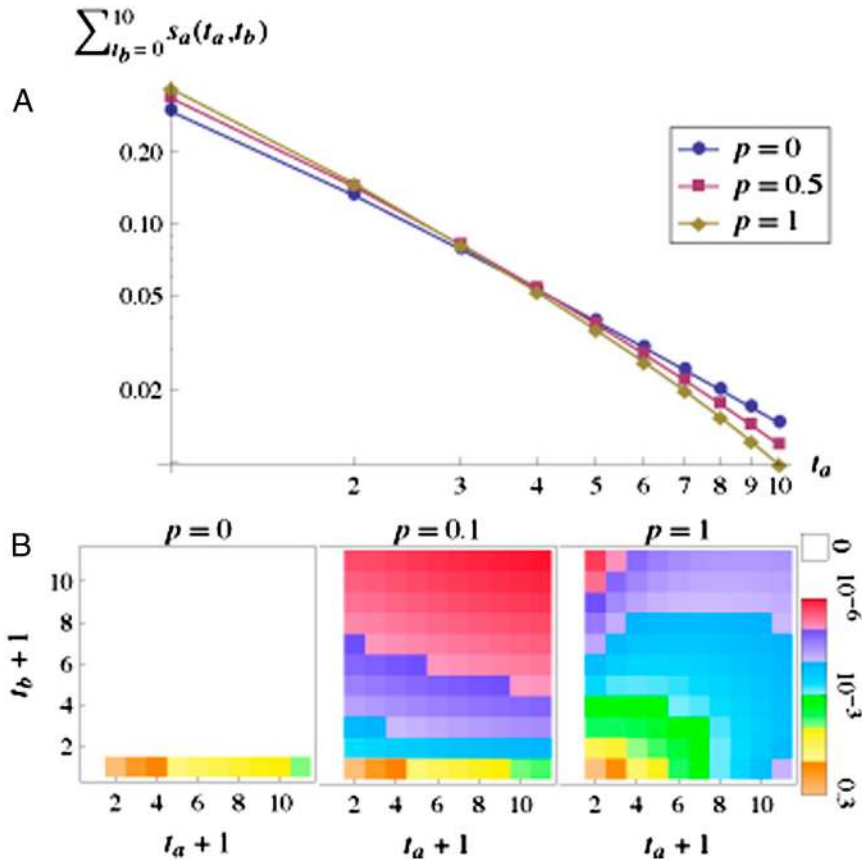


Fig. S10. (A) Increasing the interconnectivity p between two random three-regular graphs appears to mitigate the tails of the local avalanche size distribution $\sum_{t_b} s_a(t_a, t_b)$. Plotted are the first 11×11 coefficients of δ_a computed using multidimensional Lagrange inversion (1) for $p = 0, 0.5, 1$. (B) Increasing the interconnectivity p between two random three-regular graphs “smears out” the avalanche size distribution $s_a(t_a, t_b)$, so that cascades large in both networks become more likely. Plotted are $s_a(t_a, t_b)$ for $0 \leq t_a, t_b \leq 10$, computed using multidimensional Lagrange inversion (1) for $p = 0, 0.1, 1$. We plot zero probabilities in white and positive probabilities in a logarithmic color scale from orange (high probability) to bright red (low probability).

1. Good IJ (1960) Generalizations to several variables of Lagrange's expansion, with applications to stochastic processes. *Proc Cambridge Philos Soc* 56:367–380.

Table S1. Summary statistics of power grids c and d (in isolation and coupled together) and of a random three-regular graph with the same number of nodes as d

	c	d	c and d	three-regular
No. nodes	439	504	943	504
No. internal edges	527	734	1,261	756
No. external edges	8	8	—	—
$\langle k_{\text{internal}} \rangle$	2.40	2.91	2.69	3
$\langle k_{\text{external}} \rangle$	0.0205	0.0179	—	—
$\langle C \rangle$	0.0109	0.0821	0.0488	0.003(2)
$\langle \ell \rangle$	9.32	8.26	11.42	7.09(3)

The statistics for the random three-regular graph are averages over 1,000 realizations, with the standard deviation in parentheses to convey the fluctuation in the last digit.