nature

Catastrophic cascade of failures in interdependent networks

Sergey V. Buldyrev[1,2], Roni Parshani[3], Gerald Paul[2], H. Eugene Stanley[2] & Shlomo Havlin[3]

Complex networks have been studied intensively for a decade, but research still focuses on the limited case of a single, non-interacting network[1–14]. Modern systems are coupled together[15–19] and therefore should be modelled as interdependent networks. A fundamental property of interdependent networks is that failure of nodes in one network may lead to failure of dependent nodes in other networks; if this happens recursively, a cascade of failures can result. In extreme cases, the failure of a very small fraction of nodes in one network may lead to the complete fragmentation of a system of several interdependent networks. A dramatic real-world example of a cascade of failures ('concurrent malfunction') is the electrical blackout that affected much of Italy on 28 September 2003: the shutdown of power stations directly led to the failure of nodes in the Internet communication network, which in turn caused further breakdown of power stations[20]. Here we develop a framework for understanding the robustness of interacting networks subject to such cascading failures. We present exact analytical solutions for the critical fraction of nodes that, on removal, will lead to a failure cascade and to a complete fragmentation of two interdependent networks. Surprisingly, a broader degree distribution increases the vulnerability of interdependent networks to random failure, which is opposite to how a single network behaves. Our findings highlight the need to consider interdependent network properties in designing robust networks.

Today's networks are becoming increasingly dependent on one another. Diverse infrastructures such as water supply, transportation, fuel and power stations are coupled together. We show that owing to this coupling, interdependent networks are extremely sensitive to random failure, such that a random removal of a small fraction of nodes from one network can produce an iterative cascade of failures in several interdependent networks. Electrical blackouts frequently result from a cascade of failures between interdependent networks, and the problem has been dramatically exemplified by the several large-scale blackouts that have occurred in recent years. In this Letter, we demonstrate a cascade of failures using real-world data from a power network and an Internet network (a supervisory control and data acquisition system) that were implicated in the blackout that affected much of Italy on 28 September 2003[20]. These two networks feature a bidirectional dependence such that power stations depend on communication nodes for control and communication nodes depend on power stations for their electricity supply.

Figure 1 shows the two networks and the connections between them, based on the real geographical locations. The figure exemplifies a situation in which an initial failure of only one power station may lead to an iterative cascade of failures that causes both networks to become fragmented. For an isolated single network, a significant number of the network nodes must be randomly removed before the network breaks down. However, when taking into account the

dependencies between the networks, removal of only a small fraction of nodes can result in the complete fragmentation of the entire system.

To model interdependent networks, we consider for simplicity, and without loss of generality, two networks, A and B, with the same number of nodes, N. The functioning of node A_i ($i = 1, 2, …, N$), in network A, depends on the ability of node B_i, in network B, to supply a critical resource, and vice versa. If node A_i stops functioning owing to attack or failure, node B_i stops functioning. Similarly, if node B_i stops functioning then node A_i stops functioning. We denote such a dependence by a bidirectional link, $A_i \leftrightarrow B_i$, that defines a one-to-one correspondence between nodes of network A and nodes of network B. Nodes from network A and B are randomly connected. Within network A, the nodes are randomly connected by A-links with degree distribution $P_A(k)$, where the degree, k, of each node is defined as the number of A-links connected to that node in network A. Analogously, within network B, the nodes are randomly connected by B-links with degree distribution $P_B(k)$ (Fig. 2).

We begin by randomly removing a fraction, $1 - p$, of the nodes of network A and removing all the A-links connected to these removed
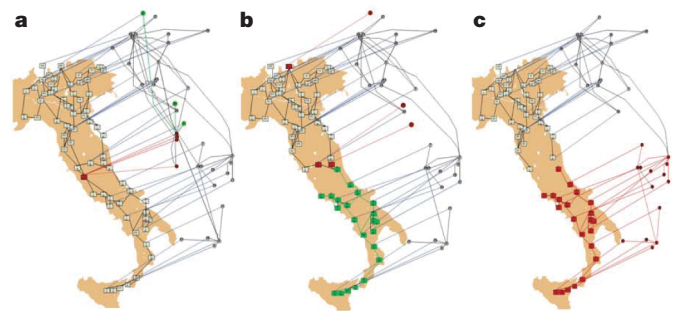

Figure 1 | Modelling a blackout in Italy. Illustration of an iterative process of a cascade of failures using real-world data from a power network (located on the map of Italy) and an Internet network (located above the map) that were implicated in an electrical blackout that occurred in Italy in September 2003[20]. The networks are drawn using the real geographical locations and every Internet server is connected to the geographically nearest power station. **a,** One power station is removed (red node on map) from the power network and as a result the Internet nodes depending on it are removed from the Internet network (red nodes above the map). The nodes that will be disconnected from the giant cluster (a cluster that spans the entire network) at the next step are marked in green. **b,** Additional nodes that were disconnected from the Internet communication network giant component are removed (red nodes above map). As a result the power stations depending on them are removed from the power network (red nodes on map). Again, the nodes that will be disconnected from the giant cluster at the next step are marked in green. **c,** Additional nodes that were disconnected from the giant component of the power network are removed (red nodes on map) as well as the nodes in the Internet network that depend on them (red nodes above map).

[1]Department of Physics, Yeshiva University, 500 West 185th Street, New York, New York 10033, USA. [2]Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA. [3]Minerva Center and Department of Physics, Bar-Ilan University, 52900 Ramat-Gan, Israel.
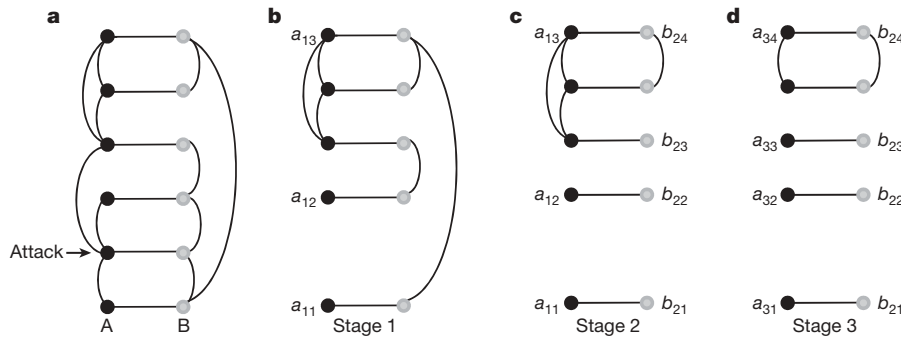
Figure 2 | Modelling an iterative process of a cascade of failures. Each node in network A depends on one and only one node in network B, and vice versa. Links between the networks are shown as horizontal straight lines, and A-links and B-links are shown as arcs. **a**, One node from network A is removed ('attack'). **b**, Stage 1: a dependent node in network B is also eliminated and network A breaks into three a_1-clusters, namely a_{11}, a_{12} and a_{13}. **c**, Stage 2: B-links that link sets of B-nodes connected to separate a_1-clusters are eliminated and network B breaks into four b_2-clusters, namely b_{21}, b_{22}, b_{23} and b_{24}. **d**, Stage 3: A-links that link sets of A-nodes connected to separate b_2-clusters are eliminated and network A breaks into four a_3-clusters, namely a_{31}, a_{32}, a_{33} and a_{34}. These coincide with the clusters b_{21}, b_{22}, b_{23} and b_{24}, and no further link elimination and network breaking occurs. Therefore, each connected b_2-cluster/a_3-cluster pair is a mutually connected cluster and the clusters b_{24} and a_{34}, which are the largest among them, constitute the giant mutually connected component.

nodes (Fig. 2a). Owing to the dependence between the networks, all the nodes in network B that are connected to the removed A-nodes by A↔B links must also be removed (Fig. 2b). Any B-links connected to the removed B-nodes are then also removed. As nodes and links are sequentially removed, each network begins to fragment into connected components, which we call clusters. The clusters in network A and the clusters in network B are different because each network is connected differently. A set of nodes, a, in network A and the corresponding set of nodes, b, in network B form a mutually connected set if (1) each pair of nodes in a is connected by a path that consists of nodes belonging to a and links of network A, and (2) each pair of nodes in b is connected by a path that consists of nodes belonging to b and links of network B. We call a mutually connected set a mutually connected cluster if it cannot be enlarged by adding other nodes and still satisfy the conditions above. Only mutually connected clusters are potentially functional.

To identify these mutually connected clusters, we first define the a_1-clusters as the clusters of network A remaining after a fraction $1 − p$ of the A-nodes are removed as the result of an attack or malfunction (Fig. 2b). This state of the networks is the first stage in the cascade of failures. Next we define the b_1-sets as the sets of B-nodes that are connected to a_1-clusters by A↔B links. According to the definition of mutually connected clusters, all the B-links connecting different b_1-sets must be removed. Because the two networks are connected differently, each b_1-set may split into several clusters, which we define as b_2-clusters (Fig. 2c). The b_1-sets that do not split, and hence coincide with a_1-clusters, are mutually connected. This state of the networks is the second stage in the cascade of failures. In the third stage, we determine all the a_3-clusters (Fig. 2d), in a similar way, and in the fourth stage we determine all the b_4-clusters. We continue this process until no further splitting and link removal can occur (Fig. 2d). We find that this process leads to a percolation phase transition for the two interdependent networks at a critical threshold, $p = p_c$, which is significantly larger than the equivalent threshold for a single network. As in classical network theory[21–25], we define the giant mutually connected component to be the mutually connected cluster spanning the entire network. Below p_c there is no giant mutually connected component, whereas above p_c a giant mutually connected cluster exists.

Our insight based on percolation theory is that when the network is fragmented, the nodes belonging to the giant component connecting a finite fraction of the network are still functional, whereas the nodes that are part of the remaining small clusters become non-functional. Therefore, for interdependent networks only the giant mutually connected cluster is of interest. The probability that two A-linked A-nodes are connected by A↔B links to two B-linked

B-nodes scales as $1/N$ (Supplementary Information). Hence, at the end of the cascade process of failures, above p_c only very small mutually connected clusters and one giant mutually connected cluster exist, in contrast to traditional percolation, wherein the cluster size distribution obeys a power law. When the giant component exists, the interdependent networks preserve their functionality; if it does not exist, the networks split into small fragments that cannot function on their own because they are not connected to any other cluster.

We apply our model first to the case of two Erdős–Rényi networks[21–23] with average degrees $\langle k_A \rangle$ and $\langle k_B \rangle$. We remove a random fraction, $1 − p$, of the nodes in network A and follow the iterative process of forming a_1-, b_2-, a_3-, …, b_{2k}- and a_{2k+1}-clusters as described above. After each stage, n, in the cascade of failures, we determine the fraction of nodes, μ_n, in the largest a_n- or b_n-cluster. At the end of the process, μ_n converges to μ_∞, the probability that a randomly chosen node belongs to the mutually connected largest cluster. As $N \to \infty$, the probability μ_∞ converges to a well defined function, $\mu_\infty(p)$, which has a single step discontinuity at the threshold $p = p_c$, where it changes from zero for $p < p_c$ to $\mu_\infty(p_c) > 0$ for $p = p_c$ (Supplementary Information). This behaviour is characteristic of a first-order phase transition, quite different from a second-order phase transition such as that characterizing percolation of a single network, where $\mu_\infty(p)$ is a continuous function at $p = p_c$. For a finite N and p close to p_c, the giant mutually connected component exists in a particular network realization with probability $P_\infty(p, N)$. As $N \to \infty$, $P_\infty(p, N)$ converges to a Heaviside step function, $\Theta(p − p_c)$, which discontinuously changes value from zero for $p < p_c$ to one for $p > p_c$ (Fig. 3a). Our simulation results for the value of p_c (Fig. 3a) agree with the analytical results presented below.

For two interdependent scale-free networks[2] with power-law degree distributions, $P_A(k) = P_B(k) \propto k^{−\lambda}$, we find that the existence criteria for the giant component are quite different from those in a single network. For a single scale-free network with $\lambda < 3$, a giant component exists for every non-zero value of p. However, for interdependent scale-free networks, the giant component does not exist below the critical value $p_c \neq 0$, even for $2 < \lambda < 3$.

In the case of a single network, p_c is smaller for a broader degree distribution. In sharp contrast, for interdependent networks a broader degree distribution results in a larger value of p_c because high-degree nodes of one network can depend on low-degree nodes of the other. The hubs (defined as nodes of exceptionally large degree) that have a dominant role in the robustness of a single network become vulnerable when a cascade of failures occurs in two interdependent networks. Moreover, a broader distribution with the same average degree implies that there are more low-degree nodes. Because the low-degree nodes are more easily disconnected, the
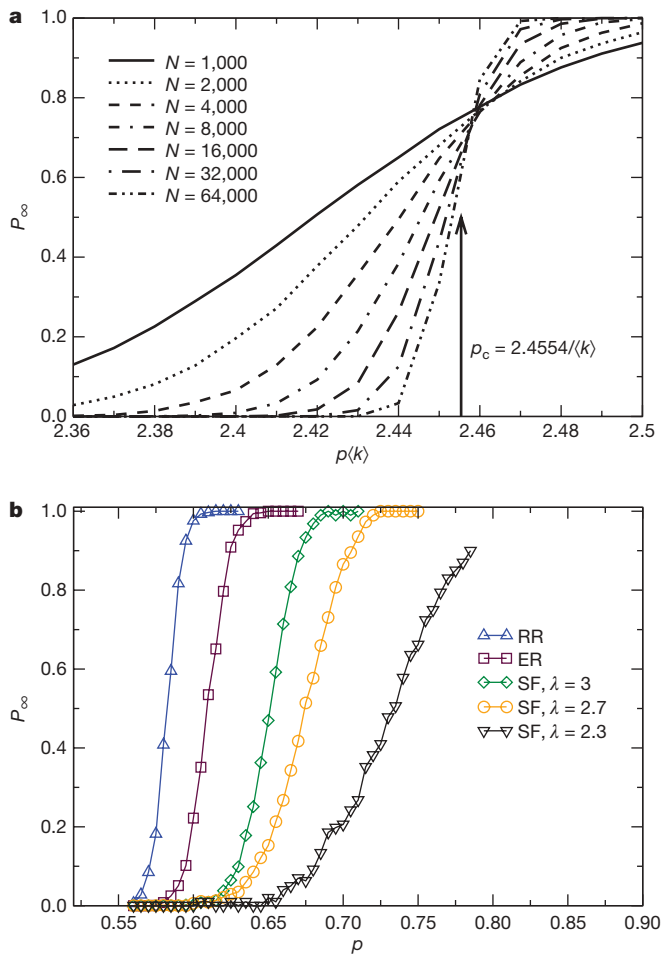
Figure 3 | Numerical validation of theoretical results. a, Numerical simulations of coupled Erdős–Rényi networks with $\langle k \rangle = \langle k_A \rangle = \langle k_B \rangle$ and a finite number of nodes, N. The probability of existence of the giant mutually connected component, P_∞, is shown as function of p for different values of N. As $N \to \infty$, the curves converge to a step function. The theoretical prediction of p_c is shown by the arrow. **b,** Simulation results for P_∞ as a function of p for coupled scale-free (SF) networks with $\lambda = 3, 2.7, 2.3$, coupled Erdős–Rényi (ER) networks and coupled random regular (RR) networks, with an average degree of $\langle k \rangle = 4$ and $N = 50,000$. The simulation results agree with our analytical results. We note that the broader the distribution, the higher the value of p_c.

advantage of a broad distribution for a single network becomes a disadvantage for interdependent networks. In Fig. 3b, we demonstrate this behaviour by comparing simulation results for several scale-free networks with different λ values, an Erdős–Rényi network and a random regular network, all with an average degree of $\langle k \rangle = 4$. The simulation results are in full agreement with our analytical results and show that p_c is indeed higher for a broader distribution.

Next we analytically solve our model of interconnected networks using the mathematical technique of generating functions. We will define generating functions for network A; similar equations describe network B. As in refs 24, 25, we will introduce generating functions of the degree distributions, $G_{A0}(z) = \sum_k P_A(k)z^k$. Analogously, we will introduce generating functions of the underlying branching processes, $G_{A1}(z) = G'_{A0}(z)/G'_{A0}(1)$.

Random removal of fraction, $1 - p$, of nodes will change the degree distribution of the remaining nodes, so the generating function of the new distribution is equal to the generating function of the original distribution with argument $1 - p(1 - z)$ (ref. 24). We denote the subsets of nodes remaining after the random removal of $1 - p$ nodes as $A_0 \subset A$ and $B_0 \subset B$, and note that there is one-to-one correspondence between nodes in A_0 and nodes in B_0, established by

$A \leftrightarrow B$ links. As the total number of nodes in network A is N, the number of nodes in A_0 and B_0 is $N_0 = pN$. The fraction of nodes that belong to the giant component of network A_0 is $g_A(p) = 1 - G_{A0}[1 - p(1 - f_A)]$ (ref. 25), where f_A is a function of p that satisfies the transcendental equation $f_A = G_{A1}[1 - p(1 - f_A)]$. Analogous equations exist for network B.

Using the generating function approach, we find that the fraction, μ_n, of the nodes in the giant component after stage n in the cascade of failures obeys a simple recursion relation (Supplementary Information). We find good agreement between simulations and theory (Supplementary Fig. 1).

To determine the final size of the giant mutually connected component, we recall that the fraction of nodes in the giant mutually connected component, μ_∞, is the limit of the sequence μ_n as $n \to \infty$. This limit must satisfy the equations $\mu_{2m+1} = \mu_{2m} = \mu_{2m-1}$ because the cluster is not further fragmented. This condition leads to the following system of two unknowns, x and y (Supplementary Information), where $\mu_\infty = x g_B(x) = y g_A(y)$:

$$x = g_A(y)p$$
$$y = g_B(x)p \qquad (1)$$

The system of equations (1) has one trivial solution, $x = 0$ and $y = 0$, for any p value, corresponding to the giant mutually connected component being of zero size. If p is large enough, there also exists a solution such that the giant mutually connected component is of non-zero size. We can easily exclude y from these equations and obtain a single equation

$$x = g_A[g_B(x)p]p \qquad (2)$$

which can be solved graphically (Supplementary Fig. 2) as the intersection of the straight line $y = x$ and the curve $y = g_A[g_B(x)p]p$. When p is small enough, the curve increases very slowly and does not intersect the straight line (except at the origin, which intersection corresponds to the trivial solution). A nontrivial solution first emerges in the critical case ($p = p_c$), in which the line touches the curve at a single point, $x = x_c$, where they have equal derivatives. Therefore, we have the condition

$$1 = p^2 \frac{dg_A}{dx}[pg_B(x)] \frac{dg_B}{dx}(x)\Big|_{x=x_c, p=p_c} \qquad (3)$$

which, together with equation (2), yields the solution for p_c and the critical size of the giant mutually connected component, $\mu_\infty(p_c) = x_c g_B(x_c)$.

In the case of two Erdős–Rényi networks[21–23], the problem can be solved explicitly. Then, $G_{A1}(x) = G_{A0} = \exp[\langle k_A \rangle(x - 1)]$, $G_{B1} = G_{B0} = \exp[\langle k_B \rangle(x - 1)]$ and the system of transcendental equations (2) and (3) for the critical value of $p = p_c$ can be expressed in terms of elementary functions (Supplementary Information). In the simple case with $\langle k_A \rangle = \langle k_B \rangle = \langle k \rangle$, the critical parameters can be expressed in terms of the nontrivial root $f = f_A = f_B = 0.28467$ of the equation $f = \exp[(f - 1)/2f]$. We find that $p_c = [2\langle k \rangle f(1 - f)]^{-1} = 2.4554/\langle k \rangle$ and that $\mu_\infty(p_c) = (1 - f)/2\langle k \rangle f = 1.2564/\langle k \rangle$. Our simulations of Erdős–Rényi networks agree with our theory (Fig. 3a).

We also find that the known result for a single scale-free network, namely that $p_c \to 0$ as $N \to \infty$ for $\lambda \leq 3$, is not valid for two scale-free interdependent networks, where instead p_c is finite for any $\lambda > 2$. Analysis of the behaviour of the generating functions as $z \to 1$ shows that as $x \to 0$ the right-hand side of equation (2) can be approximated by a power law, Cx^η (see Supplementary Information for a detailed derivation), where C is constant and

$$\eta = 1/(3 - \lambda_A)(3 - \lambda_B)$$

For $2 < \lambda_A < 3$ and $2 < \lambda_B < 3$, $\eta > 1$. Thus, the curve $y = g_A[g_B(x)p]p$ always passes below $y = x$ as $x \to 0$ and for sufficiently small p values we therefore do not have a non-trivial solution (Supplementary Fig. 2), which means that the giant mutually

3

connected component is absent. Hence, we have a percolation phase transition at some finite $p = p_c > 0$ (Fig. 3b).

The model presented here captures the important phenomenon of a cascade of failures that results in a percolation phase transition. The model can be generalized to the case of three or more interdependent networks, to the case in which the A $\leftrightarrow$ B links connecting the networks are unidirectional rather than bidirectional, and to the case in which a node from network A can depend on more than one node from network B. All these generalizations can be treated analytically by using generating functions, provided the networks are randomly connected and uncorrelated.

1. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
2. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
3. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
4. Pastor-Satorras, R. & Vespignani, A. *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge Univ. Press, 2006).
5. Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford Univ. Press, 2003).
6. Barrat, A., Barthelemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2009).
7. Cohen, R. & Erez, K. ben-Avraham, D. & Havlin, S. Resilience of the Internet to random breakdown. *Phys. Rev. Lett.* **85**, 4626–4628 (2000).
8. Newman, M. E. J., Barabasi, A.-L. & Watts, D. J. (eds) *The Structure and Dynamics of Networks* (Princeton Univ. Press, 2006).
9. Newman, M. E. J. The structure and function of complex networks. *Phys. Rev. E* **64**, 026118 (2001).
10. Goh, K. I., Kahng, B. & Kim, D. Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* **87**, 278701 (2001).
11. Callaway, D. S., Newman, M. E. J., Strogatz, S. H. & Watts, D. J. Network robustness and fragility: percolation on random graphs. *Phys. Rev. Lett.* **85**, 5468–5471 (2000).
12. Song, C. *et al.* Self-similarity of complex networks. *Nature* **433**, 392–395 (2005).
13. Motter, A. E. & Lai, Y. C. Cascade-based attacks on complex networks. *Phys. Rev. E* **66**, 065102(R) (2002).
14. Boccaletti, S. *et al.* Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2004).
15. Peerenboom, J., Fischer, R. & Whitfield, R. Recovering from disruptions of interdependent critical infrastructures, in *Proc. CRIS/DRM/IIIT/NSF Workshop Mitigat. Vulnerab. Crit. Infrastruct. Catastr. Failures* (2001).
16. Rinaldi, S., Peerenboom, J. & Kelly, T. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Contr. Syst. Mag.* **21**, 11–25 (2001).
17. Laprie, J. C., Kanoun, K. & Kaniche, M. Modeling interdependencies between the electricity and information infrastructures. *SAFECOMP-2007* **4680**, 54–67 (2007).
18. Kurant, M. & Thiran, P. Layered complex networks. *Phys. Rev. Lett.* **96**, 138701 (2006).
19. Panzieri, S. & Setola, R. Failures propagation in critical interdependent infrastructures. *Int. J. Model. Ident. Contr.* **3**, 69–78 (2008).
20. Rosato, V. *et al.* Modelling interdependent infrastructures using interacting dynamical models. *Int. J. Crit. Infrastruct.* **4**, 63–79 (2008).
21. Erdős, P. & Rényi, A. On random graphs I. *Publ. Math.* **6**, 290–297 (1959).
22. Erdős, P. & Rényi, A. The evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–61 (1960).
23. Bollobás, B. *Random Graphs* (Academic, 1985).
24. Newman, M. E. J. Spread of epidemic disease on networks. *Phys. Rev. E* **66**, 016128 (2002).
25. Shao, J. *et al.* Fractal boundaries of complex networks. *Europhys. Lett.* **84**, 48004 (2008).

Author Queries

Journal: **Nature**
Paper: **nature08932**
Title: **Catastrophic cascade of failures in interdependent networks**

Query Reference	Query
1	AUTHOR: When you receive the PDF proofs, please check that the display items are as follows (doi:10.1038/nature08932): Figs 2 (black & white), 1, 3 (colour). Please check all figures very carefully as they have been re-labelled, re-sized and adjusted to Nature's style. Please check the title and the first paragraph with care, as they may have been re-written to aid accessibility for non-specialist readers.
2	AUTHOR: please check that there has been no corruption of mathematical symbols.

For Nature office use only:

Layout	☐	Figures/Tables/Boxes	☐	References	☐
DOI	☐	Error bars	☐	Supp info (if applicable)	☐
Title	☐	Colour	☐	Acknowledgements	☐
Authors	☐	Text	☐	Author contribs (if applicable)	☐
Addresses	☐	Methods (if applicable)	☐	COI	☐
First para	☐	Received/Accepted	☐	Correspondence	☐
Display items	☐	AOP (if applicable)	☐	Author corrx	☐