SUPPLEMENTARY INFORMATION

A. Numerical simulation of the cascade process

As an example of the cascade process of $N_A = N_B = 128000$ we simulate ER networks with the same average degree $\langle k\rangle_A = \langle k\rangle_B = \langle k\rangle$. Figure (1 in SI) shows the sequences of the fractions μ_n of nodes in the largest cluster in networks A and B after stage n in the cascade of failures. The value of p placed close to the critical value

$$p_c = \frac{2.5445}{\langle k\rangle}. \tag{1}$$

We see that due to statistical fluctuations of $\langle k\rangle$ and p in finite networks, some simulations converge to $\mu_\infty > 0$, while others struggle for survival near μ_∞ over a long sequence of failures but eventually disintegrate ($\mu_\infty = 0$).

B. The size of finite clusters for coupled networks

In this section, we discuss the surprising result that, unlike clusters in percolation for single networks where the size distribution is a power law with a p-dependent cutoff, the mutual clusters as $N \to \infty$ are either isolated nodes or a single cluster constituting a finite fraction of the entire network. In each *finite* network configuration with N nodes there are many very small mutual clusters and one giant component. To understand this, note that if we have two nodes that are connected by an A-edge, the probability that they are connected by a B-edge can be approximated for $N \to \infty$ as follows

$$1 - \sum_k P_B(k)\left(1 - \frac{1}{N}\right)^k \approx 1 - \sum_k P_B(k)\left(1 - \frac{k}{N}\right) = \frac{\sum_k P_B(k)k}{N} = \frac{\langle k\rangle_B}{N} \to 0. \tag{2}$$

The probability that three nodes connected by A-edges are also connected by B edges scales as $1/N^2$.

Note that for the mutual percolation model, p_c is much greater than the percolation threshold for a single network. Therefore, at each stage of the process all a_i and b_i mutual clusters are very small except one which constitutes the giant component present at each particular stage. Therefore nodes in these small clusters are not viable and will collapse soon after splitting from the giant component.

C. The cascade of failures for coupled networks

After the initial attack which removes a $(1-p)$-fraction of nodes from both networks, the first-stage failure is caused by the fragmentation of the subset A_0 which consists of $N_0 \equiv pN$ nodes. The giant component A_1 of A_0 will constitute a $g_A(p)$-fraction of A_0. Thus, the number of nodes in A_1 is

$$N_1 = N_0 g_A(p) = p g_A(p) N \equiv \mu_1 N. \tag{3}$$

Since each node in network B depends on exactly one node from network A the same fraction of nodes remains functional in network B. Furthermore, the nodes in network B that depend on the nodes that were disconnected from A_1 in network A are totally random with respect to the connections in network B. Thus we can again apply the apparatus of generating functions and find the fraction $g_B(\mu_1)$ of the giant component B_2 of network B with respect to the subset A_1. The number of nodes in the giant component $\mathsf{B}_2 \subset \mathsf{A}_1$ is

$$N_2 \equiv \mu_2 N = g_B(\mu_1) N_1 = g_B(\mu_1)\mu_1 N = p g_A(p) g_B(\mu_1) N. \tag{4}$$

Thus, the fraction of functioning nodes after the second stage failure is

$$\mu_2 = p g_A(p) g_B(\mu_1) = \mu_1 g_B(\mu_1). \tag{5}$$

Now we will analyze what happens during the third-stage failure which is caused by further fragmentation of the giant component A_1 by removal $N_1 - N_2 = [1 - g_B(\mu_1)]N_1$, nodes which do not belong to B_2. The removal of these nodes from A_1 is equivalent to the removal of the same fraction of nodes from A_0 (because all the nodes that were removed at the stage of the initial attack do not belong to B_2, A_1, and A_0) . Thus, the total number of nodes that must be removed from network A is $[1 - g_B(\mu_1)]N_0$ nodes from A_0 plus the number of the initially attacked nodes $(1-p)N$ which is $[1-pg_B(\mu_1)]N$. Hence the effect of the third-stage failure on network A is equivalent to a random attack in which p is replaced by $\mu_2' = p g_B(\mu_1)$. Accordingly the number of nodes in the giant component $\mathsf{A}_3 \subset \mathsf{B}_2$ is

$$N_3 \equiv \mu_3 N = \mu_2' g_A(\mu_2') N. \tag{6}$$

Following this approach we can construct the sequence of giant components in the cascade of failures: $\mathsf{A} = \mathsf{B} \supset \mathsf{A}_0 = \mathsf{B}_0 \supset \mathsf{A}_1 \supset \mathsf{B}_2 \supset \mathsf{A}_3 \supset \ldots \mathsf{A}_{2m-1} \supset \mathsf{B}_{2m} \supset \mathsf{A}_{2m+1}$. The number of

nodes in each giant component of this sequence is $N > pN \equiv N\mu_0 > N\mu_1 > \ldots N\mu_{2m+1} \ldots$, where the numbers μ_n can be obtained by recursion relations (see Fig. 2) (see Fig. 1 of SI):

$$\begin{aligned}
\mu_0 &\equiv \mu'_0 \equiv p, \\
\mu_1 &\equiv \mu'_1 \equiv \mu'_0 g_A(\mu'_0), \\
\mu_2 &= \mu'_1 g_B(\mu'_1), \\
\mu'_2 &= p g_B(\mu'_1), \\
\mu_3 &= \mu'_2 g_A(\mu'_2), \\
\mu'_3 &= p g_A(\mu'_2) \ldots \mu'_{2m} = p g_B(\mu'_{2m-1}), \\
\mu_{2m} &= \mu'_{2m-1} g_B(\mu'_{2m-1}), \\
\mu'_{2m} &= p g_B(\mu'_{2m-1}), \\
\mu_{2m+1} &= \mu'_{2m} g_A(\mu'_{2m}), \\
\mu'_{2m+1} &= p g_A(\mu'_{2m}).
\end{aligned} \tag{7}$$

To summarize, removal of the $1 - g_A$ fraction of nodes that do not belong to the giant component of the remnants (fraction μ'_{2m}) of network A at the $2m$-stage of the cascade is equivalent to random removal of the same fraction of nodes from the nodes of network B that survived the original attack (fraction p). Accordingly, the $2m+1$-stage is equivalent to finding the giant component of network B after a random attack that leaves only $\mu'_{2m+1} = p g_A(\mu'_{2m})$ fraction of B-nodes intact. This procedure repeats recursively at each stage with networks A and B swapping their roles.

D. Graphical solution for SF networks

To illustrate the graphical solution of Eq. (2) in the main paper, we plot

$$y = p g_A[p g_B(x)] \tag{8}$$

for scale-free networks with $\lambda = \lambda_A = \lambda_B = 3$ (see Fig. 2 in SI). We numerically compute g_A and g_B by summing the power series for the following analytical representation of

$$P_A(k) = P_B(k) = \left(\frac{2}{k}\right)^2 - \left(\frac{2}{k+1}\right)^2 \qquad (k = 2, 3, \ldots). \tag{9}$$

Note that these distributions do not have isolated nodes ($k = 0$) and nodes with one link ($k = 1$). We can see that for $p < p_c = 0.752$ the straight line and the curve cross only at

the origin, while for $p > p_c$ a non-trivial solution emerges. Exactly at p_c the line and the curve touch each other at a single point where the slopes coincide and are equal to 1.0 [see Eq. (3) in the main paper].

E. Analytic solution for ER networks

In the case of ER networks, whose degrees are Poisson-distributed [1, 1, 3], the problem can be solved explicitly. Suppose that the average degree of the network A is $\langle k \rangle_A = a$ and the average degree of the network B is $\langle k \rangle_B = b$. Then,

$$G_{A1}(x) = G_{A0} = \exp[a(x-1)], \tag{10}$$

and

$$G_{B1} = G_{B0} = \exp[b(x-1)]. \tag{11}$$

Accordingly system (1) (in the main paper) becomes

$$\left\{ \begin{array}{l} x = p[1 - f_A] \\ y = p[1 - f_B] \end{array} \right\}, \tag{12}$$

where

$$\left\{ \begin{array}{l} f_A = \exp[ay(f_A - 1)] \\ f_B = \exp[bx(f_B - 1)] \end{array} \right\}. \tag{13}$$

Excluding x and y, we get a system with respect to f_A and f_B:

$$\left\{ \begin{array}{l} f_A = e^{-ap(f_A-1)(f_B-1)} \\ f_B = e^{-bp(f_A-1)(f_B-1)} \end{array} \right\}. \tag{14}$$

Introducing a new variable $r = f_A^{1/a} = f_B^{1/b}$, we reduce system (14) to a single equation

$$r = e^{-p(r^a-1)(r^b-1)}, \tag{15}$$

which can be solved graphically for any p. The critical case corresponds to the tangential condition

$$1 = \frac{d}{dr} e^{-p(r^a-1)(r^b-1)} = p[ar^a + br^b - (a+b)r^{a+b}], \tag{16}$$

at which the critical value of $r = r_c$ satisfies the transcendental equation

$$r = \exp\left[-\frac{(1-r^a)(1-r^b)}{ar^a + br^b - (a+b)r^{a+b}} \right], \tag{17}$$

and the critical value of $p = p_c$ can be found from Eq. (16).

$$p_c = \left[a r_c^a + b r_c^b - (a+b) r_c^{a+b} \right]^{-1}. \tag{18}$$

The values of p_c and $\mu_\infty(p_c) = p_c(1 - r_c^a)(1 - r_c^b)$ for different a and $b > a$ are presented in Fig. 3 as function of a/b. In case $a = b$, $f_A = f_B = f$, and the critical value of $f = f_c$ satisfies equation

$$f_c = \exp\left[\frac{f_c - 1}{2 f_c} \right], \tag{19}$$

which yields a solution $f_c = 0.28467$, $p_c = 2.4554/a$, and the critical fraction of nodes in the mutual giant component

$$\mu_\infty(p_c) = p_c(1 - f_c)^2 = \frac{1.2564}{a}. \tag{20}$$

Numerical simulations of the ER networks are in excellent agreement with the theory (Fig. 3).

F. Analytic solution for scale-free networks

For regular percolation in a scale-free network with a power law degree distribution $P_A(k) \sim k^{-\lambda_A}$, it is known that $p_c \to 0$, as $N \to \infty$ for $\lambda_A \leq 3$. Surprisingly, for mutual percolation this is not the case and p_c remains finite for $\lambda_A > 2$. To see this, we can find analytical approximation for $P_A(p)$. Equation (4) can be simplified by substitution $z_A = 1 - p(1 - f_A)$

$$1 - \frac{1}{p} + \frac{z_A}{p} = G_{A1}(z_A). \tag{21}$$

Then Eq. (3) becomes

$$g_A(p) = 1 - G_{A0}(z_A), \tag{22}$$

According to the Tauberian theorems, for $\lambda_A \leq 3$, $G_{A1}(x)$ has a singularity at $x = 1$ of the sort $1 - \tilde{\kappa_A}(1 - x)^{\lambda_A - 2}$ where $\tilde{\kappa_A}$ is a constant. Therefore for $\lambda_A \leq 3$ the derivative of $G_{A1}(x)$ diverges at $x \to 1$. To solve Eq. (21), we must find the intersection of the straight line $y = 1 - 1/p + z_A/p$ and the curve $y = G_{A1}(z_A)$. The straight line passes through the point $y = 1, z_A = 1$ with the slope $1/p$. Thus there is always a trivial solution $z_A = 1$, which corresponds to the absence of percolation. If, $G'_{A1}(1) = \tilde{\kappa}_A$ is finite, we do not have another solution for $p < 1/\tilde{\kappa}_A$, but for $\lambda_A \leq 3$, we always have a non trivial solution $z_A = 1 - (p\kappa_A)^{1/(3-\lambda_A)}$. Since $G'_{A0}(1) = \langle k_A \rangle$, which is finite for $\lambda_A > 2$, Eq. (22) yields

$$g_A(p) = (p\tilde{\kappa_A})^{1/(3-\lambda_A)} \langle k_A \rangle. \tag{23}$$

Finally Eq. (6) of the main paper becomes

$$x = p\langle k_A\rangle \left[p\tilde{\kappa_A}\langle k_B\rangle(\tilde{\kappa_B}x)^{1/(3-\lambda_B)}\right]^{1/(3-\lambda_A)}. \tag{24}$$

The right-hand side (r.h.s) of this equation behaves as x^η, where

$$\eta = \frac{1}{(3-\lambda_A)(3-\lambda_B)} > 1, \tag{25}$$

for $2 < \lambda_A, \lambda_B < 3$. The r.h.s curve always goes below $y = x$ for $x \to 0$, so for sufficiently small p we do not have a non-trivial solution, which means the absence of the mutual giant component. Thus we have a percolation transition at some finite $p = p_c > 0$ (see Fig. 2 in SI).

G. Finite size effects

Our considerations are rigorous for $N \to \infty$. For a finite network, the relative fluctuations of all fractions decrease as $1/\sqrt{N}$. Thus, for the finite network, there is a range of values of p for which the mutual giant component exists with probability $P_\infty(p, N)$ (Fig. 1 in SI). As $N \to \infty$ its derivative diverges as

$$\left.\frac{dP_\infty(p,N)}{dp}\right|_{p=p_c} \sim \sqrt{N}, \tag{26}$$

and for $N \to \infty$, $P_\infty(p, N)$ becomes a step function $P_\infty(p) = 0$ for $p < p_c$ and $P_\infty(p) = 1$ for $p > p_c$. The inverse square-root scaling with N of the width of the interval p for which we can have a complete fragmentation for some realizations of networks and a giant component for the other realizations of the networks can be justified by the following arguments. The actual fraction of the remaining nodes p_a in a finite network of size N will be normally distributed around a given p with the standard deviation inverse proportional to $1/\sqrt{N}$. Thus $P_\infty(p, N)$ is equal to the probability that $p_a > p_c$, which is equal to the integral of the normal probability density with zero mean and the same standard deviation from $p_c - p$ to infinity. Therefore the derivative $dP_\infty(p, N)/dp$ has a Gaussian shape with standard deviation proportional to $1/\sqrt{N}$.

The average number of stages $\langle n\rangle$ in a cascade of failures for $p > p_c$ diverges proportionally to $\ln N/\sqrt{p - p_c}$. This follows from the properties of the iterative process Eq. (2) (in the main paper). This can be seen from the fact that near $p = p_c$, Eq. (2) (in the main paper)

has two roots produced by the intersection of the curved line which can be approximated by a parabola $y = a(p)x^2 + b(p)x + c(p)$ and a straight line $y = x$ (Fig. 2 in SI). This is equivalent to solving a quadratic equation

$$a(p)x^2 + [b(p) - 1]x + c(p) = 0. \tag{27}$$

The value $p = p_c$ is given by the discriminant of this equation equal to zero: $d(p_c) \equiv (b(p_c) - 1)^2 - 4a(p_c)b(p_c) = 0$. In the general case, all three parameters, $a(p)$, $b(p)$, and $c(p)$, have non-zero derivatives at $p = p_c$. Therefore, in the general case $d(p)$ has also a non-zero derivative at $p = p_c$, and hence the difference between the roots scales as $\sqrt{p - p_c}$. Thus, the derivative of the curve at the largest root, which corresponds to the limit of the iterative process scales as $f' = 1 - \alpha\sqrt{p - p_c}$, where α is some positive constant. For Eq. (2) (in the main paper) the iterations converge to the root as $f'_n = \exp(-\alpha\sqrt{p - p_c}n)$. In a real network, they will stop when the difference between two successive iterations will be smaller than one node, which yields a condition $\exp(-\alpha\sqrt{p - p_c}n) \sim 1/N$. Hence indeed $\langle n\rangle \sim \ln N/\sqrt{p - p_c}$.

For $p < p_c$ the solution does not exist and the curve misses the line with the distance proportional to the negative descriminante. As the curve comes close to the line the steps are proportional to $(x - x_c)^2 + d$, where $d \sim p_c - p$ is the minimal distance between the curve and the line. The number of such steps per dx is $dx/[(x - x_c)^2 + d]$. The total number of steps are thus the integral of this quantity between $x = p$ and $x = 0$, which in the limit $d \to 0$ gives $\langle n\rangle = \pi/\sqrt{d} \sim 1/\sqrt{p_c - p}$.

Exactly at the critical point $p = p_c$ the straight line touches the curve at a single point and the sequence of iterations converges as

$$x_{n+1} - x_c = x_n - x_c - a(x_n - x_c)^2. \tag{28}$$

These iterations converge to x_c as $1/n$ which can be seen by plugging into this equation $x_n - x_c = C/n^\beta + o(n^{-\beta})$ where C and β are some unknown constants. Expanding $(n+1)^{-\beta}$ in Taylor series and equating coefficients for equal powers, one can see that $\beta = 1$. However, in a real network, due to Gaussian spread in μ_a, we are never at criticality, and the typical $\mu_a - p_c \sim 1/\sqrt{N}$. Therefore the distributions of the number of stages in the cascade has an exponential tail $\exp[-\alpha n\sqrt{\mu_a - p_c}]$, in which $(\mu_a - p_c)$ must be replaced by its typical value $1/\sqrt{N}$. Therefore, the distribution of $P(n)$ must have an exponential tail $P(n) \sim$

$\exp[-\alpha' n/N^{1/4}]$, where α' is some positive constant (Fig. 4). Thus at criticality, we expect that $\langle n \rangle \sim N^{1/4}$ as supported by our simulations (Fig. 4).

[1] Erdős, P. & Rényi, A. On random graphs. *I. Publ. Math.* **6**, 290–297 (1959).

[2] Erdős, P. & Rényi, A. The evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17–61 (1960).

[3] Bollobás, B. *Random Graphs* (Academic, London, 1985).

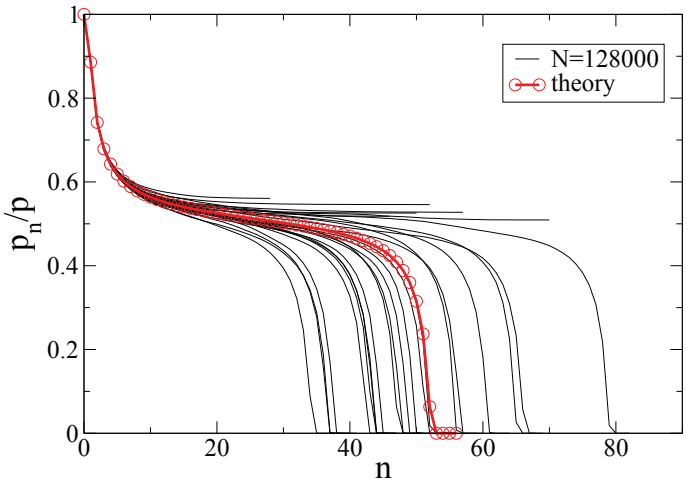

FIG. 1: (Color online) Comparison of the fraction of the giant components after n stages of the cascade failures for several random realizations of ER networks with $\langle k_A \rangle = \langle k_B \rangle$, $N = 128000$ and $\langle k \rangle_A p = 2.45 < \langle k \rangle_A p_c = 2.4554$ and theoretical prediction of Eq. (12 in SI). One can see that for the initial stages the agreement is perfect, however at larger n the deviations due to random fluctuations of the order of $1/\sqrt{N}$ in the actual fraction of the remaining nodes μ_n start to increase. The theoretical prediction after a plateau around the the critical value $2.4554/1.2564 \approx 0.51$ drops to zero, corresponding to the complete fragmentation of the network. The random realizations separate into two classes: one that converge to a mutual giant component and the other that results in a complete fragmentation.

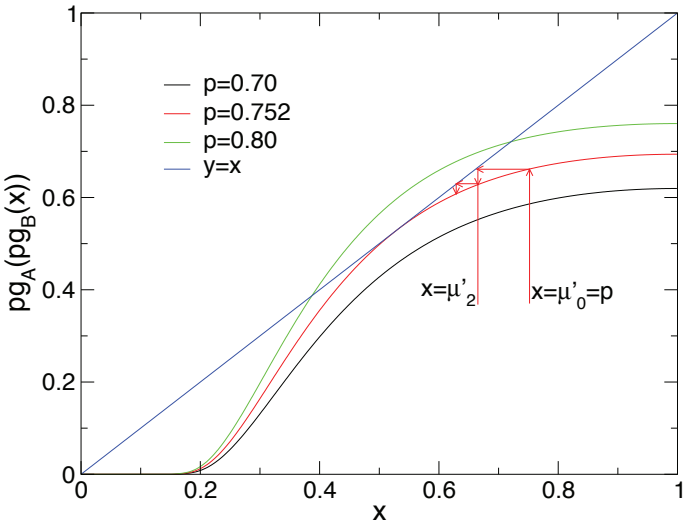

FIG. 2: (Color online) (top) Iterative process described by Eq. (2 in the main paper) for the case of the scale-free distribution $P_A(k) = P_B(k) = (2/k)^2 - [2/(k+1)]^2$ for $k = 2, 3, ...$ For $k \to \infty$, this distribution scales as $k^{-\lambda}$, where $\lambda = \lambda_A = \lambda_B = 3$. Three curves corresponding to $p = 0.70 < p_c$ (black), $p = 0.752 \approx p_c$ (red) and $p = 0.80 > p_c$ (green). One can see that for $p \geq p_c$, the iterations (red arrows) starting from $\mu'_0 = p$, converge to the largest of the two roots of Eq. (6). For $p < p_c$, the iterations converg

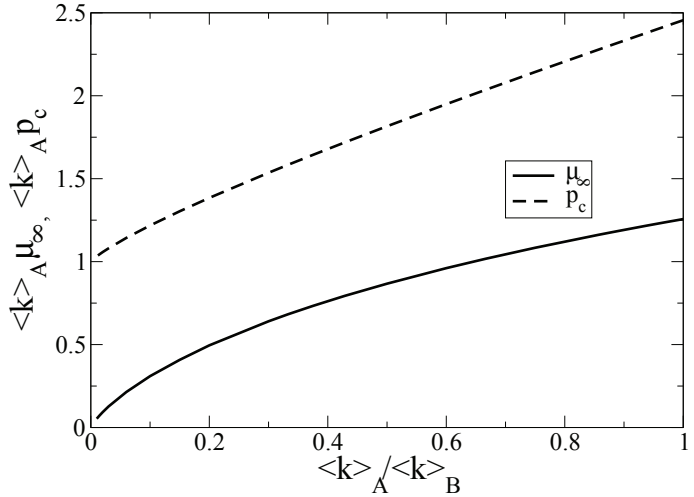

FIG. 3: For ER networks, the critical fraction p_c and the fraction of nodes in the mutual giant component at criticality μ_∞ as function of the ratio $\langle k \rangle_A / \langle k_B \rangle$, where $\langle k \rangle_A = a$ and $\langle k_B \rangle = b$ are the average degrees of networks A and B respectively. The results are obtained using Eq. (11) and (12).

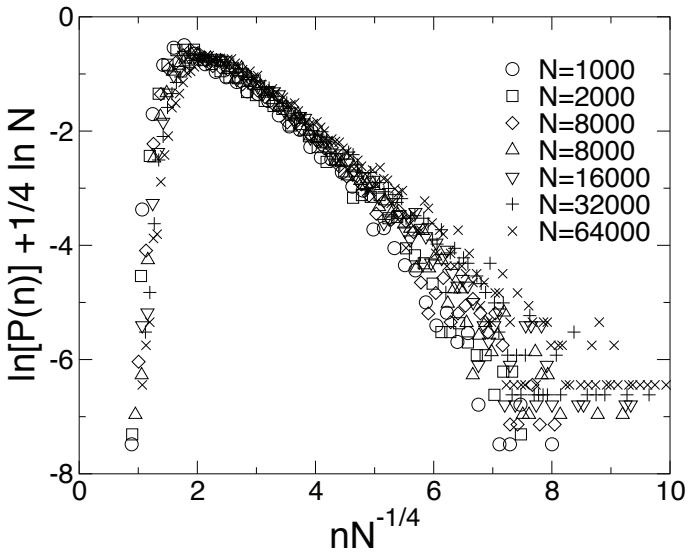

FIG. 4: Scaled distribution of the number of stages in the cascade failures for ER graphs with $a = b$ at criticality ($pa = 2.4554$) for different values of N.