Correction notice
Nature Phys. **10,** 762–767 (2014).

Avoiding catastrophic failure in correlated networks of networks

Saulo D. S. Reis, Yanqing Hu, Andrés Babino, José S. Andrade Jr, Santiago Canals, Mariano Sigman & Hernán A. Makse

In the version of this Supplementary Information file originally published equations (17)–(20) were missing terms. These equations should have read:

$$\mathcal{G}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) = (1-X_{\rm A})^{k_{\rm in}^{\rm A}-1} + (1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} - (1-X_{\rm A})^{k_{\rm in}^{\rm A}-1}(1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} + \delta_{k_{out}^{A},0}[(1-X_A)^{k_{in}^{A}-1} - 1] \quad (17)$$

$$\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) = (1-X_{\rm A})^{k_{\rm in}^{\rm A}} + (1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} - (1-X_{\rm A})^{k_{\rm in}^{\rm A}}(1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} + \delta_{k_{out}^{A},0}[(1-X_A)^{k_{in}^{A}} - 1] \quad (18)$$

$$\mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) = (1-X_{\rm B})^{k_{\rm in}^{\rm B}-1} + (1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} - (1-X_{\rm B})^{k_{\rm in}^{\rm B}-1}(1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} + \delta_{k_{out}^{B},0}[(1-X_B)^{k_{in}^{B}-1} - 1] \quad (19)$$

$$\mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) = (1-X_{\rm B})^{k_{\rm in}^{\rm B}} + (1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} - (1-X_{\rm B})^{k_{\rm in}^{\rm B}}(1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} + \delta_{k_{out}^{B},0}[(1-X_B)^{k_{in}^{B}} - 1] \quad (20)$$

where $\delta_{i,j}$ is the Kronecker delta. These errors have been corrected in this file 27 May 2015.

nature physics
SUPPLEMENTARY INFORMATION
DOI: 10.1038/NPHYS3081

Avoiding catastrophic failure in correlated networks of networks

Reis, Hu, Babino, Andrade, Canals, Sigman, Makse

I. THEORY OF CORRELATED NETWORK OF NETWORKS

We first illustrate the theory to calculate the percolation threshold for a single uncorrelated network following the standard calculations done by Moore and Newman [1]. We then generalize this theory to the case of two correlated interconnected networks to calculate p_c under redundant and conditional modes of failures.

A. Calculation of percolation threshold for a single network [1]

The percolation problem of a single network can be solved by the calculation of the probability X to reach the giant component by following a randomly chosen link [1]. First, choose a link of a single network at random. After that, select one of its ends with equal probability. The probability $1 - X$ is the probability that, by following this link using the chosen direction, we do not arrive at the giant component, but instead we connect to a finite component.

Since the degree distribution of an end node of a chosen link is given by $kP(k)/\langle k\rangle$, one can write down a recursive equation for X as:

$$X = 1 - \sum_k \frac{kP(k)}{\langle k\rangle}(1 - X)^{k-1}. \tag{1}$$

The sum is for the probability that, by following the chosen link, we arrive at a node with degree k which is not attached to the giant component through its remaining $k - 1$ connections. We rewrite the previous equation as follows:

$$X = 1 - \sum_k \frac{kP(k)}{\langle k\rangle}\mathcal{G}(X), \tag{2}$$

where

$$\mathcal{G}(X) = (1 - X)^{k-1}. \tag{3}$$

Once the probability X is known, we can use it to write the probability $1 - S$ that a randomly chosen node does not belong to the giant component. Again, this is a sum of

DOI: 10.1038/NPHYS3081

probabilities: the probability that this node has no links attached to it, plus the probability that this node has one link and this link does not lead to the giant component, plus the probability that this node has two links and none of them leads to the giant component, and so on. In other words:

$$1 - S = \sum_k P(k)(1 - X)^k. \tag{4}$$

Again, we can rewrite this equation as:

$$S = 1 - \sum_k P(k)\mathcal{H}(X), \tag{5}$$

where

$$\mathcal{H}(X) = (1 - X)^k. \tag{6}$$

Note that the probability S not only stands for the probability of choosing one node from the giant component at random, but also provides the fraction of nodes in the network occupied by the giant component. SI-Equation (5) provides the probability of a node to belong to the giant component and is the main quantity to be calculated by the theory from where the value of the percolation threshold can be calculated as the largest value of p_c such that $S(p_c) = 0$.

B. Analytical approach for two interconnected networks with correlations

Now, we present a generalization of the above approach suited to both problems studied in our work, namely, the redundant and conditional interactions of two interconnected networks with generic correlations. We have also developed an analogous theoretical framework based on the generating approach used in Ref. [2]. However, we find that the generating function approach [2] is more mathematically cumbersome if one wants to take into account the correlations between the networks to calculate the mutually connected giant component. Since the size of the giant component is the only quantity needed in this study, we find that the approach of Moore and Newman is more transparent and, furthermore, allows us to take into account both modes of failure in a single theory. Indeed, the whole theory can be cast into a few number of equations, while the generating function approach is more involved.

We define two probabilities for network A (and their equivalents for network B). As we did for the case of a single network, we will take advantage of functions similar to $\mathcal{G}(X)$ and

$\mathcal{H}(X)$. By doing this, the following recursive equations are general and can be applied to the redundant and to the conditional interaction cases depending of the way the functions $\mathcal{G}(\bullet)$ and $\mathcal{H}(\bullet)$ are written for each case. Therefore, below we develop the theory for both modes of failure and later we specialize on each interaction.

First, we define the probability $X_{\rm A}$, as the probability that, by following a randomly chosen link of network A, we reach a node from the largest connected component of network A. The second probability, $Y_{k_{\rm in}^{\rm A}}$, is the probability of choosing at random a node from network A with in-degree $k_{\rm in}^{\rm A}$ connected with a node from the largest component of network B. Analogously, we define probabilities $X_{\rm B}$ and $Y_{k_{\rm in}^{\rm B}}$ for network B.

Thus, if we initially remove a fraction $1-p_{_A}$ of nodes from network A chosen at random, and a fraction $1-p_{_B}$ of nodes from network B, we can write $X_{\rm A}$ and $X_{\rm B}$ in analogy with SI-Eq. (2) [we note that when network A and network B have the same number of nodes, $p=(p_{_A}+p_{_B})/2$]:

$$X_{\rm A}=p_{_A}\left[1-\sum_{k_{\rm in}^{\rm A},k_{\rm out}^{\rm A}}\frac{k_{\rm in}^{\rm A}P\left(k_{\rm in}^{\rm A},k_{\rm out}^{\rm A}\right)}{\langle k_{\rm in}^{\rm A}\rangle}\mathcal{G}(X_{\rm A},Y_{k_{\rm in}^{\rm A}},k_{\rm in}^{\rm A},k_{\rm out}^{\rm A})\right]. \tag{7}$$

Here, the correlations between $k_{\rm in}^{\rm A}$ and $k_{\rm out}^{\rm A}$ from Eq. (1) in the main text are quantified by $P(k_{\rm in}^{\rm A},k_{\rm out}^{\rm A})$, which is the joint probability distribution of in- and out-degrees of nodes from network A from where Eq. (1) in the main text can be derived. The probability function $\mathcal{G}(X_{\rm A},Y_{k_{\rm in}^{\rm A}},k_{\rm in}^{\rm A},k_{\rm out}^{\rm A})$ in SI-Eq. (7) is analogous to SI-Eq. (3). It stands for the probability that, by following a randomly chosen link from network A, we reach a node which is not part of the giant component of network A, which has in-degree $k_{\rm in}^{\rm A}$ and out-degree $k_{\rm out}^{\rm A}$ and/or is not connected with a node from the giant component network B (here and in what follows, "and/or" refers to the nature of the two cases of study: the redundant and conditional interactions, respectively). To write down SI-Eq. (7) we use the joint in- and out-degree distribution of an end node of a randomly chosen in-link $k_{\rm in}^{\rm A}P(k_{\rm in}^{\rm A},k_{\rm out}^{\rm A})/\langle k_{\rm in}^{\rm A}\rangle$. Finally, the terms in the squared brackets stand for the probability $X_{\rm A}=X_{\rm A}(p_{_A}=1)$ before removing the fraction $1-p_{_A}$, which is the generalization of SI-Eq. (2). Thus, after the removal of a fraction $1-p_{_A}$, the probability of following a randomly selected in-link to reach a node which belongs to the giant cluster of A is $X_{\rm A}(p_{_A}=1)$ times the probability $p_{_A}$ for this node being a survival node. In a similar fashion, we write the probability X_B, the joint degree distribution $P(k_{\rm in}^{\rm B},k_{\rm out}^{\rm B})$ and the probability function $\mathcal{G}(X_{\rm B},Y_{k_{\rm in}^{\rm B}},k_{\rm in}^{\rm B},k_{\rm out}^{\rm B})$ for network B:

DOI: 10.1038/NPHYS3081

$$X_{\rm B} = p_B \left[1 - \sum_{k_{\rm in}^B, k_{\rm out}^{\rm B}} \frac{k_{\rm in}^{\rm B} P\left(k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}\right)}{\langle k_{\rm in}^{\rm B} \rangle} \mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) \right] . \tag{8}$$

For the probability $Y_{k_{\rm in}^{\rm A}}$ of choosing at random a node from the network A with degree $k_{\rm in}^{\rm A}$ connected through an out-link with a node from the giant component of B, we write down the following expression:

$$Y_{k_{\rm in}^{\rm A}} = p_B \left[1 - \sum_{k_{\rm in}^{\rm B}} P\left(k_{\rm in}^{\rm B} | k_{\rm in}^{\rm A}\right) \left(1 - X_{\rm B}\right)^{k_{\rm in}^{\rm B}} \right] . \tag{9}$$

The term inside the squared brackets is the probability of choosing a node from network B which is not part of the giant component of B and it is connected with a node from network A of in-degree $k_{\rm in}^{\rm A}$. Naturally, $Y_{k_{\rm in}^{\rm A}}$ is this probability times the probability p_B of the B-node being a survival node after the removal of a fraction $1-p_B$ of nodes from network B. To write down this equation, we use the conditional probability $P(k_{\rm in}^{\rm B}|k_{\rm in}^{\rm A})$ of a node from network B with in-degree $k_{\rm in}^{\rm B}$ being connected with a node with in-degree $k_{\rm in}^{\rm A}$ from network A, and the probability that, by following an in-link from B, we do not reach the giant component of B, $(1 - X_{\rm B})$. The conditional probability $P\left(k_{\rm in}^{\rm B}|k_{\rm in}^{\rm A}\right)$ quantify the correlations expressed by Eq. (2) in the main text. Similar equation can be written for $Y_{k_{\rm in}^{\rm B}}$:

$$Y_{k_{\rm in}^{\rm B}} = p_A \left[1 - \sum_{k_{\rm in}^{\rm A}} P\left(k_{\rm in}^{\rm A} | k_{\rm in}^{\rm B}\right) \left(1 - X_{\rm A}\right)^{k_{\rm in}^{\rm A}} \right] . \tag{10}$$

With $X_{\rm A}$, $X_{\rm B}$, $Y_{k_{\rm in}^{\rm A}}$, and $Y_{k_{\rm in}^{\rm B}}$ on hand, it is possible to compute the fraction of survival nodes in the giant component of network A, $S_{\rm A}$, and in network B, $S_{\rm B}$, through the relations analogous to SI-Eq. (5):

$$S_{\rm A} = p_A \left[1 - \sum_{k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}} P(k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) \mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) \right] , \tag{11}$$

and

$$S_{\rm B} = p_A \left[1 - \sum_{k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}} P(k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) \mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) \right] . \tag{12}$$

The probability function $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ generalizes SI-Eq. (6), and stands for the probability of randomly selecting a node from network A with in-degree $k_{\rm in}^{\rm A}$ and out-degree $k_{\rm out}^{\rm A}$, which is not in the giant component of A and/or it is not connected with the giant

component of B (again, and/or refers to redundant and conditional modes of interaction, respectively).

Due to the different meanings that the probability function $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ may assume depending of the mode of interaction, for this general approach the nature of the quantities S_A and S_B differ conceptually from the quantity S presented by SI-Eq. (5) for a single network. See SI-Fig. 1 for more details. For the conditional mode, a node, or a set of nodes from network A, for example, will fail if (i) it loses connection with the largest component of network A, or if (ii) it loses connection with the largest component of network B. Thus $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ is the probability function that describes the probability of picking a node at random from network A that is not part of the largest component of A (due to condition (i) this node will fail) **or** that is not connected to the largest cluster of network B (due to condition (ii) this node will also fail). Thus, S_A (and its counterpart S_B for network B) is the fraction occupied by the largest component of survival node in network A. For a finite size network, $S_A = n_A/N_A$, where n_A is the number of nodes in the largest component and N_A the number of nodes in network A. It is important to note that due to the condition (ii) this fraction is necessarily the same as the size of the giant connected component of network A. S_A may be interpreted also as the fraction from network A that is part of the mutually connected giant component $S_{\rm AB}$, as in Ref. [2]. The same applies to network B. In other words, the number of nodes in the mutually connected giant component belonging to B is the same as the number of nodes in the giant connected component of B as calculated after the attack as if B was a single network.

For the redundant mode, since there is no cascading propagation of damage due to the failure of a neighbor, $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ is the function that describes the probability of picking a node at random, for example from network A, which is not connected to the largest component from its own network, network A, **and** is not connected to the largest component of network B via an out-going link. Therefore, the quantity S_A provides the fraction of "active" nodes, or in other words, the fraction of survival nodes that may be part of the largest component of network A, and in addition a fraction from network A that are disconnected from that largest component of network A, but are not failed because they are still connected to the largest component of network B via an out-going link. Thus, the mutually connected giant component $S_{\rm AB}$ has a different structure in this mode compared to the conditional mode. This situation is illustrated in Fig. 1c in the main text and SI-

DOI: 10.1038/NPHYS3081

Fig. 1. At the end of the attack process, there is a remaining node in network A which is not connected to the giant component of A calculated as if it is a single network. Such a node is still "on" since it is connected to B via an out-going link. Thus, the mutually connected giant component contains this node.

Furthermore, a node that has lost all its out-going link will fail in the conditional interaction, even if it is still connected to its own giant component. However, in the redundant mode, a node without out-going links may still function as long as it is still connected to the giant component of its own single network. For instance, many nodes are still functioning in Fig. 1c main text, redundant mode, even though they are not interconnected. However, in conditional interaction Fig. 1b main text, all stable nodes needs to have out-going links. That is, in redundant mode, the nodes can still receive power via the same network or the other network, while in the conditional node, they need out-going connectivity all the time. Taking into account these considerations, the value of p_c is obtained from the behavior of the giant component of either of the networks in the conditional mode, while in the redundant mode, the value of p_c is obtained from the size of the mutually giant connected component. However, in this last case, it is statistically the same to obtain p_c from the giant components of one of the networks as well. In what follows the calculations of the giant components are done by considering two networks of equal size N and damaging each network with a fraction $1-p$ of nodes.

Next, we explicitly write the probability functions $\mathcal{G}$ and $\mathcal{H}$ for both, conditional and redundant interactions, respectively, to occur on interactive networks after a random failure of $1-p_{_A}$ and $1-p_{_B}$ nodes. It is important to note that the probabilities $\mathcal{G}$ and $\mathcal{H}$ describe the probability of randomly choosing a node which is not part of the giant component of one network and/or is not connected to a node from the giant component of the adjacent network. In other words, this node picked at random is not part of the giant component of the whole network. We test the general case where both networks are attacked: $p_{_A} \neq 1$ and $p_{_B} \neq 1$. The theory can be used to attacking only one network by setting $p_{_B} = 1$.

Redundant interaction: We consider the total fraction $1-p$ of nodes removed from the two networks. If network A and network B have the same number of nodes, then $p = (p_{_A} + p_{_B})/2$. For redundant interaction two events are important. Both events are defined as follows. The first is the probability that, by following a randomly chosen link of a network, we do not reach the giant component of that network. For network A, this

probability can be written as $(1-X_{\rm A})$. The second is the probability of choosing at random a node from one network, say network A, with in-degree $k_{\rm in}^{\rm A}$ which is not connected with a node from the giant component of network B. This probability can be written as $(1-Y_{k_{\rm in}^{\rm A}})$. In the case of redundant interaction (with no cascading due to conditional mode) these two probabilities are independent, since the lack of connectivity with network B does not imply failure of a node from network A. Thus, the probability function $\mathcal{G}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ that, by following a randomly selected link we arrive at a node with in-degree $k_{\rm in}^{\rm A}$ and out-degree $k_{\rm out}^{\rm A}$ which is not part of the giant cluster of its own network and is not connected with a node from the giant cluster of the adjacent network can be written as:

$$\mathcal{G}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) = (1-X_{\rm A})^{k_{\rm in}^{\rm A}-1}(1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}}. \tag{13}$$

Similarly, the probability function $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ of picking a node, at random, with in-degree $k_{\rm in}^{\rm A}$ and out-degree $k_{\rm out}^{\rm A}$ from one network which is not part of the giant cluster of its own network and is not connected with a node from the giant cluster from the adjacent network is:

$$\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) = (1-X_{\rm A})^{k_{\rm in}^{\rm A}}(1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}}. \tag{14}$$

Again, we can write equivalent expressions for $\mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B})$ and $\mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B})$ as

$$\mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) = (1-X_{\rm B})^{k_{\rm in}^{\rm B}-1}(1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}}, \tag{15}$$

and

$$\mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) = (1-X_{\rm B})^{k_{\rm in}^{\rm B}}(1-Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}}. \tag{16}$$

Conditional interaction: This interaction leads to cascading processes. In the conditional interaction process, we are interested in the cascading effects on the coupled networks, A and B, due to an initial random failure of a portion of nodes in both networks, where $p_A \neq 1$ and $p_B \neq 1$. In the case of attacking network A only, the fraction p_B is set to be equal to one, such that a node from network B can only fail due to the conditional interaction.

For the conditional interaction, $\mathcal{G}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ depends on the probability that, by following a link from network A, we do not arrive at a node with in-degree $k_{\rm in}$ connected to the giant component of its own network, $(1-X_{\rm A})^{k_{\rm in}-1}$, and on the probability of randomly choosing a node from network A with $k_{\rm out}$ outgoing links towards network B, $(1-Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}}$. Also, we have the probability $\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ of picking up a node from one network

DOI: 10.1038/NPHYS3081

which is not part of the giant component of its own network or picking up one node from one network which is not connected with one node from the giant component of the adjacent network, which is also dependent of the probabilities $(1 - X_{\rm A})$ and $(1 - Y_{k_{\rm in}^{\rm A}})$.

Different from the redundant mode, these probabilities, $(1 - X_{\rm A})$ and $(1 - Y_{k_{\rm in}^{\rm A}})$, are not mutually exclusive in the conditional interaction. Thus:

$$\begin{aligned}\mathcal{G}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) &= (1 - X_{\rm A})^{k_{\rm in}^{\rm A}-1} + (1 - Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} - (1 - X_{\rm A})^{k_{\rm in}^{\rm A}-1}(1 - Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} \\ &\quad +\delta_{k_{\rm out}^{\rm A},0}[(1 - X_{\rm A})^{k_{\rm in}^{\rm A}-1} - 1],\end{aligned} \tag{17}$$

and

$$\begin{aligned}\mathcal{H}(X_{\rm A}, Y_{k_{\rm in}^{\rm A}}, k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A}) &= (1 - X_{\rm A})^{k_{\rm in}^{\rm A}} + (1 - Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} - (1 - X_{\rm A})^{k_{\rm in}^{\rm A}}(1 - Y_{k_{\rm in}^{\rm A}})^{k_{\rm out}^{\rm A}} \\ &\quad +\delta_{k_{\rm out}^{\rm A},0}[(1 - X_{\rm A})^{k_{\rm in}^{\rm A}} - 1],\end{aligned} \tag{18}$$

where $\delta_{i,j}$ is the Kronecker delta.

We can write the equivalent expressions for $\mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B})$ and $\mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B})$ as follows:

$$\begin{aligned}\mathcal{G}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) &= (1 - X_{\rm B})^{k_{\rm in}^{\rm B}-1} + (1 - Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} - (1 - X_{\rm B})^{k_{\rm in}^{\rm B}-1}(1 - Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} \\ &\quad +\delta_{k_{\rm out}^{\rm B},0}[(1 - X_{\rm B})^{k_{\rm in}^{\rm B}-1} - 1],\end{aligned} \tag{19}$$

and

$$\begin{aligned}\mathcal{H}(X_{\rm B}, Y_{k_{\rm in}^{\rm B}}, k_{\rm in}^{\rm B}, k_{\rm out}^{\rm B}) &= (1 - X_{\rm B})^{k_{\rm in}^{\rm B}} + (1 - Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} - (1 - X_{\rm B})^{k_{\rm in}^{\rm B}}(1 - Y_{k_{\rm in}^{\rm B}})^{k_{\rm out}^{\rm B}} \\ &\quad +\delta_{k_{\rm out}^{\rm B},0}[(1 - X_{\rm B})^{k_{\rm in}^{\rm B}} - 1].\end{aligned} \tag{20}$$

With the set of SI-Eqs. (13)-(14) and SI-Eqs. (18)-(19), and their equivalents for network B, SI-Eqs. (15)-(16) and SI-Eqs. (20)-(21), it is possible to solve both problems, the redundant and the conditional interactions, on a system of two coupled networks interconnected through degree-degree correlated outgoing nodes. The correlation between the coupled networks is represented by the in- out-degree distribution $P(k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ and by the conditional probability $P\left(k_{\rm in}^{\rm B}|k_{\rm in}^{\rm A}\right)$. In the following section, we present the network model used to generate a system of two networks interconnected with correlations described by power law functions with the exponents α and β. These networks are used on the calculations of

the distribution $P(k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ and $P\left(k_{\rm in}^{\rm B}|k_{\rm in}^{\rm A}\right)$ for each pair of (α, β). The final result is the probability for a node to belong to the giant component of network A or B – as given by SI-Eq. (11) and SI-Eq. (12) – as a function of the fraction of removed nodes $1-p$ (with $p_A = p_B = p$) from where the percolation threshold p_c can be evaluated from $S_A(p_c) = 0$ and $S_B(p_c) = 0$ as a function of the three exponents defining the networks: γ, α and β, and the cutoff in the degree distribution $k_{\rm max}$. We use two networks of equal size $N = 1500$ nodes, each.

C. Network model. Test of theory

In order to test the percolation theory using the above formalism, we need to generate a system of interacting networks with the prescribed set of exponents and degree cutoff. The first step of our network model is to generate two networks, A and B, with the same number N of nodes and with the desired in-degree distribution $P(k_{\rm in})$ as defined by γ and the maximum degree $k_{\rm max}$. To do this we use the standard "configuration model" which has been extensively used to generate different network topologies with arbitrary degree distribution [3]. The algorithm of the configuration model basically consists of assigning a randomly chosen degree sequence to the N nodes of the networks in such a way that this sequence is distributed as $P(k_{\rm in}) \sim k_{\rm in}^{-\gamma}$ with $1 \leq k_{\rm in} \leq k_{\rm max}$ and $P(k_{\rm in}) = 0$ for $k_{\rm in} > k_{\rm max}$. After that, we select a pair of nodes at random, both with $k_{\rm in} > 0$, and we connect them.

The next step of the model is to connect networks A and B in such a way that their outgoing nodes have degree-degree correlations that can be described by the parameters α and β as defined in Eqs. (1) and (2) in the main text. In order to do this, we use an algorithm inspired by the configuration model. First, we assign a sequence of out-degrees $k_{\rm out}$ to the nodes of each network. This process is performed independently to each network by adding the same number of outgoing links. Each outgoing link is added individually to nodes chosen at random with a probability that is proportional to $k_{\rm in}^{\alpha}$. Thus, an out-degree sequence is assigned to the nodes in each network in such a way that $k_{\rm out} \sim k_{\rm in}^{\alpha}$ according to Eq. (1) main text. This process results in a set of outgoing stubs attached to every node in network A and B. The next step is to join these stubs in such a way that we satisfy the correlations given by Eq. (2) main text.

The next step is to choose two nodes, one from each network, such that $\langle k_{\rm in}^{\rm nn}\rangle = A \times k_{\rm in}^{\beta}$,

DOI: 10.1038/NPHYS3081

and then, we connect them if they have available outgoing links. Here, we choose the factor A such that $\langle k_{\rm in}^{\rm nn}\rangle = 1$ for $k_{\rm in} = 1$ when $\beta = 1$, and $\langle k_{\rm in}^{\rm nn}\rangle = k_{\rm max}$ for $k_{\rm in} = 1$ and $\beta = -1$. Thus, we write the value of the factor as $A = A(k_{\rm max}, \beta) = k_{\rm max}^{(1-\beta)/2}$.

The algorithm works as follows: we randomly choose one node i from one network. After that, we choose another node j, from the second network, with in-degree $k_{\rm in}^{j}$ with probability that follows a Poisson distribution $P(k_{\rm in}^{j}, \lambda)$, where the mean value $\lambda = \langle k_{\rm in}^{\rm nn}\rangle$. We connect nodes i and j if they are not connected yet.

It should be noted that Eqs. (1) and (2) in the main text may not be self-consistent for all values of α, β. For instance, for very low values of β, e.g., $\beta = -1$, the degree correlations between coupled networks are not always self-consistent with the structural relations between $k_{\rm in}$ and $k_{\rm out}$ described by α. Since β measures the convergence of connections between networks, when β is negative hubs prefer to connect with low-degree nodes. To better understand these features, consider $\beta = -1$, and for nodes with $k_{\rm in} = 1$ and $k_{\rm in} = k_{\rm max}$. With this configuration, nodes with $k_{\rm in} = 1$ are likely to be connected with nodes from the adjacent network with $k_{\rm in} = k_{\rm max}$. When $\alpha = 1$, most of the links are attached to the highly active nodes, notably, nodes with $k_{\rm in} = k_{\rm max}$, and less likely to nodes with $k_{\rm in} = 1$. In this regime, there are not enough low-degree nodes with outgoing links to be connected with the high-degree nodes, thus the desired relation between $k_{\rm in}^{\rm nn}$ versus $k_{\rm in}$ cannot be realized. The other possible situation is when α is negative. In this regime, most of the outgoing links are attached to low-degree nodes, consequently, the few hubs from the network are unlikely to receive an outgoing link, and even when it happens, one hub does not have enough outgoing links to be connected to the stubs of the low-degree nodes. For these reasons we limit our study to $\alpha > -1$ and $\beta > -0.5$ where the relations are found to be self-consistent.

For every initial pair (α, β), we generate a network with the above algorithm and then we recalculate the effective values of (α, β) which are then used to plot the phase diagram $p_c(\alpha, \beta)$ in Figs. 2 and 4 in the main text.

D. Calculation of the giant components and percolation threshold $p_c(\gamma, \alpha, \beta, k_{\rm max})$

With the networks generated in the previous section we are able to compute the functions $P(k_{\rm in}^{\rm A}, k_{\rm out}^{\rm A})$ and $P\left(k_{\rm in}^{\rm B}|k_{\rm in}^{\rm A}\right)$. Then we apply the recursive equations derived previously to calculate the size of the giant components $S_{\rm A}$ and $S_{\rm B}$ from SI-Eqs. (11) and (12). We do

this calculation for different values of p for cases of study and then extract the percolation threshold p_c at which the giant components S_A and S_B vanish in conditional mode.

SI-Figure 2 shows the predictions of the theory in the conditional mode for a network with $\gamma = 2.5$, $\alpha = 0.5$, $\beta = 0.5$ and $k_{\max} = 100$. We plot the relative size of the giant components in A and B, $S_{\rm A}$ and $S_{\rm B}$, as predicted by SI-Eqs. (11) and (12). As one can see in SI-Fig. 2, there is a well-defined critical value at which the A-giant component vanishes which defines the percolation threshold $p_c(\gamma, \alpha, \beta, k_{\max}) = 0.335$ for these particular parameters.

SI-Figure 2 also presents the comparison between theoretical results and direct simulations. We test the theory by attacking randomly the generated correlated networks and calculating numerically the giant components versus the fraction of removed nodes $1 - p$. The results show a good agreement corroborating the theory.

After testing the theory, a full analysis is done spanning a large parameter space by changing the four parameters defining the theory: $(\gamma, \alpha, \beta, k_{\max})$. The results are plotted in the main text Fig. 2 and 4 for the stated values of the parameters. Beyond the calculation of $p_c(\alpha, \beta)$, we also identify regimes of first-order phase transitions in the conditional interaction, found specially when p_c is high, beyond the standard second-order percolation transition; a result that will be expanded in subsequent papers.

II. EXPERIMENTS: ANALYSIS OF INTERCONNECTED BRAIN NETWORKS

Our functional brain networks are based on functional magnetic resonance imaging (fMRI). The fMRI data consists of temporal series, known as the blood oxygen level-dependent (BOLD) signals, from different brain regions. The brain regions are represented by voxels. In this work we use data sets gathered in two different and independent experiments. The first is the NYU public data set from resting state humans participants. The NYU CSC TestRetest resource is available at `http://www.nitrc.org/projects/nyu_trt/`. The second data set was gathered in a dual-task experiment on humans previously produced by our group [4] and recently analyzed in Ref. [5]. The brain networks analyzed here can be found at: `http://lev.ccny.cuny.edu/~hmakse/soft_data.html`. Both datasets were collected in healthy volunteers and using 3.0T MRI systems equipped with echoplanar imaging (EPI). The first study was approved by the institutional review boards of the New York University School of Medicine and New York University. The second study

DOI: 10.1038/NPHYS3081

is part of a larger neuroimaging research program headed by Denis Le Bihan and approved by the Comité Consultatif pour la Protection des Personnes dans la Recherche Biomédicale, Hôpital de Bicêtre (Le Kremlin-Bicêtre, France).

Resting state experiments: A total of 12 right-handed participants were included (8 women and 4 men, mean age 27, ranging from 21 to 49). During the scan, participants were instructed to rest with their eyes open while the word Relax was centrally projected in white, against a black background. A total of 197 brain volumes were acquired. For fMRI a gradient echo (GE) EPI was used with the following parameters: repetition time (TR) = 2.0 s; echo time (TE) = 25 ms; angle = 90°; field of view (FOV) = 192 × 192 mm; matrix = 64 × 64; 39 slices 3 mm thick. For spatial normalization and localization, a high-resolution T1-weighted anatomical image was also acquired using a magnetization prepared gradient echo sequence (MP-RAGE, TR = 2500 ms; TE = 4.35 ms; inversion time (TI) = 900 ms; flip angle = 8°; FOV = 256 mm; 176 slices). Data were processed using both AFNI (version AFNI_2011_12_21_1014, `http://afni.nimh.nih.gov/afni`) and FSL (version 5.0, `www.fmrib.ox.ac.uk`) and the help of the `www.nitrc.org/projects/fcon_1000` batch scripts for preprocessing. The preprocessing consisted on: motion correcting (AFNI) using Fourier interpolation, spatial smoothing (fsl) with gaussian kernel (FWHM=6mm), mean intensity normalization (fsl), FFT band-pass filtering (AFNI) with 0.08Hz and 0.01Hz bounds, linear and quadratic trends removing, transformation into MIN152 space (fsl) with a 12 degrees of freedom affin transformation, (AFNI) and extraction of global, white matter and cerebrospinal fluid nuisance signals.

Dual task experiments: Sixteen participants (7 women and 9 men, mean age, 23, ranging from 20 to 28) were asked to perform two consecutive tasks with the instruction of providing fast and accurate responses to each of them. The first task was a visual task of comparing a given number (target T1) to a fixed reference, and, second, an auditory task of judging the pitch of an auditory tone (target T2) [4]. The two stimuli are presented with a stimulus onset asynchrony (SOA) varying from: 0, 300, 900 and 1200 ms. Subjects had to respond with a key press using right and left hands, whether the number flashed on the screen or the tone were above or below a target number or frequency, respectively. Full details and preliminary statistical analysis of this experiment have been reported elsewhere [4, 5].

Subjects performed a total of 160 trials (40 for each SOA value) with a 12 s inter-trial

interval in five blocks of 384 s with a resting time of $\sim$ 5 min between blocks. In our analysis we use all scans, that is, scans coming from all SOA. Since each of the 16 subjects perform four SOA experiments, we have a total of 64 brain scans. The experiments were performed on a 3T fMRI system (Bruker). Functional images were obtained with a T2*-weighted gradient echoplanar imaging sequence [repetition time (TR) 1.5 s; echo time 40 ms; angle 90; field of view (FOV) 192 $\times$ 256 mm; matrix 64 $\times$ 64]. The whole brain was acquired in 24 slices with a slice thickness of 5 mm. Volumes were realigned using the first volume as reference, corrected for slice acquisition timing differences, normalized to the standard template of the Montreal Neurological Institute (MNI) using a 12 degree affine transformation, and spatially smoothed (FWHM = 6mm). High-resolution images (three-dimensional GE inversion-recovery sequence, TI = 700 mm; FOV = 192 $\times$ 256 $\times$ 256 mm; matrix = 256 $\times$ 128 $\times$ 256; slice thickness = 1 mm) were also acquired. We computed the phase and amplitude of the hemodynamic response of each trial as explained in M. Sigman, A. Jobert, S. Dehaene, Parsing a sequence of brain activations of psychological times using fMRI. *Neuroimage* **35,** 655-668 (2007). We note that the present data contains a standard preprocessing spatial smoothing with gaussian kernel (FWHM=6mm), which was not applied in Ref. [5]. Such smoothing produces smaller percolation thresholds as compared with those obtained in Ref. [5].

Construction of brain networks: In order to build brain networks in both experiments, we follow standard procedures in the literature [5–7]. We first compute the correlations C_{ij} between the BOLD signals of any pair of voxels i and j from the fMRI images. Each element of the resulting matrix has value on the range $-1 \leq C_{ij} \leq 1$. If one considers that each voxel represents a node from the brain network in question, it is possible to assume that the correlations C_{ij} are proportional to the probability of nodes i and j being functionally connected. Therefore, one can define a threshold T, such that if $T < C_{ij}$ the nodes i and j are connected. We begin to add the links from higher values to lower values of T. This growing process can be compared to the bond percolation process. As we lower the value of T, different clusters of connected nodes appear, and as the threshold T approaches a critical value of T_c, multiple components merge forming a giant component.

In random networks, the size of the largest component increases rapidly and continuously through a critical phase transition at T_c, in which a single incipient cluster dominates and spans over the system [8]. Instead, since the connections in brain networks are highly

DOI: 10.1038/NPHYS3081

correlated rather than random, the size of the largest component increases progressively with a series of sharp jumps. These jumps have been previously reported in Ref. [5]. This process reveals the multiplicity of percolation transitions: percolating networks subsequently merge in each discrete transition as T decreases further. We observe this structure in the two datasets investigated in this study: for the human resting sate in Fig. 3a main text and for the human dual task in Fig. 3b main text.

For each dataset we identify the critical value of T, namely T_c, in which the two largest components merge, as one can notice in Fig. 3 in the main text. While the anatomical projection of the largest component varied across experiments, this merging pattern at T_c was clearly observed in each participant of the two experiments analyzed here, two examples are shown in Figs. 3a-b main text. The transition is confirmed by the measurement of the second largest cluster which shows a peak at T_c, see SI-Fig. 3.

For T values larger than T_c the two largest brain clusters are disconnected, forming two independent networks. Each network is internally connected by a set of strong-links, which correspond to $k_{\rm in}$ [5] in the notation of systems of networks. By lowering T to values smaller than T_c, the two networks connect by a set of weak-links, which correspond to $k_{\rm out}$ [5], i.e. the set of links connecting the two networks.

Our analysis of the structural organization of weak links connecting different clusters is performed with $T_0 < T < T_c$. Here, T_0 is chosen in such a way that the average $\langle k_{\rm out}\rangle$ of outgoing degrees of the nodes on the two largest clusters is $\langle k_{\rm out}\rangle = 1$. For lower values of T_0, where $\langle k_{\rm out}\rangle = 2$ and $= 5$, we found no relevant difference with the studied case of $\langle k_{\rm out}\rangle = 1$.

As done in previous network experiments based on the dual task data [5] we create a mask where we keep voxels which were activated in more than 75% of the cases, i.e., in at least 48 instances out of the 64 total cases considered. The obtained number of activated voxels in the whole brain is $N \approx 60,000$, varying slightly for different individuals and stimuli. The 'activated or functional map' exhibits phases consistently falling within the expected response latency for a task-induced activation [4]. As expected for an experiment involving visual and auditory stimuli and bi-manual responses, the responsive regions included bilateral visual occipito-temporal cortices, bilateral auditory cortices, motor, premotor and cerebellar cortices, and a large-scale bilateral parieto-frontal structure. In the present analysis we follow [5] and we do not explore the differences in networks between different SOA

conditions. Rather, we consider them as independent equivalent experiments, generating a total of 64 different scans, one for each condition of temporal gap and subject.

The following emergent clusters are seen in resting state: medial prefrontal cortex, posterior cingulate, and lateral temporoparietal regions, all of them part of the default mode network (DMN) typically seen in resting state data and specifically found in our NYU dataset [9].

A. Computation of parameters γ, α, β, and $k_{\max}$

Once T_c is determined, we are able to compute the degree distribution of the brain networks. For a given brain scan we search for all connected components of strong links with $C_{ij} > T_c$, where T_c is the first jump in the largest connected component as seen in Fig. 3 main text. We then calculate $P(k_{\rm in})$ using all brain networks for a given experiment; the results are plotted in Fig. 3 main text. We consider all nodes with $k_{\rm in} \geq 1$ at T_c from all the connected clusters. As one can see in Fig. 3b main text, for all data sets, we found degree distributions which can be described by power laws $P(k_{\rm in}) \sim k_{\rm in}^{-\gamma}$ with a given cut-off $k_{\max}$. For the resting state , we found $\gamma = 2.85 \pm 0.04$ and $k_{\max} = 133$ while for the dual task we found $\gamma = 2.25 \pm 0.07$, $k_{\max} = 139$ (see Table I in main text). We use a statistical test based on maximum likelihood methods and bootstrap analysis to determine the distribution of degree of the networks. We follow the method of Clauset, Shalizi, Newman, SIAM Review **51**, 661 (2009) of maximum likelihood estimator for discrete variables which was already used in our previous analysis of the dual task data [5].

We fit the degree-distribution assuming a power law within a given interval. For this, we use a generalized power-law form

$$P(k; k_{\min}, k_{\max}) = \frac{k^{-\gamma}}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max})}, \tag{21}$$

where $k_{\min}$ and $k_{\max}$ are the boundaries of the fitting interval and the Hurwitz ζ function is given by $\zeta(\gamma, \alpha) = \sum_i (i + \alpha)^{-\gamma}$. We set $k_{\min} = 1$.

We calculate the slopes in successive intervals by continuously increasing $k_{\max}$. For each one of them we calculate the maximum likelihood estimator through the numerical solution of

$$\gamma = \text{argmax} \left(-\gamma \sum_{i=1}^{M} \ln k_i - M \ln \left[\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max}) \right] \right), \tag{22}$$

DOI: 10.1038/NPHYS3081

where k_i are all the degrees that fall within the fitting interval and M is the total number of nodes with degrees in this interval. The optimum interval was determined through the Kolmogorov-Smirnov test.

For the goodness-of-fit test, we use KS test generating 10,000 synthetic random distributions following the best-fit power law. Analogous analysis is performed to test for a possible exponential distribution to describe the data. We use KS statistics to determine the optimum fitting intervals and also the goodness-of-fit. In all the cases where the power law was accepted we ruled out the possibility of an exponential distribution, see [5].

In order to compute the correlation of $k_{\rm in}$, $k_{\rm out}$ and $k_{\rm in}^{\rm nn}$ we consider the following statistics for the weak links and the degrees of the external nearest neighbors of an outgoing node. This correlation is gathered from the calculation of the average in-degree, $\langle k_{\rm in}^{\rm nn}\rangle$ of the external neighbors of a node with in-degree $k_{\rm in}$. The strong-links are those links added to the network for $T > T_c$. The weak links are those added to the network for values of $T_0 < T < T_c$ until the average out-degree reaches $\langle k_{\rm out}\rangle = 1$. For statistical determination of the scaling properties of weak-links, we consider that they connect two nodes in different networks, or even nodes in the same component. To calculate the statistical scaling properties of weak links, we consider the out-weak-degree $k_{\rm out}$ of a node as the number of all links added for $T_0 < T < T_c$.

Figure 3f main text shows that the scenario for the correlation between $\langle k_{\rm in}^{\rm nn}\rangle$ and $k_{\rm in}$ is consistent with Eq. (2) in the main text. For the resting state experiments (Fig. 3f in the main text) there is a positive correlation between the $k_{\rm in}$ of outgoing nodes placed in different functional networks. For the dual-task human subjects (Fig. 3f main text) the correlation is also positive.

Moreover, when analyzing the relation between $k_{\rm in}$ and $k_{\rm out}$ for the same outgoing nodes, they are described by the correlations presented in Fig. 3e main text using power laws. Figures 3e-f in the main text depict the power-law fits using Ordinary Least Square method within a given interval of degree. We assess the goodness of fitting in each interval via the coefficient of determination R^2. We accept fittings where $R^2 \gtrsim 0.9$. The exponents measured are presented in Table I in the main text.

Figures 4a and b in the main text show the results we found when we apply the theory presented in Section I of this Supplementary Information on two coupled networks of degree exponent $\gamma = 2.85$ and 2.25, respectively with the cut-off given by $k_{\rm max} = 133, 139$,

respectively as given by the values for human resting state and dual task. For $\gamma = 2.25$ and $\gamma = 2.85$, the value associated with the data gathered from humans, the results are similar with those presented on Fig. 2 main text in both theoretical cases, the conditional (left panels) and redundant (right panels) interactions. The main differences between the results for $\gamma = 2.25$ and 2.85 are the values found for p_c, where the values found for $\gamma = 2.25$ are systematically smaller than the values found for $\gamma = 2.85$, going from $p_c \approx 0.1$ to ≈ 0.6 for $\gamma = 2.25$, and from $p_c \approx 0.1$ to ≈ 0.8 for $\gamma = 2.85$. These results can be understood from the knowledge gathered on the percolation of single networks [10]. For lower values of the degree exponent γ the hubs on scale-free networks become more frequent, protecting the network from breaking apart. When comparing the two cases of Fig. 4 main text with the theoretical case of $\gamma = 2.5$ (Fig. 2 main text), one can notice that the broader the distribution (as lower the value of γ), the more robust is the system of coupled networks. There general trends are consistent with the calculations of p_c for unstructured interconnected networks with one-to-one connections done in Ref. [2]. The white circles in Fig. 4 in the main text correspond to the values of α and β measured from real data. As one can see, the experimental values are placed on the region that represents the best compromise between the predictions for optimal stability under conditional and redundant interactions.

It is also interesting to note that the extreme vulnerability predicted in Ref. [2] can be somehow mitigated by decreasing the number of one-to-one interconnections as shown in Ref. [11]. However, in this case, the system of networks may be rendered non-operational due to the lack of interconnections. Indeed, by connecting both networks with one-to-one outgoing links and by making these interconnections at random, there is a high probability that a hub in one network will be connected with a low degree node in the other network. These low degree nodes are highly probable to be chosen in a random attack, thus the hubs become very vulnerable due to the conditional interaction with a low degree node in the other network. This effect leads to the catastrophic cascading behavior found in [2].

Another way to protect a network in the conditional mode is to increase the number of out-going links per nodes, since the failure of a node occurs when all its inter-linked nodes have failed. Thus, by just increasing the number of interlinks from one to many out-going links emanating from a given node, larger resilience is obtained. If these links are distributed at random, then this situation corresponds to $\alpha = \beta = 0$ in our model. However, in this random conditional case, the network may be rendered non-operational due to the random

DOI: 10.1038/NPHYS3081

nature of the interlink connectivity. A functional real network is expected to be operating with correlations and therefore the most efficient structure when there are many correlated links connecting the networks is the one found for the brain networks investigated in the present work. In other words, assuming that a natural system like the brain functions with intrinsic correlations in inter-network connectivity, then the solution found here (large α and $\beta > 0$) seems to be the natural optimal structure for global stability and avoidance of systemic catastrophic cascading effects.

Another problem of interest is the targeted attack of interdependent networks as treated in Ref. [12]. It would be of interest to determine how the present correlations affect the targeted attack to, for instance, the highly connected nodes.

[1] Moore, C. & Newman, M. E. J. Exact solution of site and bond percolation on small-world networks. *Phys. Rev. E* **62**, 7059-7064 (2000).

[2] Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025-1028 (2010).

[3] Dorogovtsev, S. N. *Lectures on Complex Networks* (Oxford Univ. Press, Oxford, 2010).

[4] Sigman, M. & Dehaene, S. Brain mechanisms of serial and parallel processing during dual-task performance. *J. Neurosci.* **28**, 7585-7598 (2008).

[5] Gallos, L. K., Makse, H. A. & Sigman, M. A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks. *Proc. Natl. Acad. Sci. USA* **109**, 2825-2830 (2012).

[6] Eguiluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M. & Apkarian, A. V. Scale-free brain functional networks. *Phys. Rev. Lett.* **94**, 018102 (2005).

[7] Bullmore E. & Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10**, 186-198 (2009).

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

[9] Shehzad, Z., Kelly, A. M. C. & Reiss, P. T. The resting brain: unconstrained yet reliable. *Cereb. Cortex* **10**, 2209-2229 (2009).

[10] Cohen, R., Ben-Avraham, D. & Havlin, S. Percolation critical exponents in scale-free networks. *Phys. Rev. E* **66**, 036113 (2002).

[11] Parshani, R., Buldyrev, S. V. & Havlin, S. Interdependent networks: reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys. Rev. Lett.* **105**, 048701 (2010).

[12] Huang, X., Gao, J., Buldyrev, S. V., Havlin, S. & Stanley, H. E. Robustness of interdependent networks under targeted attack. *Phys. Rev. E* **83**, 065101 (2011).

DOI: 10.1038/NPHYS3081

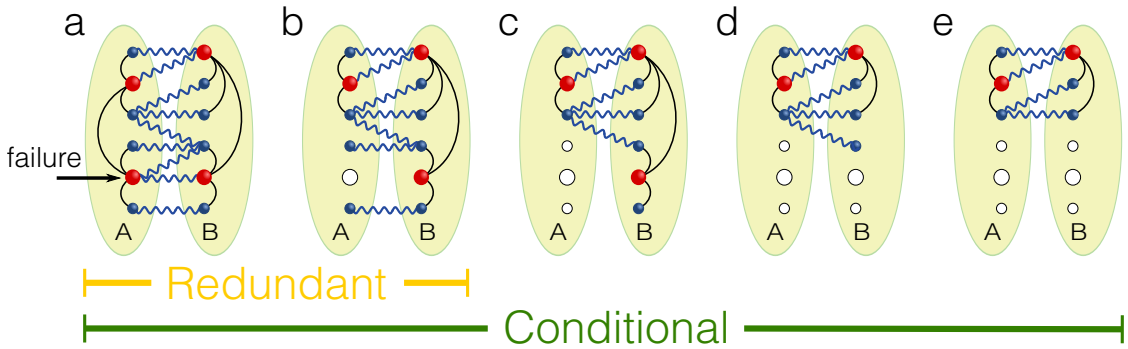

FIG. 1: Pictorial representation of the **a-e** conditional and **a-b** redundant modes of interaction. **a,** One node is removed, or fails, in network A, **b,** as in a regular percolation process this node is removed together with its links. In the redundant mode of interaction, the neighbors of this node are not removed, because they still maintain connection with the giant component from network B, but **c,** for the conditional mode of interaction the two nodes are removed, since they do not belong to the giant component of network A. **d,** As a consequence of the removal of the nodes in network A all the nodes from network B that lose connectivity with network A are also removed. **e,** Finally, the last node from network B is removed once it loses connectivity with the giant component of network B. In the end, for the conditional mode of interaction, only the mutually connected component remains.

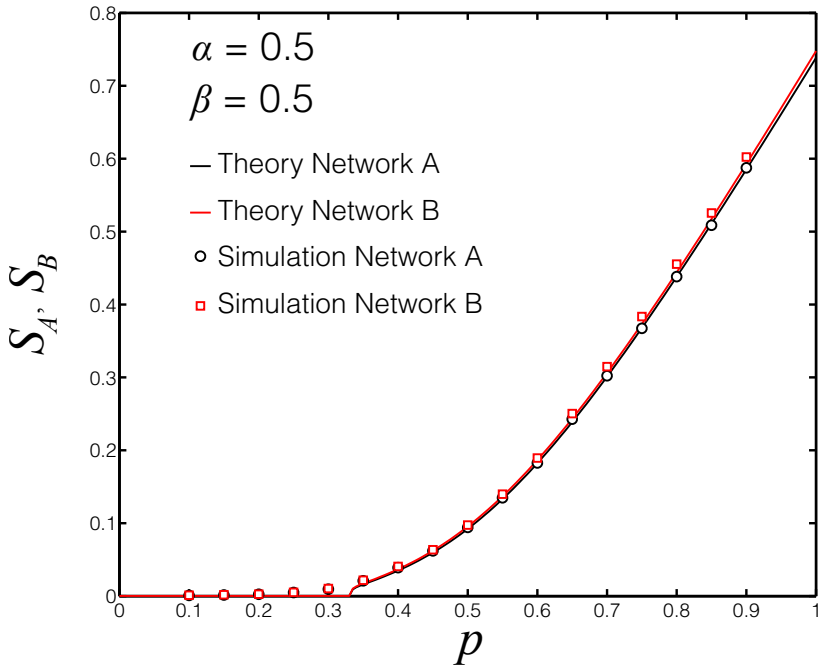

FIG. 2: Giant component of network A and B in the conditional mode of failure. We present the prediction of the theory for values of $N_A = N_B = 1500$, $\gamma = 2.5$, $\alpha = 0.5$, $\beta = 0.5$ and $k_{\max} = 100$ and compare with computer simulations of the giant component obtained numerically by attacking the same network. We perform average over 100 different realizations. We attack a fraction $1 - p$ of both networks and calculate the fraction of nodes belonging to the corresponding giant components. The results show a very good agreement between theory and simulations.

DOI: 10.1038/NPHYS3081

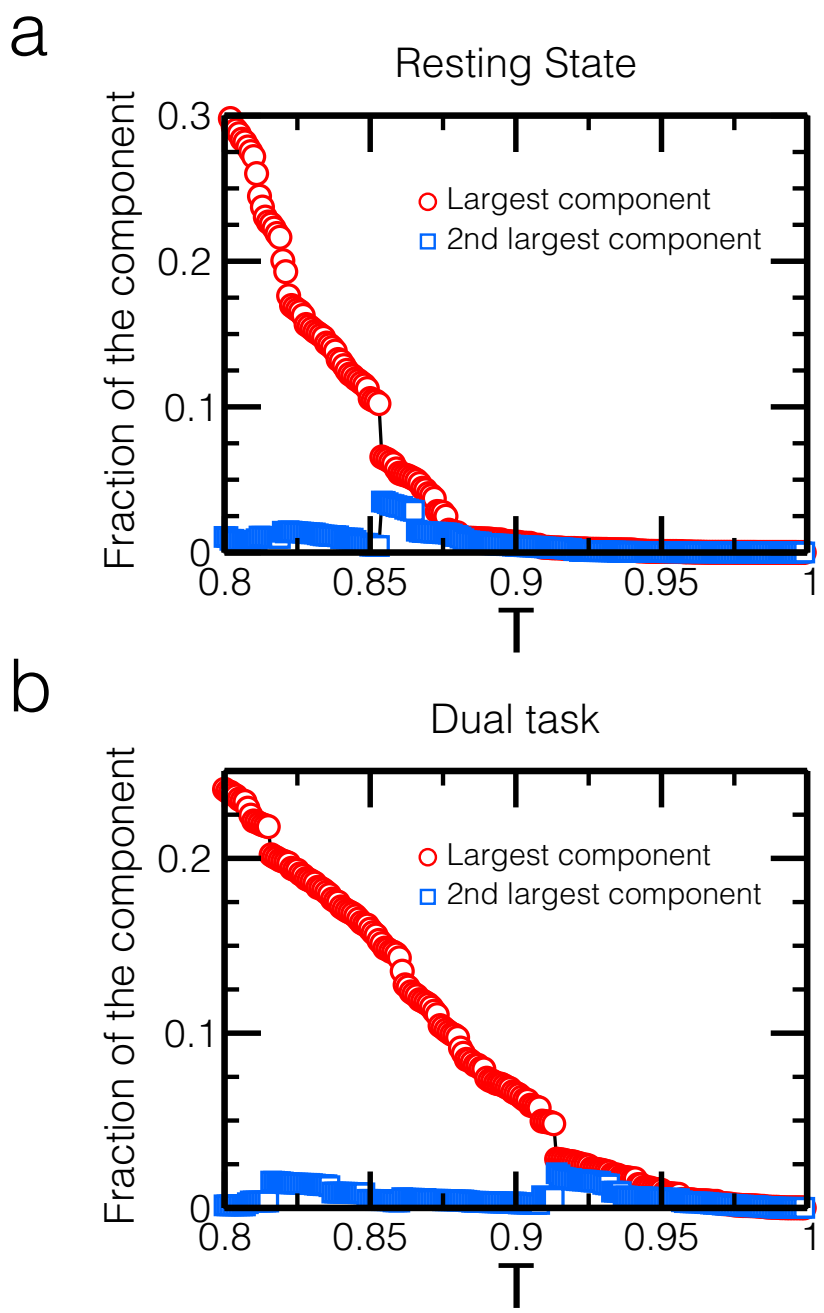

FIG. 3: First and second largest component in the brain networks corresponding to resting state and dual task. The largest component shows a jump while the second largest component shows a peak, indicating a percolation transition at T_c. **a,** Resting state. **b,** Dual task.