

Hidden Connectivity in Networks with Vulnerable Classes of Nodes

Sebastian M. Krause,^{1,2} Michael M. Danziger,³ and Vinko Zlatić¹

¹*Theoretical Physics Division, Rudjer Bošković Institute, Bijenicka c. 54, 10000 Zagreb, Croatia*

²*Faculty of Physics, University of Duisburg-Essen, Lotharstr. 1, 47048 Duisburg, Germany*

³*Department of Physics, Bar-Ilan University, Ramat Gan 5290002, Israel*

(Received 15 April 2016; revised manuscript received 20 July 2016; published 27 October 2016)

In many complex systems representable as networks, nodes can be separated into different classes. Often these classes can be linked to a mutually shared vulnerability. Shared vulnerabilities may be due to a shared eavesdropper or correlated failures. In this paper, we show the impact of shared vulnerabilities on robust connectivity and how the heterogeneity of node classes can be exploited to maintain functionality by utilizing multiple paths. Percolation is the field of statistical physics that is generally used to analyze connectivity in complex networks, but in its existing forms, it cannot treat the heterogeneity of multiple vulnerable classes. To analyze the connectivity under these constraints, we describe each class as a color and develop a “color-avoiding” percolation. We present an analytic theory for random networks and a numerical algorithm for all networks, with which we can determine which nodes are color-avoiding connected and whether the maximal set percolates in the system. We find that the interaction of topology and color distribution implies a rich critical behavior, with critical values and critical exponents depending both on the topology and on the color distribution. Applying our physics-based theory to the Internet, we show how color-avoiding percolation can be used as the basis for new topologically aware secure communication protocols. Beyond applications to cybersecurity, our framework reveals a new layer of hidden structure in a wide range of natural and technological systems.

DOI: [10.1103/PhysRevX.6.041022](https://doi.org/10.1103/PhysRevX.6.041022)

Subject Areas: Complex Systems, Statistical Physics

I. INTRODUCTION

In studies of security and robustness on complex networks, the nodes are typically assumed to be homogeneous with respect to their vulnerabilities to eavesdropping or failure. However, many real-world networks are heterogeneous [1], with large sets of nodes vulnerable to particular adversaries or failures. These vulnerabilities make networks far less secure and less robust than they seem. By developing tools based on statistical physics of networks, we show how this problem can be overcome, obtaining secure communication and optimal redundancy in networks with vulnerable classes of nodes.

A single vulnerability can affect many nodes in a number of different scenarios. For instance, in communication networks, servers under the same jurisdiction are vulnerable to eavesdropping by their controlling entity (company or government). A similar vulnerability structure emerges when communication software is vulnerable to version-specific bugs, like the Heartbleed bug of 2014 [2,3]. In each of these cases, secure communication can be obtained by splitting the message into pieces via secret sharing (so that it cannot be reconstructed without all of its fragments)

[4–6] and transmitting them on different paths, each avoiding a particular vulnerable class of nodes (controlling entities or software versions). Depending on the topology of the network, and the distribution of the vulnerable node classes, this approach may or may not be possible. We illustrate this problem in Fig. 1, where we compare two schematic communication networks: one in which both adversaries can be avoided with multiple paths and one in which one of the adversaries cannot be avoided.

Clearly, a knowledge of network topology is needed in order to obtain secure communication in this way. Unfortunately, network topology has generally not been considered when assessing the security of communication networks. Those examples that do include topology [8–12] do not consider the possibility of multiple simultaneous adversaries, nor do they consider the new connectivity structure that can emerge: Though connected to the network, certain nodes are unable to communicate securely. Percolation is the field of statistical physics that is generally used to analyze connectivity in complex networks, but in its existing forms, it cannot treat the heterogeneity of multiple vulnerable classes. Here, we develop a new type of percolation theory, which provides a general framework for assessing secure connectivity in systems with multiple adversaries or multiple classes of vulnerable nodes.

Beyond the application to secure messaging in networks with multiple adversaries, our framework can discover optimal redundancy in cases where there are correlated

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

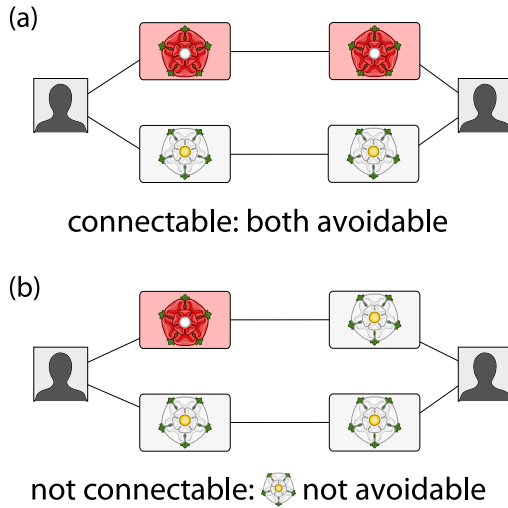


FIG. 1. Avoidable and nonavoidable adversaries. In this scenario, the sender must transmit their message through nodes that are controlled by either Lancaster (red rose) or York (white rose). In diagram (a), a message can be split in two and transmitted on two different paths: The upper path avoids York’s nodes, while the lower path avoids Lancaster’s nodes. In diagram (b), both of the paths pass through York’s nodes, and York cannot be avoided. Lancaster and York were opponents in the Wars of the Roses in England in the 15th century. Artwork after Ref. [7].

failures or other vulnerabilities that affect specific classes of nodes. Though redundant connections via multiple paths are assumed to improve robustness [13–15], if a given vulnerability affects a large set of nodes, the redundant links may not improve the robustness. Indeed, in any case where node failure is more likely to occur in groups instead of individually, effective robustness requires *optimally* redundant paths, each avoiding one (or more) of the vulnerable classes.

Here, we present a new framework for analyzing networks with vulnerable classes of nodes and show the conditions for which—even if every node is vulnerable—robust and secure connectivity can be maintained. We represent the heterogeneity of the network by assigning every node in the network exactly one color, representing a particular vulnerability. The color may represent ownership, geographical location, reliance on a critical material, or some other vulnerability. We then develop a “color-avoiding percolation” theory, which allows us to determine the maximal set of nodes that are mutually connectable under the removal of *any* color—i.e., the nodes that are connectable via a collection of paths that avoids *every* color. The existence of this giant color-avoiding component indicates whether or not secure and robust connectivity can be obtained by avoiding the vulnerable node classes.

II. COLOR-AVOIDING PERCOLATION

On a noncolored network, if node or link failures occur with a given probability, percolation theory can be used to

determine overall connectivity [16,17]. Percolation on complex networks has a rich history [16–20]. It has been used to study the resilience of the Internet [21,22] and its susceptibility to virus spreading [23,24], and even in probabilistic routing algorithms [25]. It has also been used to understand word-of-mouth processes in social networks [26,27], and the robustness of many biological networks including neural networks [28], metabolic networks [29], and mitochondrial networks [30]. Here, we develop a new framework based on percolation theory but not equivalent to any previous percolation problems. In this framework, connectivity corresponds to the ability to avoid vulnerable sets of nodes via multiple paths.

We begin with an undirected unweighted network G with N nodes and adjacency matrix A_{ij} . Every vertex i is assigned a color $c_i \in \{1, 2, \dots, C\}$, where C denotes the total number of colors. In noncolored graphs, a single path provides connectivity; in k -core percolation, any k paths are sufficient [31,32]; and in k -connectivity, k -independent paths are required [33]. Faced with the possible vulnerability or insecurity of all nodes of a single color, we seek a set of paths between two nodes such that *no color* is required for *all paths*. In general, splitting the message into more pieces and sending them along more paths would improve the possibility of secure communication between the source and receiver, provided a secret-sharing algorithm was used [4–6]. However, if the source and target become disconnected when all the nodes of a certain color are removed, then that color cannot be avoided no matter how many paths are utilized. If we knew that only one color c^* was problematic, we could determine the securely connectable nodes by simply removing all nodes of color c^* and checking which components remain, essentially equivalent to standard percolation. Since we do not know which colors are problematic, we need to examine the connectivity upon the removal of every color.

We now define a pair of nodes as color-avoiding connected if, for every color c , there exists a path connecting this pair and *avoiding* all nodes of color c [in Fig. 2(a), nodes R and S are CAC]. The color-avoiding paths are not necessarily unique: Often, one path can avoid multiple colors. We do not include the colors of the source and target nodes in the definition of CAC (see Fig. 2). This assumption captures a number of realistic scenarios: The source and receiver may have special knowledge of the security of their own nodes, the vulnerability may affect the nodes probabilistically so the nodes seek to minimize their exposure, or the vulnerability may affect the nodes as retransmitters but not as end points of the path. After analyzing this assumption in detail, we show how our theory can be extended to other trust models, in particular, where the sender and receiver trust one another’s colors. We return again to this distinction in our discussion of the AS-level Internet.

Formally, we define a “color-avoiding connected component” as a maximal set of nodes, where every node pair in

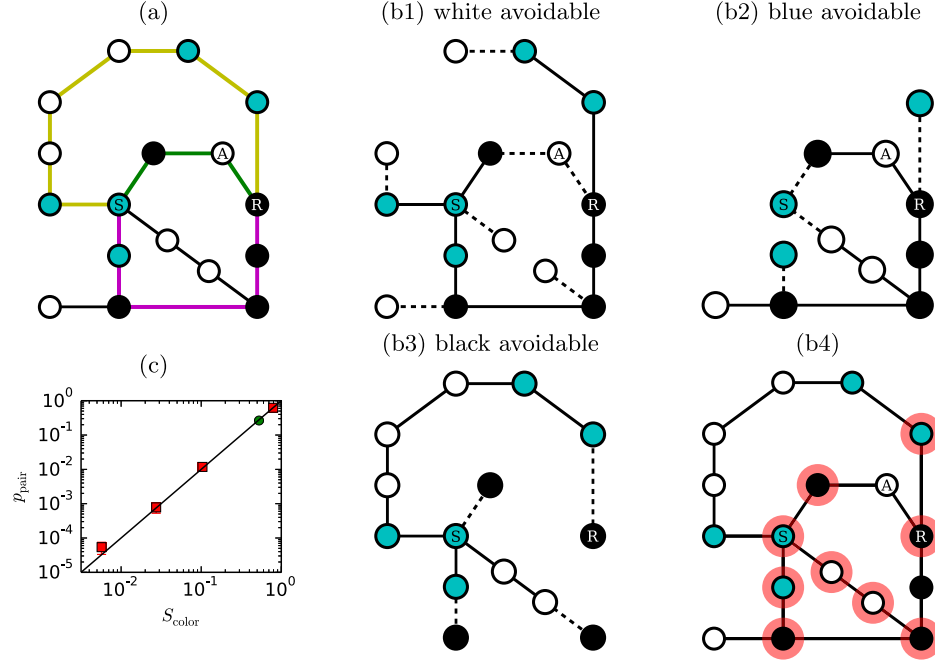


FIG. 2. Illustration of color-avoiding connectivity. (a) In this network, the sender S and the receiver R are color-avoiding connected (CAC), as the green path avoids blue nodes, the purple path avoids white nodes, and the yellow path avoids black nodes. (b) Finding $\mathcal{L}_{\text{color}}$, the largest color-avoiding connected component. (b1) The largest component without white nodes ($\mathcal{L}_{\text{white}}$) and its direct but white neighbors (links to white neighbors are dashed) form $\mathcal{L}_{\text{white}}^+$. Each pair of nodes in this component ($\mathcal{L}_{\text{white}}^+$) can communicate, avoiding white nodes on the path between sender and receiver. We repeat the process in (b2) for blue and in (b3) for black nodes. (b4) Nodes that belong to all three components (in red) constitute the largest CAC component, $\mathcal{L}_{\text{color}}$. Note that some nodes are not CAC but are necessary to form the largest color-avoiding component, for example, node A . This node is not in $\mathcal{L}_{\text{black}}^+$ [see (b3)]. However, it is still required to form a path avoiding blue nodes between nodes S and R [see (a)]. (c) Total CAC pairs p_{pair} compared to the size of the giant CAC component (S_{color}) for quenched graphs. Red squares show Poisson graphs with $N = 10^5$ nodes, average degrees $\bar{k} = 1.6, 1.7, 1.9, 4.0$ (from left to right) and $C = 3$ colors; the green circle shows the AS-level Internet [34,35], with colors representing countries as described in Fig. 4. The black line $S_{\text{color}}^2 = p_{\text{pair}}$ represents all CAC occurring within the giant CAC component. Note that p_{pair} was approximated with samples of up to 5×10^5 pairs; error bars are smaller than the visible range.

the set is color-avoiding connected. Several examples of CAC components are shown in Fig. 2(b4), and in Fig. 6 in the Appendix. In this way, we transform the path-finding problem into a percolation problem, which allows us to study it in a fundamental way, analytically and numerically. Note that there are nodes that are not themselves part of the CAC component but are necessary for the color-avoiding connectivity of nodes that are in the component. This occurs, for example, when all of the neighbors that lead from a node to the CAC component are of the same color, as for node A in Figs. 2(a) and 2(b). In such a case, the node itself is not CAC to the system as a whole because it must pass through nodes of a certain color before it can reach elsewhere. However, in general, this node will still be necessary to form paths that avoid other colors. The fact that non-CAC nodes may be needed to create overall system color-avoiding connectivity is one indication that a new kind of percolation theory is needed to uncover this hidden structure.

By studying the largest CAC component, we obtain a clear quantitative measure of the feasibility of using

multiple paths to avoid vulnerable classes of nodes and information on where those paths should be routed. Furthermore, this gives us a way to measure the effect of changes in network topology, link density, and color distribution.

To find the largest set of color-avoiding connected nodes in any network with any color distribution, we propose the following algorithm. First, for every color c , we remove all nodes with color c and find the largest component in the remaining graph, $\mathcal{L}_{\bar{c}}$. This component represents the largest set of nodes that are connected without requiring any nodes of color c . Then, with $\mathcal{L}_{\bar{c}}$, we define the set of all nodes that have a direct neighbor in $\mathcal{L}_{\bar{c}}$, trivially including $\mathcal{L}_{\bar{c}}$ and additionally nodes of color c that are *directly* connected to $\mathcal{L}_{\bar{c}}$. These “dangling” c -colored nodes represent the nodes that can communicate via $\mathcal{L}_{\bar{c}}$, without requiring any c -colored nodes aside from themselves. In Fig. 2(b1), we see $\mathcal{L}_{\text{white}}^+$, including $\mathcal{L}_{\text{white}}$ and some white “dangling” nodes as node A , which also have access to paths avoiding further white nodes. Every pair of nodes in $\mathcal{L}_{\text{white}}^+$ is connected by a path-avoiding color c . As the second step of the algorithm,

we define $\mathcal{L}_{\text{color}} = \bigcap_c \mathcal{L}_c^+$. This set consists of all the nodes that belong to \mathcal{L}_c^+ for all colors c at the same time. This algorithm generates an implicit multilayer structure, which we explore in Sec. III B and compare with other multilayer network formalisms [36–42]. In Fig. 2(b), we illustrate this method, and further technical details are discussed in Appendix A.

It is conceivable that $\mathcal{L}_{\text{color}}$ does not represent the overall color-avoiding connectivity of the system. For example, the bulk of the CAC nodes might be in smaller components, without a single dominant set. However, if $\mathcal{L}_{\text{color}}$ scales with system size and the size of the smaller color-avoiding connected components does not, then in the limit of large systems, the overall color-avoiding connectivity is determined by $\mathcal{L}_{\text{color}}$, just like the overall connectivity is determined by the size of the giant component in non-colored graphs. With S_{color} defined as the fraction of nodes that are in $\mathcal{L}_{\text{color}}$ among all N nodes, and p_{pair} defined as the total fraction of color-avoiding connected pairs among all node pairs, we can test if $\mathcal{L}_{\text{color}}$ accounts for the bulk of color-avoiding connectivity [$p_{\text{pair}} = 2n_{\text{pair}}/N(N-1)$, where n_{pair} is the number of CAC pairs and N is the size of the network]. In Fig. 2(c), we see, for random networks and a real-world network, that color-avoiding connectivity is indeed dominated by $\mathcal{L}_{\text{color}}$. When S_{color} is small, non-giant clusters and the trivial color-avoiding connectivity that accompanies individual links leads to deviations between p_{pair} and the connectivity predicted with S_{color} , but these deviations rapidly disappear as the system size increases. This validates the treatment of $\mathcal{L}_{\text{color}}$ as a proxy for color-avoiding connectivity.

By transforming the problem of connectivity via multiple color-avoiding paths into a percolation problem, we have provided a method to study the hidden connectivity that emerges in networks with vulnerable classes of nodes. This algorithm can be used to analyze any network with vulnerable classes of nodes, as we show below with the AS-level Internet. We proceed to develop analytical results based on percolation theory for random networks, where we will study the effects of the network topology and color distribution on the color-avoiding connectivity.

III. ANALYTIC THEORY FOR RANDOM NETWORKS

For the analytical treatment, we use the annealed approximation of networks of size N described through the configuration model [17], in which a degree distribution $p(k)$ is a conserved quantity from which an ensemble of network realizations is drawn. For a more comprehensive treatment, see Appendix B.

Every node i is assigned a color $c_i \in \{1, 2, \dots, C\}$. The color sequence $\{c_i\}$ has probability $\prod_i r_{c_i}$. If we assume the colors are distributed uniformly at random, we get $r_{c_i} = r_c = 1/C$ for all c_i .

We calculate S_{color} in the limit of $N \rightarrow \infty$ as the probability that a single node belongs to $\mathcal{L}_{\text{color}}$, $S_{\text{color}} = \text{Prob}(\mathcal{L}_{\text{color}}) = \text{Prob}(\mathcal{L}_1^+ \cap \mathcal{L}_2^+ \cap \dots \cap \mathcal{L}_C^+)$. Because $\mathcal{L}_{\text{color}}$ is a subset of the regular giant component by construction, we begin by obtaining the solution for standard percolation on random graphs [17,43,44]. The size of the giant component in a noncolored random graph is $S = 1 - g_0(u)$, where $g_0(z) = \sum p_k z^k$ is the generating function of the probability distribution p_k . Here, u is the probability that a node is not connected to the giant component over one particular link, and it is computed as the solution of $u = g_1(u)$, where $g_1(z) = g'_0(z)/g'_0(1)$ is the generating function of excess degree [17]. Second, we let κ_c be the expected number of a randomly chosen node's neighbors of color c , which are connected to the giant component of standard percolation. Considering κ_c for all colors, we obtain the vector $\vec{\kappa} = (\kappa_1, \dots, \kappa_C)$, with $k' = \sum_c \kappa_c$ being the total number of links to the normal giant component. Third, the conditional probability $P_{\vec{\kappa}}$ that the links suffice to connect to $\mathcal{L}_{\text{color}}$, given that they belong to distribution $\vec{\kappa}$ and that they already belong to the normal giant component, is

$$P_{\vec{\kappa}} = \prod_{c=1}^C (1 - U_{\bar{c}}^{k' - \kappa_c}), \quad (1)$$

$$U_{\bar{c}} = 1 - \frac{1 - u_{\bar{c}}}{(1 - u)(1 - r_c)}, \quad (2)$$

in which $U_{\bar{c}}$ denotes the conditional probability that a link fails to connect to $\mathcal{L}_{\bar{c}}$ given that it does connect to the normal giant component via a node having a color $c' \neq c$. We define $U_{\bar{c}} = 1$ if $u = 1$. The probability $u_{\bar{c}}$ that a single link does not connect to a giant $\mathcal{L}_{\bar{c}}$ is calculated with $u_{\bar{c}} = r_c + (1 - r_c)g_1(u_{\bar{c}})$ (site percolation with a surviving fraction of nodes of $1 - r_c$ [17]). Notice that $(1 - U_{\bar{c}}^{k' - \kappa_c})$ is the probability that the randomly chosen node is in $\mathcal{L}_{\bar{c}}^+$, given that it has exactly $k' - \kappa_c$ links to the normal giant component over nodes without color c . Combining these terms, we obtain a formula for S_{color} :

$$S_{\text{color}} = \sum_{k=0}^{\infty} p_k \sum_{k'=0}^k B_{k,k'} \sum_{\kappa_1, \dots, \kappa_C=0}^{k'} M_{k', \vec{\kappa}} P_{\vec{\kappa}}, \quad (3)$$

where the binomial factor $B_{k,k'}$ [Appendix B, Eq. (B7)] accounts for the probability that, out of k links, the k' links connect to the normal giant component. The multinomial factor $M_{k', \vec{\kappa}}$ [Appendix B, Eq. (B8)] gives the multinomial probability of having the color distribution $\vec{\kappa}$ among the neighbors belonging to the normal giant component.

To obtain a closed-form solution for S_{color} , we now assume that every color occurs with equal probability:

$r_c = 1/C$. With $U_{\bar{1}} = U_{\bar{c}}$ being identical for all colors, we have [Appendix C, Eq. (C11)]

$$S_{\text{color},C} = \sum_{j=0}^C (-1)^j \binom{C}{j} \times g_0 \left\{ u + (1-u) \left[\frac{j}{C} U_{\bar{1}}^{j-1} + \frac{C-j}{C} U_{\bar{1}}^j \right] \right\}. \quad (4)$$

An alternative derivation of this equation based on combinatorics and set theory is presented in Appendix D. This approach is equivalent to the theory presented here and can be more tractable to some readers, depending on their background.

We now discuss the limiting cases $C = 2$ and $C \rightarrow \infty$. The result for two colors can be simplified to [Appendix C, Eq. (C8)]

$$S_{\text{color},2} = 1 - 2g_0(u_{\bar{1}}) + g_0(2u_{\bar{1}} - 1), \quad (5)$$

which directly depends on $u_{\bar{1}}$ only. As the number of colors tends to infinity, standard percolation is *not* recovered. S_{color} remains smaller than the relative size of the giant component S , and in fact, $S_{\text{color},\infty}$ is identical to the giant component in k -core percolation with $k = 2$ [31,32]. The reason that $S_{\text{color},\infty}$ is equivalent to two-core percolation is that—even if every node is a different color—if a node were connected via only one link, it would not be able to avoid the color of its sole neighbor. We demonstrate this directly by deriving an asymptotic form for S_{color} as $C \rightarrow \infty$ [Appendix C, Eq. (C14)]:

$$S_{\text{color},\infty} = S - (1-u) \frac{dg_0(z)}{dz} \Big|_{z=u}, \quad (6)$$

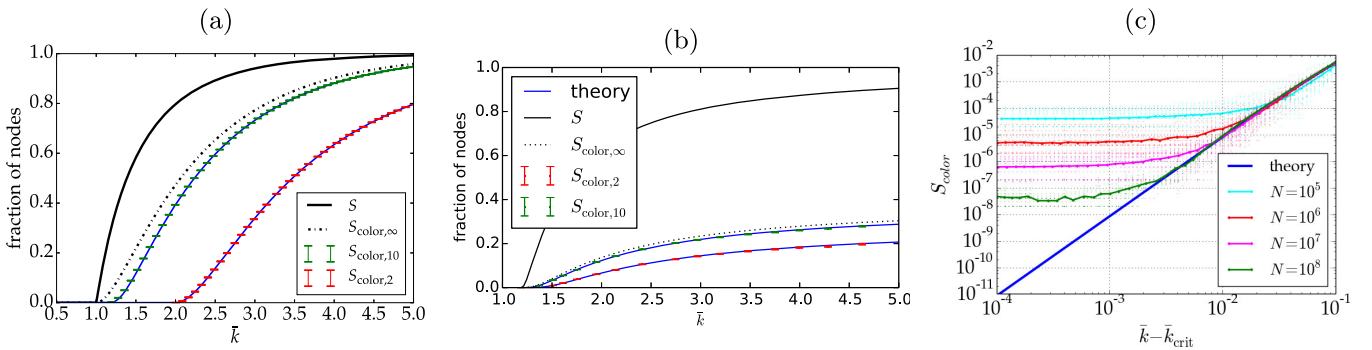


FIG. 3. Size of the giant color-avoiding component S_{color} in random networks with uniformly distributed colors. We show the dependence of S_{color} on average degree \bar{k} (a) for Erdős-Rényi networks and (b) scale-free networks with different numbers of colors. Standard-deviation error bars are shown but barely visible for networks of size $N = 10^6$. The blue lines show the corresponding analytical results. For comparison, we include the giant component size of standard percolation S (black solid line) and the limiting case of a system with an infinite number of colors, $S_{\text{color},\infty}$ (black dashed line). As mentioned in the text, $S_{\text{color},\infty}$ is the same as the giant component in two-core percolation. (c) Critical exponent and finite-size scaling for Erdős-Rényi networks with $C = 3$. Note that in the critical region, the theory and simulations show a slope of almost exactly 3 as predicted by Eq. (8). Finite-size scaling is shown, with results of > 150 realizations per size plotted individually and averaged.

which is the same result as in two-core percolation. In Fig. 3(a), we see that $S_{\text{color},C}$ comes close to $S_{\text{color},\infty}$ even for $C = 10$, indicating that even moderate color diversity comes close to the infinite color case. We can thus use two-core percolation as an envelope for maximal color-avoiding connectivity: If a node is not in the two-core giant component, it will not be in the CAC giant component, regardless of the specific coloring.

We now discuss graphs with broad degree distributions with $p_k \sim k^{-\alpha}$ ($k > 0$) and generating functions $g_0(z) = \text{Li}_\alpha(z)/\zeta(\alpha)$ and $g_1(z) = \text{Li}_{\alpha-1}(z)/[z\zeta(\alpha-1)]$, with $\text{Li}_\alpha(z)$ the polylogarithm function. In Fig. 3(b), we see results for $C = 2$ and $C = 10$ depending on the average degree $\bar{k} = \zeta(\alpha-1)/\zeta(\alpha)$ [17]. The limiting cases are diverging \bar{k} for $\alpha = 2$ and $\bar{k} = 1$ for $\alpha \rightarrow \infty$. We see that \bar{k}_{crit} is not strongly affected by the number of colors but that the size of the giant CAC component is substantially smaller than in the case of Erdős-Rényi networks [see Figs. 3(a) and 3(b)]. The critical connectivity can be calculated using Cohen's criterion for site percolation [21]. With the fraction $1 - r_c$ of nodes surviving random removal, we obtain $1 - r_c = 1 - 1/C = \bar{k}/(\langle k^2 \rangle - \bar{k})$. Since $\langle k^2 \rangle = \zeta(\alpha-2)/\zeta(\alpha)$, we have $\zeta(\alpha-2)/\zeta(\alpha-1) = 1 + C/(C-1)$. Accordingly, $\bar{k} \approx 1.254$ for two colors, and it converges to $\bar{k} \approx 1.195$ for $C \rightarrow \infty$.

We find that Erdős-Rényi (ER) networks are more color-avoiding connected than scale-free networks of equal average degree, the opposite of the results for resilience to random failures [17,22,45]. This follows from the fact that if the average degree is conserved while the degree distribution is widened, there is a larger proportion of very low degree nodes, and we see the same effect in the two-core envelopes; compare Figs. 3(a) and 3(b).

We now turn to the critical behavior of S_{color} in Erdős-Rényi networks with C uniformly distributed colors. Similar to standard percolation, we find that the size of the largest color-avoiding connected component S_{color} undergoes a phase transition at a specific \bar{k}_{crit} , which is now determined by the number of colors (see Fig. 3). For $\bar{k} < \bar{k}_{\text{crit}}$, color-avoiding connectivity is confined to clusters of finite size (with vanishing relative size in the limit of large N), and for $\bar{k} > \bar{k}_{\text{crit}}$, we have a largest color-avoiding connected component of relative size S_{color} , which scales with system size. We find that the value of \bar{k}_{crit} decreases as C increases and approaches the standard percolation threshold as $C \rightarrow \infty$. Since color-avoiding connectivity requires that the giant component not be destroyed after the removal of any single color, we require that $\bar{k}_{\text{crit}}^{\text{ER}} = \bar{k}_{\text{crit}}[(C-1)/C]$, where $\bar{k}_{\text{crit}}^{\text{ER}} = 1$ is the percolation threshold for ER graphs and $[(C-1)/C]$ is the fraction of links remaining after the removal of $1/C$ nodes. Therefore, $\bar{k}_{\text{crit}} = C/(C-1)$.

To discuss the scaling and critical exponents, we return to the definition of $P_{\bar{k}}$, Eq. (1). We consider the region just above \bar{k}_{crit} by defining $\varepsilon \equiv 1 - U_{\bar{1}} \approx C(\bar{k} - \bar{k}_{\text{crit}})$, which holds as long as $(\bar{k} - \bar{k}_{\text{crit}}) \ll 1/C$ [Appendix C, Eq. (C18)]. We analyze the behavior of $P_{\bar{k}}$ for small ε by expanding $(1 - (U_{\bar{1}})^{k' - \kappa_c}) \approx (k' - \kappa_c)\varepsilon$. Plugging this approximation into Eqs. (1) and (3), we obtain

$$S_{\text{color}} \propto (\bar{k} - \bar{k}_{\text{crit}})^\beta \quad (7)$$

$$\beta = C, \quad \bar{k}_{\text{crit}} = C/(C-1). \quad (8)$$

We confirm the value of \bar{k}_{crit} and the scaling of S_{color} numerically in Fig. 3(c) for $C = 3$ colors. As $C \rightarrow \infty$, we need to resolve the seeming contradiction of a divergent critical exponent $\beta = C$ and convergence towards $S_{\text{color},\infty}$ as it appears in Eq. (6). For ER networks, we show [Appendix C, Eq. (C22)] that $S_{\text{color},\infty} \propto (\bar{k} - 1)^2$ for \bar{k} near 1, implying $\beta = 2$. The reason that we do not observe $\beta \rightarrow \infty$ as described in Eq. (8) is that the approximation used to obtain Eq. (7) is only valid in a critical region defined as $(\bar{k} - \bar{k}_{\text{crit}}) \ll 1/C$. As $C \rightarrow \infty$, S_{color} increases with the high exponent $\beta = C$. However, the shrinking critical region overpowers the diverging critical exponent, and $S_{\text{color}} \sim 0$ takes on unobservable small values and crosses over to $\beta = 2$ scaling outside the critical region.

As mentioned above, in color-avoiding percolation, there are transmission nodes that are needed to maintain connectivity of the largest CAC component but are themselves excluded from it. The number of these nodes can be much larger than the number of nodes they are connecting near criticality. We test this performing numeric simulations for Poisson graphs with $C = 3$ colors and $r_c = 1/3$ for all colors. We take networks of size $N = 5 \times 10^6$ and connectivity $\bar{k} = \bar{k}_{\text{crit}} + 0.02 = 1.52$. All nodes in $\cup_c \mathcal{L}_{\bar{c}}$ might be needed for other nodes in order to connect. Averaging

over 50 network realizations, we find, for the relative size of this component, $|\cup_c \mathcal{L}_{\bar{c}}|/N = 5 \times 10^{-2}$, while $S_{\text{color}} = 7 \times 10^{-5}$ is more than 2 orders of magnitude smaller. We note that transmission nodes are not present in other percolation concepts as k -core percolation [31] or percolation in multiplex and interdependent networks [36–42,46], though a similar feature arises in k connectivity [33].

A. Generalization of the theory

While we presented closed-form results only for homogeneous color frequencies [see Eq. (4)], Eq. (3) holds for any set of color frequencies r_c . For heterogeneous color frequencies r_c , Eq. (3) can be evaluated directly. Because of the sum on k , this becomes cumbersome when k is large. For broad degree distributions p_k , the full summation over \bar{k} can be replaced with a sampling method. The critical behavior can be determined for heterogeneous color distributions r_c , as demonstrated above, using an expansion in Eq. (1).

An important generalization of color-avoiding percolation is to allow for trusted colors. In particular, in many scenarios, it is reasonable to trust the colors of the sender and receiver nodes, as they cannot be avoided anyhow. For this, we define \mathcal{A} as the set of colors that have to be avoided. This includes all colors $1, \dots, C$, except for trusted colors. We assume \mathcal{A} not to be empty. For the algorithm, we have $\mathcal{L}_{\text{color}} = \cap_{c \in \mathcal{A}} \mathcal{L}_{\bar{c}}^+$. For analytical results, we replace Eq. (1) with

$$P_{\bar{k}} = \prod_{c \in \mathcal{A}} (1 - U_{\bar{c}}^{k' - \kappa_c}) \quad (9)$$

and use it together with Eqs. (2) and (3). S_{color} can be compared to the standard percolation result, if only one color has to be avoided, and all other colors are trusted, in the limit of $\mathcal{A} = \{c\}$. We have $S_{\text{color}} = 1 - g_0(u_{\bar{c}})$, while the standard percolation result reads $(1 - r_c)[1 - g_0(u_{\bar{c}})]$. S_{color} is larger than standard percolation in this case because nodes of the avoided color c are formally allowed as sender or receiver nodes. However, restricting to nodes of colors other than the avoided color c , standard percolation is recovered. Notice that with trusted colors, results for S_{color} can exceed the two-core. This is because a single link connecting to a node with a trusted color can be enough to avoid all colors in \mathcal{A} .

B. Comparison with other percolation types

Color-avoiding percolation includes two properties of heterogeneous complex networks: the topological connectivity, and metadata of node coloring, which we combine with a multipath concept. Our theory for random graph ensembles treats these two properties in two distinct steps, using conditional probabilities. Probabilities for colors to

be avoidable are conditioned on the overall connectivity (via the giant component), which is treated separately. The comparison with numerical results in Fig. 3 demonstrates that this procedure works: Topology and coloring can indeed be treated separately, and conditional probabilities for avoiding colors are independent even when requiring multiple color-avoiding paths. Furthermore, the interaction of topology and color distribution implies a rich critical behavior, with critical values and critical exponents depending both on the topology and on the color distribution. To our knowledge, the only reported use of colorings in percolation is for polychromatic percolation [47,48], which was developed to model polymerization via simultaneous nonoverlapping largest components in colored lattices. Polychromatic percolation does not consider effects of network topology or multiple paths, and it is designed to answer different questions compared to our study, such as conductivity in heterogeneous lattices.

Multiplex networks are characterized by several layers of links between the same nodes. In this context, there are two main percolation approaches: one that assumes that connections can move from layer to layer via the shared nodes, and one that treats the connectivity of each layer separately and requires the connectedness of the node in all of the layers. The first approach is associated with interconnected networks and is well suited to model multimodal transportation networks, while the second approach is associated with interdependent networks and is better suited to model critical infrastructure. The first approach determines overall connectivity as the union of the giant components in each layer, while the second approach uses the intersection of giant components, with the additional requirement that the reduced set of nodes continue to define a giant component in each layer (this gives rise to the iterative solution and cascade process).

Our percolation process can be understood as a transformation from a single-layer node-colored network with C colors to a C -layered multiplex noncolored network. Each color defines a layer through its removal: After removing all the nodes in the network of color c , we determine the largest connected component and add to it nodes of color c that are directly linked to it. This set \mathcal{L}_c^+ of nodes and links constitutes the layer corresponding to color c . Notice that layers for two different avoidable colors c and c' can have many links in common, which connect nodes of colors other than c or c' . Such overlapping links have strong implications for multiplex connectivity [46,49].

Even after this transformation, the percolation process is distinct from other multilayer percolation processes. After obtaining layers for all colors, we take the intersection of the giant component of all the layers. This collapsed giant component corresponds to the maximal set of nodes that are

color-avoiding connected. This is not the same as the mutual giant connected component of interdependent networks because we do not require that the nodes in the intersection of the giant components remain connected in each layer. Allowing the final set of nodes to not be directly connected to one another in each layer leads to two qualitative differences: There are no cascades because the giant component is not recalculated repeatedly [50,51], and there are transmission nodes that are required for color-avoiding connectivity but are not color-avoiding connected themselves.

This result is not merely a technical difference; it reflects a qualitative difference in the reality that we are modeling: Instead of looking at the set of nodes that remain in the giant components of each layer (as in interdependent networks) or that can be connected via paths that bridge the layers (in interconnected networks), color-avoiding connectivity looks for the pairs of nodes that have a path in every layer, even if the nodes that compose the path do not share this property. To our knowledge, this kind of analysis has never been performed in a single-layer or a multilayer context. The nearest analog is in studies of k components [33], where connectivity is limited to nodes that are connectable via k -independent paths.

IV. APPLICATIONS AND DISCUSSION

One immediate application of our framework is to secure communication in a network where every node is controlled by some entity and thereby subject to eavesdropping. Assuming C node owners, each of which eavesdrops on its routers' traffic, we can securely communicate if messages are split with a secret-sharing protocol [4,5,52] and transmitted along multiple color-avoiding paths [53,54]. The nodes that can take advantage of this method are exactly the elements of the largest CAC component ($\mathcal{L}_{\text{color}}$).

To study the hidden CAC structure of the Internet, we use a symmetrized version of the AS-level Internet prepared by Ref. [34], which was generated using data from the CAIDA project [35] up to December 2013. We then color every router according to the country to which the router is registered, reflecting the assumption that every country is eavesdropping on its traffic but that no countries share information [Fig. 4(a)]. If two countries do share information, we would simply recolor the collaborating node colors to have the same color. For simplicity, we confine our analysis to the case where no countries share information. Using the algorithm for finding the largest CAC component, we can determine which nodes are securely connectable and which are not [Fig. 4(b)].

We find that 26 228 out of 49 743 ($\approx 52.73\%$) of the routers are in the largest CAC component and that this accounts for the vast majority of CAC connected nodes [Fig. 2(c)]. However, we also find that these results vary greatly from country to country. For instance, only 25% of

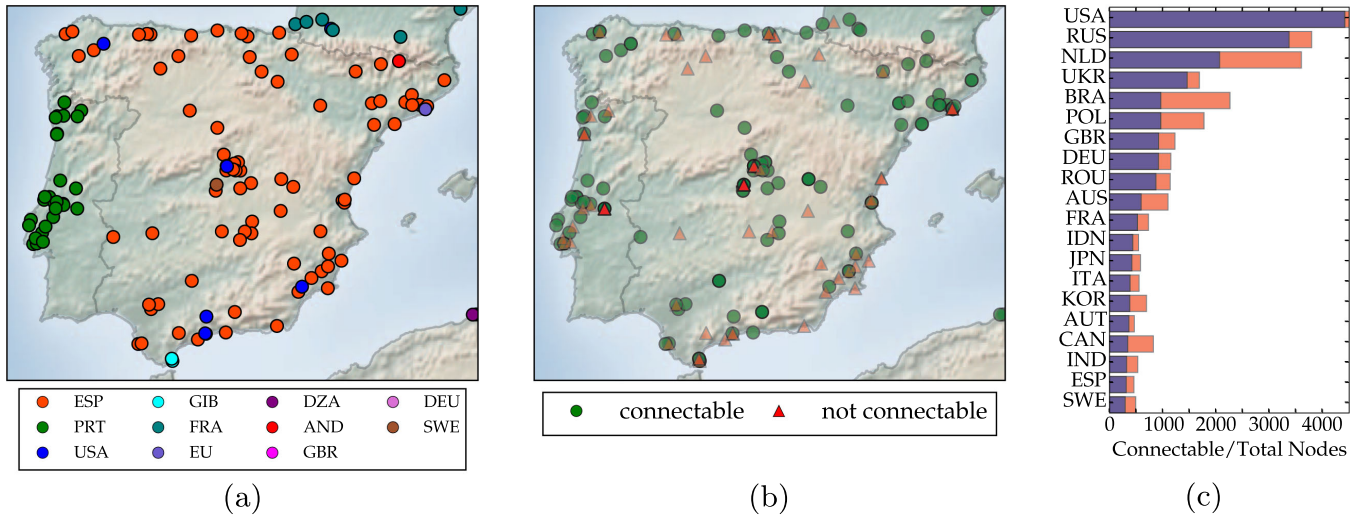


FIG. 4. Color-avoiding connectivity of the AS-level Internet. (a) Here, we show the routers of the AS-level Internet in the Iberian peninsula as a disjointly vulnerable network, with the colors determined by the country to which the router is registered. (b) The green nodes are members of this set, while the red are not. This means that these routers can take advantage of multiple paths to maintain security, as described in the main text. Panel (c) shows the number of CAC routers (nodes in L_{color}) compared to the total number of routers for the top 20 countries worldwide, in terms of total number of AS routers registered to that country in L_{color} . Data for the USA have been truncated for visibility; the total number of AS routers is 17 690. We use a symmetrized version of the network of Ref. [34], which was generated using data from the CAIDA project [35] up to December 2013.

the routers registered to the USA are in the largest CAC component compared to 89% of routers registered to Russia [Fig. 4(c)]. This is partly because of the density of routers in the USA, which is much higher than Russia, and indicates that US eavesdroppers have far greater capacity to intercept communication than their Russian counterparts.

Since the analytic calculations we presented in Sec. III assumed that there was no correlation between the topology and the node coloring, it is of interest to understand how the deviations from that assumption affect color-avoiding connectivity in a real-world network. Details for the countries with the most nodes, together with a comparison with theoretical results, are presented in Fig. 5(a). For theoretical results, the degree distribution was used as measured from the network, and the nodes were given uniform color frequencies, where the frequencies r_c are determined from the shares of AS of each color. We find three main reasons for the breakdown of color connectivity in the real-world network compared to the equivalent random system: (i) Their own country is not avoidable (USA, Brazil, Poland, Sweden), (ii) USA is not avoidable (Australia, Canada, Korea, India), or (iii) the nodes are under-represented in the two-core largest component (Netherlands).

We further consider the possibility that a pair of nodes trust their own colors in Figs. 5(b) and 5(c). We find that for type (i) countries, trusting their own color substantially improves their connectivity, while for type (ii) countries, it only helps when communicating with the USA; for

type (iii) countries, it does not make a substantial difference at all.

A similar analysis can be made of technological communication networks where the nodes are colored by version, or in human communication networks like spy networks with agents operating under different flags.

If classes of nodes are vulnerable to failure (as opposed to eavesdropping), avoiding them with multiple paths as determined by color-avoiding percolation leads to optimal redundancy: Each backup path is optimized for the failure affecting a different node class. When percolation is used to describe spreading phenomena or robustness in networks, it is commonly assumed that the nodes fail at random or preferentially according to a topological property like degree or betweenness [21,22]. However, in many systems, certain sets of nodes are liable to fail all at once: pipe networks with above-ground water pipes freezing over all at once, smart grids where multiple power stations depend on one communication station [54,55], logistical networks where one contractor is responsible for several connections, metabolic networks with different metabolic pathways [56,57] depending on particular biochemicals, and any other case where a specific set of nodes is liable to fail at once [58,59]. Color-avoiding percolation provides a general and extensible tool for analyzing effective robustness and optimal redundancy in any of these realistic scenarios.

In economics and operations research, it has been argued that global market competition today is essentially between supply chains instead of between individual companies

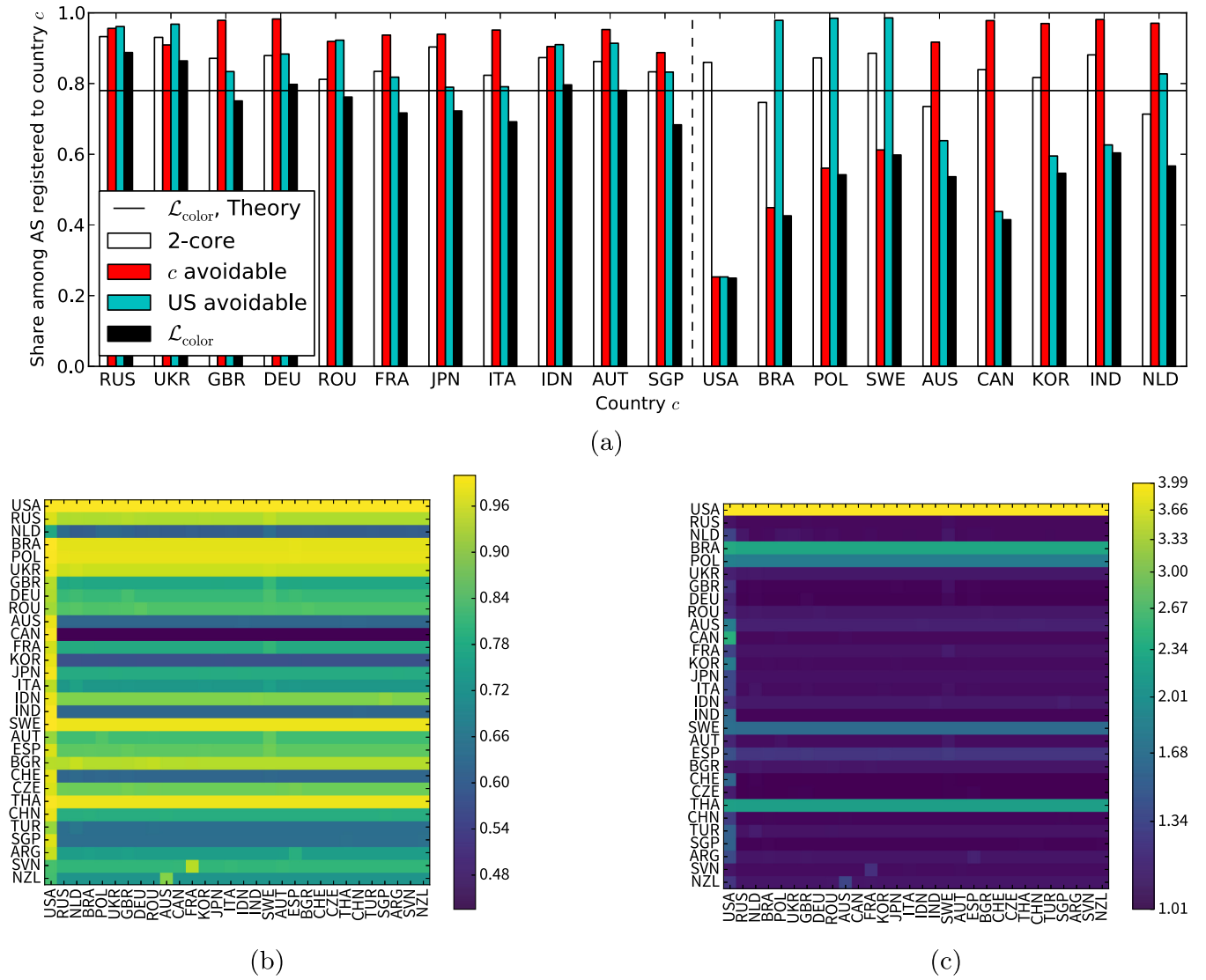


FIG. 5. Details for the AS-level Internet. (a) For comparing with theoretical results $S_{\text{color}} = |\mathcal{L}_{\text{color}}|/N$ (black line), we calculate shares $|\mathcal{L}_{\text{color}} \cap \{c_i = c\}|/|\{c_i = c\}|$ for 20 countries with the most AS. Here, $\{c_i = c\}$ is the set of all AS assigned to country c , and $||$ is the number of nodes in a set or component. Results are shown with black bars and compare well with theoretical results for the group of eleven countries shown to the left of the vertical dashed line. To understand why color connectivity breaks down for the other countries, we also plot the shares of two-core AS in the respective countries (white bars). The fraction $|\mathcal{L}_{\text{color}}^+ \cap \{c_i = c\}|/|\{c_i = c\}|$ (shown with red bars) is the fraction of AS in a country c which can communicate while avoiding further AS of that country c . The fraction $|\mathcal{L}_{\text{color}}^+ \cap \{c_i = c\}|/|\{c_i = c\}|$ (shown with cyan bars) is the fraction of AS in a country c which can communicate while avoiding AS registered to the USA. Color-avoiding connectivity when sender and receiver colors are trusted: (b) The fraction of nodes of row color that are in the largest CAC component when trusting all nodes of row color and column color. The value indicates the probability that a node of color i can communicate with the largest CAC set of nodes of color j . In panel (c), we show the relative improvement compared to not trusting the sender-receiver colors. For USA nodes, this makes them 4 times more likely to be color-avoiding connected, while for most countries, it makes a less substantial difference, unless they are communicating with the USA.

[60,61]. Network studies about risks to supply chains have shown that supply chains are often highly vulnerable, in general [62,63]. With the exception of case studies [64], risks associated with correlated failures due to geographical proximity [65–67] or shared ownership have not been studied. However, the effect of correlated failures can cause

serious supply-chain disruptions, such as those that occurred with the hard-drive industry following the 2011 Thailand floods. Color-avoiding percolation can be directly used to compute the portion of the world trade network between entities, which is CAC secure with respect to correlated failures. Calculating the fraction of a supply

chain that is in the giant CAC component provides an estimate of the resilience of different economic sectors to correlated failures.

Another application that fits neither the rubric of correlated failures nor that of security against eavesdroppers is in determining the maximal infectable set of individuals in a multistrain epidemic [68]. Coloring nodes by strain immunity, color-avoiding percolation can be used to evaluate the population's susceptibility to a multistrain infection. For example, let us consider the case in which C different strain immunities are present in the population. Let us assume that we do not know *a priori* which of the strains will emerge next and that the probability of infection for an immunized individual is very small compared to the probability of infection of nonimmunized individuals. Using color-avoiding percolation, we can compute the largest set of infection-prone individuals without knowledge of the exact type of strain that will emerge. This component represents the lower bound of the component affected by the epidemic and could be used as a lower bound of the amount of resources needed to contain the spread of disease, whatever strain emerges.

Color-avoiding percolation can also be applied to ecology. In Ref. [69], Barter and Gross propose a model for an ecological system in which species are spread over a network of spatial patches (islands). If different species can not survive in certain patches, we can color the patches depending on which species they *do not support*. The model presented here could estimate a portion of islands where it is possible to find all the species, assuming that species can occasionally visit neighboring islands even if they are uninhabitable for that species.

Vulnerability colorings in real-world networks have not been measured systematically in the past. In light of our findings that this heterogeneity of vulnerabilities can be used for improving security and robustness of complex systems, we hope that the collection of vulnerability-colored network data will be pursued in the future.

Here, we have presented the first systematic study of complex networks with vulnerabilities affecting classes of nodes and a way to maintain network robustness by utilizing multiple paths. This led us to develop color-avoiding percolation, a new type of percolation with a number of unique properties. We have shown that even a small diversity of colors can enable color-avoiding connectivity to a large fraction of nodes in a random network but that, in real-world networks, uneven distribution of vulnerabilities can undermine this effect. The framework and metrics uncover a hidden structure that underlies any complex network with classes of vulnerable nodes and can be used to devise new network design principles and protocols for improving robustness and security in real-world networks.

ACKNOWLEDGMENTS

We acknowledge financial support from the European Commission FET-Proactive project MULTIPLEX (Grant No. 317532) and the Italy-Israel NECST project. M. D. is grateful to the Azrieli Foundation for support. M. D. thanks Alan Danziger for first suggesting router software versions as a percolation problem. V. Z. acknowledges support by the H2020 CSA Twinning Project No. 692194, RBI-T-WINNING, and Croatian centers of excellence Quantix and Center of Research Excellence for Data Science and Cooperative Systems. We also express gratitude to Shlomo Havlin, Damir Vukičević, Marko Popović, Hrvoje Štefančić, and Damir Korančić for helpful comments in the preparation of this manuscript.

All authors contributed to the idea, discussion of results, and writing of the paper. S. K. and M. D. performed simulations, and S. K. developed the analytical treatment.

APPENDIX A: COLOR-AVOIDING CONNECTED COMPONENTS AND THE MAXIMALITY OF $\mathcal{L}_{\text{color}}$

As shown in Fig. 6, CAC components can have a broad variety of forms, and they can overlap. Compared to standard component analysis, this complicates the identification of CAC components. However, it is possible to show that $\mathcal{L}_{\text{color}}$ is always a CAC component (if it is not empty). For this, it has to be shown that $\mathcal{L}_{\text{color}}$ is maximal.

To show that $\mathcal{L}_{\text{color}}$ is a maximal component in the sense that there can be no other nodes that are CAC with all the nodes in $\mathcal{L}_{\text{color}}$ and are *not* a member of $\mathcal{L}_{\text{color}}$, we provide the following argument. First, let us assume that $\mathcal{L}_{\text{color}}$ is not maximal. That means that there exists a node v in the network which is (a) CAC to every node in $\mathcal{L}_{\text{color}}$ and (b) excluded from $\mathcal{L}_{\text{color}}$. Condition (b) means that there exists a color c' for which the node v is *not* a member of $\mathcal{L}_{c'}^+$. Consequently the node v cannot connect to the nodes in $\mathcal{L}_{\text{color}}$ which is by construction a subset of $\mathcal{L}_{c'}^+$, which contradicts (a).

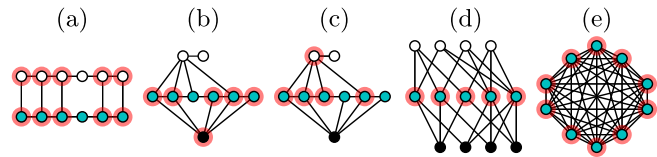


FIG. 6. Variety of color-avoiding components. Color-avoiding components may overlap, as shown in diagrams (b) and (c). Color-avoiding components can assume diverse forms. In a chain (a), paths between nodes of one color exist and can be reached by connections between nodes of different colors. In diagram (b), the black node serves as an alternative path provider for the blue nodes. Graph (d) does not need any connection among nodes of the same color, but there is a massive overhead of nodes and connections to achieve color-avoiding connectivity of the blue nodes. Graph (e) shows that a clique is a color-avoiding component.

<i>List of variables</i>	
Networks	
N	Number of nodes
\bar{k}	Average degree
k_i	Degree of node i
p_k	Degree distribution
α	Exponent of scale-free degree distribution
g_0	Generating function of degree
g_1	Generating function of excess degree
Colors	
C	Number of colors
$c \in 1, 2, \dots, C$	A color
r_c	Color distribution
Standard percolation ingredients	
\mathcal{L}	Set of nodes in the largest component (color blind)
u	Probability of not being connected to a giant comp. over a link
S	Size of giant component
$\mathcal{L}_{\bar{c}}$	Set of nodes in the largest component, after nodes of color c deleted
$u_{\bar{c}}$	Probability of not being connected to giant $\mathcal{L}_{\bar{c}}$ over a link
$S_{\bar{c}}$	Size of giant $\mathcal{L}_{\bar{c}}$
$\mathcal{L}_{\bar{c}}^+$	Union of $\mathcal{L}_{\bar{c}}$ with all nodes of color c being direct neighbors of nodes in $\mathcal{L}_{\bar{c}}$
Percolation over color-avoiding paths	
$\mathcal{L}_{\text{color}}$	Candidate set of nodes for the largest avoidable color component
S_{color}	Size of giant $\mathcal{L}_{\text{color}}$
$B_{k,k'}$	Probability that, out of k links, k' connect to a giant component
$M_{k',\bar{\kappa}}$	Probability that, out of k' links, κ_1 connect to color 1, etc.
$P_{\bar{\kappa}}$	Success probability having neighbors of colors acc. to $\bar{\kappa}$
$U_{\bar{c}}$	Probability that a link fails to connect to $\mathcal{L}_{\text{color}}$, which already connects to \mathcal{L} and a node not having color c
$S_{\text{color},\infty}$	Size of the set of all nodes being connected to giant components over two links or more
β	Critical exponent
\bar{k}_{crit}	Critical value of average degree

APPENDIX B: SIZE OF GIANT AVOIDABLE-COLORS COMPONENT IN THE CONFIGURATION MODEL

We can find analytical results for S_{color} for random graph ensembles with randomly distributed colors in the limit of infinite graphs. These results can be used to estimate the situation in finite quenched networks. We are able to gain a general understanding including phase transitions. This knowledge can guide our understanding of real-world networks.

We use the generalized-configuration-model graph ensemble with N nodes, where each degree sequence $\{k_i\}$ occurs with probability $\prod_i p_{k_i}$, with the degree distribution p_k . Additionally, we assign to every node i a color $c_i \in 1, 2, \dots, C$. The color sequence $\{c_i\}$ has probability $\prod_i r_{c_i}$ with the color distribution r_c . For a graph G_N out of the graph ensemble, $\mathcal{L}_{\text{color}}$ has a certain size $N_{\text{color}}(G_N)$. For the whole graph ensemble, we have to use the average value. By considering only giant contributions growing with network size, we have

$$S_{\text{color}} = \lim_{N \rightarrow \infty} \sum_{G_N} P(G_N) \frac{N_{\text{color}}(G_N)}{N}, \quad (\text{B1})$$

where $P(G_N) = \prod_i p_{k_i} \omega \prod_i r_{c_i}$ is the probability of having the graph G_N of size N , including ω , the probability of the connection scheme of G_N as a matching of half edges.

1. Question and connection to percolation theory

For calculating S_{color} in the random graph ensemble, we follow the ideas of Erdős and Rényi [43] and Newman [44]. For calculating the size of the giant component, they used probabilities of connections for a single node in the graph ensemble. As we have to extend the method to a gradual procedure with conditional probabilities, it is useful to introduce the original method in detail with a shifted viewpoint.

Let us denote with \mathcal{L} the set of all nodes belonging to the largest component. In the top panels of Fig. 7, a possible situation is illustrated. The largest component contains a large part of the network, and the remaining nodes belong to smaller components. We have to calculate the size S of the giant component, meaning the average relative size of \mathcal{L} in the network ensemble in the limit of infinite network size. For this calculation, we can define the average probability u that a node fails to connect to \mathcal{L} over one particular link. This is illustrated in the left part of the figure. Again, the thermodynamic limit $N \rightarrow \infty$ is implied. With the definition of u at hand, we can calculate S in two steps: First, using a self-consistency equation, u is calculated. The probability u is identical to the probability that the neighbor does not connect to the giant component over any of the remaining links,

$$u = g_1(u), \quad g_1(z) = \sum_k q_k z^k. \quad (\text{B2})$$

In this equation, g_1 is the generating function of excess degree $q_k = (k+1)p_{k+1}/\bar{k}$. For important degree distributions, e.g., Poisson or scale-free, the equation for u can only be solved numerically. The second step is an averaging over nodes with various degrees k . The probability to connect to the giant component over any of k links is $(1 - u^k)$, meaning that not all links fail at the same time.

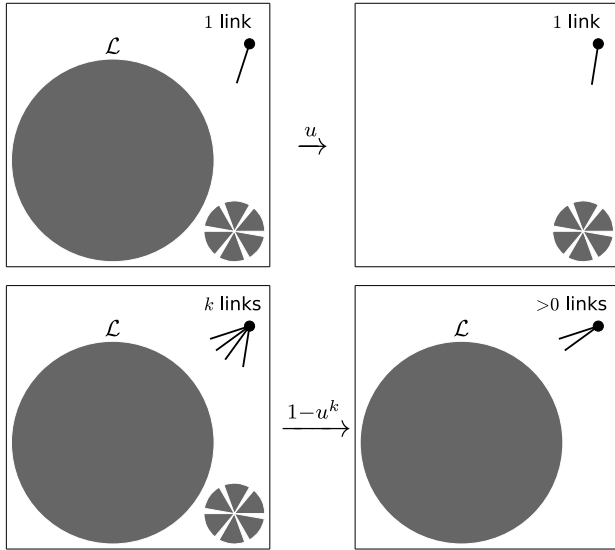


FIG. 7. We base our theory on the method to calculate the size of normal giant components, as illustrated in this figure. Using a self-consistency equation, the probability u can be calculated. This is the probability that a node is not connected to the giant component over a single link (see top). On the bottom, the probability for a node with k links is illustrated to have at least one link connecting to the giant component. Note that u^k is the probability that all links fail.

This is illustrated in the bottom part of the figure. As a node that connects to the giant component belongs to it,

$$S = \sum_{k=0}^{\infty} p_k (1 - u^k) = 1 - g_0(u), \quad g_0(z) = \sum_k p_k z^k. \quad (\text{B3})$$

In analogy to the procedure described above, we calculate S_{color} as the probability that a randomly chosen node belongs to $\mathcal{L}_{\text{color}}$. This has to be evaluated in the graph ensemble of infinite size. As we perform an averaging over nodes with various degrees k , the following question has to be answered: What is the probability that a node with k links connects to a giant $\mathcal{L}_{\bar{c}}$ for all colors c at the same time. This is illustrated in Fig. 8. On the left, the situation for a graph with colors on the nodes is illustrated. Nodes of all colors might be in the largest component. After deleting all nodes of one color c , the remaining largest component $\mathcal{L}_{\bar{c}}$ might still contain a large part of all nodes in \mathcal{L} . The condition for the node belonging to $\mathcal{L}_{\text{color}}$ is illustrated in the right part of the figure.

We use the same two steps to attack this problem, as described for calculating the giant component above. First, we provide some single link probabilities that can be used as primitives for further calculations. Second, we combine the single link probabilities to calculate S_{color} .

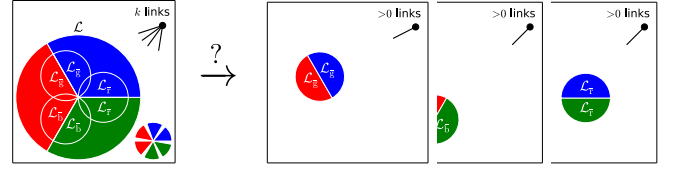


FIG. 8. We have to calculate the probability, if a node with k links is, for every color c , connected to the giant component $\mathcal{L}_{\bar{c}}$ with deleted color c . All connections over at least one link have to exist at the same time. We illustrate this question with red ($c = r$), green ($c = g$), and blue ($c = b$). If a link connects to $\mathcal{L}_{\bar{g}}$, it definitely does not connect to $\mathcal{L}_{\bar{r}}$ for one of the other colors. This kind of dependence forces us to use a stepwise calculation with conditional probabilities.

2. Single link probabilities

We have already shown Eq. (B2) for calculating the probability u . In the case of colors on the nodes, as illustrated in the left part of Fig. 9, the colors can simply be ignored. We further need the probability to connect to a node of color c , which is simply r_c . This is illustrated in the second column of the figure. We further introduce $u_{\bar{c}}$, the probability that a single link does not connect to a giant $\mathcal{L}_{\bar{c}}$. See the third column of the figure for an illustration. This can be calculated using percolation theory for random attack by solving

$$u_{\bar{c}} = r_c + (1 - r_c)g_1(u_{\bar{c}}). \quad (\text{B4})$$

Unfortunately, $u_{\bar{c}}$ cannot be used directly for calculating S_{color} . If we look at the same link, the probabilities $u_{\bar{c}}$ are dependent for different colors. The most obvious argument is that we always have $\Pi_c(1 - u_{\bar{c}}) = 0$, as a link must at least miss one of the $\mathcal{L}_{\bar{c}}$. Instead, we use the conditional probability $U_{\bar{c}}$, as illustrated with the outer right column of the figure. The precondition is that a link connects to the giant component, and the node it connects to has no color c . Here, $U_{\bar{c}}$ is the probability that such a link connects to $\mathcal{L}_{\bar{c}}$. For calculating it, we use the primitives introduced so far, as illustrated in Fig. 10. Assuming independence of the probabilities $(1 - u)$ for connecting to the giant component and $(1 - r_c)$ for not connecting to a node of color c , the precondition of $U_{\bar{c}}$ can be constructed. In this way, we can construct $(1 - u_{\bar{c}})$ using the probability we are searching for: $(1 - u_{\bar{c}}) = (1 - u)(1 - r_c)(1 - U_{\bar{c}})$. With this result, we find

$$U_{\bar{c}} = 1 - \frac{1 - u_{\bar{c}}}{(1 - u)(1 - r_c)}. \quad (\text{B5})$$

If $(1 - u)(1 - r_c) = 0$, the precondition holds for an empty set of nodes. In this case, we define $U_{\bar{c}} = 1$. Notice that the additional information of the explicit color, instead of only stating that the color is not c , does not alter the results, as a

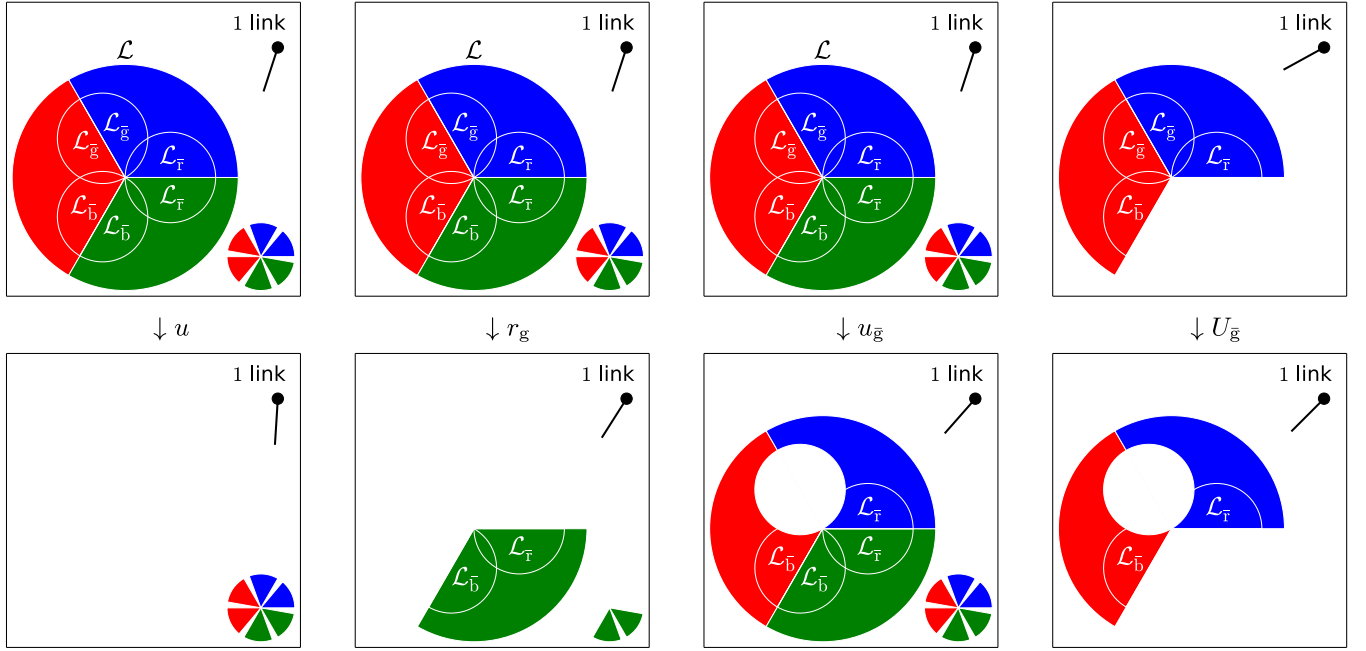


FIG. 9. Probabilities for a single link to connect to different parts of the network. We use these probabilities as primitives to calculate the probability for many links. While u , r_g , and $u_{\bar{g}}$ can be calculated with standard methods invented for the configuration model, the conditional probability $U_{\bar{c}}$ can be calculated as a combination of the others.

further restriction of the colors would decrease the numerator and the denominator by the same factor.

3. Averaging over link distributions

As in Eq. (B3) for the giant component, we want to get an analytical result for S_{color} by averaging over possible link

constellations of a randomly chosen node. Let us give the whole result and then explain it step by step afterwards:

$$S_{\text{color}} = \sum_{k=0}^{\infty} p_k \sum_{k'=0}^k B_{k,k'} \sum_{\kappa_1, \dots, \kappa_C=0}^{k'} M_{k', \vec{\kappa}} P_{\vec{\kappa}}, \quad (\text{B6})$$

$$B_{k,k'} = \binom{k}{k'} (1-u)^{k'} u^{k-k'}, \quad (\text{B7})$$

$$M_{k', \vec{\kappa}} = \frac{k'!}{\kappa_1! \dots \kappa_C!} (r_1)^{\kappa_1} \dots (r_C)^{\kappa_C} \delta_{k', \kappa_1 + \dots + \kappa_C}, \quad (\text{B8})$$

$$P_{\vec{\kappa}} = \prod_{c=1}^C [1 - (U_{\bar{c}})^{k' - \kappa_c}]. \quad (\text{B9})$$

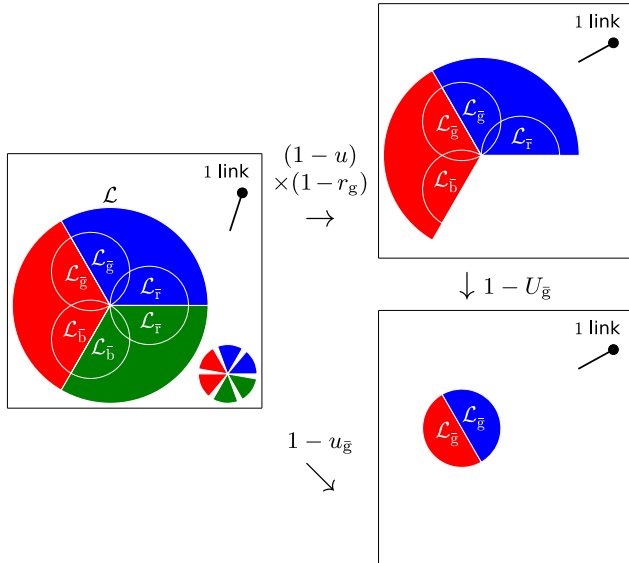


FIG. 10. Calculation of $U_{\bar{g}}$ using the equality $(1-u)(1-r_g)(1-U_{\bar{g}}) = 1-u_{\bar{g}}$. For this calculation, we have assumed independence of the qualities of the link under consideration, especially of the color it connects to and if it connects to the giant component.

The formulas include the single link probabilities r_c , u [Eq. (B2)] and $U_{\bar{c}}$ [Eq. (B5) with (B4)]. An illustration of the procedure can be seen in Fig. 11. Here, $B_{k,k'}$ is the binomial probability that, out of the k links, k' links connect to the giant component. Note that $M_{k', \vec{\kappa}}$ gives the multinomial probability for a certain color distribution among the k' links connecting to the giant component. We assume that this second step is independent of the first step, which is confirmed with the final results. The numbers κ_c count the links that connect to a node of color c in the giant component. Finally, $P_{\vec{\kappa}}$ gives the joint probability that for the color distribution given by $\vec{\kappa}$, all giant $\mathcal{L}_{\bar{c}}$ are connected at the same time. There is at least one link connecting to $\mathcal{L}_{\bar{c}}$ with probability $1 - (U_{\bar{c}})^{k' - \kappa_c}$. If this is successful, the

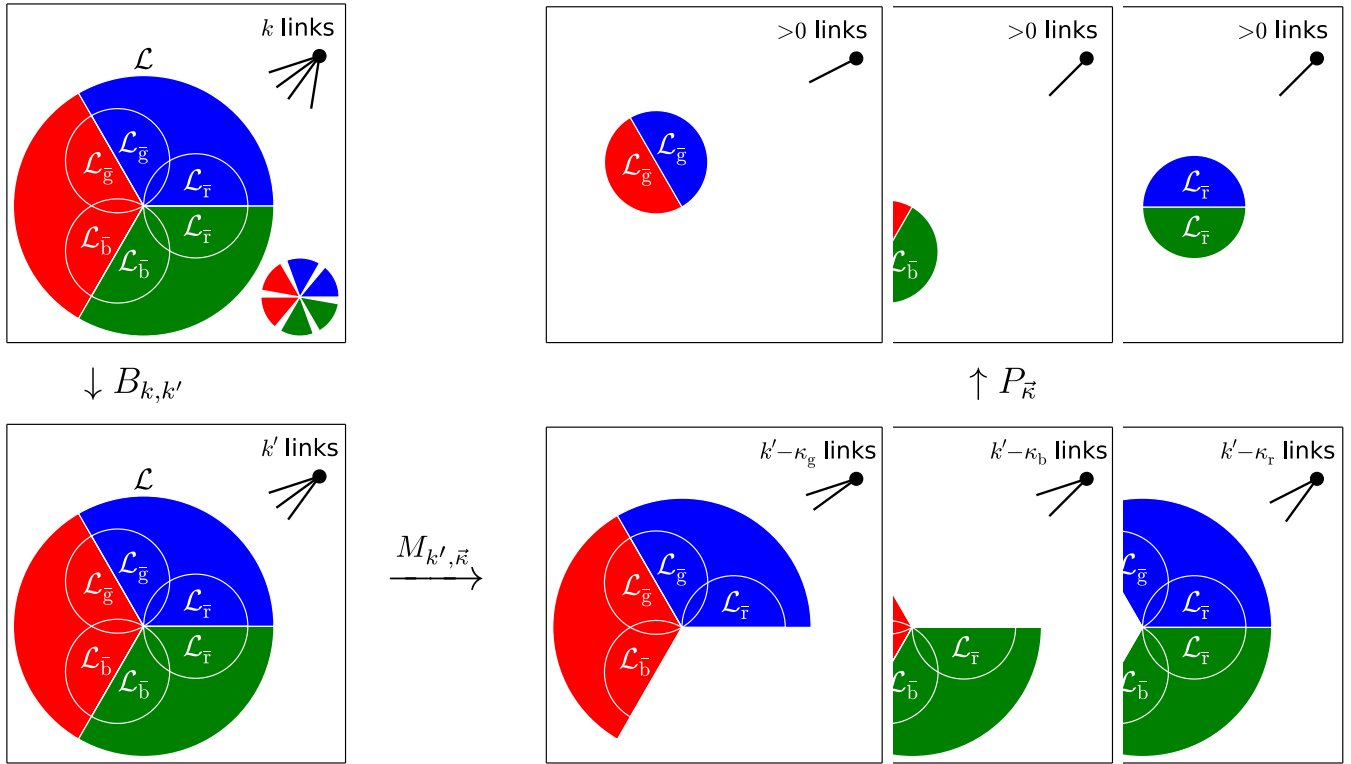


FIG. 11. For calculating the probability of a node with k links to belong to $\mathcal{L}_{\text{color}}$, we have to average over different link constellations that this node might show. First, $B_{k,k'}$ is the probability that, out of the k links, k' connect to the giant component. It is calculated using u (compare the left side of Fig. 9). Second, $M_{k',\vec{x}}$ gives the probability for a certain color distribution among the links. It is calculated using r_g , etc. (compare Fig. 9, second from left). We assume that this second step is independent of the first step, which is confirmed with the final results. Third, $P_{\vec{x}}$ gives the joint probability that for this color distribution, $\mathcal{L}_{\vec{r}}$, $\mathcal{L}_{\vec{b}}$, and $\mathcal{L}_{\vec{g}}$ are connected at the same time. This is calculated using $U_{\vec{r}}$, etc. (compare the right side of Fig. 9).

randomly chosen node belongs to \mathcal{L}_c^+ . The success probabilities for different colors have to be multiplied, as all $\mathcal{L}_{\bar{c}}$ have to be simultaneously reached.

APPENDIX C: EXAMINATION OF S_{color}

1. Closed-form solutions

We now calculate closed-form solutions for S_{color} for special cases. This calculation is done to demonstrate how the extensive summations over k' , k , and \vec{k} can be performed analytically. In cases where this is not possible, a sampling of values \vec{k} has to be performed. The results can be tested against the analytically tractable situations and by comparing with numerical results. The closed-form solutions presented here were used to calculate analytical results for the main article as well.

For evaluating Eq. (B6) with two colors, we first rewrite

$$\sigma_{k'} \equiv \sum_{K_1, K_2=0}^{k'} M_{k', \vec{K}} P_{\vec{K}} \quad (\text{C1})$$

$$= \sum_{\kappa_1=0}^{k'} \binom{k'}{\kappa_1} (r_1)^{\kappa_1} (r_2)^{k'-\kappa_1} [1 - (U_1)^{k'-\kappa_1}] [1 - (U_2)^{\kappa_1}] \quad (\text{C2})$$

$$= \sum_{\kappa_1=0}^{k'} \binom{k'}{\kappa_1} [(r_1)^{\kappa_1} (r_2)^{k'-\kappa_1} - (r_1 U_2)^{\kappa_1} (r_2)^{k'-\kappa_1}] \quad (\text{C3})$$

$$\begin{aligned}
& - (r_1)^{\kappa_1} (r_2 U_{\bar{1}})^{k' - \kappa_1} + (r_1 U_{\bar{2}})^{\kappa_1} (r_2 U_{\bar{1}})^{k' - \kappa_1} \\
& = 1 - (r_1 + r_2 U_{\bar{1}})^{k'} - (r_2 + r_1 U_{\bar{2}})^{k'} + (r_1 U_{\bar{2}} + r_2 U_{\bar{1}})^{k'}.
\end{aligned} \tag{C4}$$

In the last step, the binomial formula was used backward. We can use this procedure once more, and with Eq. (B5) and $r_1 + r_2 = 1$, we find

$$S_{\text{color}} = \sum_k p_k \sum_{k'=0}^k B_{k,k'} \sigma_{k'} \quad (\text{C5})$$

$$= \sum_k p_k \sum_{k'=0}^k \binom{k}{k'}' u^{k-k'} \times [(1-u)^{k'} - ((1-u)(r_1 + r_2 U_{\bar{1}}))^{k'} - \dots] \quad (\text{C6})$$

$$= \sum_k p_k [1 - (u_{\bar{1}})^k - (u_{\bar{2}})^k + (u_{\bar{1}} + u_{\bar{2}} - 1)^k] \quad (\text{C7})$$

$$= 1 - g_0(u_{\bar{1}}) - g_0(u_{\bar{2}}) + g_0(u_{\bar{1}} + u_{\bar{2}} - 1). \quad (\text{C8})$$

This result holds for any degree distribution and color distribution. Notice that $r_c \leq u_{\bar{c}} \leq 1$. The result for two colors only depends on the probabilities $u_{\bar{c}}$, while conditional probabilities such as $U_{\bar{c}}$ were eliminated. This was possible as $\mathcal{L}_{\bar{1}}$ and $\mathcal{L}_{\bar{2}}$ are not overlapping for two colors. For Poisson graphs, we find, with the according generating function,

$$g_0(z) = g_1(z) = e^{\bar{k}(z-1)}, \quad (\text{C9})$$

$$S_{\text{color}} = [1 - g_0(u_{\bar{1}})][1 - g_0(u_{\bar{2}})]. \quad (\text{C10})$$

For more than two colors, $\mathcal{L}_{\bar{c}}$ overlap. For homogeneous color distributions $r_c = 1/C$, a closed-form solution can be found in the same way as for two colors with the binomial formula. We find

$$S_{\text{color}} = \sum_{j=0}^C (-1)^j \binom{C}{j} \times g_0 \left[u + (1-u) \left(\frac{j}{C} U_{\bar{1}}^{j-1} + \frac{C-j}{C} U_{\bar{1}}^j \right) \right]. \quad (\text{C11})$$

Let us finally discuss the behavior for $C \rightarrow \infty$. This can be done by utilizing the term $\sigma_{k'}$, the probability that a node connecting over k' links to the giant component belongs to S_{color} . As can be seen with Eq. (B9), $\sigma_0 = \sigma_1 = 0$. On the other hand, with $r_c \rightarrow 0$, Eq. (B4) converges to Eq. (B2), and therefore $U_{\bar{1}} \rightarrow 0$. This means that $\sigma_{k' > 1} \rightarrow 1$. We finally find, with Eq. (B6), that

$$S_{\text{color}, \infty} \equiv \lim_{C \rightarrow \infty} S_{\text{color}} \quad (\text{C12})$$

$$= 1 - \sum_{k=0}^{\infty} p_k [u^k + k(1-u)u^{k-1}] \quad (\text{C13})$$

$$= 1 - g_0(u) - (1-u) \frac{dg_0(z)}{dz} \Big|_{z=u}. \quad (\text{C14})$$

2. Critical behavior for Poisson graphs

With Eq. (B6), vanishing $\sigma_{k'}$ causes $S_{\text{color}} = 0$. According to

$$\sigma_{k'} = \sum_{\kappa_1, \dots, \kappa_C=0}^{k'} M_{k', \vec{\kappa}} \prod_{c=1}^C [1 - (U_{\bar{c}})^{k' - \kappa_c}], \quad (\text{C15})$$

this is the case if $U_{\bar{c}} = 1$ for any color c . With Eq. (B5), we find that $U_{\bar{c}} = 1$ whenever $u_{\bar{c}} = 1$. Examining Eq. (B4) for $u_{\bar{c}}$, we can relate it to site percolation (random removal of nodes). For Poisson graphs, we have $r_{\text{crit}} = (\bar{k} - 1)/\bar{k}$. With homogeneous color distribution $r_c = 1/C$, we can resolve the critical connectivity given the number of colors

$$\bar{k}_{\text{crit}} = C/(C-1). \quad (\text{C16})$$

The normal giant component size S shows a special critical behavior just above the transition point; it scales linearly with $\bar{k} - 1$. Here, we are interested in the behavior of S_{color} , which is a function of $1 - u_{\bar{1}}$, which itself can be related to $1 - u = S$. By inserting into Eq. (B4), it can be shown that $u_{\bar{c}}(\bar{k}) = r_c + (1 - r_c)u((1 - r_c)\bar{k})$. For small arguments $(\bar{k} - \bar{k}_{\text{crit}})$,

$$1 - u_{\bar{1}}(\bar{k} > \bar{k}_{\text{crit}}) \approx (1 - r_1)^2 \frac{d(1-u)}{d\bar{k}} \Big|_{\bar{k}=1+0} (\bar{k} - \bar{k}_{\text{crit}}). \quad (\text{C17})$$

Inserting this into Eq. (B5), we find, using $1 - u(\bar{k} > 1) \approx [d(1-u)/d\bar{k}]|_{\bar{k}=1+0}(\bar{k} - 1)$,

$$\varepsilon \equiv 1 - U_{\bar{1}} \approx C(\bar{k} - \bar{k}_{\text{crit}}), \quad (\text{C18})$$

if additionally $\bar{k} - \bar{k}_{\text{crit}} \ll \bar{k} - 1$ holds ($1 - u_{\bar{1}}$ small compared to $1 - u$).

For calculating $\sigma_{k'}$, we first need to evaluate $P_{\vec{\kappa}}$, including expressions $1 - (U_{\bar{1}})^{k' - \kappa_c}$. Replacing this with ε and applying an approximation, we find $1 - (U_{\bar{1}})^{k' - \kappa_c} = 1 - (1 - \varepsilon)^{k' - \kappa_c} \approx (k' - \kappa_c)\varepsilon$. This is true at least as long as $k'\varepsilon \ll 1$. With this result, we find $P_{\vec{\kappa}} \propto (\bar{k} - \bar{k}_{\text{crit}})^C$ independent of $\vec{\kappa}$, and

$$\sigma_{k'} \propto (\bar{k} - \bar{k}_{\text{crit}})^C. \quad (\text{C19})$$

Finally, we find

$$S_{\text{color}} \propto (\bar{k} - \bar{k}_{\text{crit}})^\beta, \quad (\text{C20})$$

$$\beta = C. \quad (\text{C21})$$

The critical behavior of $S_{\text{color}, \infty}$, for Poisson graphs can be evaluated with the generating function Eq. (C9) and $S = 1 - u$. We find $S_{\text{color}, \infty} = S - \bar{k}S(1 - S)$, and for small positive $\bar{k} - 1$, the giant component grows approximately with $S \approx 2(\bar{k} - 1)/\bar{k}^2$. Therefore,

$$S_{\text{color},\infty} \approx (\bar{k} - 1)^2(4/\bar{k}^2 - 2/\bar{k}^3) \propto (\bar{k} - 1)^2. \quad (\text{C22})$$

APPENDIX D: FINDING S_{color} IN A DIFFERENT WAY

So far, we have constructed S_{color} by averaging over possible link configurations of a randomly chosen node. There is a more direct way to calculate S_{color} , by understanding \mathcal{L}_c^+ as events and using rules for probabilities of intersections and unions of events. We start by rewriting with the complementary event

$$S_{\text{color}} = \text{Prob}(\mathcal{L}_1^+ \cap \mathcal{L}_2^+ \cap \dots \cap \mathcal{L}_C^+) \quad (\text{D1})$$

$$= 1 - \text{Prob}(\overline{\mathcal{L}_1^+} \cup \dots \cup \overline{\mathcal{L}_C^+}). \quad (\text{D2})$$

Negation turns the intersection into a union. The union is hard to calculate using single link probabilities. Therefore, we use the inclusion exclusion principle [70]

$$\begin{aligned} \text{Prob}(\overline{\mathcal{L}_1^+} \cup \dots \cup \overline{\mathcal{L}_C^+}) &= \sum_c \text{Prob}(\overline{\mathcal{L}_c^+}) \\ &\quad - \sum_{c < q} \text{Prob}(\overline{\mathcal{L}_c^+} \cap \overline{\mathcal{L}_q^+}) + \dots \end{aligned} \quad (\text{D3})$$

The probability of a union of events equals the sum of probabilities of the events, corrected for pair overlaps of events, which are further corrected for triplet overlaps and so on. Furthermore, we can calculate the needed probabilities using generating functions and single link probabilities. We use $u_{\bar{1},\dots,\bar{j}}$ as the probability that a node does not connect to component $\mathcal{L}_{\bar{1}}^+$, and that at the same time, over the same link, it does not connect to $\mathcal{L}_{\bar{2}}^+$, ..., and that over the same link, it does not connect to $\mathcal{L}_{\bar{j}}^+$. We have

$$\text{Prob}(\overline{\mathcal{L}_1^+} \cap \dots \cap \overline{\mathcal{L}_j^+}) = g_0(u_{\bar{1},\dots,\bar{j}}). \quad (\text{D4})$$

If all colors have the same frequency, it holds that $u_{\bar{1},\bar{2}} = u_{\bar{c},\bar{q}}$ for any $c < q \leq C$, etc. Therefore, many terms can be grouped together, and we find

$$S_{\text{color}} = \sum_{j=0}^C (-1)^j \binom{C}{j} g_0(u_{\bar{1},\dots,\bar{j}}). \quad (\text{D5})$$

For $u_{\bar{1},\dots,\bar{j}}$, we can give an approximate result, which finally leads to what we found in Eq. (C11):

$$u_{\bar{1},\dots,\bar{j}} \approx u + (1 - u) \left[\frac{j}{C} U_1^{j-1} + \frac{C-j}{C} U_1^j \right]. \quad (\text{D6})$$

With probability u , a link does not connect to the giant component, so $u_{\bar{1},\dots,\bar{j}} \geq u$. If it connects to the giant

component [probability $(1 - u)$], then the node on the other side has either one of the j colors (probability j/C), leaving a simultaneous failure probability of U_1^{j-1} , or it has a different color with probability $(C - j)/C$, leaving a simultaneous failure probability of U_1^j . In this approximation, we only neglect dependencies among the conditional probabilities U_1, U_2 , etc.

-
- [1] D. Helbing, *Globally Networked Risks and How to Respond*, *Nature (London)* **497**, 51 (2013).
 - [2] Heartbleedbug, <http://heartbleed.com> (2014).
 - [3] Z. Durumeric, J. Kasten, D. Adrian, J. A. Halderman, M. Bailey, F. Li, N. Weaver, J. Amann, J. Beekman, M. Payer, and V. Paxson, *The Matter of Heartbleed*, in *Proceedings of the 2014 Conference on Internet Measurement Conference* (ACM, New York, NY, 2014), pp. 475–488.
 - [4] G. R. Blakley, *Safeguarding Cryptographic Keys*, in *International Workshop on Managing Requirements Knowledge* (IEEE Computer Society, New York, 1979), pp. 313–317.
 - [5] A. Shamir, *How to Share a Secret*, *Commun. ACM* **22**, 612 (1979).
 - [6] A. Beimel, *Secret-Sharing Schemes: A Survey*, in *Coding and Cryptology, Lecture Notes in Computer Science*, edited by Y. Chee, Z. Guo, S. Ling, F. Shao, Y. Tang, H. Wang, and C. Xing (Springer, Berlin, Heidelberg, 2011), Vol. 6639, pp. 11–46.
 - [7] Wikimedia commons artwork CC BY-SA 3.0, (2015), <http://tinyurl.com/WikiCommonsCitation>.
 - [8] N. Cai and R. W. Yeung, *Secure Network Coding*, in *Proceedings of the 2002 IEEE International Symposium on Information Theory* (IEEE, New York, 2002), pp. 323.
 - [9] K. Jain, *Security Based on Network Topology Against the Wiretapping Attack*, *IEEE Wireless Comm.* **11**, 68 (2004).
 - [10] Y. Wang and Y. Desmedt, *Edge-Colored Graphs with Applications to Homogeneous Faults*, *Inf. Proc. Lett.* **111**, 634 (2011).
 - [11] R. Du, C. Zhao, F. Zhao, and S. Li, *Strategies of Network Coding Against Nodes Conspiracy Attack*, *Security Commun. Networks* **8**, 2396 (2015).
 - [12] R. Du, C. Zhao, S. Li, and J. Li, *Efficient Weakly Secure Network Coding Scheme Against Node Conspiracy Attack Based on Network Segmentation*, *EURASIP J. Wirel. Commun. Network.* **2014**, 5 (2014).
 - [13] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle III, and J. Doyle, *Robustness of Cellular Functions*, *Cell* **118**, 675 (2004).
 - [14] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, *A Model of Internet Topology Using K-Shell Decomposition*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 11150 (2007).
 - [15] C. H. Yeung, D. Saad, and K. Y. Michael Wong, *From the Physics of Interacting Polymers to Optimizing Routes on the London Underground*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13717 (2013).
 - [16] R. Cohen and S. Havlin, *Complex Networks: Structure, Robustness and Function* (Cambridge University Press, Cambridge, England, 2010).

- [17] M. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
- [18] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Complex Networks: Structure and Dynamics*, *Phys. Rep.* **424**, 175 (2006).
- [19] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, Oxford, 2007).
- [20] D. Achlioptas, R. M. D'Souza, and J. Spencer, *Explosive Percolation in Random Networks*, *Science* **323**, 1453 (2009).
- [21] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, *Resilience of the Internet to Random Breakdowns*, *Phys. Rev. Lett.* **85**, 4626 (2000).
- [22] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin, *Breakdown of the Internet under Intentional Attack*, *Phys. Rev. Lett.* **86**, 3682 (2001).
- [23] R. Pastor-Satorras and A. Vespignani, *Epidemic Spreading in Scale-Free Networks*, *Phys. Rev. Lett.* **86**, 3200 (2001).
- [24] F. Morone and H. A. Makse, *Influence Maximization in Complex Networks through Optimal Percolation*, *Nature (London)* **524**, 65 (2015).
- [25] Y. Sasson, D. Cavin, and A. Schiper, *Probabilistic Broadcast for Flooding in Wireless Mobile Ad Hoc Networks*, in *Wireless Communications and Networking* (IEEE, New York, 2003), Vol. **2**, pp. 1124–1130.
- [26] J. Goldenberg, B. Libai, S. Solomon, N. Jan, and D. Stauffer, *Marketing Percolation*, *Physica (Amsterdam)* **284A**, 335 (2000).
- [27] S. Solomon, G. Weisbuch, L. de Arcangelis, N. Jan, and D. Stauffer, *Social Percolation Models*, *Physica (Amsterdam)* **277A**, 239 (2000).
- [28] I. Breskin, J. Soriano, E. Moses, and T. Tlusty, *Percolation in Living Neural Networks*, *Phys. Rev. Lett.* **97**, 188102 (2006).
- [29] A. G. Smart, L. A. N. Amaral, and J. M. Ottino, *Cascading Failure and Robustness in Metabolic Networks*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 13223 (2008).
- [30] M. A. Aon, S. Cortassa, and B. O'Rourke, *Percolation and Criticality in a Mitochondrial Network*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4447 (2004).
- [31] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *k-Core Organization of Complex Networks*, *Phys. Rev. Lett.* **96**, 040601 (2006).
- [32] A. V. Goltsev, S. N. Dorogovtsev, and J. F. F. Mendes, *k-Core (Bootstrap) Percolation on Complex Networks: Critical Phenomena and Nonlocal Effects*, *Phys. Rev. E* **73**, 056101 (2006).
- [33] M. E. J. Newman and G. Ghoshal, *Bicomponents and the Robustness of Networks to Failure*, *Phys. Rev. Lett.* **100**, 138701 (2008).
- [34] T. P. Peixoto, *Hierarchical Block Structures and High-Resolution Model Selection in Large Networks*, *Phys. Rev. X* **4**, 011047 (2014).
- [35] Caida Project, <http://www.caida.org/data>.
- [36] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, *Catastrophic Cascade of Failures in Interdependent Networks*, *Nature (London)* **464**, 1025 (2010).
- [37] J. Gao, S. V. Buldyrev, H. E. Stanley, and S. Havlin, *Networks Formed from Interdependent Networks*, *Nat. Phys.* **8**, 40 (2012).
- [38] G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, *Avalanche Collapse of Interdependent Networks*, *Phys. Rev. Lett.* **109**, 248701 (2012).
- [39] M. Kivelä, A. Arenas, M. Barthélemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, *Multilayer Networks*, *J. Complex Netw.* **2**, 203 (2014).
- [40] M. M. Danziger, A. Bashan, Y. Berezin, L. M. Shekhtman, and S. Havlin, *An Introduction to Interdependent Networks*, in *Nonlinear Dynamics of Electronic Systems*, Communications in Computer and Information Science, edited by V. M. Mladenov and P. C. Ivanov (Springer International Publishing, Cham, Switzerland, 2014), Vol. 438, pp. 189–202.
- [41] J. Gao, D. Li, and S. Havlin, *From a Single Network to a Network of Networks*, *Natl. Sci. Rev.* **1**, 346 (2014).
- [42] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin, *The Structure and Dynamics of Multilayer Networks*, *Phys. Rep.* **544**, 1 (2014).
- [43] P. Erdős and A. Rényi, *On Random Graphs*, *Publ. Math.* **6**, 290 (1959).
- [44] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, *Random Graphs with Arbitrary Degree Distributions and Their Applications*, *Phys. Rev. E* **64**, 026118 (2001).
- [45] R. Albert, H. Jeong, and A.-L. Barabási, *Error and Attack Tolerance of Complex Networks*, *Nature (London)* **406**, 378 (2000).
- [46] D. Cellai, S. N. Dorogovtsev, and G. Bianconi, *Message Passing Theory for Percolation Models on Multiplex Networks with Link Overlap*, *Phys. Rev. E* **94**, 032301 (2016).
- [47] R. Zallen, *Polychromatic Percolation: Coexistence of Percolating Species in Highly Connected Lattices*, *Phys. Rev. B* **16**, 1426 (1977).
- [48] J. C. Wierman, *AB Percolation: A Brief Survey*, *Banach Center Publications* **25**, 241 (1989).
- [49] D. Cellai, E. López, J. Zhou, J. P. Gleeson, and G. Bianconi, *Percolation in Multiplex Networks with Overlap*, *Phys. Rev. E* **88**, 052811 (2013).
- [50] D. Zhou, A. Bashan, R. Cohen, Y. Berezin, N. Shnerb, and S. Havlin, *Simultaneous First- and Second-Order Percolation Transitions in Interdependent Networks*, *Phys. Rev. E* **90**, 012803 (2014).
- [51] M. M. Danziger, A. Bashan, Y. Berezin, and S. Havlin, *Percolation and Cascade Dynamics of Spatial Networks with Partial Dependency*, *J. Complex Netw.* **2**, 460 (2014).
- [52] D. Dolev, C. Dwork, O. Waarts, and M. Yung, *Perfectly Secure Message Transmission*, *J. Assoc. Comput. Mach.* **40**, 17 (1993).
- [53] M. Radi, B. Dezfouli, K. Abu Bakar, and M. Lee, *Multipath Routing in Wireless Sensor Networks: Survey and Research Challenges*, *Sensors* **12**, 650 (2012).
- [54] Y. Hong, D. Kim, D. Li, L. Guo, J. Son, and A. O. Tokuta, *Two New Multi-path Routing Algorithms for Fault-Tolerant Communications in Smart Grid*, *Ad Hoc Networks* **22**, 3 (2014), Special Issue on Routing in Smart Grid Communication Networks.
- [55] D. T. Nguyen, Y. Shen, and M. T. Thai, *Detecting Critical Nodes in Interdependent Power Networks for Vulnerability Assessment*, *IEEE Trans. Smart Grid* **4**, 151 (2013).
- [56] S. Schuster, D. A. Fell, and T. Dandekar, *A General Definition of Metabolic Pathways Useful for Systematic*

- Organization and Analysis of Complex Metabolic Networks*, *Nat. Biotechnol.* **18**, 326 (2000).
- [57] R. Feil and M. F. Fraga, *Epigenetics and the Environment: Emerging Patterns and Implications*, *Nat. Rev. Genet.* **13**, 97 (2012).
- [58] R. Parshani, S. V. Buldyrev, and S. Havlin, *Critical Effect of Dependency Groups on the Function of Networks*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1007 (2011).
- [59] A. Bashan, R. Parshani, and S. Havlin, *Percolation in Networks Composed of Connectivity and Dependency Links*, *Phys. Rev. E* **83**, 051127 (2011).
- [60] P. Trkman, M. I. Štemberger, J. Jaklič, and A. Groznik, *Process Approach to Supply Chain Integration*, *Supply Chain Manage.* **12**, 116 (2007).
- [61] S. Li, S. S. Rao, T. S. Ragu-Nathan, and B. Ragu-Nathan, *Development and Validation of a Measurement Instrument for Studying Supply Chain Management Practices*, *J. Operations Manage.* **23**, 618 (2005).
- [62] P. Trkman and K. McCormack, *Supply Chain Risk in Turbulent Environments—A Conceptual Model for Managing Supply Chain Network Risk*, *Intl. J. Prod. Econ.* **119**, 247 (2009).
- [63] L. Tang, K. Jing, J. He, and H. E. Stanley, *Complex Interdependent Supply Chain Networks: Cascading Failure and Robustness*, *Physica (Amsterdam)* **443A**, 58 (2016).
- [64] A. Punter, *Supply Chain Failures: A Study of the Nature, Causes and Complexity of Supply Chain Disruptions*, www.riskmethods.net/resources/research/supply_chain_failures_2013_final_web.pdf.
- [65] S. Neumayer, G. Zussman, R. Cohen, and E. Modiano, *Assessing the Impact of Geographically Correlated Network Failures*, in *Military Communications Conference* (IEEE, New York, 2008), pp. 1–6.
- [66] P. K. Agarwal, A. Efrat, S. Ganjugunte, D. Hay, S. Sankararaman, and G. Zussman, *The Resilience of WDM Networks to Probabilistic Geographical Failures*, in *Proceedings of INFOCOM* (IEEE, New York, 2011), pp. 1521–1529.
- [67] Y. Berezin, A. Bashan, M. M. Danziger, D. Li, and S. Havlin, *Localized Attacks on Spatially Embedded Networks with Dependencies*, *Sci. Rep.* **5**, 8934 (2015).
- [68] N. Masuda and N. Konno, *Multi-State Epidemic Processes on Complex Networks*, *J. Theor. Biol.* **243**, 64 (2006).
- [69] E. Barter and T. Gross, *Meta-Food-Chains as a Many-Layer Epidemic Process on Networks*, *Phys. Rev. E* **93**, 022303 (2016).
- [70] C. M. Grinstead and L. J. Snell, *Introduction to Probability*, 2006 ed. (American Mathematical Society, Providence, 2006).