Smeared phase transitions in percolation on real complex networks

Laurent Hébert-Dufresne[1,2,*] and Antoine Allard[2,3]

[1]*Department of Computer Science and Vermont Complex Systems Center, University of Vermont, Burlington, Vermont 05405, USA*
[2]*Département de physique, de génie physique et d'optique, Université Laval, Québec G1V 0A6, Canada*
[3]*Centre de modélisation mathématique, Université Laval, Québec G1V 0A6, Canada*

Percolation on complex networks is used both as a model for dynamics *on* networks, such as network robustness or epidemic spreading, and as a benchmark for our models *of* networks, where our ability to predict percolation measures our ability to describe the networks themselves. In many applications, correctly identifying the phase transition of percolation on real-world networks is of critical importance. Unfortunately, this phase transition is obfuscated by the finite size of real systems, making it hard to distinguish finite-size effects from the inaccuracy of a given approach that fails to capture important structural features. Here, we borrow the perspective of smeared phase transitions and argue that observed discrepancies may be due to the complex mesoscopic structure of real networks rather than to finite-size effects only. We support and illustrate this claim by studying real and synthetic networks through the lens of local order parameters, message passing, and local susceptibility. Our results not only shed light on the nature of the percolation transition in complex networks but also provide two important insights on the numerical and analytical tools we use to study them. First, we propose a measure of local susceptibility to better detect both clean and smeared phase transitions by looking at the topological variability of the order parameter. Second, we discuss a shortcoming in state-of-the-art analytical approaches such as message passing, which can detect smeared transitions but not characterize their nature.

I. INTRODUCTION

Percolation on networks is simple to define [1]. Given an original network structure, predict the size distribution of connected components if edges (bond percolation) or nodes (site percolation) are randomly removed such that, on average, only a fraction p remain and are said to be "occupied." Connected components are groups of nodes that are reachable from one another by following occupied edges, and the relative size of the largest component is a natural order parameter for the connectivity of the system. While percolation is simple to define, it is widely used to model complex systems. For example, we can model epidemic spreading with percolation by assuming that occupied edges are contacts that would transmit a disease should one of the two nodes at its ends becomes infected [2]. Because of the breadth and depth of its applications, percolation has become a canonical problem of network science where it reflects the current state of the field: Percolation can be solved on many ordered lattices or random networks, but it is much more complicated on the real, complex networks that exist between order and randomness.

One of the most salient features of percolation is its phase transition in infinite systems, i.e., in the thermodynamic limit

where the number of nodes formally goes to infinity. When p is close to zero, and few edges or nodes are occupied, the system is almost fully disconnected. As p increases, connected components grow. Eventually, at $p = p_c$, we see the emergence of a macroscopic connected component called the giant connected component (GCC) whose size scales linearly with the size of the system. That is, if we double the size of the system by doubling the number of nodes, there exists a component that will also double in size only if $p > p_c$, while the size of all other components will remain virtually independent of system size. Therefore, by using the relative size of the GCC as the order parameter, the percolation threshold p_c marks the transition between two phases: (1) A disconnected phase where connectivity does not scale with system size such that the size of the largest connected component vanishes compared to the size of the system and (2) a connected phase where connectivity scales with system size such that the GCC contains a nonvanishing fraction of all nodes even in an infinite system. In applications, the percolation threshold can, for example, help determine whether a disease can invade a contact network.

Predicting the percolation threshold of a complex network is a highly nontrivial task because it depends on the topology of the network at all scales. In fact, even in direct simulations, *detecting* the phase transition can be complicated. While in theory, there is a clean phase transition where the order parameter goes from zero to a nonzero value at p_c, in practice this transition is masked by noise due to the finite size of real networks. Formally speaking, the definition of the GCC involves taking a thermodynamic limit and following its size as we increase the system size. For real, finite data, doing

*Corresponding author: laurent.hebert-dufresne@uvm.edu

so would require assuming some specific growth mechanism, whose specifics can hardly be inferred from the data itself. Hence, the transition in the connectivity of finite networks does not qualify as a genuine phase transition. That being said, several concepts from phase transition theory are still useful in finite-size networks, and detecting the finite-size analog of a phase transition remains an important problem.

This paper studies our ability to detect and characterize the percolation phase transition in complex networks as follows. In Sec. II, we outline and test existing methods to numerically detect phase transitions using representative real complex networks as case studies. In Sec. III, we interpret the results of the previous section using the perspective of smeared phase transitions. We show how the phase transition may consist in a sum of sequential phase transitions within inhomogeneities such as modules, core-periphery structures, or degree classes, thus differing significantly from the expected behavior of a clean transition. In Sec. IV, we briefly discuss our results on finite-size effects through the lens of the message-passing approach. We also propose a measure of local susceptibility to potentially identify smeared transitions.

II. PHASE TRANSITION DETECTION

To numerically investigate percolation on real networks, we have to adapt the theoretical definition of the order parameter. The relative size of the GCC can be approximated in practice by the ratio S_1/N, where S_1 is the size of the largest connected component (LCC) and N is the number of nodes in the network (i.e., system size). Moreover, to avoid confusing subcritical but large components with a supercritical component, we consider S_1 as nonextensive whenever it is smaller than $0.01N$ [3–6] (other authors use $\sqrt{N}$ with similar results [7]).

We consider three methods to detect the position of the percolation threshold which are based on quantities that are known to peak in homogeneous phase transitions (see Fig. 1): the susceptibility and the size of small, or nonextensive, components. The susceptibility measures the expected response of the system as the fraction p of occupied edges (or nodes) is varied. Since the derivative of the order parameter is discontinuous at the phase transition, the susceptibility diverges. The methods we consider to detect the percolation transition are the following:

(1) Method 1: We denote $S_{1,i}$ as the size of the LCC in the ith realization of the percolation process such that S_1 is the average of $S_{1,i}$ over all simulations i. The susceptibility χ_1 of S_1 can be written as

$$\chi_1 = \frac{\sum_i (S_{1,i} - S_1)^2}{\sum_i S_{1,i}} . \tag{1}$$

As per classic percolation theory, χ_1 peaks, or diverges in the limit of infinite system size, at the phase transition [8,9].

(2) Method 2: Before the phase transition, the expected size of a small component $\langle s \rangle$ in which we might find a random node should increase monotonously with p until it grows very large and forms the GCC. After the phase transition, the largest small components are increasingly absorbed by the GCC such that $\langle s \rangle$ decreases. One can therefore look for a peak in $\langle s \rangle$—excluding the LCC—to identify the phase transi-
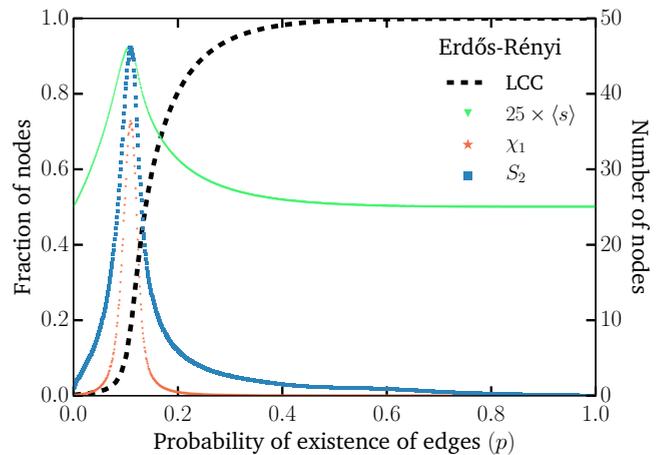
FIG. 1. Detection of the phase transition on a small random network. We study the phase transition, as we vary the probability of existence of edges p, in the relative size of the giant connected component (GCC) as approximated by the largest connected component (LCC) when greater than 1% of the network size. We also follow three different metrics—the susceptibility χ_1 of the LCC, the expected size of small components $\langle s \rangle$ (enlarged by a factor 25), and the size of the second-largest connected component S_2—which should all peak at the phase transition. We use a single realization of an Erdős-Rényi (ER) random network of size 1000 with an average degree of 10. The three methods all peak together to suggest a transition very close to the theoretical value of $p_c = 1/10$. The effects of the finite size of the network are reflected in the width of the peaks.

tion [10,11]. Because $\langle s \rangle$ is calculated by selecting a random node and because a random node is s times more likely to be in a component of size s than a component of size 1, it is proportional to the second moment of the true component size distribution P_s, that is, $\langle s \rangle = \sum_s s^2 P_s / \sum_{s'} s' P_{s'}$. It is therefore equivalent to a classic definition of susceptibility [12].

(3) Method 3: The third method is based on the previous one but only looks at the size S_2 of the second-largest connected components (SLCC) [13]. It typically leads to a refined and more evident peak because of its greater amplitude (i.e., $S_2 \geqslant \langle s \rangle$ for $p > p_c$) and because of its sharper decrease since the SLCC is quickly absorbed by the LCC.

Contrary to Fig. 1, Fig. 2 shows that these three methods may not identify the same clean phase transition when applied to real network datasets and can behave in unexpected ways. First and foremost, neither do they always agree nor do they always rank consistently or provide bounds on what we would consider the actual phase transition, as suggested in the supplemental material of Ref. [18]. Second, the peak can be very wide, with the width at half-amplitude spanning over 10% of the parameter space in some cases. Third, and most surprisingly, method 2 does not always peak, and some methods can peak more than once.

III. SMEARED PHASE TRANSITIONS

The first two problems observed at the end of the previous section are related—the peaks of the observables do not align
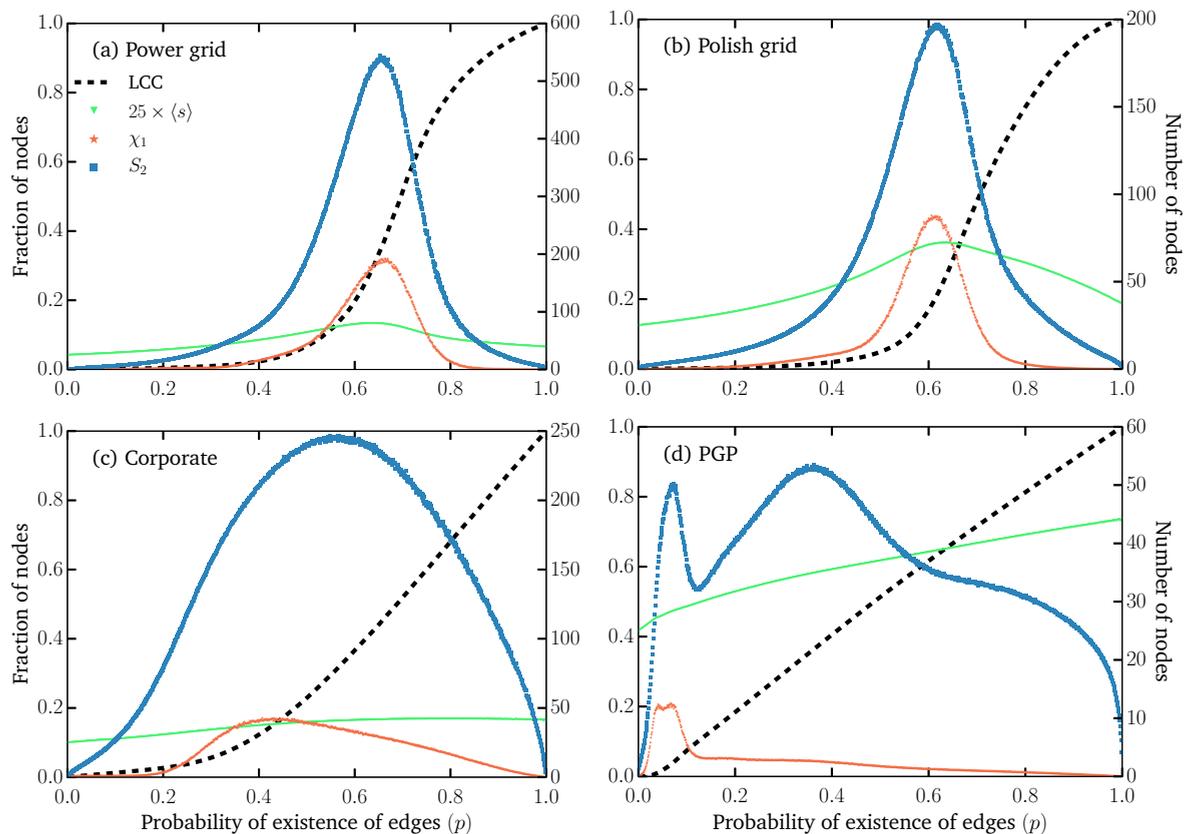
FIG. 2. Detection of the phase transition on real complex networks. We study the phase transition as we did in Fig. 1 but now considering four real networks: an American power grid [14], a Polish power grid [15], a social network among the boards of directors of public Norwegian companies [16], and the web of trust of the pretty good privacy (PGP) encryption algorithm [17]. These networks were chosen to highlight the common problems with standard methods to detect the phase transition in real complex systems: (i) They do not necessarily agree, (ii) they can peak once the order parameter is already very large, and (iii) they can peak more than once.

and are not sharp—and correspond to what we would expect given strong finite-size effects. Indeed, the finite size of real systems tends to smooth out phase transitions of all nature. In the case of percolation, this happens because large but nonextensive components are hard to distinguish from a GCC if we cannot change the system size to investigate the scaling relations. Just below the threshold, small components larger than our LCC criteria (1% of system size) exist such that a nonzero order parameter below p_c is numerically observed. These effects are inherent to the use of the phase transition framework to real finite-size systems, which in theory only applies in the infinite-size limit. While there are methods available to account for some finite-size effects [19] and other rare fluctuations [20], the basic phenomenology of the transition, averaged over all possible realizations, remains the same. Finite-size effects therefore fall short of explaining the enormous width of the susceptibility peaks in Fig. 2 and cannot explain why they would peak more than once.

A. Empirical results

Another possible explanation is that we are dealing with smeared phase transitions. One classic example is that of the Ising model in systems with defects [21]. The Ising model considers spins, which can take a $+1$ or -1 value, laying on

the nodes of a regular lattice. At high temperature, the spins are independent of each other and free to take either value such that the global (or average) magnetization of the system is zero; this is the disordered or paramagnetic phase. As temperature decreases, the interactions force correlations and the system eventually enters a correlated state with nonzero global magnetization; this is the ordered or ferromagnetic phase. Because of the regular structure of lattices and because all interactions are equal, the system is highly homogeneous. This homogeneity leads to a clean phase transition: There is no spatial variation in thermodynamic observables and all mesoscopic domains undergo an identical phase transition. This also means that in the thermodynamic limit, we see a vanishing width or variance in the distribution of the order parameter across domains. This property is called self-averaging.

One of the most powerful aspects of phase transitions is their resilience to microscopic details in the structure or rules of our models. For example, we can introduce significant defects in the lattice on which the Ising model occurs without destroying the sharpness of its phase transitions. Defects such as weakening/strengthening or removing/adding edges do not affect the phenomenology of the model as long as they are not strongly correlated (e.g., random micro- or mesoscopic noise in space).
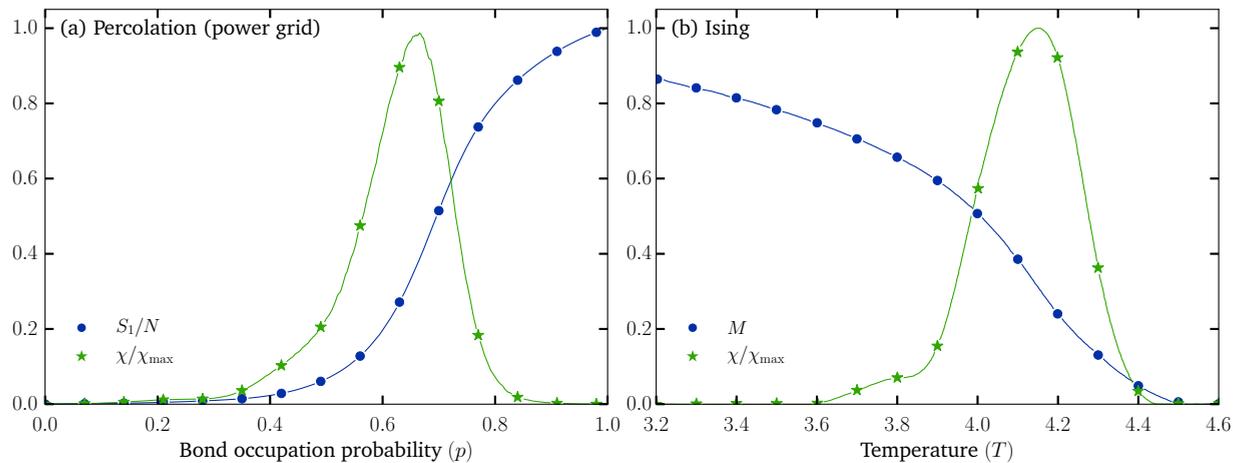
FIG. 3. Smeared phase transitions in two different models. (a) Simulations of percolation on the American power grid of size $N = 4941$ of Fig. 2 where the relative size of the LCC is again followed as the external parameter p (bond occupation probability) is varied. (b) Simulations of a three-dimensional Ising model in a cubic lattice of size $N = 20\,000$ with planar defects where the average magnetization is followed as the external parameter T (temperature) is varied, reproduced from Ref. [21]. In both cases, we attempt to detect the phase transition using the susceptibility of the system. Unfortunately, in both cases, susceptibility peaks when the order parameter is roughly equal to a third of its maximal value.

However, the observed phase transition can change drastically when strongly correlated defects are introduced. A classic example is that of the Ising model in a three-dimensional cubic lattice with planar defects creating weaker bonds. The phase transition observed in that model is compared to percolation on a power grid on Fig. 3, and the phenomenological similarity is striking. This smearing is related, at least physically, to the Griffiths phenomenon [22] but is different through the long-range order established by correlated defects [21].

If we accept that our inability to accurately locate the percolation phase transition in real complex networks is not solely due to finite-size effects but potentially also to correlated defects, the question becomes the following: What is the source of this disorder? First, since complex networks are not regular systems like lattices, "disorder" is the norm and we instead look for correlated *inhomogeneities*. Second, we can detect these inhomogeneities by using the definition of the smeared phase transition. We thus look for subsets of nodes around which the local order parameter deviates from the global order parameter. More specifically, we wish to identify sets of nodes $\{i\}$ such that the probability $P_i(p)$ that node i is in the LCC under occupation probability p deviates from $S_1(p)/N$ in a way that cannot be explained by degree inhomogeneities [23].

In Fig. 4, we show the curves $P_i(p)$ for all nodes i in a homogeneous random network and in a Polish power grid network and compare them to the corresponding average $\langle P_i(p) \rangle = S_1(p)/N$. The main conclusion that can be drawn from this experiment is that we can expect significant variance in the distribution of the order parameter in both random and real complex networks, but significantly more in the latter. Without an idea of how the variance would scale with system size, it is impossible to rule out the possibility of a finite-size effect, but it does seem to be more in line with the smeared phase transition interpretation. We can investigate the nature

of these inhomogeneities by averaging $P_i(p)$ over subset of nodes i who share a same given structural property.

In Fig. 5, the curves $P_i(p)$ are compared to the averages obtained over different centrality classes: Degree centrality given by the number of edges on a given node, k-core centrality given by the largest k such that a node is in the maximal subset of nodes with degree at least k among one another [24], and the onion decomposition which assigns a layer centrality to every node based on when they are removed in the k-core decomposition [25]. Degree heterogeneity is sufficient to capture all the inhomogeneities observed in random networks, as shown in Ref. [23], but not in the power grid. However, using the onion decomposition, we now capture much more directly the fact that the system is divided into two (or more) subsystems composed of nodes with very different centrality, not necessarily related to degree but to their position in the mesoscale organization of the network structure.

B. Synthetic results

There are several mesoscopic features that are known to lead to exotic behaviors in the phase transitions of complex networks: modular structure where nodes are grouped in densely connected modules with sparse connectivity across modules [8,26–28], core-periphery structure where a subset of nodes form a densely connected nodes linked to a sparsely interconnected periphery [8,11], and strong heterogeneity in the degree distribution where star nodes drive a weak initial phase transition followed by sequential activation of different degree classes [29–31].

To highlight the role of these mesoscopic features on the percolation phase transition, we study the set of probabilities $P_i(p)$ introduced in the previous section within two toy networks. Both are based on the traditional Erdős-Rényi (ER) random network considered in Fig. 4: A simple network of
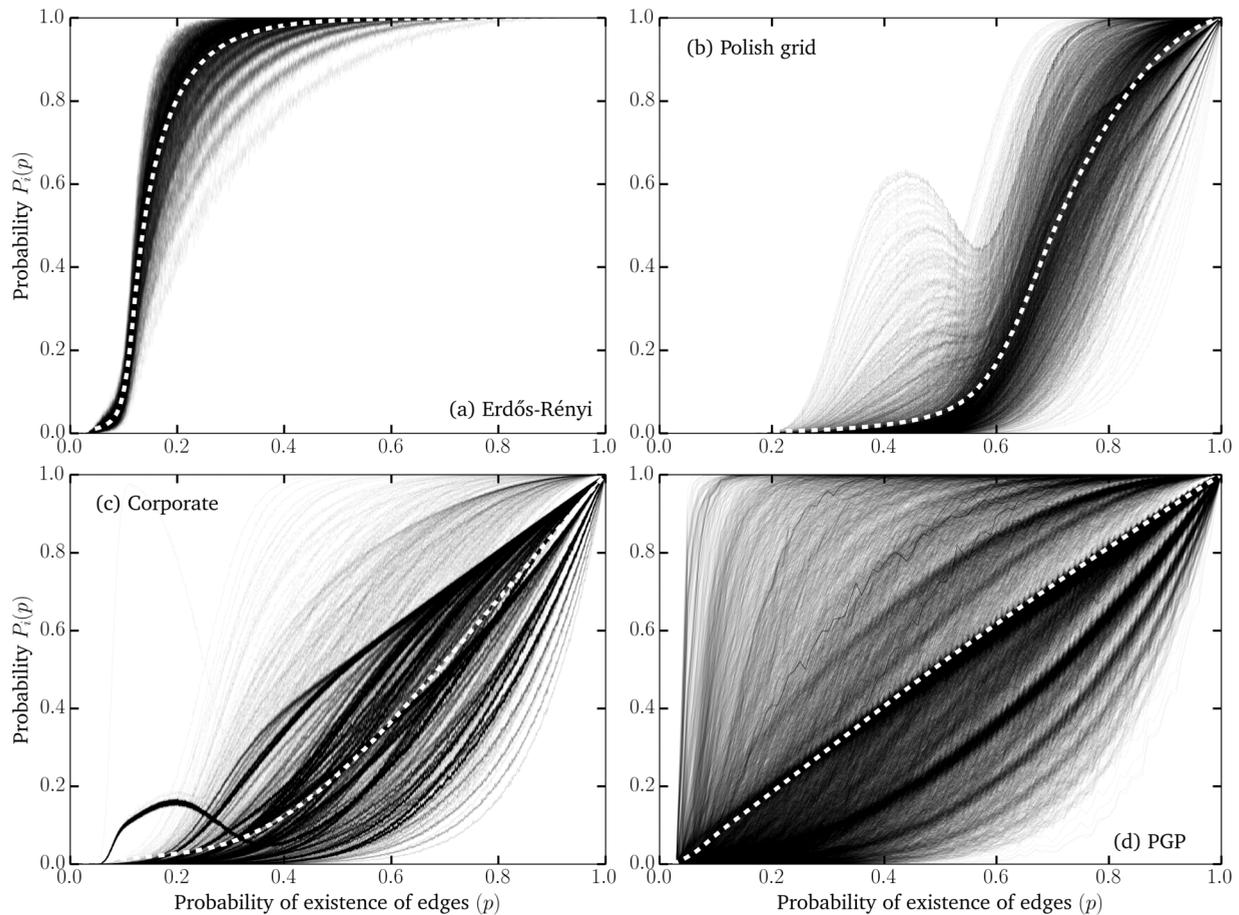
FIG. 4. Local order parameters. Results of percolation simulations on (a) an Erdős-Rényi (ER) random network of size 1000 with an average degree of 10, (b) the Polish power grid network, (c) the corporate network, and (d) the PGP encryption network. We show the global order parameter (i.e., the relative size of the LCC) in white, while the underlying gray curves show the probability $P_i(p)$ that every individual node i is found in the LCC. In panel (a), despite the small size of the network, we find a relatively clean phase transition, meaning that the global order parameter accurately describes the behavior observed around every individual nodes. Importantly, all curves of the local order parameter in panel (a) are the steepest at the same point (i.e., the maxima of $dP_i(p)/dp$ correspond to the phase transition) and well approximated by the global order parameter. In panels (b)–(d), we find that while the global order parameter represents the average of all $P_i(p)$ (by definition), it is not representative of the behavior of every $P_i(p)$.

N nodes where every unique pair of nodes is independently connected with probability ρ. Our first toy model produces a core-periphery structure using two nested ER networks: An inner smaller network of size N_1 with density ρ_1 and a larger outer network of size $N_2 > N_1$, where nodes are connected among themselves and to the inner network with density $\rho_2 < \rho_1$. The second toy network is produced by a set of independent ER networks of different size and/or density which are then connected by a single edge; this results in a network with a strong modular structure. The top row of Fig. 6 provides a typical realization of these two toy models.

The bottom row of Fig. 6 shows the $P_i(p)$ curves obtained with the two toy models alongside the average value over all nodes. In the case of the core-periphery structure, the percolation process nucleates in the core at an occupation probability p much lower than the known threshold $1/3$ of the periphery. Peripheral nodes are still "activated" below their threshold due to subcritical spillover from the core into the periphery. Eventually, the periphery activates and creates a sudden rise

in the $P_i(p)$ curves of peripheral nodes. Importantly, we find that all curves increase monotonously as p increases, with the LCC spreading progressively from the core outward.

In the modular networks, the percolation process again nucleates in the denser modules, but eventually the $P_i(p)$ curves corresponding to nodes in denser modules decrease. This is due to the weak coupling between modules: there exist a regime where large connected components exist in each module but are unlikely to connect and are therefore in competition for the title of *largest* connected component. Hence, whether there is a subset of nonmonotonous $P_i(p)$ curves can be used to distinguish core-periphery and modular structures. For example, the set of $P_i(p)$ curves observed with the modular networks are strongly reminiscent of the results on the Polish power grid which, in turn, suggests a high modularity in its inner cores.

We further test the robustness of these inhomogeneities in the distribution of local order parameter $P_i(p)$ by measuring their standard deviation as we increase system size.
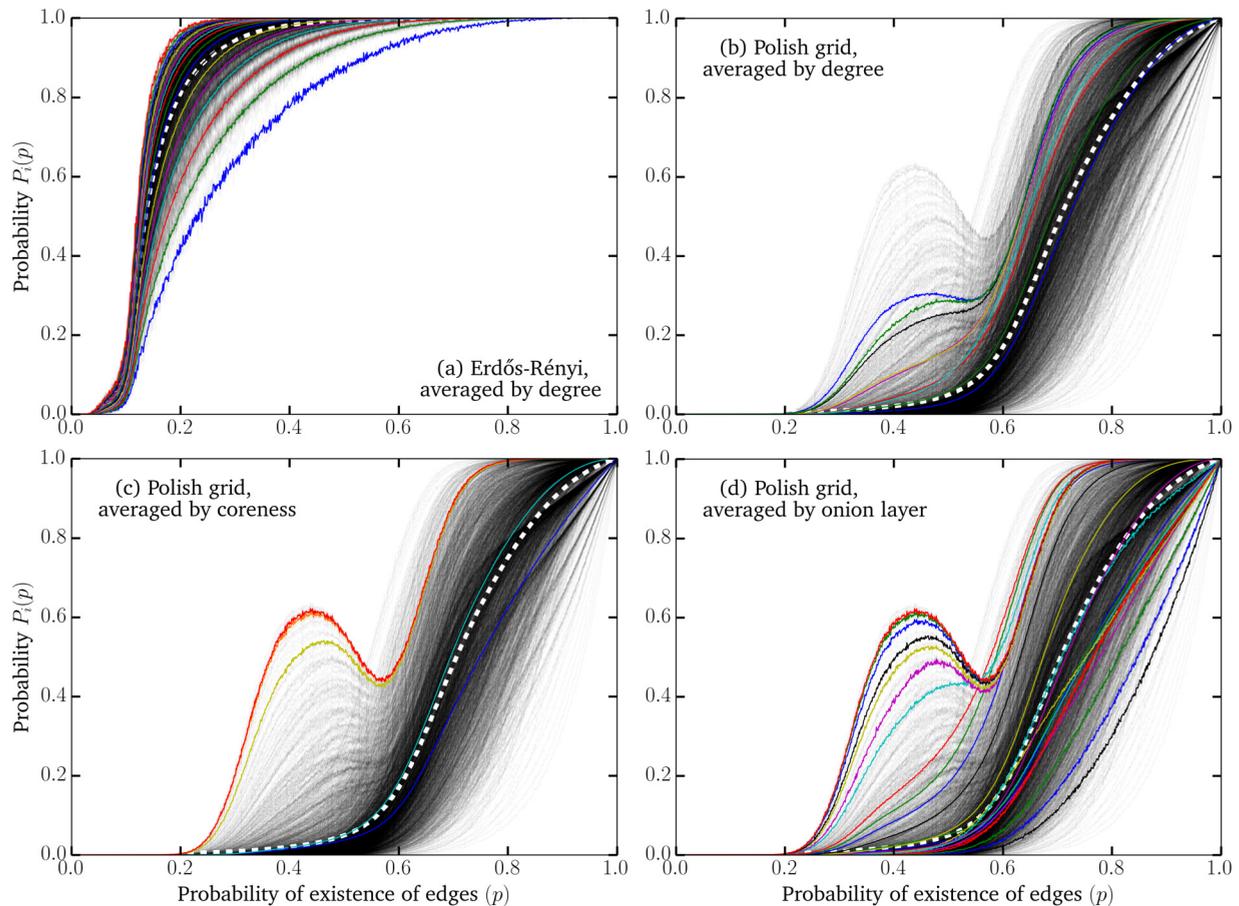
FIG. 5. Local order parameters combined according to different centrality classes. [(a), (b)] We compare the curves shown in Figs. 4(a) and 4(b) to averages (additional colored curves) based on the degree of nodes. The new curves show the average probability that a particular node is found in the LCC based on its degree (i.e., from the bottom up, we plot averages over degrees 1, 2, 3, and so on). [(c), (d)] For the Polish grid only, we show color curves now representing the average of $P_i(p)$ over nodes i belonging to given (c) k shells of the k-core decomposition and (d) layers of the onion decomposition. These more complicated centrality metrics capture progressively more topological diversity, especially around the most and least central nodes.

The results are shown in Fig. 7. There are different ways to increase the size of these random networks, and we investigate three options. (i) We increase the size of all subnetworks at the same time. (ii) We only increase the size of the largest subnetwork. (iii) We increase the size of the largest subnetwork, and add more subnetworks of a fixed smaller size. Doing so we find that the smeared transition is only preserved when all subnetworks are scaled simultaneously. Interestingly, adding more inhomogeneities (smaller denser modules) of fixed size collapses the transition even faster than not scaling the inhomogeneities at all. Again, this is due to increased competition between inhomogeneities which might all have large connected components competing for the title of the LCC. Importantly, this result stresses the need for *correlated* inhomogeneities to produce smeared phase transitions.

IV. DISCUSSION

A. The thermodynamic limit and message passing approaches to percolation

To sum up the results obtained so far, one can characterize a phase transition by looking at the set of probabilities $P_i(p)$

of finding node i in the LCC at occupation probability p. The curves $P_i(p)$ can help identify the mesoscopic organization of the structure at hand. Using two simple toy models, we showed that core periphery leads to monotonously increasing $P_i(p)$ where the core activates first and gradually invades the periphery. Modular structure, on the other hand, can lead to nonmonotonous $P_i(p)$ where the LCC appears in a dense module first, and eventually jumps to a sparser but larger second module. In other words, the strength of the coupling between cores or modules distinguishes these different results.

To better understand the role of coupling between substructures, we follow the language of Sknepnek and Volta and distinguish three types of mesoscopic network inhomogeneities [21]. The first one, called vanishing randomness, corresponds to where the distribution of local thermodynamic observables collapses in the thermodynamic limit, meaning that the system is self-averaging and does not produce true smeared phase transitions. The second and third types of network inhomogeneities are finite randomness and infinite randomness, meaning that the distribution of local thermodynamic observables are *not* self-averaging but instead lead to finite or infinite width in the distributions of local order
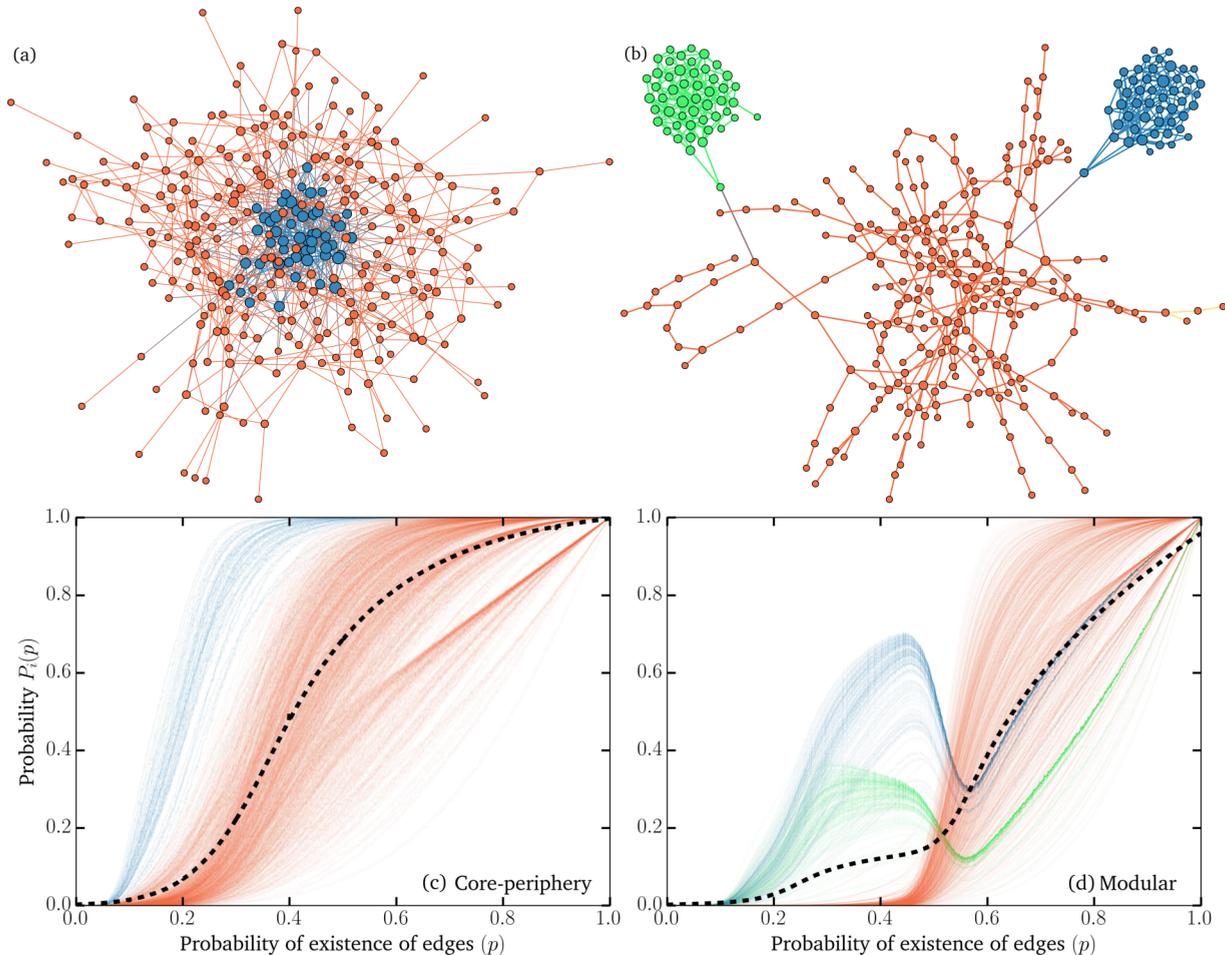
FIG. 6. Local order parameters in two toy networks with core-periphery and modular structures. (a) Core periphery: The core is an Erdős-Rényi (ER) random network of N_1 nodes with density ρ_1 and the periphery is an ER random network whose $N_2 > N_1$ nodes are connected to each other and to the core with density $\rho_2 < \rho_1$. (b) Modular structure: There are three distinct ER random networks of sizes N_1, N_2, and N_3 with densities ρ_1, ρ_2, and ρ_3 respectively. Modules of size N_1 and N_2 are both connected to N_3 with a single random edge. The curves $P_i(p)$ (c) in a core-periphery produced with $N_1 = 51$ and $N_2 = 1001$ with densities $\rho_1 = 0.15$ and $\rho_2 = 0.003$ and (d) in a modular structure with $N_1 = N_2 = 201$ and $N_3 = 1101$ with densities $\rho_1 = \rho_2 = 0.025$ and $\rho_3 = 0.002$. Each thin curve correspond to the local order parameter around a single node, and the color of the curve indicates the module in which that node can be found in the network above. The black dashed line corresponds to the average of the local order parameters, as in Fig. 5.

parameters in the thermodynamic limit. Vanishing and finite randomness both appear in Fig. 7 depending on how the network grows when taking the thermodynamic limit.

Since this categorization depends on how we take the thermodynamic limit of a system, there are no way to apply our intuition to real networks. Indeed, there are no way to know whether the dense cores observed in power grids should scale with the size of the system or not. For example, if the power grid was twice as large, would we find a core of similar size, twice as many cores, or a single core twice as large? Distinguishing important mesoscale structures from finite-size effects is therefore context dependent and must be done on a case-by-case basis. What we can say, however, is that many of current mathematical approaches are making that decision for us.

Current state-of-the-art analytical approaches to percolation are based on the message-passing approximation (MPA)

[32]. The MPA takes the entire network structure as an input but then ignores loops when solving the percolation process. Because of this approximation, the MPA is effectively solving percolation not on the true network but on an infinite network where there exist an infinite number of copies of every node [33,34]. In practice, this means that any modular structure is mapped to a core-periphery structure. For example, when using the MPA on a network with two modules of size N with uniform degrees k_1 and $k_2 < k_1$ and connected by a single edge, we are solving percolation on a core-periphery structure where the core is a k_1-core, and the periphery a k_2-core, interconnected by a bridge between a fraction $1/N$ of the nodes (which, in this case, nevertheless corresponds to an infinite number of bridges).

In practice, this means that any core-periphery structure will be captured by the MPA, but that modular structure will be mapped to an equivalent core-periphery structure [33].
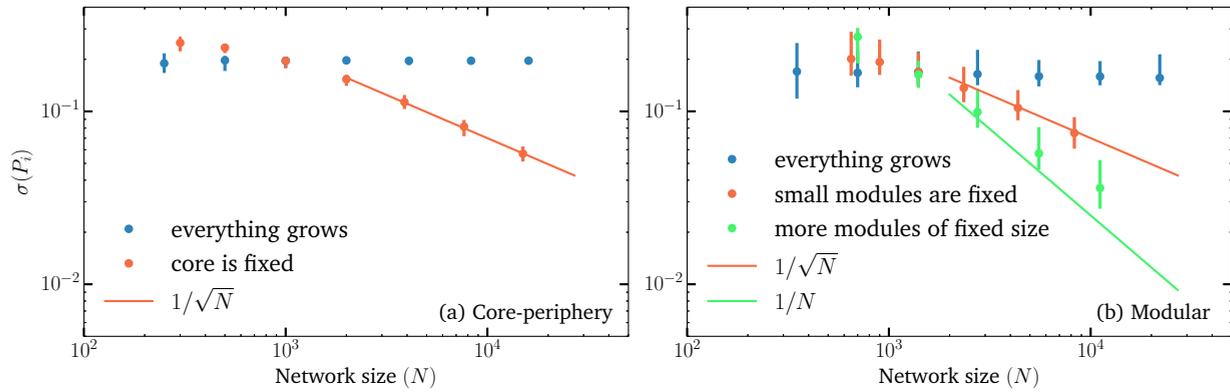
FIG. 7. Vanishing and finite randomness in two toy networks with core-periphery and modular structures. We use the core-periphery (a) and modular (b) structures presented in Fig. 6. We follow the standard deviation of $P_i(p)$ at $p = 0.3$ when the network size is increased while keeping the average degree of subsystems fixed by varying their densities. In the core periphery, growing a network means either that the size of both the core and the periphery grow together (option 1, in blue) or that only the periphery grows and the size of the core remains fixed (option 2, in orange). In the modular structure, growing a network means, again, growing all modules together (option 1, in blue), growing only the largest module with fixed number and size for small modules (option 2, in orange), or growing the largest module while adding smaller modules of fixed size (option 3, in green). The markers correspond to the average and error bars to the full range of observed values. Lines showing theoretical bounds of $N^{-1/2}$ and N^{-1} have been added to guide the eye. The $N^{-1/2}$ behavior corresponds to typical vanishing finite size effects. The faster decay in N^{-1} corresponds to competing finite-size effects; for example, adding small modules makes it less likely for any given module to contain the LCC.

For the local order parameter, this implies that the MPA will always predict monotonously increasing $P_i(p)$ curves. Consequently, while it is tempting to dismiss dense but small substructures, it is important to know that the MPA, and other analytical approaches that account for node centrality [6,35], will capture smeared phase transitions [23] but not necessarily their nature. This result is illustrated in Fig. 8(a).

B. Local susceptibility

Finally, we propose a measure of local susceptibility that extracts some of the important information contained in the set of $P_i(p)$ curves. We want this local susceptibility to be a single curve describing the response function of how the spatial distribution of the local order parameter changes following a small variation in occupation probability p. Since many linear combinations of the $P_i(p)$ curves, or their derivatives, can be written as functions of the global order parameter, $S = \sum_i P_i(p)/N$, averages over $P_i(p)$ are unlikely to capture small localized responses. We instead look at a second order property—the standard deviation $\sigma(P_i(p))$—which captures the heterogeneity of the $P_i(p)$ curves.

Local maxima of $\sigma(P_i(p))$ correspond to points where some substructures have supercritical behavior while others may not. However, it would not capture the initial percolation transition where all $P_i(p)$ are still close to zero. Local maxima of the first derivative in p of $\sigma(P_i(p))$ correspond to points where the spatial heterogeneity of the order parameter changes the most with varying p. This measure will therefore detect sequential phase transitions but will typically peak *after* the transitions. Local maxima of the second derivative in p of $\sigma(P_i(p))$ correspond to points of maximum curvature where the response of the order parameter changes rapidly with varying p. These can be caused by both regular phase transitions or sequential transitions in different subsystems.

We thus propose

$$\chi_{\text{local}} = \frac{d^2}{dp^2}\sigma(P_i(p)) , \qquad (2)$$

where the derivative is to be performed numerically. Figure 8 shows the local susceptibility and compares it with the global and local order parameters of the Polish power grid, the network of corporate board of directors, and the PGP encryption network. As desired, the maxima of local susceptibility accurately capture the emergence of the LCC as well as secondary transitions that produce inflection points in the growth of the LCC.

C. Conclusion

We observed that percolation transition in real complex networks can be often described through the lens of smeared phase transitions. We illustrated with a few case studies that the nature of the inhomogeneities can be studied by looking at the local order parameters and then by determining whether subdivisions of nodes by degrees, modules, or centrality classes best explain the observed variations. We showed that modular and core-periphery structure can both be responsible for the observed smeared phase transitions and we discussed the qualitative difference between the two types of mesoscopic structure. Importantly, looking at the topological distribution of the order parameter through the set of $P_i(p)$ curves can help identify the nature (or cause) of the smeared phase transition. While in theory these results might all be due to the finite size of real networks, they are all captured by the state-of-the-art analytical approaches to percolation on complex networks. It is therefore an important feature to consider when comparing how well different analytical models predict the percolation transition. Similarly, it is critical to investigate the potential for smeared transitions in real systems before applying results from clean phase transition to percolation-
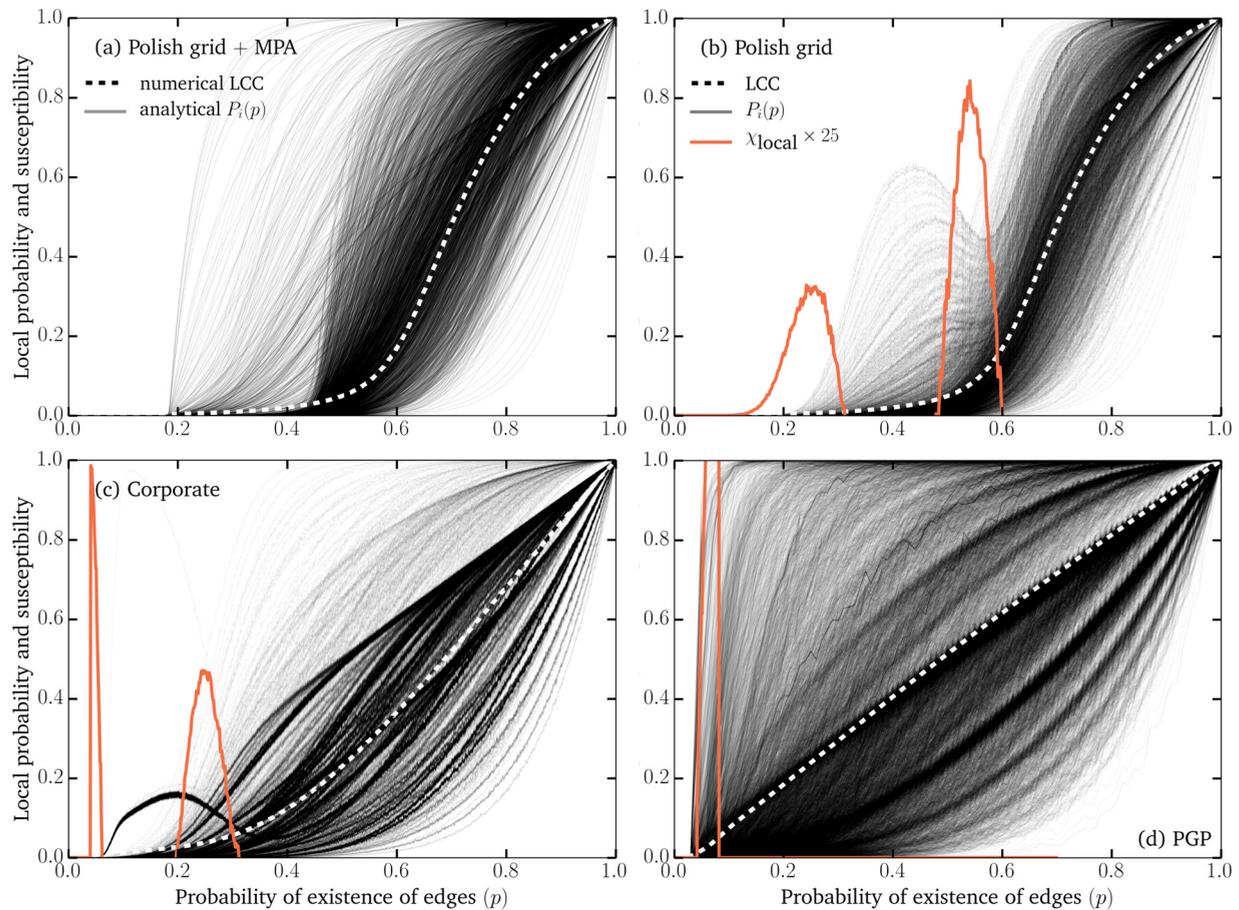
FIG. 8. Using message passing and local susceptibility to detect smeared phase transitions. (a) Solving the local order parameters in a system of message passing equations defined over the structure of the Polish power grid allows us to detect the smeared phase transition found in the simulated LCC. As discussed in Sec. IV A, because of the treelike approximation used in message passing, all inhomogeneities are mapped to an effective core-periphery structure and message passing therefore always leads to monotonous local order parameters that do not capture the cause or nature of the smeared transition. Here, message passing correctly detects the two transitions but would lead us to conclude that they stem from a core-periphery structure as it misses the nonmonotonic regime of competition between modules. The equations used to obtain $P_i(p)$ as predicted by the MPA are explained in the Appendix. [(b)–(d)] The results of Eq. (2) in orange are compared to both the global and local order parameters (i.e., S_1/N in white and $P_i(p)$ for all nodes i in gray). Not only is the local susceptibility able to capture the first transition, unlike traditional susceptibility shown in Fig. 2, but it also better detects the second transition around the point of maximum curvature in the global order parameter for the Polish power grid in panel (b) and for the network of corporate boards of directors in panel (c). The PGP encryption network in panel (d) does not appear to have a second transition.

like models of system resilience or disease spread. For that purpose, we developed a measure of local susceptibility which can help identify smeared or sequential phase transitions at a glance.

APPENDIX: THE MESSAGE-PASSING APPROXIMATION

The message-passing approximation (MPA) relies on the assumption that the structure of networks does not contain short loops—that it is locally treelike—to predict the outcome of percolation (see Ref. [32] for a detailed derivation of the equations). It defines u_{ij} as the probability that following the edge from node i to node j does not lead to the giant connected component. This situation occurs if the edge has been removed (probability $1 - p$) or if the node at the other end of the edge does not itself lead to the giant connected component via its other neighbors. In other words, the $\{u_{ij}\}$ are the solution of the self-consistency equation

$$u_{ij} = (1 - p) + p \prod_{l \in \mathcal{N}_j \backslash i} u_{jl} , \qquad (A1)$$

where $i, j = 1, \ldots, N$ and where $\mathcal{N}_j \backslash i$ corresponds to the neighbors of node j excluding node i.

Having solved Eq. (A1) for both directions of every edges (i.e., for each ordered pair ij for which an edge exists), the probability that a node i is part of the giant connected component is simply the probability $P_i(p)$ that at least one of

its neighbors leads to it:

$$P_i(p) = 1 - \prod_{j \in \mathcal{N}_i} u_{ij} . \qquad (A2)$$

[1] M. E. J. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford, UK, 2018), p. 800.

[2] M. E. J. Newman, Spread of epidemic disease on networks, Phys. Rev. E **66**, 016128 (2002).

[3] A. Allard, L. Hébert-Dufresne, P.-A. Noël, V. Marceau, and L. J. Dubé, Bond percolation on a class of correlated and clustered random graphs, J. Phys. A **45**, 405005 (2012).

[4] A. Allard, L. Hébert-Dufresne, J.-G. Young, and L. J. Dubé, General and exact approach to percolation on random graphs, Phys. Rev. E **92**, 062807 (2015).

[5] L. Zdeborová, P. Zhang, and H.-J. Zhou, Fast and simple decycling and dismantling of networks, Sci. Rep. **6**, 37954 (2016).

[6] A. Allard and L. Hébert-Dufresne, Percolation and the Effective Structure of Complex Networks, Phys. Rev. X **9**, 011023 (2019).

[7] F. Coghi, F. Radicchi, and G. Bianconi, Controlling the uncertain response of real multiplex networks to random damage, Phys. Rev. E **98**, 062317 (2018).

[8] P. Colomer-de-Simón and M. Boguñá, Double Percolation Phase Transition in Clustered Complex Networks, Phys. Rev. X **4**, 041020 (2014).

[9] F. Radicchi, Predicting percolation thresholds in networks, Phys. Rev. E **91**, 010801(R) (2015).

[10] 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).

[11] A. Allard, B. M. Althouse, S. V. Scarpino, and L. Hébert-Dufresne, Asymmetric percolation drives a double transition in sexual contact networks, Proc. Natl. Acad. Sci. USA **114**, 8969 (2017).

[12] K. Christensen and N. R. Moloney, *Complexity and Criticality* (World Scientific, Singapore, 2005).

[13] P. Zhang, Spectral estimation of the percolation transition in clustered networks, Phys. Rev. E **96**, 042303 (2017).

[14] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, Nature (London) **393**, 440 (1998).

[15] R. D. Zimmerman, C. E. Murillo-sánchez, and R. J. Thomas, MATPOWER: Steady-state operations, systems research, and education, IEEE Trans. Power Syst. **26**, 12 (2011).

[16] C. Seierstad and T. Opsahl, For the few not the many? The effects of affirmative action on presence, prominence, and social capital of women directors in Norway, Scand. J. Manag. **27**, 44 (2011).

[17] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, Models of social networks based on social distance attachment, Phys. Rev. E **70**, 056122 (2004).

[18] F. Radicchi and C. Castellano, Beyond the locally treelike approximation for percolation on real networks, Phys. Rev. E **93**, 030302(R) (2016).

[19] G. Bianconi, Rare events and discontinuous percolation transitions, Phys. Rev. E **97**, 022314 (2018).

[20] G. Bianconi, Fluctuations in percolation of sparse complex networks, Phys. Rev. E **96**, 012302 (2017).

[21] R. Sknepnek and T. Vojta, Smeared phase transition in a three-dimensional Ising model with planar defects: Monte Carlo simulations, Phys. Rev. B **69**, 174410 (2004).

[22] M. A. Muñoz, R. Juhász, C. Castellano, and G. Ódor, Griffiths Phases on Complex Networks, Phys. Rev. Lett. **105**, 128701 (2010).

[23] R. Kühn and T. Rogers, Heterogeneous micro-structure of percolation in sparse networks, EPL **118**, 68003 (2017).

[24] V. Batagelj and M. Zaveršnik, Fast algorithms for determining (generalized) core groups in social networks, Adv. Data Anal. Classification **5**, 129 (2011).

[25] L. Hébert-Dufresne, J. A. Grochow, and A. Allard, Multiscale structure and topological anomaly detection via a new network statistic: The onion decomposition, Sci. Rep. **6**, 31708 (2016).

[26] J. P. Gleeson, Cascades on correlated and modular random networks, Phys. Rev. E **77**, 046117 (2008).

[27] L. K. Gallos, H. A. Makse, and M. Sigman, 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 (2012).

[28] U. Bhat, M. Shrestha, and L. Hébert-Dufresne, Exotic phase transitions of k-cores in clustered networks, Phys. Rev. E **95**, 012314 (2017).

[29] D.-S. Lee, K.-I. Goh, B. Kahng, and D. Kim, Evolution of scale-free random graphs: Potts model formulation, Nucl. Phys. B **696**, 351 (2004).

[30] A. V. Goltsev, S. N. Dorogovtsev, J. G. Oliveira, and J. F. F. Mendes, Localization and Spreading of Diseases in Complex Networks, Phys. Rev. Lett. **109**, 128702 (2012).

[31] G. St-Onge, J.-G. Young, E. Laurence, C. Murphy, and L. J. Dubé, Phase transition of the susceptible-infected-susceptible dynamics on time-varying configuration model networks, Phys. Rev. E **97**, 022305 (2018).

[32] B. Karrer, M. E. J. Newman, and L. Zdeborová, Percolation on Sparse Networks, Phys. Rev. Lett. **113**, 208702 (2014).

[33] A. Allard and L. Hébert-Dufresne, On the accuracy of message-passing approaches to percolation in complex networks, arXiv:1906.10377.

[34] A. Faqeeh, S. Melnik, and J. P. Gleeson, Network cloning unfolds the effect of clustering on dynamical processes, Phys. Rev. E **91**, 052807 (2015).

[35] L. Hébert-Dufresne, A. Allard, J.-G. Young, and L. J. Dubé, Percolation on random networks with arbitrary k-core structure, Phys. Rev. E **88**, 062820 (2013).