Indirect influence in social networks as an induced percolation phenomenon

Jiarong Xie[a,1] , Xiangrong Wang[b,c,1] , Ling Feng[d,e] , Jin-Hua Zhao[f,g] , Wenyuan Liu[h], Yamir Moreno[i,j,k] , and Yanqing Hu[l,2]

[a]School of Computer Science and Engineering, Sun Yat-sen University, 510006 Guangzhou, China; [b]Institute of Future Networks, Southern University of Science and Technology, 518055 Shenzhen, China; [c]Peng Cheng Laboratory, 518066 Shenzhen, China; [d]Institute of High Performance Computing, A*STAR, 138632 Singapore; [e]Department of Physics, National University of Singapore, 117551 Singapore; [f]Guangdong Provincial Key Laboratory of Nuclear Science, Institute of Quantum Matter, South China Normal University, 510006 Guangzhou, China; [g]Guangdong-Hong Kong Joint Laboratory of Quantum Matter, Southern Nuclear Science Computing Center, South China Normal University, 510006 Guangzhou, China; [h]Division of Physics and Applied Physics, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore; [i]Institute for Biocomputation and Physics of Complex Systems, University of Zaragoza, 50018 Zaragoza, Spain; [j]Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain; [k]ISI Foundation, 10126 Torino, Italy; and [l]Department of Statistics and Data Science, College of Science, Southern University of Science and Technology, 518055 Shenzhen, China

Percolation theory has been widely used to study phase transitions in network systems. It has also successfully explained various macroscopic spreading phenomena across different fields. Yet, the theoretical frameworks have been focusing on direct interactions among nodes, while recent empirical observations have shown that indirect interactions are common in many network systems like social and ecological networks, among others. By investigating the detailed mechanism of both direct and indirect influence on scientific collaboration networks, here we show that indirect influence can play the dominant role in behavioral influence. To address the lack of theoretical understanding of such indirect influence on the macroscopic behavior of the system, we propose a percolation mechanism of indirect interactions called induced percolation. Surprisingly, our model exhibits a unique anisotropy property. Specifically, directed networks show first-order abrupt transitions as opposed to the second-order continuous transition in the same network structure but with undirected links. A mix of directed and undirected links leads to rich hybrid phase transitions. Furthermore, a unique feature of the nonmonotonic pattern is observed in network connectivities near the critical point. We also present an analytical framework to characterize the proposed induced percolation, paving the way to further understanding network dynamics with indirect interactions.

percolation | indirect interactions | social network | phase transition | behavioral contagion

Percolation theory (1) is one of the most prominent frameworks within statistical physics. Initially developed (2, 3) to explain the chemical formation of large macromolecules, it has been recently used to study various dynamical processes in complex networks (4–9). Examples include the use of bond percolation (9, 10) to study the wide spread of rumors over online social media and outbreaks of infectious diseases on structured populations. Site percolation (4, 5, 11) has been employed to study the cascading failures of infrastructure networks (6, 12–16) and the resilience of protein–protein interaction networks (17). Likewise, bootstrap percolation (18), k-core (19–21), and linear threshold percolation (7, 22–24) have enabled the study of the spreading of behaviors over social networks. Finally, the so-called explosive percolation (25) has allowed a better characterization of systems' structural transitions when they are growing or can adapt, whereas core percolation (26, 27) has contributed significantly to insights into nondeterministic polynomial problems. Common to all these percolation models is that they have successfully described various important dynamical phenomena by considering different direct interactions (8, 9, 28) among network nodes; in particular, they have captured the behavior of network systems as given by phase transitions (4, 8, 9, 28, 29).

Our study is motivated by recent evidence that there are many systems in which indirect interactions play a major role in their spreading dynamics (30–35). Such underlying indirect interactions have important implications not only on the dynamics of the system but also on the evolution and the emergence of network structures. For example, Christakis and Fowler (30, 31) found that for the spreading of many social behaviors, such as drug (36) and alcohol addictions (37) and obesity (30), an individual can span their influence to their friends around three degrees of separation (friend of a friend's friend). This phenomenon is also widely known as "three degrees of influence" in social science. In ecological networks, Guimarães et al. (32, 33) discovered in 2017 that indirect effects contribute strongly to the trait coevolution among reciprocal species, which can alter environmental selection and promote the evolution of species.

Despite the ubiquity of indirect influence in various real-world systems, few studies have examined the exact mechanisms by which the indirect influences occur, or the relative strengths between direct and indirect influences. Here, based on empirical

Significance

Increasing empirical evidence in diverse social and ecological systems has shown that indirect interactions play a pivotal role in shaping systems' dynamical behavior. Our empirical study on collaboration networks of scientists further reveals that an indirect effect can dominate over direct influence in behavioral spreading. However, almost all models in existence focus on direct interactions, and the general impact of indirect interactions has not been studied. We propose a new percolation process, termed induced percolation, to characterize indirect interactions and find that indirect interactions raise a plethora of new phenomena, including the wide range of possible phase transitions. Such an indirect mechanism leads to very different spreading outcomes from that of direct influences.

Table 1. Description of bibliographic datasets used in the empirical studies

Established field (PACS)	Emerging field (PACS)	No. of. authors	No. of edges	Constructing network period	Field emerging period	Observation period
Chaos (05.45)	Complex networks (89.75)	1,833	3,128	1999–2003	2001–2003	2004–2006
Phase transitions (64.60)	Complex networks (89.75)	1,265	2,864			
EPLDS (73.20)	OPLDS (78.67)	2,069	5,900			
Carbon nanotubes (39).	Graphene (39).	20,011	110,041	2009–2013	2011–2013	2014–2016

The first and second columns are names and Physics and Astronomy Classification Scheme (PACS) numbers (except carbon nanotubes and graphene) for the four pairs of research fields. The third to fifth columns are the number of authors and edges and the period used to construct collaboration networks, respectively. The sixth column is the period during which a new field emerges and "focused" scientists are specified. The seventh column is the period to observe scientists' behavioral change and to calculate the indicator Q_i.

analyses of scientific collaboration networks, we reveal that indirect influence occurs through next-nearest neighbors and can be the dominant mechanism through which research interests change; on the contrary, evidence of direct (nearest) influence is relatively weak.

However, on the theoretical front, up to now there has been no percolation-based theoretical model to describe the underlying mechanism of indirect influence or its distinctions with existing percolation models in terms of the macroscopic behaviors. For either regular networks or complex networks, various percolation models like bond, site, bootstrap, k-core, linear threshold and core, etc., are always based on direct interactions (8, 9, 28) among nodes. In essence, all of these models only take into account the existence and the strength of directly connected nodes, regardless of any indirect influences of other nodes. Hence, they are not suitable for describing the indirect mechanism. Here, we propose a percolation framework called induced percolation to theoretically study the impact of such an indirect mechanism on the whole system.

Our results show that indirect interactions lead to a unique macroscopic behavior characterized by anisotropy and phase transitions and different spreading outcomes compared to the direct influence mechanisms. Specifically, we study the most general scenario in which links can have directions and report that varying the links' directionality could change the order of the phase transition. This is in sharp contrast to previous percolation models, for which the nature of the phase transitions is not affected by the directionality of links. Such rich phase transition behavior is further illustrated in our simulations on empirical networks. To the best of our knowledge, the phenomenon of directionality-related order of the phase transitions only exists in some special cases of core percolation (27), whereas it is shown to be a generic feature in our indirect interaction model.

Results

Empirical Indirect Influence Mechanisms. To investigate the exact mechanism of neighboring influence and its direct/indirect nature in empirical networks, we study collaboration networks of scientists. Here the "behavior" is meant as the research field(s) of a scientist, and the "spreading of behavior" is defined as the propensity of the scientist to stay in his/her established field or shift to an emerging field. We then study how scientists' research fields are influenced by their direct (nearest) neighbors and indirect (next-nearest) neighbors. We choose four pairs of fields in physics that have large numbers of scientists involved: chaos vs. complex networks, phase transitions vs. complex networks, electrical properties of low-dimensional structures vs. optical properties of low-dimensional structures (hereinafter referred to as EPLDS vs. OPLDS), and carbon nanotubes vs. graphene. For each pair of fields, the latter field is the emerging field (new field) that attracts scientists from the former (old, already established) field.

Specifically, we analyze the datasets of articles published by the American Physical Society (APS) (38) and Web of Science (39), considered as representative data sources for the studied fields (we share all the data of this study at https://github.com/Jia-Rong-Xie). Based on articles in each pair of fields, covering in total 5 y around the emergence of a new field (see Table 1 and *SI Appendix*, Fig. S1), we construct a collaboration network. The nodes are the scientists, and a link is established between any pair of scientists who have at least one joint publication. Scientists who have published multiple articles (at least two in APS and five in Web of Science dataset; refer to extended discussion for parameter robustness in *SI Appendix*, Figs. S7–S13) in the old field yet have not published any articles in the new field are defined as focused scientists in the old field. They are assumed to be the influencers in the networks and labeled as state 1. For any other nodes (influenced) in the networks, we calculate the number of direct and indirect "influencers" for each node. The number of direct influencers of node i is simply the number of its nearest neighbors with state 1, and we denote it as $\tilde{k}_i$. For the number of indirect influencers of node i, we first identify its state 1 neighbors. For each of the direct influencers (direct state 1 neighbors), we then count the number of their own state 1 neighbors, and the maximum count is defined as the number of indirect influencers m_i (also called induced index). On each of the direct influencers of node i, we further count its degree and define the maximum degree of them as degree index d_i. A visual illustration of the definitions is shown in Fig. 1*A*.
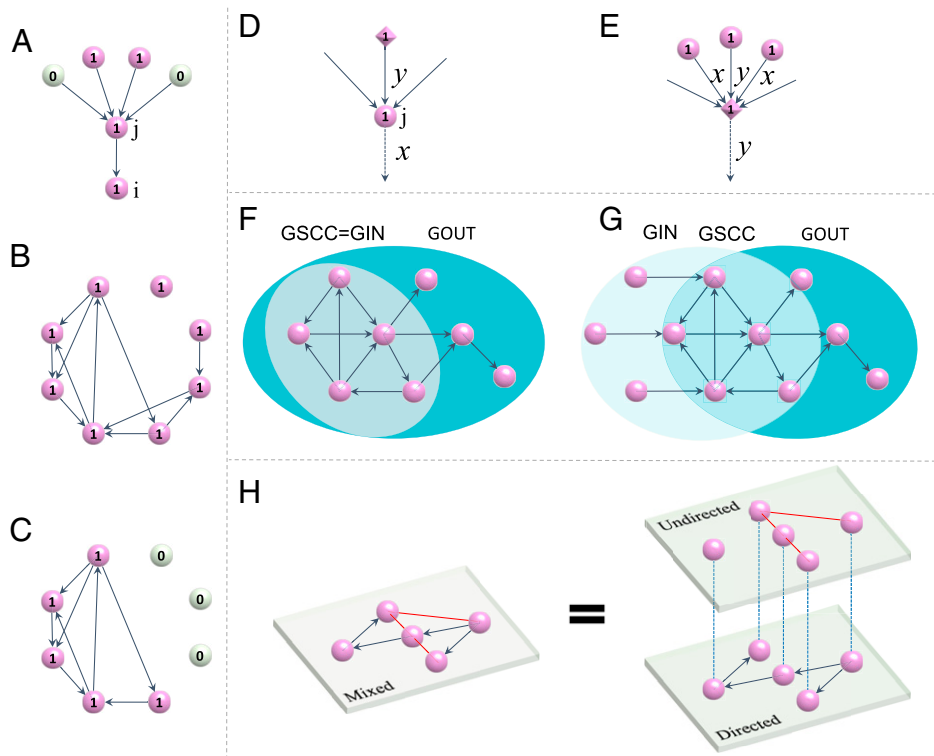
Within the next 3 y (see Table 1 and *SI Appendix*, Fig. S1), we count each influenced i's publications and calculate the proportion Q_i of articles in the old field by the following expression:

$$Q_i = \frac{P_i^{\text{old}}}{P_i}, \quad \textbf{[1]}$$

where P_i^{old} and P_i represent the number of papers in the old field and the total number of papers published by scientist i during the observation period, respectively. A higher Q_i value of the influenced i then indicates that it receives more influence by the "influencers" with state 1, either directly or indirectly. Our results in Fig. 1 *B*, *C*, and *E* (and *SI Appendix*, Figs. S4–S6) clearly show that Q_i increases with the indirect influence index m_i, yet not so much with direct influence index $\tilde{k}_i$ (also called k-core index; see Fig. 1 *C* and *D*) or the degree index d_i (see Fig. 1 *E* and *F*) and the second-nearest degree index κ_i (see *SI Appendix*, Fig. S6). This indicates that rather than direct influence, indirect influence plays a dominant role in the choice of research focus among scientists. Indeed, we find that nodes i and h, via node j (see Fig. 1*A*), are more likely to coauthor publications in the old field, which means that the quantitative correlation between Q_i and m_i does mediate the collaboration relationship (see *SI Appendix*, Fig. S14).
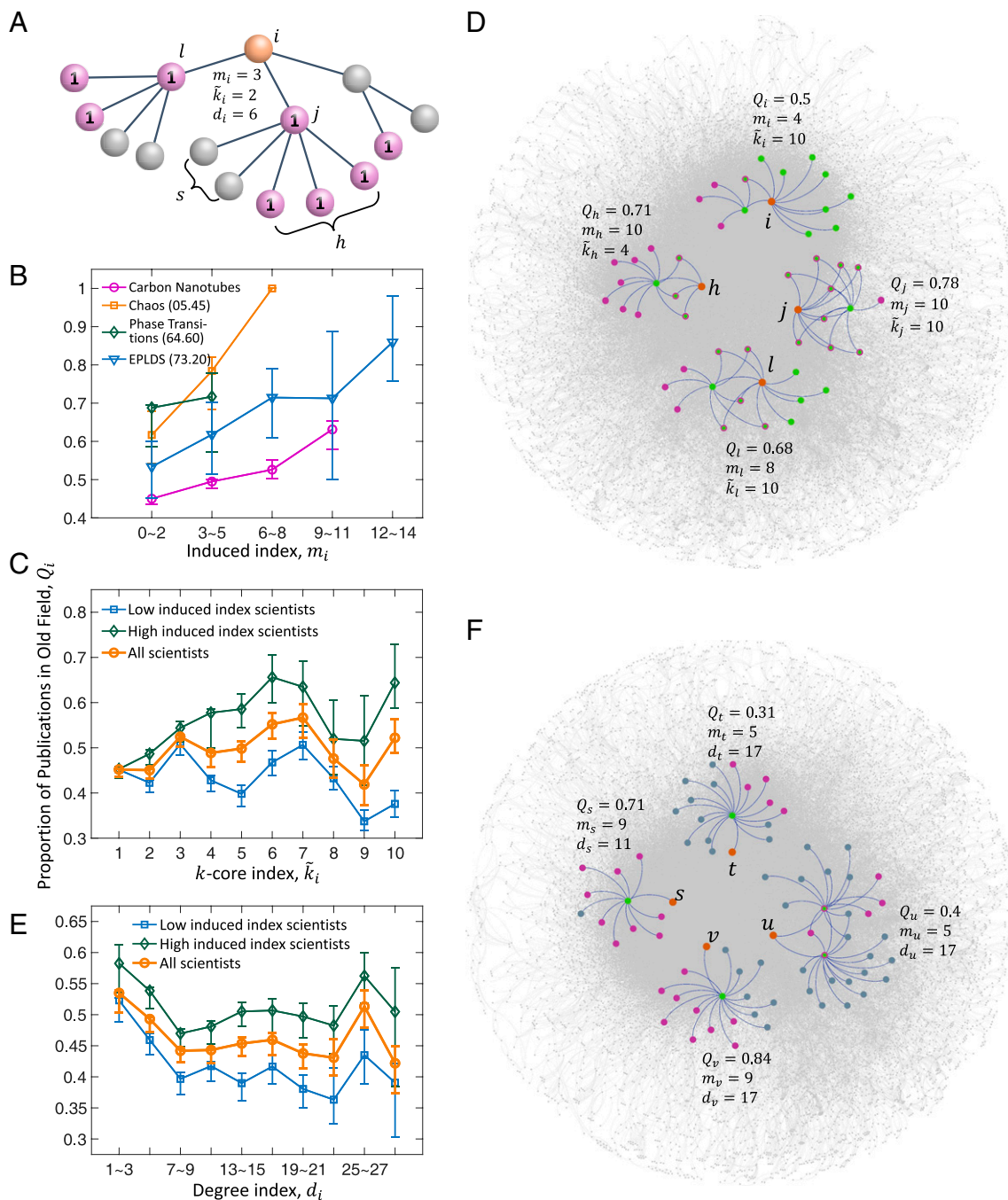
Fig. 1. Indirect influence mechanism in empirical collaboration networks. (*A*) Schematic representation for the induced index m_i, the k-core index $\tilde{k}_i$, and the degree index d_i. In this example, "focused" scientists in the established field are denoted as state 1. Node i has an induced index $m_i = 3$ because among all direct neighbors in state 1, node j has the maximum neighbors in state 1 (i.e., three excluding node i). The degree index of node i is $d_i = 6$, which is the degree of node j. The k-core index of node i is $\tilde{k}_i = 2$, which is the number of direct neighbors in state 1 (nodes l and j). (*B*) Empirical evidence of indirect influence. It shows a clear indirect influence mechanism in four pairs of established and emerging fields in physics that have large numbers of scientists involved. The proportion Q_i of publications in the established fields significantly increases with the scientists' induced index m_i in all the datasets. To compare with direct influence, the orange lines in *C* show that the value of Q_i is hardly affected by the direct influence measured through the k-core index, while the scientists with higher indirect influence index (top 50% of m_i values) clearly have a higher Q_i value than that of the lower indirect influence (bottom 50% of m_i values), indicating a strong indirect influence. *D* highlights four sample scientists (nodes) labeled as h, i, j, l. Each orange node is a node of interest, its connected green nodes are the neighbors of state 1, pink nodes are green nodes' state-1 neighbors used to calculated induced index m_i. Higher induced index nodes h and j ($m_h = m_j = 10$) publish a proportion $Q_h = 0.71$ and $Q_j = 0.78$ of old field articles, much higher than that of node i with a lower induced index ($m_i = 4$), although i's k-core index is higher ($\tilde{k}_i = 10$) than h and the same as j. Comparing node l and i, j again indicates that the influence is stronger through induced index m than that of $\tilde{k}$. A similar comparison in *E* and *F* shows that the proportion Q_i is hardly affected by the degree index but clearly affected by the induced index. *C–F* show results performed on the collaboration network of carbon nanotubes vs. graphene. Note that in *F* we also show the state 0 nodes labeled in blue, since the calculation of degree index considers them.

Xie et al.
Indirect influence in social networks as an induced percolation phenomenon

PNAS | 3 of 10
https://doi.org/10.1073/pnas.2100151119

Fig. 2. Induced percolation on directed networks. *A* illustrates the proposed mechanism of induced percolation for the case $m = 2$. In order for a node i to remain in state 1, at least one node (j) at the other end of an incoming link should be in state 1. In its turn, j should also have at least m (= 2 in the example) incoming links from neighbors that are in state 1. *B* shows a directed graph of eight nodes all in state 1. *C* shows the GOUT at equilibrium state when the graph on panel *B* is pruned according to the induced percolation rules. *D* and *E* illustrate the variables x and y defined in the main text by Eqs. **3** and **4**. *F* and *G* show the relationship between the order parameters GSCC, GIN, and GOUT for induced percolation and typical bond percolation processes, respectively. *H* schematically represents the multilayer representation employed to derive the order parameter P_∞ when there are directed and undirected links in the substrate network.

The observed indirect influence mechanism in empirical collaboration networks is possibly due to the following two factors. First, the fact that a scientist has a high value of induced index means he/she collaborates with a highly active scientist (a state 1 neighbor on its own connecting to a large amount of state 1 neighbors). This active scientist could strongly influence collaborators. Second, researchers who collaborate with highly active scientists have better chances to find new potential collaborators through their connections with respect to researchers who have no highly active neighbors. In other words, scientists with high induced index can interact with researchers of the same field indirectly, through their highly active neighbors.

Induced Percolation Model. We now define a percolation model based only on this indirect influence mechanism characterized by the indirect index m_i. As empirically shown before, the indirect influence increases with m_i. Here we present the most simplified version of this influence mechanism that assumes a deterministic influence outcome, i.e., a node i is influenced to state 1 with probability $h(m_i) = 1$ if its indirect induced index m_i is not smaller than a threshold m (see *SI Appendix*, Fig. S20 for a slightly more complicated case):

$$h(m_i) = \begin{cases} 0, & m_i < m, \\ 1, & m_i \geq m. \end{cases} \quad \text{[2]}$$

More formally, induced percolation can be defined on directed networks as follows. Let us assume that the state of the nodes is characterized by an integer value, 0 or 1. Initially, we set the state of all nodes in the network to 1. A node i remains in state 1 if at least one of its incoming links comes from a node,

say j, with state 1, and in turn the node j has at least m other incoming links from nodes that are in state 1; see Fig. 2*A* for an illustration of the case $m = 2$. Otherwise, node i changes to state 0 at the next time step. The influence of the m nodes on the node i defines the indirect interactions among them. Under this mechanism, certain nodes will change their states from 1 to 0 at each time step until no more changes are possible; see Fig. 2 *B* and *C* for an example. Compared with bond, bootstrap, or k-core percolation, the fundamental difference of induced percolation is that the current state of a node is affected not only by its nearest neighbors but also by a number of its next-nearest neighbors. The mechanism for induced percolation through a network captures the observation that there are behaviors whose influence reaches nodes beyond the first shell.

In network percolation theory, the giant strongly connected component (GSCC), giant in-component (GIN), and giant out-component (GOUT) are three main order parameters. In particular, GSCC refers to the largest strongly connected component whose size is comparable to the entire network. GIN is the group of nodes from which any node in GSCC can be reached, while GOUT is the group of nodes that can be reached from any node in GSCC. For various types of propagation dynamics on networks, the GOUT corresponds to the largest spreading coverage, and it serves as an indicator of network connectivity under a given propagation mechanism. The size of the GOUT in the empirical studies corresponds to the number of scientists who stay in the old field. Therefore, in induced percolation, the main quantity of interest is GOUT (8, 28, 29) and the size of GOUT is the order parameter, i.e., the macroscopic quantity that characterizes phase transitions. In addition, we also examine the size distribution of small outgoing components.

In undirected networks, each link can be viewed as two directed links with opposite directions. Therefore, induced percolation can be studied on fully directed networks, and then the methodology can be extended to either undirected (i.e., fully bidirectional) networks or to networks in which there are both bidirectional and unidirectional links. Note that the GOUT of undirected networks is the same as the GIN and the GSCC. We schematically illustrate the proposed induced percolation mechanism on directed networks in Fig. 2, where we also show the order parameter as compared with the one typically used in bond percolation. Similar diagrams for undirected and mixed networks can be found in *SI Appendix*, Figs. S15 and S16.

The phase transition that characterizes the induced percolation process can be analytically studied on random networks. The class of random directed networks is constructed by independently connecting two arbitrary nodes with a directed link with a fixed probability. The network can be described by the joint degree distribution $P(k_{in}, k_{out})$, which is the probability that a randomly selected node has in-degree k_{in} and out-degree k_{out}. For random directed networks, the size of GOUT is derived through the following recursive equations. We first define two recursive variables x and y (see Fig. 2 *D* and *E*): x represents the probability that when selecting at random a directed link, the node at the origin of the link is active (in state 1), whereas y represents the probability that a random link enables its end node to be in an active state. According to the definitions of x and y, we have

$$x = \sum_{k_{in}, k_{out}}^{+\infty} \frac{k_{out} P(k_{in}, k_{out})}{\langle k \rangle} \left[1 - (1-y)^{k_{in}} \right], \qquad [3]$$

where $(1-y)^{k_{in}}$ is the probability that none of the incoming k_{in} links can keep node j in state 1 (see Fig. 2 *A* and *D*), while $1 - (1-y)^{k_{in}}$ represents the probability that at least one

of the incoming k_{in} links can keep node j in state 1. The term $\frac{k_{out} P(k_{in}, k_{out})}{\langle k \rangle}$ is the excess incoming degree distribution (28, 29) for the node at the origin of an arbitrary directed link. This is because the likelihood of a node's being the origin of a randomly chosen directed link is proportional to the node's out-degree.
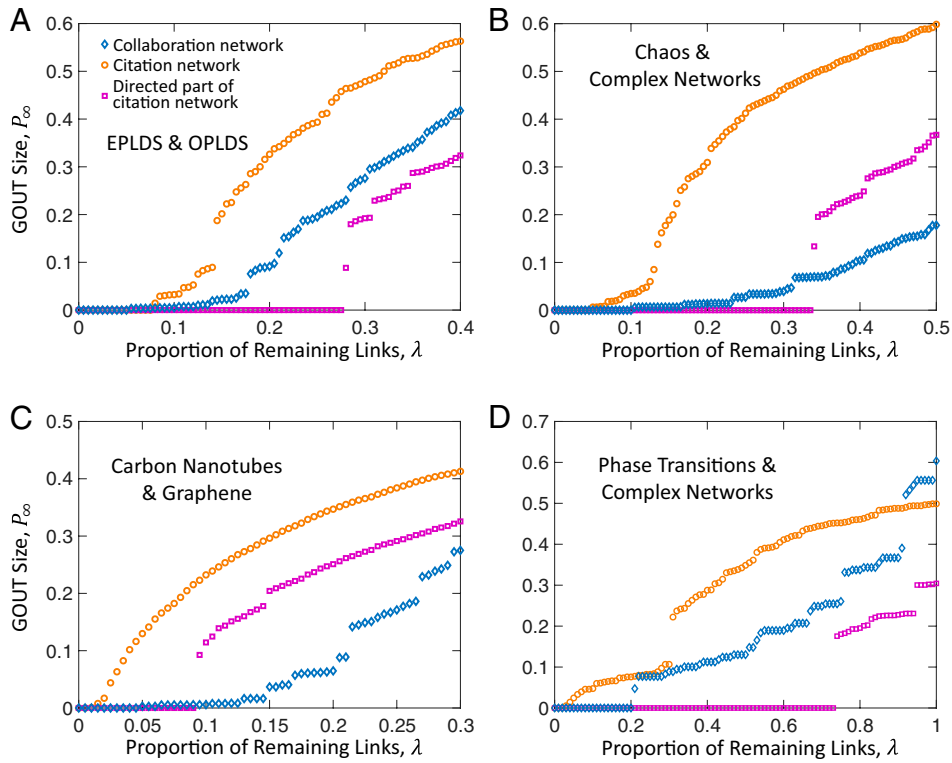
Calculating the probability y is a little more involved. The definition of the induced percolation process implies that even if the starting node of a directed link is active (which happens with probability x), it is not guaranteed that the end node of this directed link remains active (which happens with probability y). However, if the starting node of this directed link is itself active, and at the same time at least m neighbors pointing to the starting node are active, then this directed link can keep its end node active. Conversely, if a directed link can keep the node it points to active (corresponding to y), then the starting node of this directed link must be active (corresponding to x). Therefore, it must hold $x > y$ when $m > 1$ ($x = y$ when $m = 1$ which corresponds to bond percolation). The above analysis yields the expression of y as

$$y = \sum_{k_{in}, k_{out}}^{+\infty} \frac{k_{out} P(k_{in}, k_{out})}{\langle k \rangle}$$
$$\sum_{s=m}^{k_{in}} \binom{k_{in}}{s} x^s (1-x)^{k_{in}-s} \left[1 - \left(1 - \frac{y}{x} \right)^s \right], \qquad [4]$$

where $\binom{k_{in}}{s} x^s (1-x)^{k_{in}-s}$ gives the probability that for a node of incoming degree k_{in}, s out of k_{in} neighbors are active. y/x represents the conditional probability that a directed link keeps its end node active, given the starting node is active. Therefore, $1 - (1 - y/x)^s$ is the probability that at least 1 out of the s active incoming neighbors keeps this node active.
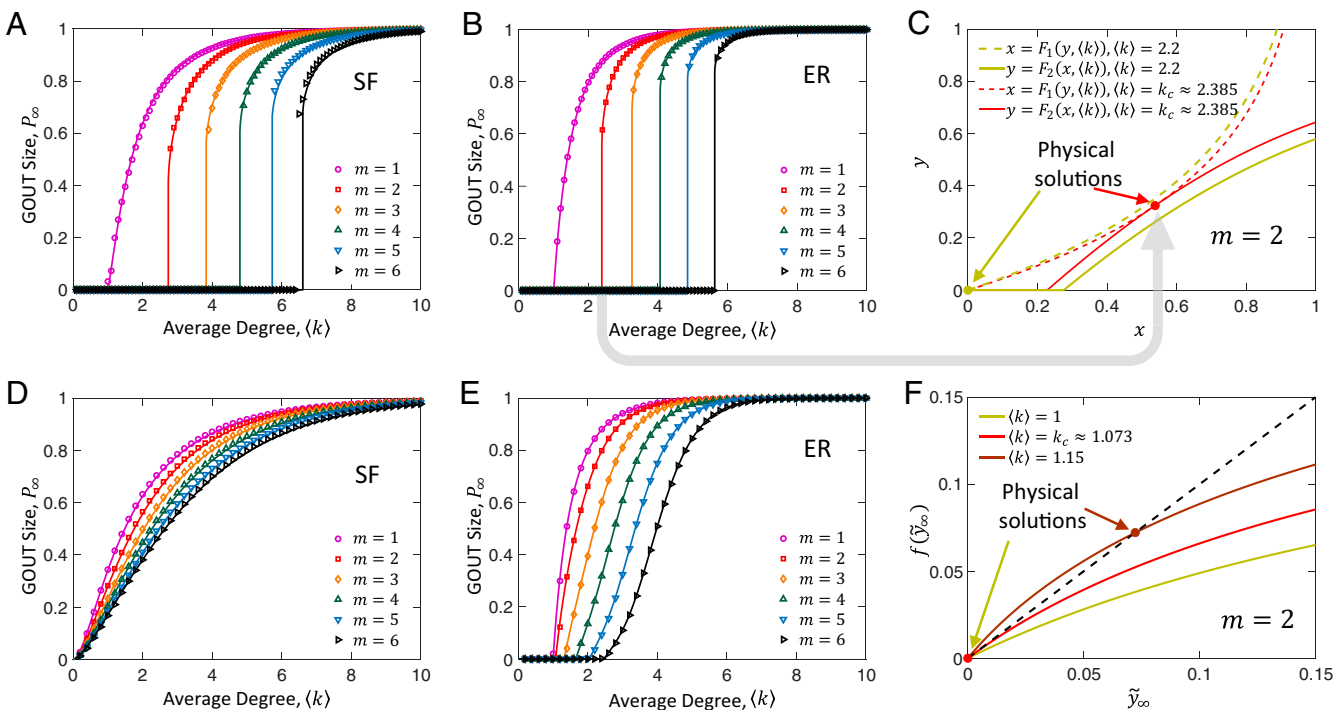
Fig. 3. Order parameter GOUT for induced percolation on directed and undirected random networks. The symbols represent simulation results and the curves are corresponding theoretical results. *A* and *B* show GOUT for induced percolation ($m = 2, \ldots, 6$) on directed scale-free (SF) and Erdős–Rényi (ER) networks as a function of the average degree $\langle k \rangle$. Results are compared with the behavior of the same order parameter for bond percolation (equivalent to setting $m = 1$). *C* shows the graphical solution of Eq. **4** for induced percolation ($m = 2$) on directed ER graphs, where k_c is the critical average degree at which a first-order phase transition takes place. *D* and *E* show results for undirected networks, whereas the graphical solution shown in *F* is derived from Eq. **13** (see *Methods*) for induced percolation ($m = 2$) on undirected ER graphs. Directed SF networks are generated by the static model (41, 42) with exponents $\gamma_{in} = 2.5$ and $\gamma_{out} = 3.0$ for the incoming and outgoing degree distributions, respectively. Undirected SF networks are generated with the exponent $\gamma = 2.5$.

Xie et al.
Indirect influence in social networks as an induced percolation phenomenon

PNAS | 5 of 10
https://doi.org/10.1073/pnas.2100151119

Fig. 4. Order parameter GOUT for induced percolation on empirical networks based on datasets in Table 1. *A–D* show GOUT as a function of the proportion λ of remaining links. Each point of GOUT is computed as the steady state of induced percolation ($m = 3$) on real-world networks after randomly removing a fraction of $1 - \lambda$ links. The collaboration network is constructed based on published articles within the first 5 y in the four pairs of fields described in Table 1. The directed part of a citation network is obtained by removing all the bidirectional links from the citation network (see *SI Appendix*, section 1E). For the studied four pairs of fields, GOUT in general well agrees with main findings: a continuous, discontinuous, and hybrid phase transition for undirected collaboration networks (blue diamonds), directed part of citation networks (purple squares), and mixed citation networks (orange circles).

Finally, the order parameter P_∞ for the size of GOUT can be calculated based on Eqs. **3** and **4** as follows:

$$P_\infty = \sum_{k_{\text{in}}, k_{\text{out}}}^{+\infty} P\left(k_{\text{in}}, k_{\text{out}}\right) \left[1 - (1 - y)^{k_{\text{in}}}\right]. \qquad [5]$$

Here P_∞ is equivalent to the probability that a randomly chosen node has at least one incoming node to keep it active. One interesting finding worth highlighting is that the GSCC coincides with the GIN for the induced percolation process on directed networks, which is not the case for classical percolation models (see Fig. 2 *F* and *G*). The theoretical analysis of the order parameter P_∞ on undirected networks is illustrated in *Methods*. We also note that the analysis of P_∞ on mixed networks can be done by mapping the structure to a multilayer network; see Fig. 2*H* and more details in *SI Appendix*, section 2.

Phase Transitions of Induced Percolation. Theoretical analyses allow us to show that the type or order of the phase transition depends on the directionality of the links for the same network connectivity pattern, i.e., the phase transition is anisotropic in nature. On directed networks, when $m > 1$ ($m = 1$ is the case of typical bond percolation), induced percolation shows discontinuous (first-order) phase transitions (see Figs. 3 *A–C* and 4 *A–D* for real-world networks). Yet, on undirected networks, the same percolation process always leads to continuous (second-order) phase transitions (see Figs. 3 *D–F* and 4 *A–D* for real-world networks). These results are in sharp contrast with previous percolation models on networks (see Table 2), for which it has never been found that the directionality of network links fundamentally alters the type of phase transitions. This means that previously studied types of percolation models might have significantly underestimated the effects of asymmetry in link directions

on the system's macroscopic behavior. An important implication of this observation is that abrupt transitions in complex systems like ecological and social networks might be way more likely to occur than previously anticipated by existing percolation models.

The anisotropy induced by the directionality of the links leads to a rich and complex behavior when the network is composed of a mixture of directed and undirected links. Specifically, a hybrid phase transition emerges with the presence of a certain amount of directed links. Fig. 5 *A* and *B* show that by increasing the fraction p of directed links in the network, the order parameter GOUT evolves, as the average degree $\langle k \rangle$ increases, from a continuous transition to a hybrid phase transition where both continuous and discontinuous transition exist, to a first-order transition for larger values of $\langle k \rangle$. In addition, in the region where the hybrid phase transition is observed, several quantities follow a set of scaling relations with critical exponents that are in line with Landau's mean-field theory.

We label the critical hybrid point where the hybrid transition first appears as point $C\left(k^*, p^*\right)$ in Fig. 5*A*. We find a set of scaling relations connecting GOUT to other quantities near C that are predicted by Landau's mean-field theory: Within the hybrid transition, the jump height of GOUT, $\Delta P_\infty(p^* + \Delta p) \coloneqq \lim_{\langle k \rangle \to k_c^+} P_\infty(\langle k \rangle, p^* + \Delta p) - \lim_{\langle k \rangle \to k_c^-} P_\infty(\langle k \rangle, p^* + \Delta p)$, where k_c is the critical point at which the first-order transition occurs, follows a scaling function of Δp with the critical exponent $\eta = 1/2$ (Fig. 5*E*):

$$\Delta P_\infty(p^* + \Delta p) \sim (\Delta p)^{\frac{1}{2}}. \qquad [6]$$

The same critical exponent holds for the jump height as a scaling function of $\langle k \rangle - k^*$ as shown in the *SI Appendix*, section 5. When fixing p at p^* and varying $\langle k \rangle$ in the vicinity of k^*, the

Table 2. Comparison of percolation models

| Percolation model	Type of phase transition		Clusters distribution near critical point	β	Hybrid phase transition at critical point
	Undirected	Directed			
Induced percolation	Second	First	Nonmonotonic	1 (second) 1/2 (first)	$\theta = 1/3$ $\eta = 1/2$
Bond percolation (5, 8, 9, 28)	Second	Second	Monotonic	1	—
Site percolation (5, 8, 9, 17, 28)	Second	Second	Monotonic	1	—
Bootstrap percolation (18)	Second/first	—	Monotonic	1 (second) 1/2 (first)	$\theta = 1/3$ $\eta = 1/2$
k-core percolation (19)	Second/first	Second/first	—	1 (second) 1/2 (first)	—
Core percolation (26, 27)	Second	Second/first	—	1 (second) 1/2 (first)	—
Explosive percolation (40, 43, 44)	Second	—	—	0.0555	—
Articulation percolation (45)	Second/first	—	—	1 (second) 1/2 (first)	—

Dashes indicate that no related research has been found.

size deviation of GOUT is quantified by the following scaling function of $\langle k \rangle - k^*$ with critical exponent $\theta = 1/3$ (Fig. 5F), reached from both below and above:
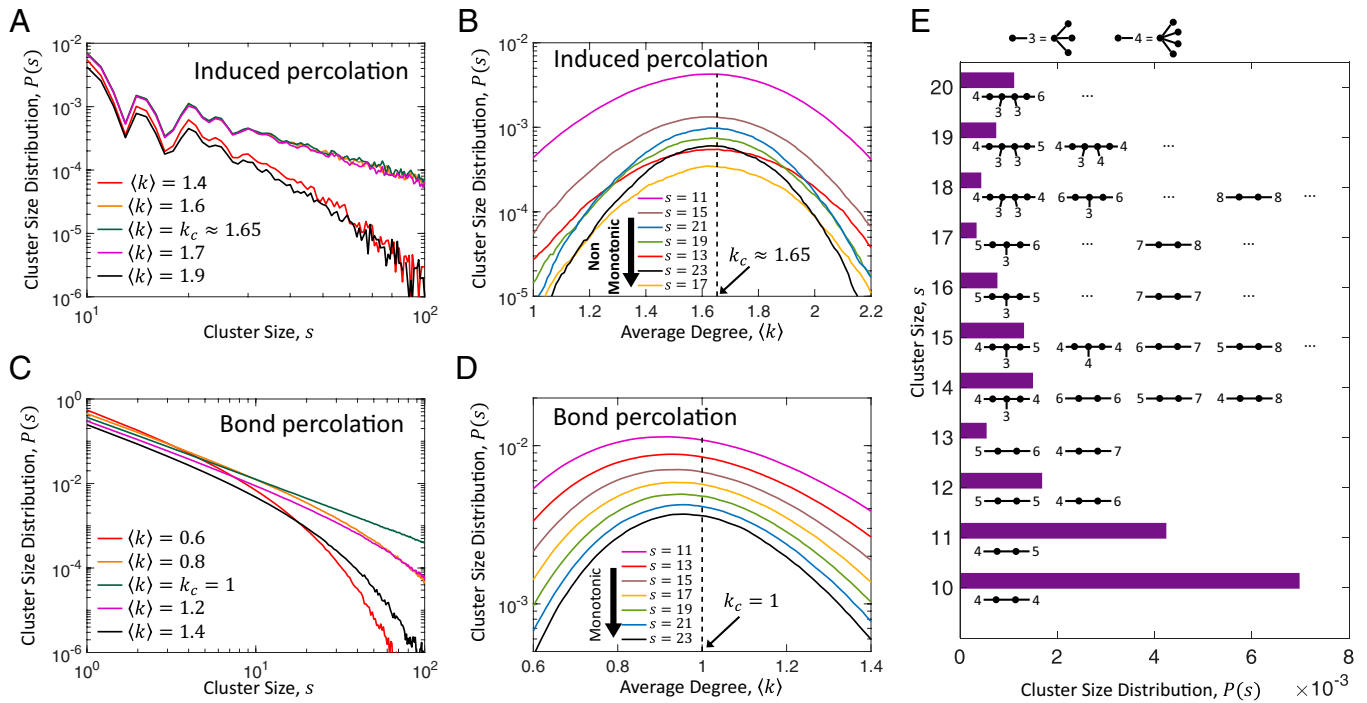
$$|P_\infty(\langle k \rangle, p^*) - P_\infty^*(k^*, p^*)| \sim |\langle k \rangle - k^*|^{\frac{1}{3}}. \qquad [7]$$

We note that Baxter et al. also find these two critical exponents in k-core percolation (20).

Another unexpected feature that distinguishes the percolation process formulated here from other percolation is the cluster size distribution near criticality. Typically, for the second-order phase

APPLIED MATHEMATICS
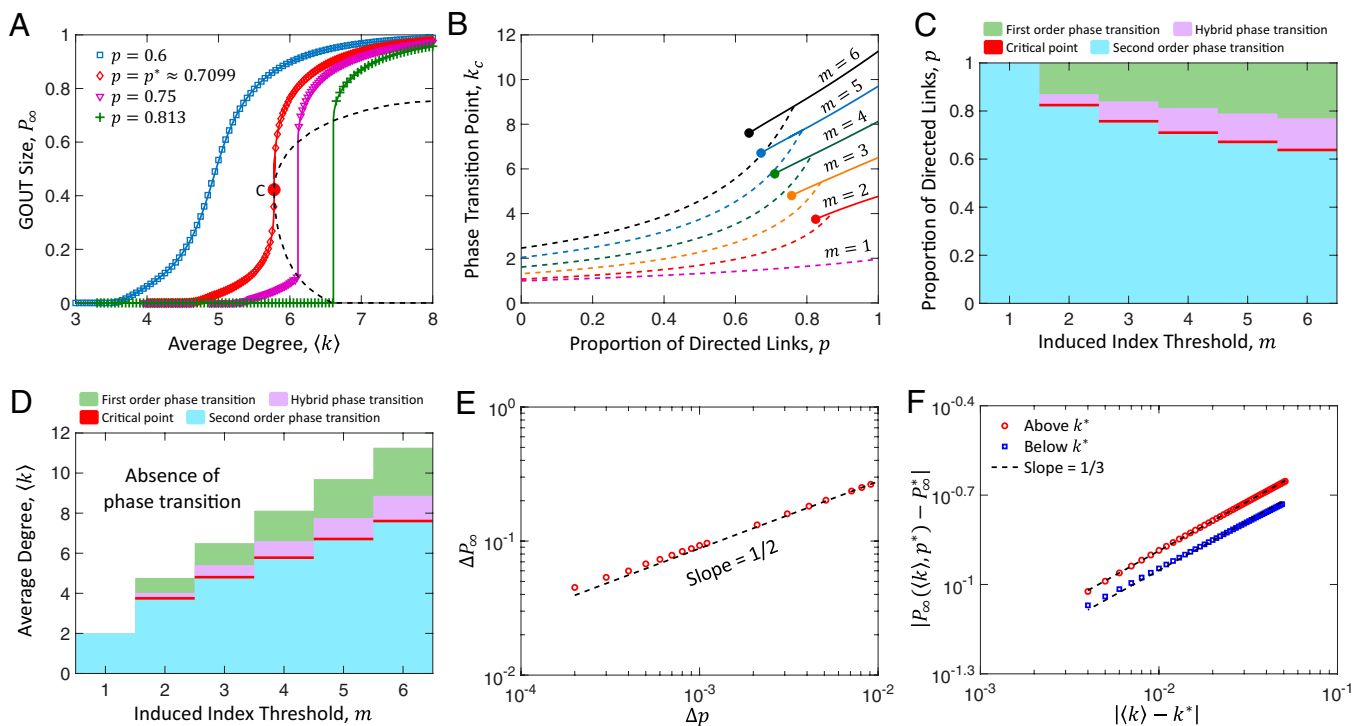
Fig. 5. Phase transitions and critical behaviors of induced percolation on mixed networks. In *A*, we show theoretical and numerical results for GOUT as a function of the average degree $\langle k \rangle$ when the fraction of directed links is varied. The point C denotes the point at which coexistence of second- and first-order phase transitions occurs for the first time. The curved dotted line represents the value of GOUT before and after the first-order phase transition. The symbols represent simulation results and the curves are corresponding theoretical results. *B* shows the values of the critical points in the parameter space made up by the average degree and the percentage of directed links; the dotted line describes the critical value at which a second-order phase transition occurs, while the solid line corresponds to the first-order phase transition. Dots correspond to critical points, C. *C* represents the types of phase transitions that can be observed in the $m - p$ plane. Blue, purple, and green colors bound the area in which second-order, hybrid, and first-order phase transitions exist, respectively. The red boundary lines between the blue and the purple areas correspond to the critical points C. When the parameters are such that they lay on the red line, the behavior of GOUT corresponds to the green line marked with point C in *A*. *D* shows the types of phase transitions shown in *C* but in the $m - \langle k \rangle$ plane. *E* presents results of the jump size, ΔP_∞, as a function of $\Delta p = p - p^*$ when the critical point C is approached either from below or from above. *F* depicts the change of P_∞ near the critical point k^* as a function of $\langle k \rangle - k^*$, when fixing $p = p^*$. The mixed network is generated by assigning a percentage p of directed links to an undirected ER network with an average degree $\langle k \rangle$ and consists of 10^6 nodes.

Xie et al.
Indirect influence in social networks as an induced percolation phenomenon

PNAS | 7 of 10
https://doi.org/10.1073/pnas.2100151119

Fig. 6. Size distribution, $P(s)$, of small clusters at the critical point of induced percolation ($m = 4$) on undirected networks. In *A* we show the size distribution, $P(s)$, which exhibits a fluctuating behavior especially for small sizes. *B* plots the same distribution $P(s)$ but as a function of the average degree $\langle k \rangle$, showing an unambiguous nonmonotonic decrease of the size distribution. *C* and *D* depict the monotonous power-law decay of the cluster size distribution in the limit of classical bond percolation. Finally, *E* displays the cluster size distribution at the critical point $k_c \approx 1.65$, also showing the structure of each cluster. Results are averaged over 10^3 independent realizations of undirected Erdős–Rényi networks (of size 10^6 nodes). As it can be clearly seen, the critical behavior of induced percolation is different from that of classical one.

transitions, in the vicinity of the phase transition point, the size distribution of small connected clusters is in general governed by the monotonous function of $P(s) \sim s^{-\tau} e^{-s/s^*}$, where s^* provides a characteristic size of the finite components (4, 10). The closer to the critical point, the larger s^* will be. At the exact phase transition point, s^* approaches infinity and $P(s)$ exhibits a monotonic power-law distribution of $P(s) \sim s^{-\tau}$, signifying a loss of characteristic scale in the distribution. However, for induced percolation on undirected networks, we find that near the critical point $P(s)$ exhibits a novel oscillatory-like behavior, i.e., it is no longer monotonically decreasing with s (see Fig. 6 *A* and *B*).

As it can be seen in the Fig. 6 *A*, *B*, and *E*, the observed oscillatory-like behavior of $P(s)$ is more pronounced for small values of s and does not change the asymptotic power law distribution for large s nor the critical exponent of the phase transition, which is the same as in bond percolation, $\beta = 1$, $\tau = 5/2$ (40). This behavior of $P(s)$ is, however, clearly distinct from the classical monotonic distribution (see Fig. 6 *C* and *D*). We note that we do not have a clear notion of what the exact impact of this pattern on the macroscopic behavior of the system is, which is a question to be further examined in future work.

Conclusion and Discussion

Let us first mention that in addition to our empirical results on collaboration networks we believe that the induced percolation mechanism could play a relevant role in other examples of behavioral influence or contagion, such as in the behavioral spreading of drug abuse, alcoholism, obesity, divorce, happiness, and loneliness, among others. These examples are usually listed to show the "three degrees of influence" mechanism. That is, one individual's influence can significantly spread out to their friends' friends'

friends. However, the specific spreading mechanisms behind this phenomenon remain unknown and with no theoretical, first-principled grounds. Although our empirical work reveals only one mechanism of influence within two degrees, we believe that it can be regarded as the first step to provide a specific spreading mechanism for the "three degrees of influence" and potentially opens new paths in the field of percolation on networked systems.

Based on our empirical discovery that indirect influence can dominate over direct influence we have proposed an induced percolation model to characterize the dynamics and outcomes of this indirect spreading mechanism. We found that such indirect interactions lead to a plethora of percolation transitions in complex networks that are rooted in the degree of anisotropy of the connectivity pattern. Specifically, we have shown that the amount of directed links in a network determines the order of the phase transition, which spans from a second order in networks without directed links to a first order when all links are directed. In between, a rich behavior associated with hybrid phase transitions emerges with the coexistence of second- and first-order phase transitions. In addition, the indirect effect makes the size distribution of small clusters near the phase transition point exhibit a nonmonotonic pattern, which has not been previously seen in other percolation models.

Our theoretical framework provides the tools to investigate the implications of having different indirect influence mechanisms in a spreading phenomenon and understand their associated dynamical process and macroscopic spreading outcomes. For instance, we have found that indirect influence can dominate over direct influence in social systems like what we found in scientific collaboration networks—if similar mechanisms in other social behaviors like drug abuse, alcoholism, etc. also hold, this

implies very different mitigation policies from that based on direct influence mechanisms.

Methods

Induced Percolation on Undirected Networks. We elaborate on the definition and the theoretical derivation of induced percolation on undirected networks. All nodes in an undirected network are initially set to state 1. A node l remains in state 1 if at least one of its undirected links has a node j in state 1, and this node j has at least m neighbors (excluding the node l) with state 1 (as illustrated in *SI Appendix*, Fig. S16A for the case of $m = 2$); otherwise, node l changes to state 0 at the next time step. To theoretically analyze the percolating probability that any node belongs to a GCC (giant connected component, equivalent to GOUT), P_∞, we start by defining six conditional probabilities as intermediate variables, whose notations are shown collectively in *SI Appendix*, Fig. S17. Without loss of generality, we denote a randomly chosen undirected link as $\{j, l\}$ and deduce the probability that node l belongs to a GCC.

According to the definition of induced percolation for undirected networks, the condition for node l to remain active is that there is at least one active neighbor j, and the number $\tilde{u}$ of active neighbors (except node l) of node j satisfies $\tilde{u} \geq m$. We refer to a node in state 1 as an active node and in state 0 as an inactive node. Unlike active neighbors in directed networks, the number $\tilde{u}$ of active neighbors in undirected networks is closely related to the degree k of node j. Specifically, if $k > m$ and node j is active, then node j can keep all its neighbors active. Conversely, if $k \leq m$, then node j cannot keep any of its neighbors active. Hereafter, we employ the degree k instead of the number of active neighbors $\tilde{u}$ to derive percolation probability.

The conditional probability $\tilde{v}$ is the probability that node j can keep node l active, given node l can keep node j active. As per the definition of induced percolation, the event of j keeping l active implies that the degree of node j satisfies $k > m$. Node j simultaneously keeps all of its neighbors active. The above analysis yields the following recursive equation:

$$\tilde{v} = \sum_{k=m+1}^{\infty} \frac{kP(k)}{\langle k \rangle}, \qquad [8]$$

where $\frac{kP(k)}{\langle k \rangle}$ represents the excess degree distribution of the end node of a randomly chosen link.

On the other hand, the conditional probability $\tilde{v}_\infty$ is defined as the probability that node j can keep node l active, and node l is connected to the GCC via node j, given that node l can keep j active. Again, as per the definition of induced percolation, the degree of node j satisfies $k > m$. Analogously, node j can keep all its neighbors active. In addition, the event that node l connects to GCC through node j is equivalent to the event that node j connects to GCC through at least one of the $k - 1$ neighbors other than l. The corresponding probability is $1 - (1 - \tilde{t}_\infty - \tilde{v}_\infty)^{k-1}$ (as shown in *SI Appendix*, Fig. S17C), where the probability $1 - \tilde{t}_\infty - \tilde{v}_\infty$ accounts for the likelihood that one of the $k - 1$ neighbors does not belong to the GCC given that node j can keep it active. Therefore, the self-consistent equation for the conditional probability $\tilde{v}_\infty$ can be written as

$$\tilde{v}_\infty = \sum_{k=m+1}^{\infty} \frac{kP(k)}{\langle k \rangle} \left[1 - (1 - \tilde{t}_\infty - \tilde{v}_\infty)^{k-1} \right]. \qquad [9]$$

In the previous definition, we made use of the conditional probability $\tilde{t}_\infty$, which is the probability that node j cannot keep node l active while node l connects to GCC through node j, under the condition that node l maintains node j in state 1. Thus, it follows that the degree of node j satisfies $k \leq m$ and that node j cannot keep any of its neighbors active. Moreover, the event in which node l connects to the GCC through node j is equivalent to the event in which node j reaches the GCC through one of the $k - 1$ neighbors other than l. The corresponding probability reads $1 - (1 - \tilde{a}_\infty - \tilde{y}_\infty)^{k-1}$ (as shown in *SI Appendix*, Fig. S17D), where the probabilities $\tilde{a}_\infty$, $\tilde{y}_\infty$ stand for cases in which node j cannot keep any neighbors in state 1 (see below). Therefore, the conditional probability $\tilde{t}_\infty$ can be calculated using

$$\tilde{t}_\infty = \sum_{k=1}^{m} \frac{kP(k)}{\langle k \rangle} \left[1 - (1 - \tilde{a}_\infty - \tilde{y}_\infty)^{k-1} \right]. \qquad [10]$$

Once the above probabilities have been defined, we can proceed with the derivation of the remaining three conditional probabilities, namely, $\tilde{y}$, $\tilde{a}_\infty$, and $\tilde{y}_\infty$, which are analogous to v, $\tilde{v}_\infty$, and $\tilde{t}_\infty$, but under the condition that node l cannot keep node j active. The derivation of the probability $\tilde{y}$ is similar to $\tilde{v}$, except that node j relies on at least one of the $k - 1$ neighbors (except l) to remain active. This probability can be expressed as

$$\tilde{y} = \sum_{k=m+1}^{\infty} \frac{kP(k)}{\langle k \rangle} \left[1 - (1 - \tilde{v})^{k-1} \right]. \qquad [11]$$

The derivation of the conditional probability $\tilde{a}_\infty$ is similar to that of $\tilde{t}_\infty$, except that one additional condition is required: of the $k - 1$ neighbors different from l, at least one can keep j active and connected to the GCC. Assuming that there are exactly s ($1 \leq s \leq k - 1$) neighbors that can keep node j active, the probability is $\binom{k-1}{s} \tilde{y}^s (1 - \tilde{y})^{k-1-s}$. The probability that node j is connected to the GCC through one of the s neighbors is $\frac{\tilde{y}_\infty}{\tilde{y}}$. For the remaining $k - 1 - s$ neighbors that cannot keep j active, the probability of j connecting to the GCC through one of them is $\frac{\tilde{a}_\infty}{1-\tilde{y}}$. Therefore, the probability that node j is not connected to the GCC through any neighbor is $(1 - \frac{\tilde{y}_\infty}{\tilde{y}})^s (1 - \frac{\tilde{a}_\infty}{1-\tilde{y}})^{k-1-s}$, as shown in *SI Appendix*, Fig. S17F. Therefore, the self-consistent equation to derive the conditional probability $\tilde{a}_\infty$ is

$$\begin{aligned}
\tilde{a}_\infty &= \sum_{k=1}^{m} \frac{kP(k)}{\langle k \rangle} \sum_{s=1}^{k-1} \binom{k-1}{s} \tilde{y}^s (1 - \tilde{y})^{k-1-s} \\
&\quad \times \left[1 - (1 - \frac{\tilde{y}_\infty}{\tilde{y}})^s (1 - \frac{\tilde{a}_\infty}{1-\tilde{y}})^{k-1-s} \right] \\
&= \sum_{k=1}^{m} \frac{kP(k)}{\langle k \rangle} \left[1 - (1 - \tilde{y})^{k-1} - (1 - \tilde{y}_\infty - \tilde{a}_\infty)^{k-1} \right. \\
&\quad \left. + (1 - \tilde{y} - \tilde{a}_\infty)^{k-1} \right]. \qquad [12]
\end{aligned}$$

Finally, the conditional probability $\tilde{y}_\infty$ can be obtained similarly to $\tilde{v}_\infty$, with the additional consideration that for $k - 1$ neighbors except l at least one can keep j active and that node j connects to the GCC via at least one of the $k - 1$ neighbors. Thus, the degree of node j satisfies $k > m$, which also implies that j keeps all its neighbors active. Therefore, the conditional probabilities $\tilde{y}$, $\tilde{y}_\infty$, and $\tilde{a}_\infty$ in Eq. **12** are replaced by probabilities $\tilde{v}$, $\tilde{v}_\infty$, and $\tilde{t}_\infty$. This leads to the following expression for the conditional probability $\tilde{y}_\infty$:

$$\begin{aligned}
\tilde{y}_\infty &= \sum_{k=m+1}^{\infty} \frac{kP(k)}{\langle k \rangle} \sum_{s=1}^{k-1} \binom{k-1}{s} \tilde{v}^s (1 - \tilde{v})^{k-1-s} \\
&\quad \left[1 - (1 - \frac{\tilde{v}_\infty}{\tilde{v}})^s (1 - \frac{\tilde{t}_\infty}{1-\tilde{v}})^{k-1-s} \right] \\
&= \sum_{k=m+1}^{\infty} \frac{kP(k)}{\langle k \rangle} \left[1 - (1 - \tilde{v})^{k-1} - (1 - \tilde{t}_\infty - \tilde{v}_\infty)^{k-1} \right. \\
&\quad \left. + (1 - \tilde{v} - \tilde{t}_\infty)^{k-1} \right], \qquad [13]
\end{aligned}$$

where s represents the number of neighbors that can keep j active. The graphical solution of the self-consistent equation $\tilde{y}_\infty$ is shown in the main text, where $f(\tilde{y}_\infty) = F(\tilde{y}_\infty) - \tilde{y}_\infty$ and $F(\tilde{y}_\infty)$ represents the expression on the right-hand side of Eq. **13**. The value of $F(\tilde{y}_\infty)$ is obtained by solving the self-consistent Eqs. **8**–**12**.

The previously defined conditional probabilities allow us to derive the order parameter, P_∞, for induced percolation on undirected networks. For an arbitrarily chosen node l to belong to the GCC, we have that 1) at least one of its neighbors should keep it active and 2) node l is attached to the GCC through at least one of its neighbors. If the degree of node l satisfies $k \leq m$, then the probability that node l belongs to the GCC is $1 - (1 - \tilde{y})^k - (1 - \tilde{y}_\infty - \tilde{a}_\infty)^k + (1 - \tilde{y} - \tilde{a}_\infty)^k$, whose derivation is similar to Eq. **12** in $\tilde{a}_\infty$. If the degree of node l satisfies $k > m$, then the probability that node l belongs to the GCC is $1 - (1 - \tilde{v})^k - (1 - \tilde{t}_\infty - \tilde{v}_\infty)^k + (1 - \tilde{v} - \tilde{t}_\infty)^k$ and the derivation is similar to $\tilde{y}_\infty$ in Eq. **13**. Therefore, the order parameter P_∞ can be computed, for undirected networks, as

$$\begin{aligned}
P_\infty &= \sum_{k=0}^{m} P(k) \left[1 - (1 - \tilde{y})^k - (1 - \tilde{y}_\infty - \tilde{a}_\infty)^k + (1 - \tilde{y} - \tilde{a}_\infty)^k \right] \\
&\quad + \sum_{k=m+1}^{\infty} P(k) \left[1 - (1 - \tilde{v})^k - (1 - \tilde{t}_\infty - \tilde{v}_\infty)^k + (1 - \tilde{v} - \tilde{t}_\infty)^k \right]. \\
&\qquad [14]
\end{aligned}$$

Xie et al.
Indirect influence in social networks as an induced percolation phenomenon

PNAS | 9 of 10
https://doi.org/10.1073/pnas.2100151119

1. D. Stauffer, A. Aharony, *Introduction to Percolation Theory* (Taylor and Francis, ed. 2nd Rev., 1994).
2. P. J. Flory, Molecular size distribution in three dimensional polymers. I. Gelation. *J. Am. Chem. Soc.* **63**, 3083–3090 (1941).
3. W. H. Stockmayer, Theory of molecular size distribution and gel formation in branched-chain polymers. *J. Chem. Phys.* **11**, 45–55 (1943).
4. M. E. J. Newman, S. H. Strogatz, D. J. Watts, Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **64**, 026118 (2001).
5. D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* **85**, 5468–5471 (2000).
6. R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Resilience of the internet to random breakdowns. *Phys. Rev. Lett.* **85**, 4626–4628 (2000).
7. D. J. Watts, A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5766–5771 (2002).
8. S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275–1335 (2008).
9. C. Castellano, S. Fortunato, V. Loreto, Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646 (2009).
10. Y. Hu *et al.*, Local structure can identify and quantify influential global spreaders in large scale social networks. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7468–7472 (2018).
11. A. Bashan, Y. Berezin, S. V. Buldyrev, S. Havlin, The extreme vulnerability of interdependent spatially embedded networks. *Nat. Phys.* **9**, 667–672 (2013).
12. R. Parshani, S. V. Buldyrev, S. Havlin, Interdependent networks: Reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys. Rev. Lett.* **105**, 048701 (2010).
13. S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, S. Havlin, Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).
14. J. Gao, S. V. Buldyrev, H. E. Stanley, S. Havlin, Networks formed from interdependent networks. *Nat. Phys.* **8**, 40–48 (2012).
15. C. D. Brummitt, R. M. D'Souza, E. A. Leicht, Suppressing cascades of load in interdependent networks. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E680–E689 (2012).
16. J. Xie *et al.*, Detecting and modelling real percolation and phase transitions of information on social media. *Nat. Hum. Behav.* **5**, 1161–1168 (2021).
17. R. Albert, H. Jeong, A. L. Barabási, Error and attack tolerance of complex networks. *Nature* **406**, 378–382 (2000).
18. G. J. Baxter, S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Bootstrap percolation on complex networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **82**, 011103 (2010).
19. S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, *k*-core organization of complex networks. *Phys. Rev. Lett.* **96**, 040601 (2006).
20. G. J. Baxter, S. N. Dorogovtsev, K. E. Lee, J. F. F. Mendes, A. V. Goltsev, Critical dynamics of the *k*-core pruning process. *Phys. Rev. X* **5**, 031017 (2015).
21. J. H. Zhao, H. J. Zhou, Y. Y. Liu, Inducing effect on the percolation transition in complex networks. *Nat. Commun.* **4**, 2412 (2013).
22. M. Granovetter, Threshold models of collective behavior. *Am. J. Sociol.* **83**, 1420–1443 (1978).
23. D. Kempe, J. Kleinberg, É. Tardos, "Maximizing the spread of influence through a social network" in *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM, 2003), pp. 137–146.
24. F. Morone, H. A. Makse, Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68 (2015).
25. D. Achlioptas, R. M. D'Souza, J. Spencer, Explosive percolation in random networks. *Science* **323**, 1453–1455 (2009).
26. M. Bauer, O. Golinelli, Core percolation in random graphs: A critical phenomena analysis. *Eur. Phys. J. B* **24**, 339–352 (2001).
27. Y. Y. Liu, E. Csóka, H. Zhou, M. Pósfai, Core percolation on complex networks. *Phys. Rev. Lett.* **109**, 205703 (2012).
28. R. Cohen, S. Havlin, *Complex Networks: Structure, Stability and Function* (Cambridge University Press, 2010).
29. M. E. J. Newman, *Networks* (Oxford University Press, 2018).
30. N. A. Christakis, J. H. Fowler, The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357**, 370–379 (2007).
31. J. H. Fowler, N. A. Christakis, Cooperative behavior cascades in human social networks. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 5334–5338 (2010).
32. P. R. Guimarães Jr., M. M. Pires, P. Jordano, J. Bascompte, J. N. Thompson, Indirect effects drive coevolution in mutualistic networks. *Nature* **550**, 511–514 (2017).
33. T. Ohgushi, O. Schmitz, R. Holt, *Trait-Mediated Indirect Interactions: Ecological and Evolutionary Perspectives* (Ecological Reviews, Cambridge University Press, 2012).
34. J. M. Lehn, Supramolecular chemistry. *Science* **260**, 1762–1763 (1993).
35. J. Gierschner *et al.*, Excitonic versus electronic couplings in molecular assemblies: The importance of non-nearest neighbor interactions. *J. Chem. Phys.* **130**, 044105 (2009).
36. A. E. Rudolph, N. D. Crawford, C. Latkin, J. H. Fowler, C. M. Fuller, Individual and neighborhood correlates of membership in drug using networks with a higher prevalence of HIV in New York City (2006-2009). *Ann. Epidemiol.* **23**, 267–274 (2013).
37. J. N. Rosenquist, J. Murabito, J. H. Fowler, N. A. Christakis, The spread of alcohol consumption behavior in a large social network. *Ann. Intern. Med.* **152**, 426–433, W141 (2010).
38. Y. Hu, S. Havlin, H. A. Makse, Conditions for viral influence spreading through multiplex correlated social networks. *Phys. Rev. X* **4**, 021031 (2014).
39. A. L. Nguyen, W. Liu, K. A. Khor, A. Nanetti, S. A. Cheong, The golden eras of graphene science and technology: Bibliographic evidences from journal and patent publications. *J. Informetrics* **14**, 101067 (2020).
40. R. A. da Costa, S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Explosive percolation transition is actually continuous. *Phys. Rev. Lett.* **105**, 255701 (2010).
41. M. Catanzaro, R. Pastor-Satorras, Analytic solution of a static scale-free network model. *Eur. Phys. J. B Cond. Matter Complex Syst.* **44**, 241–248 (2005).
42. K. I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks. *Phys. Rev. Lett.* **87**, 278701 (2001).
43. O. Riordan, L. Warnke, Explosive percolation is continuous. *Science* **333**, 322–324 (2011).
44. P. Grassberger, C. Christensen, G. Bizhani, S. W. Son, M. Paczuski, Explosive percolation is continuous, but with unusual finite size behavior. *Phys. Rev. Lett.* **106**, 225701 (2011).
45. L. Tian, A. Bashan, D. N. Shi, Y. Y. Liu, Articulation points in complex networks. *Nat. Commun.* **8**, 14223 (2017).

10 of 10 | PNAS
https://doi.org/10.1073/pnas.2100151119

Xie et al.
Indirect influence in social networks as an induced percolation phenomenon