




Emergence of nonlinear crossover under epidemic dynamics in heterogeneous networksZhen Su ^{1,2,3}, Chao Gao ^{1,3,*}, Jiming Liu,⁴ Tao Jia ³, Zhen Wang,¹ and Jürgen Kurths^{2,5,6}¹The Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xian 710072, China²The Potsdam Institute for Climate Impact Research (PIK), Potsdam 11473, Germany³College of Computer and Information Science, Southwest University, Chongqing 400715, China⁴Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong⁵The Institute of Physics, Humboldt University of Berlin, Berlin 12489, Germany⁶Nizhny Novgorod State University, Nizhny Novgorod, Russia

(Received 15 August 2020; accepted 28 October 2020; published 20 November 2020)

Potential diffusion processes of real-world systems are relevant to the underlying network structure and dynamical mechanisms. The vast majority of the existing work on spreading dynamics, in response to a large-scale network, is built on the condition of the infinite initial state, i.e., the extremely small seed size. The impact of an increasing seed size on the persistent diffusion has been less investigated. Based on classical epidemic models, this paper offers a framework for studying such impact through observing a crossover phenomenon in a two-diffusion-process dynamical system. The two diffusion processes are triggered by nodes with a high and low centrality, respectively. Specifically, given a finite initial state in networks with scale-free degree distributions, we demonstrate analytically and numerically that the diffusion process triggered by low centrality nodes pervades faster than that triggered by high centrality nodes from a certain point. The presence of the crossover phenomenon reveals that the dynamical process under the finite initial state is far more than the vertical scaling of the spreading curve under an infinite initial state. Further discussion emphasizes the persistent infection of individuals in epidemic dynamics as the essential reason rooted in the crossover, while the finite initial state is the catalyst directly leading to the emergence of this phenomenon. Our results provide valuable implications for studying the diversity of hidden dynamics on heterogeneous networks.

DOI: [10.1103/PhysRevE.102.052311](https://doi.org/10.1103/PhysRevE.102.052311)**I. INTRODUCTION**

Diffusion, running on top of varieties of real-world complex systems, has attracted increasing interest in capturing the underlying dynamics [1–3]. Many studies on modeling dynamical process are based on the infinite initial state—a quite small proportion of seed nodes to trigger a propagation [4–11]. Yet such a theoretical assumption faces challenges in realistic diffusion scenarios. For example, at the initial stage of coronavirus, our lack of knowledge on this dangerous pandemic underestimates flulike symptoms, making it difficult to precisely estimate the potential number of infected individuals [12]. Despite the importance of tracing undocumented infection sources, controlling the evolution of spreading curve is the priority under this noticeable finite initial state [13,14].

The infinite initial state would benefit, in particular, high-centrality-based diffusion, since the overlap of individual influence is weak. Under such initial state, practical applications include the network immunization [15,16], source identification [17], or influence maximization [18].

While, for the finite initial state, the overlap of individual influence cannot be omitted, due to the great impact on the subsequent dynamical spreading process. Specifically, the high centrality nodes are centralized at the core positions and

have the tendency to be highly clustered with each other, leading to the rich-club phenomenon [19]. Therefore, it is more likely that the overlap of individual influence emerges from the beginning of the high-centrality-based diffusion [18,20,21]. For the low-centrality-based diffusion, the overlap is unlikely to show up at the initial stage, since low centrality nodes scatter in the boundary of a network. In this backdrop, we might witness the low-centrality-based diffusion reaching a larger propagation scale first.

Taking Fig. 1 as an example, we define a crossover as a critical point—after which the propagation scales, triggered by nodes with a high and a low centrality, respectively, are reversed. Figure 1(b₃) shows such a crossover case where the lowest-degree-based diffusion [Fig. 1(b₁)] pervades the whole network first. Despite the extreme nature of this case, it illustrates the difficulty of using central nodes to infect those boundary ones, and the strong impact of the finite initial state on the dynamics underlying. Naturally, it is worth investigating the following questions, of which, however, a detailed inspection is still missing: (1) Since when does the finite initial state start to impose such an impact on a dynamical process? (2) How does this impact evolve over the finite initial state?

To fill this gap, we aim to unfold the emergence of the crossover phenomenon under different epidemic dynamics [e.g., susceptible-infected (SI), susceptible-infected-susceptible (SIS), and susceptible-infected-recovered (SIR)] [22], centralities [e.g., the degree (DC), betweenness (BC),

*Corresponding author: cgao@swu.edu.cn

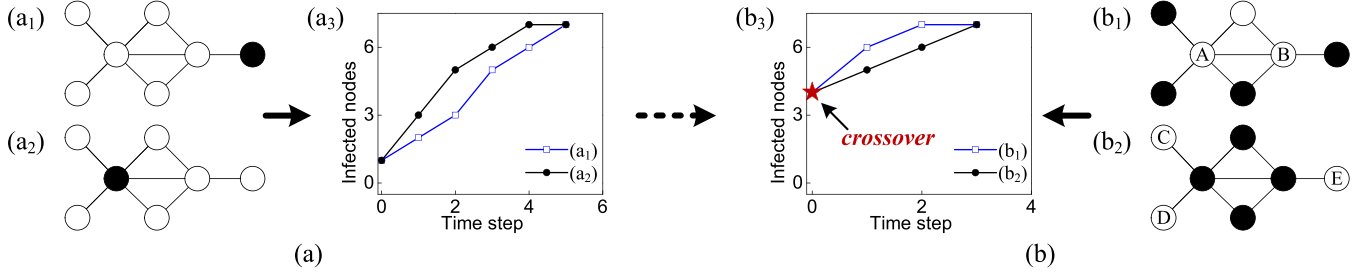


FIG. 1. A schematic illustration of the effect of the finite initial state, i.e., the increasing proportion of the initial seed nodes. Based on the SI model, a susceptible node becomes infected with a probability λ ($\lambda \ll 1$ for approximating a continuous-time (CT) infection probability as a discrete-time (DT) one [23]). (a) Two kinds of initial states with one seed node for each and (b) two kinds of initial states with four seed nodes for each, both represent a special dynamical system consisting of two independent but simultaneous diffusion processes. For (a₁), the probability that all neighbors of the seed node remain susceptible in the next time step is $1 - \lambda$, while it changes into $(1 - \lambda)^5$ in (a₂). This demonstrates the capability of high-degree nodes in accelerating diffusion. Therefore, (a₂) could reach a stable state earlier than (a₁), as given in (a₃). However, in (b₁), A and B remain susceptible in the next time step with probabilities $(1 - \lambda)^3$ and $(1 - \lambda)^2$, respectively, while in (b₂), both C and D remain susceptible with a probability $(1 - \lambda)^2$, and E is with $1 - \lambda$. It is more likely that (b₁) reaches the stable state first, leading to the critical crossover point in (b₃) and also revealing the great impact of the finite initial state.

and eigenvector centralities (EC)] [18], and simulation approaches [i.e., continuous-time (CT) and discrete-time (DT) Monte Carlo simulations] [23]. We find that disassortative connections between low-degree and high-degree nodes act as a bridge in driving a globally large-scale diffusion, and the reversibility of dynamical processes can shape the behavior of a crossover. More importantly, although the increasing proportion of seed nodes triggers the emergence of the crossover, this phenomenon essentially stems from an individual's persistent infection in epidemic models.

II. ANALYTICAL EVIDENCE OF THE EMERGENCE OF THE CROSSOVER PHENOMENON

First, dynamical processes on a network $G = (V, E)$ with $|V|$ nodes and $|E|$ edges are considered to follow classical epidemic models. In the SI dynamical system, each node has only two corresponding states, either susceptible or infected. The only transition, denoted as $S \rightarrow I$, is achieved with a probability λ by successful interactions between a susceptible node and its infected neighbors. For the SIS model, the reversible transition between the susceptible and infected states contains $S \rightarrow I$ with a probability λ similar to that in the SI model and $I \rightarrow S$ with a probability μ realized by the spontaneous recover from a I state to a S one while, in the SIR process, the second transition, i.e., $I \rightarrow R$, is implemented with a probability β , signaling a permanent immunity or becoming removed. Then, since CT approach is often combined with the infinitesimally small initial conditions, we study DT approximation to characterize two kinds of initial states. It should be noted that $\lambda \ll 1$ is a fundamental assumption for approximating the exact infection probability of a CT-based Markov process as a DT-based one [23]. Moreover, to ensure a visible diffusion for the analysis of the crossover phenomenon, we let $\mu, \beta < \lambda \ll 1$.

In what follows, a general framework for investigating the underlying crossover phenomenon (as illustrated in Fig. 1) is provided. We take the DC as an example to derive the possibility of a crossover. The analytical results can be extended to other centralities due to the correlation between them [24].

We start from the proportion of infected nodes in the k -degree set [denoted as $i_k(t)$]. The dynamical evolution of $i_k(t)$ can be written following Ref. [25] as

$$i_k(t) \equiv i_k(t-1) + \Delta i_{k(\cdot)}(t-1), \quad (1)$$

where $\Delta i_{k(\cdot)}(t-1)$ accounts for the proportion of newly generated infected nodes from $t-1$ to t . The basic quantification is with $\Delta i_{k(\text{SI})}(t-1) = \lambda k[1 - i_k(t-1)]\Theta(t-1)$ for the SI dynamical process. Additionally, in the SIS and SIR models, the recovery processes need to be taken into account, leading to $\Delta i_{k(\text{SIS})}(t-1) = \lambda k[1 - i_k(t-1)]\Theta(t-1) - \mu i_k(t-1)$ and $\Delta i_{k(\text{SIR})}(t-1) = \lambda k[1 - i_k(t-1) - r_k(t-1)]\Theta(t-1) - \beta i_k(t-1)$, respectively. Here, $r_k(t)$, the proportion of removed nodes in the k -degree set of the SIR process, is defined by $r_k(t) \equiv r_k(t-1) + \beta i_k(t-1)$ with $r_k(0) \equiv 0$. The variable $\Theta(t)$ represents the probability of targeting an infected node through an arbitrarily chosen link. Such value is given by $\Theta(t) = (1/\langle k \rangle) \sum_k k P(k) i_k(t)$ [4]. Considering the mean-field (MF) character imposed on the calculation of $\Theta(t)$ and the increasing proportion of the initial seed nodes, we assume $\Theta(t) \approx i_k(t)$ for focusing on the influence of the finite initial state and derive

$$i_k(t) = i_k(0) + \Delta i_{k(\cdot)}(0 \rightarrow t), \quad (2)$$

where $\Delta i_{k(\cdot)}(0 \rightarrow t)$ embodies the proportion of generated infected nodes during a diffusion process. In the SI process, $\Delta i_{k(\text{SI})}(0 \rightarrow t) \approx \lambda k t [1 + \frac{1}{2}(t-1)\lambda k] i_k(0) - \lambda k t [1 + \frac{3}{2}(t-1)\lambda k] i_k^2(0)$ is obtained by neglecting higher order terms with respect to λ and $i_k(0)$. Similarly, with higher order terms concerning λ, μ, β , and $i_k(0)$ being neglected, we have $\Delta i_{k(\text{SIS})}(0 \rightarrow t) \approx (\lambda k - \mu) t [1 + \frac{1}{2}(t-1)(\lambda k - \mu)] i_k(0) - \lambda k t [1 + \frac{3}{2}(t-1)(\lambda k - \mu)] i_k^2(0)$ and $\Delta i_{k(\text{SIR})}(0 \rightarrow t) \approx (\lambda k - \beta) t [1 + \frac{1}{2}(t-1)(\lambda k - \beta)] i_k(0) - \lambda k t [1 + (t-1)(\frac{3}{2}\lambda k - \beta)] i_k^2(0)$ for the SIS and SIR models, respectively.

Next, to analyze the presence of a crossover needs to consider whether the diffusion triggered by seed nodes with the lowest centrality, can propagate faster than that triggered by seed nodes with the highest centrality. For this purpose, an

auxiliary function is defined

$$\Phi(\bar{i}_k(0), t) = I^{\min}(t) - I^{\max}(t), \quad (3)$$

in which the proportion of infected nodes at t , i.e., $I(t) = \sum_k i_k(t)P(k)$ represents the propagation scale. The two independent but simultaneous diffusion processes are labeled as min and max, respectively. We let the variable $\bar{i}_k(0)$ quantify the average effect of the finite initial state. Without loss of generality, it is given by

$$\bar{i}_k(0) = \frac{\sum_k i_k(0)}{|K|}. \quad (4)$$

$|K|$ is defined as the number of different degrees of all nodes in a network and typically $\bar{i}_k(0) \in [0, 1]$. Note that the $\bar{i}_k(0)$ is defined for the k -degree node set at $t = 0$ and is less intuitive to capture the increasing proportion of the initial seed nodes, we thereby use the $I(0)$, in the next section, as the indicator of the finite initial state. The assumption of $\Theta(t) \approx i_k(t)$ in Eq. (2) also implies that $i_k(t)$ is independent of $\sum_k kP(k)$, we

adopt it in the derivation of Eq. (3) (see Supplemental Material [26]).

Finally, the emergence of a crossover will be met when $\Phi(\bar{i}_k(0), t) > 0$. It should be noted that the SF degree distribution, i.e., $P(k) \sim k^{-\gamma}$, allows the presence of $\bar{i}_k^{\min}(0) < \bar{i}_k^{\max}(0)$. We use this condition to derive the solution of crossover time, i.e., $t_{c(\cdot)}$ of Eqs. (5). Then, $t_{c(\text{SI})} > 0$, $t_{c(\text{SIS})} > 0$, and $t_{c(\text{SIR})} > 0$ hold if and only if $\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0) > \frac{1}{3}$ in the SI model, $\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(\text{SIS})(0) > \frac{\lambda^2 \langle k^2 \rangle - 2\lambda \mu \langle k \rangle + \mu^2}{3\lambda(\lambda \langle k^2 \rangle - \mu \langle k \rangle)}$ in the SIS model, and $\bar{i}_k^{\min}(\text{SIR})(0) + \bar{i}_k^{\max}(\text{SIR})(0) > \frac{\lambda^2 \langle k^2 \rangle - 2\lambda \beta \langle k \rangle + \beta^2}{3\lambda^2 \langle k^2 \rangle - 2\lambda \beta \langle k \rangle}$ in the SIR model, respectively, as demonstrated in the Supplemental Material [26].

The use of the MF assumption yields an overestimation of the propagation scale [2], further resulting in some inaccuracy of the crossover time. However, the presence of this phenomenon can still be guaranteed by (1) $t_{c(\cdot)} > 0$ (see Supplemental Material [26]) and (2) $t_{c(\cdot)}$ tends to be lower as $\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0)$ increases. That is, the tendency of this phenomenon being earlier could be witnessed with the increment of the proportion of the initial seed nodes. Thus, the focus in the next section is on the dynamic trend of this phenomenon:

$$\begin{aligned} t_{c(\text{SI})} &= \left[1 + \frac{2\langle k \rangle [1 - (\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0))]}{\lambda \langle k^2 \rangle [3(\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0)) - 1]} \right], \\ t_{c(\text{SIS})} &= \left[1 + \frac{2\{\lambda \langle k \rangle [1 - (\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0))] - \mu\}}{3\lambda(\lambda \langle k^2 \rangle - \mu \langle k \rangle)(\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0)) - (\lambda^2 \langle k^2 \rangle - 2\lambda \mu \langle k \rangle + \mu^2)} \right], \\ t_{c(\text{SIR})} &= \left[1 + \frac{2\{\lambda \langle k \rangle [1 - (\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0))] - \beta\}}{(3\lambda^2 \langle k^2 \rangle - 2\lambda \beta \langle k \rangle)(\bar{i}_k^{\min}(0) + \bar{i}_k^{\max}(0)) - (\lambda^2 \langle k^2 \rangle - 2\lambda \beta \langle k \rangle + \beta^2)} \right]. \end{aligned} \quad (5)$$

III. NUMERICAL EVIDENCE OF THE EMERGENCE OF THE CROSSOVER PHENOMENON

To reveal the emergence of the crossover phenomenon, we adopt both CT (the Gillespie algorithm [27]) and DT Monte Carlo simulations. For the DT-based simulations, the probability that a susceptible node becomes infected within each time step $\Delta(t) = 1$ depends on the number of its infected neighbors. For example, if the probability λ is for the case that there is only one infected neighbor node, then, for n infected neighbors, the probability is $1 - (1 - \lambda)^n$. However, for the CT-based simulations, $\Delta(t)$ is a random variable following an

exponential distribution of which the parameter is given by the sum of transition rates of all nodes. These transition rates are quantified based on the number of infected neighbors accordingly. For example, in the SIS dynamics, if a susceptible node has n infected neighbors, the infection and recovery rates of this node are $n\tilde{\lambda}$ and $\tilde{\mu}$, respectively. $\tilde{\lambda}$ and $\tilde{\mu}$ are instantaneous transition rates by taking the limit of $\Delta t \rightarrow 0$ [23].

According to the assumption of $\mu, \beta < \lambda \ll 1$ in the previous section, we let $\lambda = 0.01$ and $\mu = \beta = 0.005$, which are also applicable to $\tilde{\lambda}$, $\tilde{\mu}$, and $\tilde{\beta}$. To ensure stable states in different dynamical systems, the total time step t is specified with

TABLE I. Structural properties of used networks, including the number of nodes $|V|$ and edges $|E|$, the average degree $\langle k \rangle$, the exponent α of degree distribution, the average clustering coefficient c , and the degree assortativity coefficient r .

Network	LFR ₁ [28]	LFR ₂ [28]	LFR ₃ [28]	BA ₁ [29]	BA ₂ [29]	BA ₃ [29]	PB [30]	WV [31]	BY [32]	EF [33]	HEP-PH [34]	AS [35]
$ V $	1000	5000	5000	10 000	10 000	10 000	793	1300	2224	4039	11 204	22 963
$ E $	24 818	47 958	126 027	39 984	69 951	99 900	13 484	36 529	6609	88 234	117 619	48 436
$\langle k \rangle$	49.636	19.183	50.411	7.997	13.990	19.980	34.008	56.198	5.943	43.691	20.996	4.219
α	3.50	2.73	3.01	2.81	2.86	2.90	3.50	3.09	3.26	2.51	2.09	2.09
c	0.733	0.19	0.314	0.006	0.009	0.011	0.396	0.297	0.201	0.617	0.69	0.35
r	0.260	-0.096	-0.101	-0.031	-0.021	-0.018	-0.217	-0.117	-0.105	0.064	0.630	-0.198

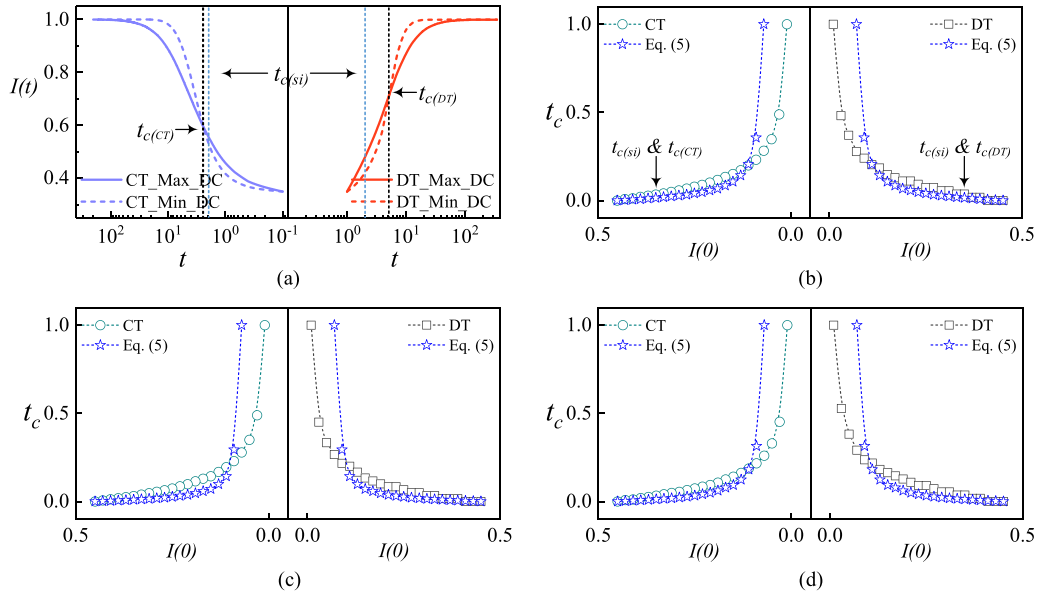


FIG. 2. The emergence of crossover phenomenon in WV network based on the SI model. (a) presents the propagation scale $I(t)$, as a function of t . When $I(0) = 0.35$, $t_{c(CT)}$ and $t_{c(DT)}$ are obtained by using continuous-time (CT) (left) and discrete-time (DT) (right) Monte Carlo simulations, respectively, while $t_{c(SI)}$ is the analytical point based on Eq. (5). The solid and dash lines represent diffusion processes triggered by seed nodes with the lowest and highest degree centrality (DC), respectively. (b)–(d) show the normalized crossover time $t_{c(-)}$, as a function of the initial proportion of seed nodes $I(0)$, under three centralities. The lower analytical and numerical $t_{c(-)}$ indicates the earlier presence of a crossover as $I(0)$ increases.

1000 in our paper. Each stochastic diffusion is obtained by averaging 100 realizations. For each realization, we choose two groups of seed nodes based on the DC (neighborhood-based), BC (path-based), and EC (iterative-refinement-based) [18] to trigger two independent but simultaneous diffusion processes. Let us consider the descending order of DC as an example. One group is selected from the first $(100 \times I(0))\%$ nodes, $I(0) \in [0.01, 0.45]$, to initiate the highest-degree-based diffu-

sion; the other is chosen from the last $(100 \times I(0))\%$ nodes to initiate the lowest-degree-based diffusion. Numerical simulations are performed in both synthetic and real-world networks whose details are given in Table I. The numerical crossover time is also obtained, when the propagation scale of the lowest-centrality-based diffusion is larger than that of the highest-centrality-based one. We normalize both the analytical and the numerical $t_{c(-)}$ for the trend analysis.

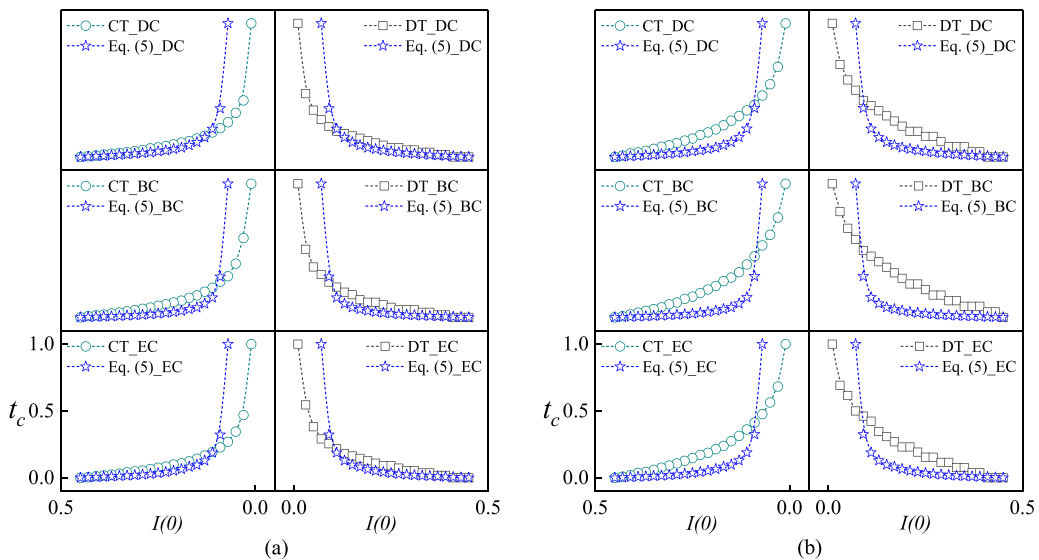


FIG. 3. The crossover time $t_{c(-)}$, as a function of the proportion of the initial seed nodes $I(0)$ in WV network in the (a) SIS and (b) SIR models. CT and DT represent continuous-time and discrete-time Monte Carlo simulations, respectively. DC, BC, and EC are the abbreviations of the degree, betweenness, and eigenvector centralities, respectively. Similar to Fig. 2, it also indicates the emergence of crossover which tends to be intensified when $I(0)$ increases.

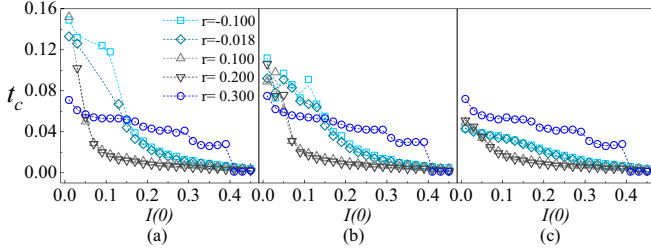


FIG. 4. The numerical crossover time $t_{c(\cdot)}$, as a function of the proportion of the initial seed nodes $I(0)$, in five networks with different degree assortativity coefficients r . (a)–(c) report discrete-time-based Monte Carlo simulation results under the degree centrality (DC) in the SI, SIS, and SIR models, respectively. A lower value of $t_{c(\cdot)}$ represents the earlier emergence of a crossover. Apart from $I(0)$, the recovery process and the degree assortativity could make a crossover earlier.

Taking the WV network (see Table I) as an example (more results in [26]), the emergence of crossover phenomenon shown in Figs. 2 and 3 is in agreement with the analytical results of Eq. (5). The dynamic trend of this phenomenon starts from a relatively sharp decrease, then to a smooth one gradually, implying the presence of an earlier crossover. Moreover, this phenomenon is independent of selected epidemic models, centralities, and simulation approaches.

A. The epidemic-model-free crossover

We observe a slower decrease trend of the numerical $t_{c(\cdot)}$ from Fig. 3(b) compared with Fig. 3(a). To explain this, let us first borrow the unnormalized results of $r = -0.018$, i.e., the BA₃ network, from Fig. 4. It shows the emergence order: SI > SIS > SIR (earliest). Such an order could be explained by the spontaneously irreversible transition of $I \rightarrow R$ in the SIR dynamical process. This transition largely restrains the highest-centrality-based diffusion by disabling part of high centrality nodes permanently while, for the SIS process, the reversible transitions between S and I slightly relieve the suppression of $I \rightarrow R$ imposed on the highest-centrality-based diffusion. It is therefore why we have this emergence order in Fig. 4 and also the slower decrease trend in Fig. 3(b).

B. The centrality-free crossover

Numerical results indicate the consistency of the dynamic trend of the crossover time under different centralities, even when the DC~BC in LFR₁, EF, and HEP-PH networks (see Table II) has a relatively low value. This implies the centrality-free feature of this phenomenon and the extensibility of the analytical results of Eq. (5) to other centralities.

TABLE II. The Pearson correlation coefficient (PCC) between the degree centrality (DC) and betweenness centrality (BC), the DC and eigenvector centrality (EC), respectively, in SF networks of Table I. This table shows that in most cases, the DC is highly correlated to the BC and EC.

Network	LFR ₁ [28]	LFR ₂ [28]	LFR ₃ [28]	BA ₁ [29]	BA ₂ [29]	BA ₃ [29]	PB [30]	WV [31]	BY [32]	EF [33]	HEP-PH [34]	AS [35]
DC ~ BC	0.47	0.92	0.97	0.90	0.90	0.90	0.76	0.76	0.88	0.45	0.48	0.96
DC ~ EC	0.92	0.93	0.96	0.91	0.95	0.97	0.94	0.96	0.85	0.78	0.91	0.85

C. The simulation-approach-free crossover

Under the assumption of $\mu, \beta < \lambda \ll 1$, both CT and DT Monte Carlo simulations show consistent results regarding the emergence of a crossover and its trend. This further demonstrates the conclusion in Ref. [23], that to ensure the accurate application of the DT infection probability to the CT process, the time step $\Delta(t)$ or parameters (e.g., λ , μ , and β) should be controlled to be as small as possible.

Considering the overestimation of Eq. (5), i.e., $t_{c(\text{SI})} < t_{c(\text{CT})}$ and $t_{c(\text{DT})}$ as illustrated in Fig. 2(a), it is still worth mentioning the reliability of the analytical and the numerical solutions. For one thing, our assumption of $\Theta(t) \approx i_k(t)$ neglects the degree correlations of networks and gives the only consideration to the average degree $\langle k \rangle$. The propagation speed during a diffusion process is then overestimated. Consequently, for example, compared to the numerical results, the analytical $t_{c(\cdot)}$ has a sharp decline when $I(0)$ is close to 0 [see Figs. 2 and 3]. However, the emergence of this phenomenon is still guaranteed by analytical solutions to the dynamic trend. For another, the numerical solutions are obtained by using Monte Carlo simulations, which can be seen as a relatively empirical demonstration for a crossover. Therefore, both the analytical and the numerical solutions provide valid evidence for the presence of this phenomenon.

IV. DISCUSSION

In this section, we further analyze the effect of assortativity and epidemiological parameters on the crossover phenomenon. More importantly, we concentrate on the persistent infection characteristic in epidemics dynamics to figure out the reason rooted in this phenomenon. As we mentioned in the previous sections, our analytical framework neglects the degree correlation. We here therefore mainly consider the numerical solutions since they provide a relatively accurate characterization for this phenomenon.

A. The effect of degree assortativity

In our analytical framework, the crossover time is relevant to the average degree $\langle k \rangle$ [see Eq. (5)]. However, $\langle k \rangle$ is not sufficient to characterize the degree correlation of a network, because any two networks sharing the same number of nodes and edges, would generate the same $\langle k \rangle$. Considering that the rich-club coefficient ϕ [19] is defined as a function of the degree k , we here choose the degree assortativity coefficient r [36] as a global indicator for our analysis. Specifically, this indicator captures the connection patterns between nodes with different degrees. For example, $r > 0$ indicates the assortative mixing pattern, e.g., high-degree nodes preferentially attaching to high-degree nodes, while $r < 0$ shows the disassortative

TABLE III. Structural properties of generated networks based on BA_3 , including the number of nodes $|V|$ and edges $|E|$, the average degree $\langle k \rangle$, the exponent α of degree distribution, and the average clustering coefficient c .

r	-0.100	-0.018 (BA_3)	0.100	0.200	0.300
$ V $	10 000	10 000	10 000	10 000	10 000
$ E $	99 900	99 900	99 900	99 900	99 900
$\langle k \rangle$	19.980	19.980	19.980	19.980	19.980
α	2.90	2.90	2.90	2.90	2.90
c	0.0076	0.0113	0.0104	0.0106	0.0194

mixing connection tendency, i.e., high-degree nodes attach to low-degree ones [36]. To obtain networks with different r , we first consider BA_3 as the basic network. By applying a series of rewiring process to any pair of randomly selected edges, r is tuned to a desirable value [37]. Note that this process preserves the number of nodes and edges, and also the degree of each node. That is, the average degree $\langle k \rangle$ and degree distribution of generated networks (see Table III) are the same as those of the BA_3 .

Similarly, we use $\lambda = 0.01$ and $\mu = \beta = 0.005$ as parameter settings in the SI, SIS, and SIR models. It is also worth noting that (1) the limitation of Eq. (5)—with a fixed average degree, stems from our assumption of neglecting the degree correlation; (2) The BC and EC, and CT Monte Carlo simulations are additional demonstrations for the crossover phenomenon. Thus, we only analyze DT Monte Carlo simulations under the DC in this part. Meanwhile, we also let $I(0) \in [0.01, 0.45]$ as the indicator characterizing the initial proportion of seed nodes. Instead of normalization for a dynamic trend analysis, numerical results are simply rescaled by $t_{c(\cdot)}/1000$ to see how different r affect $t_{c(\cdot)}$.

Figure 4 shows that the increasing values of r make $t_{c(\cdot)}$ lower until the absence of disassortative mixing pattern. Specifically, when the rewiring process increases r to be positive, e.g., from $r = -0.100$ or -0.018 to $r = 0.100$ or 0.200 , we have an increasing number of edges with assortative mixing pattern, and also higher values of $\phi(k)$ [see Fig. 5(a)]. As indicated in Figs. 4 and 5, the conditions for an earlier emergence of crossover are provided by (1) the existence of connections between low-degree nodes and high-degree nodes [see Figs. 5(b) and 5(c)] and (2) the presence of a higher $\phi(k)$ [see Fig. 5(a)]. Condition 1 ensures that even the lowest-

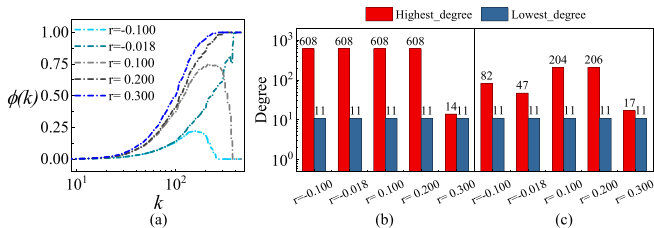


FIG. 5. Interpretations to Fig. 4(a) show the rich-club spectrum $\phi(k)$ as a function of k . For nodes with the lowest degree ($k = 10$), (b) and (c) report the highest and lowest degrees of their one-hop (except themselves) and two-hop (except themselves and one-hop neighbors) neighbors, respectively.

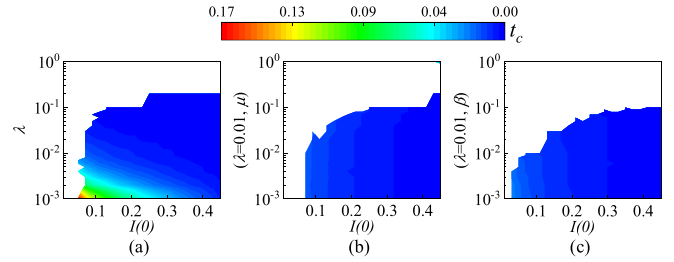


FIG. 6. The numerical crossover time $t_{c(\cdot)}$, as a function of the proportion of the initial seed nodes $I(0)$, epidemiological parameters λ , μ , and β , in WV network in the (a) SI, (b) SIS, and (c) SIR models. Based on discrete-time Monte Carlo simulations under the degree centrality (DC), (a)–(c) also report the absence of a crossover in white. For each model, we use $\Phi(\vec{i}_k(0), t) > 0.01 \times |V|$ to exclude the potential fluctuations from numerical simulations. It is clear that the behaviors of a crossover depend on whether the recovery process has the reversibility.

degree-based diffusion could reach high-degree nodes as soon as possible [see Fig. 1(b₁), for example] while condition 2 emphasizes the appearance of overlap of individual influence from the beginning of the highest-degree-based diffusion, as illustrated in Fig. 1(b₂). Therefore, the crossover emerges earlier in assortative networks. However, in a highly assortative network, the absence of connections where low-degree nodes attach to high-degree nodes, breaks the condition 1 for the lowest-degree-based diffusion [see Figs. 5(b) and 5(c)]. As a result, a high value of $t_{c(\cdot)}$ is observed in Fig. 4. This is also why the crossover phenomenon in the network with $r = 0.300$ remains almost the same under different epidemic models.

B. The effect of epidemiological parameters

Our analytical and numerical demonstrations in the previous section are given under fixed parameters with $\lambda = 0.01$, $\mu = \beta = 0.005$. We here take a wide range of μ , β , and λ , even when the assumption of $\mu, \beta < \lambda \ll 1$ is violated. Specifically, in the SI model, we let $\lambda \in [0.001, 1]$. For the SIS and SIR models, the underlying dynamics concern the effective spreading probability defined as $\frac{\lambda}{\mu}$ and $\frac{\lambda}{\beta}$, respectively. By fixing $\lambda = 0.01$, we let $\mu, \beta \in [0.001, 1]$, yielding $\frac{\lambda}{\mu}, \frac{\lambda}{\beta} \in [0.01, 10]$. Based on DT Monte Carlo simulations under the DC, this part also uses $t_{c(\cdot)}/1000$ and $I(0) \in [0.01, 0.45]$ to see how $t_{c(\cdot)}$ evolves over different effective spreading probabilities. Note that due to the potential fluctuations of numerical simulations, it is possible for $\Phi(\vec{i}_k(0), t) > 0$ to be recovered with $\Phi(\vec{i}_k(0), t) = 1$ or 2. Such cases usually appear even when the spreading process tends to be stable. While on the other hand, we may also observe $\Phi(\vec{i}_k(0), t) > 0$ even when one of the propagation scales $I(t)$ starts to decrease (see Sec. III in the Supplemental Material [26]). To avoid these situations, we further study $t_{c(\cdot)}$ when the following two conditions are satisfied: (1) $\Phi(\vec{i}_k(0), t) > 0.01 \times |V|$, where $|V|$ is the number of nodes in a network; and (2) both propagation scales (triggered by seed nodes with the highest and lowest DC, respectively) remain an increasing trend.

In Fig. 6, we observe that the reversibility of the recovery process shapes the behavior of the crossover phenomenon.

It should be pointed out that considering the assumption of $\mu, \beta < \lambda \ll 1$ and the conclusion in Ref. [23], we here mainly concern the numerical $t_{c(\cdot)}$ when λ, μ , and $\beta \in [0.001, 0.01]$.

In the irreversible transitions, when $I(0)$ is given, $t_{c(\cdot)}$ presents a decrease behavior as λ (in the SI model) increases or $\frac{\lambda}{\beta}$ (in the SIR model) decreases. Specifically, compared with low-degree nodes in the highest-degree-based diffusion, high-degree nodes in the lowest-degree-based diffusion have a relatively higher probability to be infected [see Fig. 1(b)]. This would be amplified with λ being increased during the SI process. While in the SIR dynamical process, the highest-degree-based diffusion shares relatively more suppression from the beginning than the lowest-degree-based diffusion, due to the permanent irreversible transition of $I \rightarrow R$ cutting off propagation paths from high-degree nodes to low-degree nodes. The higher β implies the higher probability to have this irreversible transition. Thus, we also witness a slight decline with β being increased in the SIR model [see Fig. 6(a)].

In the SIS model, $t_{c(\cdot)}$ has a relatively stable value under a specified $I(0)$. This could be explained by the reversible transitions between S and I . The transition of $S \rightarrow I$, recovering the propagation ability of temporarily susceptible nodes, is quite important for the highest-degree-based diffusion.

Besides, recall that Fig. 1(b) shows that compared with high-degree nodes, boundary nodes are less likely to be infected. During the lowest-degree-based diffusion, low-degree nodes take more time to finish the transition of $I \rightarrow S \rightarrow I$, because they have a quite limited number of infected neighbor nodes. While in the highest-degree-based diffusion, high-degree nodes realize this transition process relatively easily, due to the existence of the rich-club phenomenon. That is, the higher values of $\phi(k)$ ensure that those resusceptible high-degree nodes could be surrounded by a higher number of infected nodes. Therefore, Fig. 6(b) shows a different trend from Figs. 6(a) and 6(c).

C. The persistent infection characteristic in epidemic dynamics

We analyze above how the initial proportion of seed nodes $I(0)$, degree assortativity r , and epidemiological mechanism shape the behavior of a crossover. However, none of them could be portrayed as the reason rooted in the crossover phenomenon. Under the epidemic dynamics we discuss in this paper, there is a basic characteristic—persistent infection—guaranteeing the continuous propagation of a diffusion. Specifically, this condition ensures the equivalent propagation scale $I(t)$ in both the highest-degree-based and the lowest-degree-based diffusions, as illustrated in Figs. 1 and 2(a). To give validation to this assumption, we compare (1) the persistent-infection-based dynamics, i.e., the SI, SIS, and SIR models with (2) the non-persistent-infection-based dynamics, i.e., the independent cascading (IC) model [38], under the same parameters [see Fig. 7].

When $I(0) = 0.1$, all epidemic dynamical processes can reach a large-scale propagation due to the persistent infection. In the IC dynamics, individuals lacking persistent infection largely restrains the final $I(t)$ while, on the other hand, the gap between the red and blue lines in the IC dynamics also indicates the difficulty for low-degree nodes to infect high-degree nodes under the lack of persistent infection. In Fig. 7(b), the

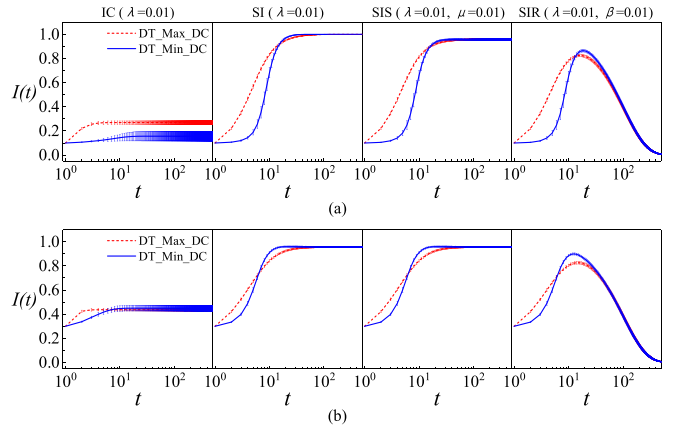


FIG. 7. The propagation scale $I(t)$, as a function of t , in the independent cascading (IC) model and epidemic models. Based on discrete-time (DT) Monte Carlo simulations, (a) and (b) report the diffusion details when the initial proportions of seed nodes $I(0) = 0.1$ and 0.3 , respectively. The solid and dash lines represent diffusion processes triggered by seed nodes with the lowest and highest degree centrality (DC), respectively. The final differences between the two processes under the IC dynamics show that the emergence of the crossover phenomenon essentially stems from the persistent infection characteristics of individuals.

increase of $I(0)$ narrows this gap to a trivial value, signaling the direct influence of the increasing $I(0)$ on the presence of a crossover. Therefore, we can conclude that the emergence of the crossover phenomenon essentially stems from the persistent infection under epidemic dynamics, and also directly induced by the increasing initial proportion of seed nodes.

V. CONCLUSION

In summary, we start from a two-diffusion-process system and focus on a potential crossover phenomenon to investigate the great impact of the finite initial state on the underlying dynamical process. We demonstrate the emergence of this phenomenon under classical epidemic dynamics numerically and analytically. The propagation scale triggered by low centrality nodes could be larger than that triggered by high centrality nodes during the diffusion process, provided that the initial seed size increases.

Furthermore, the existence of the disassortative mixing pattern and the irreversibility of the recovery process are shown to own the capability to control an earlier emergence of a crossover. Apart from the effects of the topological structure and the underlying dynamical mechanism, our work emphasizes the impact of the finite initial condition on the epidemic spreading. By analyzing the dynamical process in the IC model, we reach the conclusion that compared with the finite initial state, the persistent infection characteristic of epidemic dynamics is the essential reason for this phenomenon to arise.

Further investigation relating to this work could be concentrated on evolving structure and heterogeneous dynamics. Human contact scenarios are more complicated than we often think. Before we treat coronavirus consciously and

seriously, realistic human-to-human coronavirus transmissions are likely to happen anywhere and anytime. Therefore, instead of a static network structure, (data-driven) evolving networks provide a reasonable characterization for the dynamic changes of human contacts. Regarding the hidden dynamics, the classical epidemic models considered in this paper incorporate simplified homogeneous settings like the constant parameters, which actually should be allowed to vary over time [39,40]. Furthermore, since the SI, SIS, and SIR dynamical processes are probabilistic, the extension of our conclusion to other probability-based dynamics which incorporates similar characteristic of persistent infec-

tion would be another question to be investigated in the near future.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (No. 61976181, No. U1803263, and No. 11931015) and CQ CSTC (No. cstc2018jcyjAX0274) and Hong Kong Research Grants Council (No. HKBU12201619). Z.S. was supported by the China Scholarship Council (CSC) scholarship and J.K. was supported by the Russian Ministry of Science and Education Agreement (No. 13.1902.21.0026).

-
- [1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [2] J. P. Gleeson, *Phys. Rev. X* **3**, 021004 (2013).
- [3] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Rev. Mod. Phys.* **87**, 925 (2015).
- [4] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, *Eur. Phys. J. B* **26**, 521 (2002).
- [5] C. Castellano and R. Pastor-Satorras, *Phys. Rev. Lett.* **105**, 218701 (2010).
- [6] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, *Europhys. Lett.* **89**, 38009 (2010).
- [7] B. A. Prakash, D. Chakrabarti, N. C. Valler, M. Faloutsos, and C. Faloutsos, *Knowl. Inf. Syst.* **33**, 549 (2012).
- [8] M. Boguñá, C. Castellano, and R. Pastor-Satorras, *Phys. Rev. Lett.* **111**, 068701 (2013).
- [9] S. C. Ferreira, R. S. Sander, and R. Pastor-Satorras, *Phys. Rev. E* **93**, 032314 (2016).
- [10] C.-R. Cai, Z.-X. Wu, M. Z. Q. Chen, P. Holme, and J.-Y. Guan, *Phys. Rev. Lett.* **116**, 258301 (2016).
- [11] C. Castellano and R. Pastor-Satorras, *Phys. Rev. X* **7**, 041024 (2017).
- [12] R. Y. Li, S. Pei, B. Chen, Y. M. Song, T. Zhang, W. Yang, and J. Shaman, *Science* **368**, 489 (2020).
- [13] C. Menni, A. M. Valdes, M. B. Freidin, C. H. Sudre, L. H. Nguyen, D. A. Drew, S. Ganesh, T. Varsavsky, M. J. Cardoso, J. S. E.-S. Moustafa, A. Visconti, P. Hysi, R. C. E. Bowyer, M. Mangino, M. Falchi, J. Wolf, S. Ourselein, A. T. Chan, C. J. Steves, and T. D. Spector, *Nat. Med.* **26**, 1037 (2020).
- [14] L. Ferretti, C. Wymant, M. Kendall, L. L. Zhao, A. Nurtay, L. Abeler-Dörner, M. Parker, D. Bonsall, and C. Fraser, *Science* **368**, eabb6936 (2020).
- [15] Z. Wang, C. T. Bauch, S. Bhattacharyya, A. d'Onofrio, P. Manfredi, M. Perc, N. Perra, M. Salathé, and D. W. Zhao, *Phys. Rep.* **664**, 1 (2016).
- [16] Y. Liu, X. Wang, and J. Kurths, *IEEE Trans. Evol. Comput.* **23**, 1049 (2019).
- [17] J. J. Jiang, S. Wen, S. Yu, Y. Xiang, and W. L. Zhou, *IEEE Commun. Surveys Tuts.* **19**, 465 (2017).
- [18] L. Y. Lü, D. B. Chen, X.-L. Ren, Q.-M. Zhang, Y.-C. Zhang, and T. Zhou, *Phys. Rep.* **650**, 1 (2016).
- [19] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani, *Nat. Phys.* **2**, 110 (2006).
- [20] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, *Nat. Phys.* **6**, 888 (2010).
- [21] C. Gao, Z. Su, J. M. Liu, and J. Kurths, *Commun. ACM* **62**, 61 (2019).
- [22] R. M. Anderson, B. Anderson, and R. M. May, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1991).
- [23] P. G. Fennell, S. Melnik, and J. P. Gleeson, *Phys. Rev. E* **94**, 052125 (2016).
- [24] D. Schoch, T. W. Valente, and U. Brandes, *Soc. Netw.* **50**, 46 (2017).
- [25] C. C. Zou, D. Towsley, and W. B. Gong, *IEEE Trans. Dependable Secure Comput.* **4**, 105 (2007).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevE.102.052311> for the additional numerical and analytical demonstrations.
- [27] C. L. Vestergaard and M. Génois, *PLoS Comput. Biol.* **11**, 1 (2015).
- [28] A. Lancichinetti, S. Fortunato, and F. Radicchi, *Phys. Rev. E* **78**, 046110 (2008).
- [29] A.-L. Barabási and R. Albert, *Science* **286**, 509 (1999).
- [30] L. A. Adamic and N. Glance, in *Proceedings of the 3rd International Workshop on Link Discovery* (ACM, New York, 2005), pp. 36-43.
- [31] J. Leskovec, D. Huttenlocher, and J. Kleinberg, in *Proceedings of the 19th International Conference on World Wide Web* (ACM, New York, 2010), pp. 641-650.
- [32] D. B. Bu, Y. Zhao, L. Cai, H. Xue, X. P. Zhu, H. C. Lu, J. F. Zhang, S. W. Sun, L. J. Ling, N. Zhang, G. J. Li, and R. S. Chen, *Nucleic Acids Res.* **31**, 2443 (2003).
- [33] J. McAuley and J. Leskovec, in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Curran Associates, Inc., New York, 2012), pp. 539-547.
- [34] J. Leskovec, J. Kleinberg, and C. Faloutsos, *ACM Trans. Knowl. Discov. Data* **1**, 2 (2007).
- [35] Route Views Archive Project, <http://routeviews.org>.
- [36] M. E. J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).
- [37] R. Xulvi-Brunet and I. M. Sokolov, *Phys. Rev. E* **70**, 066102 (2004).
- [38] J. Goldenberg, B. Libai, and E. Muller, *Mark. Lett.* **12**, 211 (2001).
- [39] Q. H. Liu, M. Ajelli, A. Aleta, S. Merler, Y. Moreno, and A. Vespignani, *Proc. Natl. Acad. Sci. USA* **115**, 12680 (2018).
- [40] S. T. Ali, L. Wang, E. H. Y. Lau, X. K. Xu, Z. W. Du, Y. Wu, G. M. Leung, and B. J. Cowling, *Science* **369**, 1106 (2020).