

## Hopping in the Crowd to Unveil Network Topology

Malbor Asllani,<sup>1,\*</sup> Timoteo Carletti,<sup>1</sup> Francesca Di Patti,<sup>2</sup> Duccio Fanelli,<sup>2</sup> and Francesco Piazza<sup>3</sup>

<sup>1</sup>*naXys, Namur Institute for Complex Systems, University of Namur, rempart de la Vierge 8, B 5000 Namur, Belgium*

<sup>2</sup>*Dipartimento di Fisica e Astronomia, Università di Firenze, INFN and CSDC, Via Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy*

<sup>3</sup>*University of Orléans and Centre de Biophysique Moléculaire (CBM), CNRS UPR 4301, Rue C. Sadron, 45071 Orléans, France*

 (Received 29 November 2017; published 9 April 2018)

We introduce a nonlinear operator to model diffusion on a complex undirected network under crowded conditions. We show that the asymptotic distribution of diffusing agents is a nonlinear function of the nodes' degree and saturates to a constant value for sufficiently large connectivities, at variance with standard diffusion in the absence of excluded-volume effects. Building on this observation, we define and solve an inverse problem, aimed at reconstructing the *a priori* unknown connectivity distribution. The method gathers all the necessary information by repeating a limited number of independent measurements of the asymptotic density at a single node, which can be chosen randomly. The technique is successfully tested against both synthetic and real data and is also shown to estimate with great accuracy the total number of nodes.

DOI: [10.1103/PhysRevLett.120.158301](https://doi.org/10.1103/PhysRevLett.120.158301)

Networks are everywhere. The brain, Internet and cyber world, food webs, social contacts, and commuting fall within the vast realm of network science [1–3]. Irrespective of the specific domain of application, individual entities (e.g., material units, bits of information) belonging to a certain population may jump from one site (node) to its adjacent neighbors, following the intricate web of distinct links, which defines the architecture of a complex network [4–9]. The ability of agents to explore the network to which they are bound is customarily modeled as a standard diffusive process [4,10]. Individuals are therefore independent from each other, while mutual interference stemming from the competition for the available space is deliberately omitted. However, in real-world applications, the carrying capacity of each node is finite, an ineluctable constraint that should be accommodated for under crowded conditions [11–20]. On the other hand, the structure of the network is often unknown. Several methods have been introduced in the literature aimed at reconstructing the topology of the network from a direct assessment of the transport dynamics [21–23]. This amounts to solving an inverse problem, from functions back to structure, a task that involves formidable challenges [23–25].

Working along these lines, the aim of this Letter is twofold. On the one side, we will introduce a nonlinear operator to describe diffusion in a crowded network. The derivation originates from a microscopic stochastic framework and it is solidly grounded on first principles. As an important by-product of the analysis, we will discuss the idea of “functional degree centrality,” as opposed to the usual structural notion that is routinely invoked in network studies. Then, we will outline an innovative scheme that exploits the dynamical entanglement among walkers to reconstruct the unknown topology of the network. The method allows one to accurately determine the distribution

of connectivities (degrees) and to quantify, with unprecedented efficiency, the size of the examined network, as we shall here demonstrate for a selection of paradigmatic examples. Notably, all necessary information is gathered from just one randomly chosen node of the collection.

Let us consider a generic undirected graph composed of  $\Omega$  nodes and characterized by its adjacency matrix  $\mathbf{A}$  ( $A_{ij} = 1$ , if nodes  $i$  and  $j$  are connected, zero otherwise). The degree of each node, namely, the number of connected neighbors, is  $k_i = \sum_j A_{ij}$ . Each node is endowed with a given “carrying capacity,” i.e., is assumed to be partitioned into a large number  $N$  of compartments, which can be either occupied by an agent (a physical or abstract entity) or empty. The stochastic dynamics of a collection of particles randomly hopping on a network and competing for the available space within nodes is described by the following master equation

$$\frac{d}{dt}P(\mathbf{n}, t) = \sum_{i,j} A_{ij} [T(n_i, n_j | n_i + 1, n_j - 1)P(n_i + 1, n_j - 1, t) - T(n_i - 1, n_j + 1 | n_i, n_j)P(n_i, n_j, t)], \quad (1)$$

where  $P(\mathbf{n}, t)$  denotes the probability that the system will be in the state  $\mathbf{n} = (n_1, \dots, n_\Omega)$  at time  $t$  [15]. The scalar quantity  $n_i$  identifies the number of particles on node  $i$ . On the right-hand side of Eq. (1) we indicate explicitly only the species that are involved in the selected transition. In this stochastic process, one of the walkers sitting on node  $i$  is selected with probability  $n_i/N$  and then made to jump to one of the  $k_i$  neighboring nodes with probability  $1/k_i$ . However, this can only occur if the target node is not fully occupied, i.e., with probability  $(N - n_j)/N$ . Overall, the hopping probability  $T(\cdot | \cdot)$  from node  $i$  to node  $j$  reads

$$T(n_i - 1, n_j + 1 | n_i, n_j) = \frac{1}{k_i N} \frac{n_i (N - n_j)}{N}. \quad (2)$$

The above nonlinear transition probabilities accommodate for excluded-volume interactions [12]; i.e., the jump rate is modulated by the crowding at destination. Multiplying by  $n_i$  both sides of Eq. (1) and summing over all  $n_i$  (see Supplemental Material [26] and Ref. [27]), one obtains the rate equations for the evolution of the average density  $\langle n_i \rangle \equiv \sum_{\mathbf{n}} n_i P(\mathbf{n}, t)$ . In the thermodynamic limit  $\rho_i = \lim_{N \rightarrow \infty} \langle n_i \rangle / N$  and a straightforward manipulation yields

$$\frac{\partial}{\partial t} \rho_i = \sum_{j=1}^{\Omega} \Delta_{ij} \left( \rho_j (1 - \rho_i) - \frac{k_j}{k_i} \rho_i (1 - \rho_j) \right), \quad (3)$$

where use has been made of the fact that  $A_{ij} = A_{ji}$  and  $\langle n_i n_j \rangle = \langle n_i \rangle \langle n_j \rangle$  (a condition that holds exact when  $N \rightarrow \infty$  [27]).

In the above equation,  $\Delta_{ij} = A_{ij}/k_j - \delta_{ij}$  denotes the elements of the Laplacian  $\Delta$ , the transport operator that describes diffusion of noninteracting agents on a heterogeneous network [10]. Working under diluted conditions amounts to neglecting nonlinear terms in Eq. (3), which therefore reduces to the standard diffusion equation  $(\partial/\partial t)\rho_i = \sum_{j=1}^{\Omega} \Delta_{ij} \rho_j$  [28]. The mass (total number of individuals) is an invariant of the dynamics, which means that the quantity  $\beta = \sum_{i=1}^{\Omega} \rho_i(t) \in (0, \Omega]$  is conserved.

Owing to the inherent symmetry, which implies detailed balance, the equilibrium solution  $\rho_i^{\infty}$  of Eq. (3) can be determined analytically by setting  $\rho_j^{\infty}(1 - \rho_i^{\infty}) - \rho_i^{\infty}k_j/k_i(1 - \rho_j^{\infty}) = 0, \forall i, j$  [30]. Disregarding the trivial solution  $\rho_i^{\infty} = 1$ , one gets  $\rho_j^{\infty} = a_i k_j / (1 + a_i k_j)$ , where  $a_i = \rho_i^{\infty} [k_i(1 - \rho_i^{\infty})]^{-1}$ . By definition,  $\rho_j^{\infty}$  should not depend on  $i$ , which in turn implies  $a_i = a \forall i$ . This gives

$$\rho_i^{\infty} = \frac{a k_i}{1 + a k_i}, \quad (4)$$

where the constant  $a$  should be determined from the normalization condition  $\sum_{i=1}^{\Omega} (a k_i) / (1 + a k_i) = \beta$ . Under diluted conditions, Eq. (4) returns the standard equilibrium solution for the diffusion operator: the asymptotic distribution of walkers at node  $i$  is proportional to its connectivity  $k_i$ . Peaks in the steady-state concentration therefore identify structural hubs of the network, while peripheral nodes are associated with modest densities. In many practical applications, including page ranking schemes, the distribution of hopping walkers is hence believed to return an immediate measure of the nodes' centrality. Further, punctual information as delivered by a large, although finite, population of individual diffusing on a network is usually processed by assuming a distribution of the incoming signals, which scales linearly with the connectivity of the nodes [1]. Nonlinear interference among interacting walkers competing for space alters significantly the aforementioned simple

scenario. For sufficiently large connectivities, the predicted distribution (4) reaches a constant value. This effect is more pronounced the larger the value of  $a$  or, equivalently, the larger  $\beta$ . In Fig. 1, we report the equilibrium density obtained from Eq. (4) on a scale-free network under different crowding conditions. When  $\beta$  grows, hubs become progressively less distinct. *Structural* centrality, as revealed by the distribution of noninteracting agents (and asymptotically approached in the limit  $\beta \rightarrow 0$ ), differs significantly from *functional* centrality, which follows the equilibrium distribution when  $\beta \neq 0$  and nodes are assigned a finite carrying capacity. Elaborating on this dichotomy is particularly relevant for scale-free networks.

As anticipated, the nonlinear transport operator introduced above opens up the perspective of designing a novel scheme to access key global features of a network from direct measurements of the steady-state dynamics. We aim, in particular, at determining  $p(k)$ , the distribution of connectivities  $k$ , which conditions implicitly the equilibrium density of agents [31]. Importantly, our method also allows us to estimate the total number of nodes  $\Omega$  in the network. In the following, we will outline the mathematical steps of the procedure and then move forward to discussing a selected gallery of case studies.

For any fixed  $\beta$ , we monitor the dynamics of the system at just one node of the collection (hereafter,  $i$ ). After a sufficiently long time, one can measure the local asymptotic concentration  $\rho_i^{\infty}$ . Assuming the local connectivity  $k_i$  to be known, which is a reasonable working hypothesis given that we are sitting at node  $i$ , we can write

$$a(\beta) = \frac{\rho_i^{\infty}}{1 - \rho_i^{\infty}} \frac{1}{k_i},$$

where the dependence of  $a$  on  $\beta$  has been emphasized. We now rewrite the normalization condition so as to bring  $p(k)$  into the picture, that is,

$$\beta = \sum_k p(k) \frac{a(\beta) k}{1 + a(\beta) k}. \quad (5)$$

Since the network is limited in size, the previous sum involves a finite number of terms. Performing a

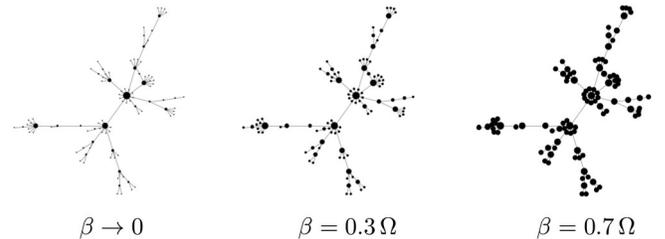


FIG. 1. Asymptotic distribution of walkers diffusing on a scale-free network under different crowding conditions, as specified by  $\beta$ . Standard diffusion is recovered in the limit  $\beta \rightarrow 0$  and displayed in the leftmost picture. Nodes are drawn with a size proportional to the corresponding asymptotic density,  $\rho_i^{\infty}$ .

independent measurements (or experiments), carried out for different values of the constant  $\beta$ , yields

$$\boldsymbol{\beta} = \mathbf{F}\mathbf{p}, \quad (6)$$

where  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_s)^T$ ,  $\mathbf{p} = (p(1), \dots, p(s))^T$ , and  $F_{lr}$ , the generic element of matrix  $\mathbf{F}$ , is given by

$$F_{lr} = \frac{ra(\beta_l)}{1 + ra(\beta_l)}.$$

Determining the components of the vector  $\mathbf{p}$  amounts to inverting matrix  $\mathbf{F}$ . Notice that  $s$ , the number of independent measurements performed, should be at least equal to the maximum degree of the inspected network, which is not known *a priori*. As it is shown in the Supplemental Material [26], one can progressively increase  $s$  until the reconstructed  $p(k)$  converges to a stable profile. Monitoring the first moments of the distribution against  $s$  provides a robust stopping criterion (see Supplemental Material [26]).

The matrix  $\mathbf{F}$  can be ill-conditioned when the size of the inspected network is too large. In this case, dedicated regularization schemes can be employed to carry out the matrix inversion and recover the degree distribution [32]. To further improve the accuracy of the reconstruction, one can also impose additional constraints on the inversion algorithm, such as requiring a positive-defined  $p(k)$ . A comment is in order before turning to test the adequacy of our methodology. Our protocol to reconstruct the unknown distribution of connectivities works if excluded-volume effects are accounted for. Under extremely diluted conditions, standard diffusion holds and  $\rho_i^\infty = bk_i$ , with  $b$  constant. In this case, Eq. (5) yields the trivial condition  $\langle k \rangle = \beta/b$ , where  $\langle k \rangle = \sum_j p(k_j)k_j$ . Hence, performing many experiments for different values of  $\beta$  simply returns independent estimates of the first moment of the distribution. It is the nonlinear nature of (4), the macroscopic blueprint of crowding, that enables the higher order moments of  $p(k)$  to be self-consistently computed from independent experiments [33]. Notice that one can readily extend the microscopic diffusion model so as to accommodate for different carrying capacities on each node. Also, in this case, it is straightforward to obtain the corresponding mean-field equation and to compute the associated stationary solution. Assuming known the capacity of the node where the measurements are performed, one can immediately adapt the reconstruction scheme to this general setting (see Supplemental Material [26]). In the remaining part of this Letter, we will validate the procedure against a limited selection of case studies that bear both theoretical and applied interest.

As a first example, we consider an Erdos-Renyi random network [34] with  $\Omega = 1000$  nodes and a probability for two random nodes to be linked  $p = 0.5$ . According to the procedure discussed above, we track the evolution of the walkers and measure the asymptotic density at a generic

node  $i$ , whose connectivity  $k_i$  is assumed to be known. Here,  $s = 600 < \Omega$  independent experiments are performed and the information processed to infer the distribution of connectivity  $p(k)$ . In Fig. 2(a), we compare the reconstructed profile with the exact distribution. The agreement is good and illustrates well the efficacy of our procedure. Integrating the mean-field deterministic dynamics, or operating under the original stochastic framework [35], returns consistent and equally accurate results (see Supplemental Material [26]). From the reconstructed distribution, one can readily calculate  $\Omega = \sum_k p(k)$ , the total number of nodes of the network. In this case, with  $s = 600$  measurements, one obtains the true value  $\Omega = 1000.0$ .

The second example is a scale-free network generated according to the preferential attachment rule [36]. In Fig. 2(b), we compare the exact distribution of connectivities (circles) with the approximated solution (squares) inferred through our algorithm. Also in this case, the agreement is satisfying: the power-law scaling is correctly reproduced and the predicted number of nodes are in excellent agreement with the true value (see caption of Fig. 2).

Finally, we turn to study two other examples that build on real networks: the well-known Karate Club network [37,38] and the *C. elegans* metabolic network (which was artificially made undirected) [38,39]. Diffusion is assumed to occur subject to finite-volume constraints and the method exemplified above is applied in order to recover the degree distribution from single-node measurements of the steady state. The results reported in Figs. 2(c) and 2(d), respectively, quantify the predictive power of the proposed methodology.

Summing up, we have introduced a novel nonlinear operator to model the process of diffusion on a complex network under crowded conditions. Agents may randomly hop from one node to another, provided free space is available at the target destination. The asymptotic density distribution in the presence of crowding differs significantly from that predicted by standard diffusion, i.e., under diluted conditions. In the latter case, the density scales linearly with the nodes' degree, while in crowded conditions, the equilibrium concentration saturates for large enough values of the connectivity. Based on this observation, we discussed the notion of functional hubs, as opposed to that based only on topology. Further, we developed an efficient scheme that enables one to infer the unknown distribution of connectivities from single-node measurements of the asymptotic diffusion dynamics. The method takes advantage of the interference that builds up among microscopic agents as a consequence of the competition for the available space. Tests are performed using both synthetic and real networks, which illustrate convincingly the power of our method.

The reconstruction scheme that we have introduced could prove useful in a number of different applications where the competition for available spatial resources is to be taken into account. For instance, crawlers that are customarily

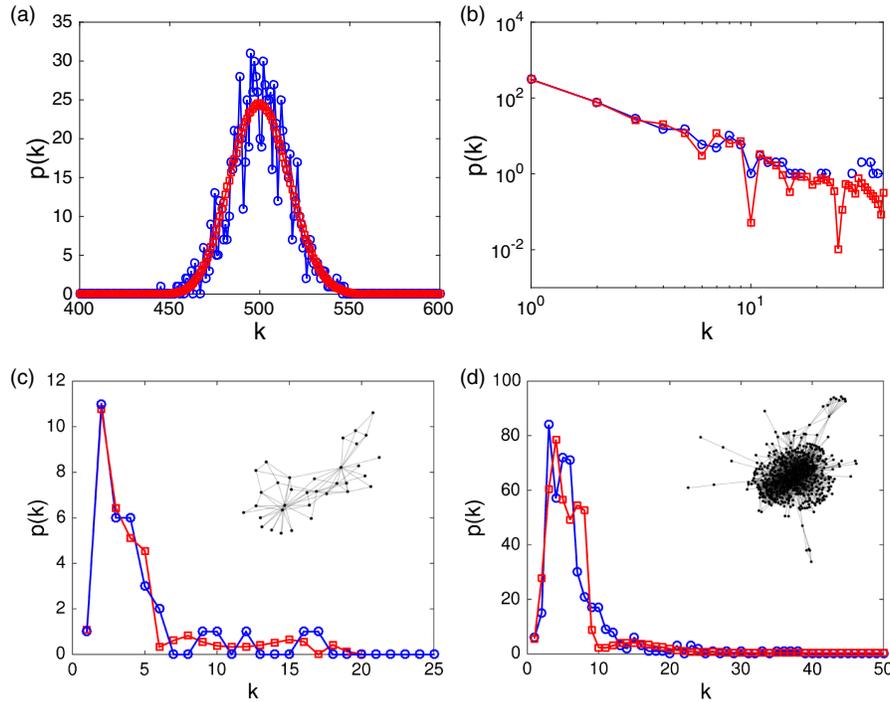


FIG. 2. Examples of network reconstructions,  $p(k)$  vs  $k$ . (a) Erdos-Renyi network with  $\Omega = 1000$ . The probability  $p$  for a link between any two pairs of nodes is set to 0.5. Exact profiles (blue circles) and profiles reconstructed with  $s = 600$  (red squares). The estimated number of nodes is  $\Omega = \sum_k p(k) = 1000$ . (b) Scale-free network with  $\Omega = 500$  and  $\gamma = 2$ . The degree distribution  $p(k)$  is plotted in log-log scale, focusing on the relevant region in  $k$ . Blue circles represent the exact distribution, while red squares identify the profile reconstructed with  $s = 150$  independent single-node measurements. The estimated number of nodes is  $\Omega = \sum_k p(k) = 500.2$ . (c) Reconstructed (red squares) vs exact (blue circles) degree distribution for the undirected Karate Club network (shown in the inset),  $\Omega = 34$ . Here,  $s = 20$ . The estimated size of the network is  $\Omega = \sum_k p(k) = 34.02$ . (d) The performance of the method is assessed by using the (symmetrized) *C. elegans* metabolic network (displayed in the inset) with  $\Omega = 453$ . Blue circles depict the exact distribution, while red squares refer to the reconstruction performed with  $s = 200$  experiments, which gives  $\Omega = \sum_k p(k) = 453.76$ .

employed for web indexing protocols are naturally exposed to mutual interferences due to the finite bandwidth. Thus, they are often treated as packets of information randomly hopping on a queuing network [40,41]. The packets' mobility depends on the degree of congestion of the nodes. Starting from these premises, it should be possible to use our method to recover structural information on the local network community by injecting a population of labeled crawlers of different size ( $\beta$ ) and monitoring their on site distribution. Other examples may include vehicle mobility and traffic flow [4], assuming that the true displacement dynamics can be described by a random walk.

More generally, our method allows one, in principle, to determine the degree distribution of networks that are *a priori* undefined and become defined, alongside their connectivity patterns, once the coarse-grained (CG) units are identified as the nodes. As such, we have introduced an effective (functional) coarse-graining procedure, suitable for large complex systems of interacting components. An example could be the diffusion of small molecules (e.g., enzyme inhibitors or competitive ligands for specific cell surface receptors) through complex tissues, such as the skin. Thus, from the local concentration, one should be able to characterize the functional connectome of the tissue, seen as

a complex network of interlinked CG components. Other applications may include the diffusion of molecules that mediate quorum sensing in multibacterial colonies, e.g., social interactions of bacteria in biofilms [42,43]. In this case, it would be instructive to adapt our algorithm to resolve the statistics of interactions among different species.

In its present formulation, our approach is suited for symmetric networks only. Future extensions are planned to reconstruct the distribution of outgoing and incoming connectivities in asymmetric crowded networks, i.e., when detailed balance may break, by examining the directed flux of particles from just one node. Finally, it would be interesting to gauge the impact of endowing nodes with a finite carrying capacity on other measures of centrality introduced in the literature.

We thank Renaud Lambiotte for insightful comments. The work of M. A. is supported by a FRS-FNRS Postdoctoral Fellowship. The work of T. C. presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its author(s). This project

has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement No 642563.

\*Corresponding author.

malbor.asllani@unamur.be

- [1] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, Oxford, 2010).
- [2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [3] R. Albert and A.-L. Barabasi, *Rev. Mod. Phys.* **74**, 47 (2002).
- [4] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks*, 1st ed. (Cambridge University Press, Cambridge, England, 2008).
- [5] J. D. Noh and H. Rieger, *Phys. Rev. Lett.* **92**, 118701 (2004).
- [6] R. Sinatra, J. Gómez-Gardeñes, R. Lambiotte, V. Nicosia, and V. Latora, *Phys. Rev. E* **83**, 030103(R) (2011).
- [7] V. Nicosia, F. Bagnoli, and V. Latora, *Europhys. Lett.* **94**, 68009 (2011).
- [8] N. Masuda, M. A. Porter, and R. Lambiotte, *Phys. Rep.* **716–717**, 1 (2017).
- [9] S. Manfredi, E. Di Tucci, and V. Latora, *Phys. Rev. Lett.* **120**, 068301 (2018).
- [10] I. Simonsen, K. A. Eriksen, S. Maslov, and K. Sneppen, *Physica (Amsterdam)* **336A**, 163 (2004).
- [11] A. P. S. de Moura, *Phys. Rev. E* **71**, 066114 (2005).
- [12] T. M. Liggett, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes*, 1st ed. (Springer-Verlag, Berlin, 1999).
- [13] E. Almaas, R. V. Kulkarni, and D. Stroud, *Phys. Rev. E* **68**, 056105 (2003).
- [14] S. Kwon and Y. Kim, *Phys. Rev. E* **84**, 041103 (2011).
- [15] D. Fanelli and A. J. McKane, *Phys. Rev. E* **82**, 021113 (2010).
- [16] M. Galanti, D. Fanelli, and F. Piazza, *Front. Phys.* **4**, 33 (2016).
- [17] A. E. Fernando, K. A. Landman, and M. J. Simpson, *Phys. Rev. E* **81**, 011903 (2010).
- [18] K. A. Landman and A. E. Fernando, *Physica (Amsterdam)* **390A**, 3742 (2011).
- [19] M. Galanti, D. Fanelli, A. Maritan, and F. Piazza, *Europhys. Lett.* **107**, 20006 (2014).
- [20] M. Galanti, S. Traytak, D. Fanelli, and F. Piazza, *Phys. Chem. Chem. Phys.* **18**, 15950 (2016).
- [21] E. Schneidman, M. J. Berry II, R. Segev, and W. Bialek, *Nature (London)* **440**, 1007 (2006).
- [22] S. Coccoa, S. Leiblerb, and R. Monasson, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14058 (2009).
- [23] R. Burioni, M. Casartelli, M. di Volo, R. Livi, and A. Vezzani, *Sci. Rep.* **4**, 4336 (2014).
- [24] S. G. Shandilya and M. Timme, *New J. Phys.* **13**, 013004 (2011).
- [25] I. Malvestio, T. Kreuz, and R. G. Andrzejak, *Phys. Rev. E* **96**, 022203 (2017).
- [26] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.120.158301> for the derivation of the main equations and the calculation of the associated fixed point, for an extension of the nonlinear operator to the setting with nodes characterized by distinct carrying capacities, for a numerical analysis on the convergence of the proposed algorithm and for its extension to the stochastic framework.
- [27] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).
- [28] As a side remark, we observe that Eq. (3) can be cast in the equivalent form  $(\partial/\partial t)\rho_i = \sum_{j=1}^{\Omega} \Delta_{ij}\rho_j - \rho_i \sum_{j=1}^{\Omega} \Delta_{ij}\rho_j + (\rho_i/k_i) \sum_{j=1}^{\Omega} \Delta_{ij}\rho_j k_j$ . When  $(k_i = k\forall i)$ , as it happens on a regular lattice subject to periodic boundary conditions, the last two terms cancel out and one recovers the linear diffusion equation. Stated differently, the signature of the microscopic exclusion rule disappears when heterogeneity gets lost. This is the generalization of an observation made by Huber in [29] (see also [16]).
- [29] D. L. Huber, *Phys. Rev. B* **15**, 533 (1977).
- [30] Indeed, this is the only stationary solution of Eq. (3). Looking for the stationary solution, (3) amounts to solving  $\Delta \vec{f}(i) = 0$ , where  $f_j(i) = \rho_j^{\infty}(1 - \rho_i^{\infty}) - \rho_i^{\infty} k_j/k_i(1 - \rho_j^{\infty})$ . For all  $i$ ,  $\vec{f}(i) = (f_1(i), \dots, f_{\Omega}(i))$  should be proportional to  $(k_1, \dots, k_{\Omega})$ , the eigenvector relative to the null eigenvalue of the random walk Laplacian  $\Delta$ . Hence,  $f_j(i) = b_i k_j$ , for some  $b_i$ . By imposing  $i = j$ , the previous equation yields  $0 = b_i k_i$ , and since  $k_i > 0$  (the network is connected), one finally gets  $b_i = 0 \forall i$ , thus proving the claim (see also Supplemental Material [26]).
- [31]  $p(k)$  is normalized so as to yield  $\sum_k p(k) = \Omega$ .
- [32] G. H. Golub and C. F. van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1996).
- [33] If the system has relaxed to the equilibrium solution, the choice of the node  $i$  where the asymptotic density  $\rho_i^{\infty}$  is measured does not impact the quality of the reconstruction. Identical distributions are in fact recovered irrespective of the specific selection operated (be the selected node a hub or a leaf). On the other hand, sampling the dynamics out of equilibrium could yield a poor estimate of the parameter  $a$ , which could spuriously depend on the node of observation. Averaging over different measurements allows one, however, to improve on the estimate of  $a$ , yielding a smooth convergence of the reconstruction procedure, as discussed in the Supplemental Material [26] with reference to the stochastic analogue of the problem.
- [34] E. N. Gilbert, *Ann. Math. Stat.* **30**, 1141 (1959).
- [35] D. T. Gillespie, *J. Phys. Chem.* **81**, 2340 (1977).
- [36] A.-L. Barabasi and R. Albert, *Science* **286**, 509 (1999).
- [37] W. W. Zachary, *Journal of anthropological research* **33**, 452 (1977).
- [38] J. Kunegis, in *Proceedings of the International World Wide Web Conference Companion, Rio de Janeiro, 2013* (ACM, New York, 2013), p. 1343.
- [39] J. Duch and A. Arenas, *Phys. Rev. E* **72**, 027104 (2005).
- [40] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, *Phys. Rev. Lett.* **89**, 248701 (2002).
- [41] D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili, *Phys. Rev. E* **79**, 015101(R) (2009).
- [42] H. C. Flemming, J. Wingender, U. Szewzyk, P. Steinberg, S. A. Rice, and S. Kjelleberg, *Nat. Rev. Microbiol.* **14**, 563 (2016).
- [43] Y. H. Li and X. L. Tian, in *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria* (John Wiley & Son, Inc., Hoboken, 2016), Vol. 2, p. 1197.