

# Nonlinear walkers and efficient exploration of congested networks

Timoteo Carletti,<sup>1</sup> Malbor Asllani,<sup>2</sup> Duccio Fanelli,<sup>3</sup> and Vito Latora<sup>4,5,6</sup>

<sup>1</sup>*naXys, Namur Institute for Complex Systems, University of Namur, Belgium*

<sup>2</sup>*MACSI, Department of Mathematics and Statistics,  
University of Limerick, Limerick V94 T9PX, Ireland*

<sup>3</sup>*Dipartimento di Fisica e Astronomia, University of Florence,  
INFN and CSDC, Via Sansone 1, 50019 Sesto Fiorentino, Florence, Italy*

<sup>4</sup>*School of Mathematical Sciences, Queen Mary University of London, Mile End Road, E1 4NS, London (UK)*

<sup>5</sup>*Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, I-95123 Catania, Italy*

<sup>6</sup>*The Alan Turing Institute, The British Library, London NW1 2DB, United Kingdom*

Random walks are the simplest way to explore or search a graph, and have revealed a very useful tool to investigate and characterize the structural properties of complex networks from the real world. For instance, they have been used to identify the modules of a given network, its most central nodes and paths, or to determine the typical times to reach a target. Although various types of random walks whose motion is biased on node properties, such as the degree, have been proposed, which are still amenable to analytical solution, most if not all of them rely on the assumption of linearity and independence of the walkers. In this work we introduce a novel class of nonlinear stochastic processes describing a system of interacting random walkers moving over networks with finite node capacities. The transition probabilities that rule the motion of the walkers in our model are modulated by nonlinear functions of the available space at the destination node, with a bias parameter that allows to tune the tendency of the walkers to avoid nodes occupied by other walkers. Firstly, we derive the master equation governing the dynamics of the system, and we determine an analytical expression for the occupation probability of the walkers at equilibrium in the most general case, and under different level of network congestions. Then, we study different type of synthetic and real-world networks, presenting numerical and analytical results for the entropy rate, a proxy for the network exploration capacities of the walkers. We find that, for each level of the nonlinear bias, there is an optimal crowding that maximises the entropy rate in a given network topology. The analysis suggests that a large fraction of real-world networks are organised in such a way as to favour exploration under congested conditions. Our work provides a general and versatile framework to model nonlinear stochastic processes whose transition probabilities vary in time depending on the current state of the system.

## I. INTRODUCTION

Random walks are basic stochastic processes, which bear universal interest in light of their widespread and cross-disciplinary usage. Since the pioneering work by Pearson and Rayleigh, back in 1905 [1, 2], the number of studies invoking the notion of random walker has grown rapidly, to eventually cover a broad spectrum of applications, from physics to engineering, via biology and economics.

Random walks have been thoroughly studied on regular lattices [3] and, more recently, on graphs displaying complex topologies [4–7]. In the simplest possible scenario, the walker moves, with a uniform probability, from a given node  $i$  to one of its neighbours  $j$ . Alternatively, when the dynamics takes place on a weighted graph, one can gauge the probability of performing the move with the weight of the link  $(i, j)$  [8, 9]. Various other classes of random walkers are however possible on complex networks [7]. The walk can be for instance biased on the topological properties of the nodes of the network, such as the node degree or the betweenness. In [10], the probability for a walker to perform a move is modulated by a power law of the degree of the target node. Tuning the scaling exponent enables one to steer the dynamics towards the hubs or favour, at variance, the motion

towards low-degree nodes. Furthermore, when the nodes are also characterised by endogenous state variables, mirroring congestion or tagging local deficiencies, these can be considered as a feedback to modify the motion of individual agents [11]. Metapopulation models of random walkers which integrate random relocation moves with local interactions depending on the node occupation probabilities have also been proposed in [12] and employed to extract information on the architecture of the underlying network. Mutual interference, as stemming from the competition for available spatial resources, is unavoidably present when many walkers are moving at the same time across the nodes of a given network [13]. In Ref. [14] a model of transport on networks which accounts for the finite carrying capacity of the nodes has been proposed. In particular, it has been shown that the equilibrium density (stationary distribution) of crowded walkers saturates for large enough values of the connectivity, while conventional non-interacting agents have a stationary distribution which depends linearly on the nodes degree.

In this work we introduce and study a novel and general class of *nonlinear* Markov chains with transition probabilities that change in time depending on the current state of the system. These describe the motion of interacting random walkers whose probability to jump to

a node of a network is a nonlinear function of the number of walkers currently at the node. Such class of *nonlinear random walkers* provides a versatile, but at the same time analytically treatable, framework to study the dynamics of active agents that modulate their motion depending on the level of perceived congestion on the network. As a special case, we will study walkers whose probability to move from node  $i$  to a neighbour  $j$  scales as a power law of the occupation density of node  $j$ , with an exponent  $\sigma \geq 0$  that measures the anti-social behaviour of the walkers, i.e. their tendency to avoid nodes already occupied by other walkers. Under this framework we will prove that, for any given network and each selected value of  $\sigma$ , there is always an optimal value of the network crowding (the total load on the network), that maximises the entropy rate, i.e. facilitates the exploration of the network. We will also show that, in many real-world networks, the maximal value of the entropy rate is larger than that in randomised networks with the same degree distributions.

## II. THE STOCHASTIC PROCESS

Consider a set of interacting agents (walkers) moving on an undirected network with  $N$  nodes, each endowed with a finite carrying capacity. For the sake of simplicity, we assume that all the nodes have the same carrying capacity, i.e. each of them can simultaneously host a maximum number of agents equal to  $M$ . The architecture of the network is described in terms of the binary adjacency matrix  $A = \{a_{ij}\}$ , with  $a_{ij} = 1$  if there is a link connecting nodes  $i$  and  $j$ , while  $a_{ij} = 0$  otherwise. At each time  $t$ , the state of the system (our set of walkers) is specified by the vector  $\mathbf{m} = (m_1, \dots, m_N)$ , where  $0 \leq m_i \leq M$  is the number of agents that belong to node  $i$ , at time  $t$ . The total number of walkers in the network is fixed in time and is a tunable parameter of the model. We can control it by introducing the *average node crowding*  $\beta = 1/N \sum_{i=1}^N m_i/M$ . By definition,  $\beta \in (0, 1]$  quantifies the average node congestion, with  $\beta \rightarrow 0$  corresponding to the idealised diluted setting. Hence, we can tune the total number of walkers in the network,  $\beta MN$ , by independently changing  $M$  and  $\beta$ . Agents perform a biased random walk hopping between neighbouring nodes, provided there is enough space at the arriving destination. Differently from Refs. [10], the motion of the agents is not biased on the topological properties of the underlying graph but on the positions of the other agents in the network. More specifically, the bias results in two distinct contributions, respectively representing the *willingness* to leave a node  $i$ , and the *attractiveness* of the target node  $j$ . The first component is a function,  $f(x_i)$ , of the density  $x_i = m_i/M$  on node  $i$ . The second term is made to depend on the available space  $1 - x_j = (M - m_j)/M$  at node  $j$ , as  $g(x_j) \equiv \hat{g}(1 - x_j)$ . As a natural constraint, we require that  $f(x)$  vanishes at zero, i.e.  $f(0) = 0$ , since no hops can take place from an empty node. Further, we assume that  $f(\cdot)$  is a non-decreasing nonlinear

function of  $x$ , a choice that amounts to modelling anti-social reactions of the walkers to enhanced crowded conditions, i.e. their tendency to avoid nodes already occupied by other walkers. Observe that the standard unconstrained random walk is eventually recovered when setting  $f(x) = x$  and  $g(x) = 1$ , for all  $x$ . The finite carrying capacity signifies that no transition towards node  $j$  can take place, if  $x_j = 1$ , namely if the arrival node is fully packed. We hence require the self-consistent condition  $g(1) \equiv \hat{g}(0) = 0$ . Any possible choice of  $f(x)$  and  $g(x)$  fulfilling the above prescription is in principle possible. Notice that the linear model studied in [14] can be obtained as a particular case of our model if we fix  $f(x) = x$  and  $g(x) = 1 - x$ . On the other hand, adopting nonlinear functions for  $g(x)$ , enables one to reveal a large plethora of interesting dynamical features, which reflect different modalities of active reaction to perceived crowding conditions, encompassing social/antisocial attitudes.

The evolution of our system of nonlinear interacting random walks is ruled by the master equation:

$$\frac{d}{dt} P(\mathbf{m}, t) = \sum_{\mathbf{m}'} [T(\mathbf{m}|\mathbf{m}')P(\mathbf{m}', t) - T(\mathbf{m}'|\mathbf{m})P(\mathbf{m}, t)]$$

where  $P(\mathbf{m}, t)$  denotes the probability to find the system in the state  $\mathbf{m}$  at time  $t$ ,  $T(\mathbf{m}'|\mathbf{m})$  is the transition probability from state  $\mathbf{m}$  to state  $\mathbf{m}'$ , and the sum is restricted to states  $\mathbf{m}'$  compatible with  $\mathbf{m}$  [15]. Because the transitions involve pairs  $(i, j)$  connected by a link, i.e. such that  $a_{ij} = 1$  and only increments and decrements by one unity are allowed, we get  $\mathbf{m}' = (\dots, m_i \pm 1, \dots, m_j \mp 1, \dots)$ . The transition probabilities read:

$$T(m_i - 1, m_j + 1 | m_i, m_j) = \frac{a_{ij}}{k_i} f\left(\frac{m_i}{M}\right) g\left(\frac{m_j}{M}\right),$$

where  $k_i = \sum_j a_{ij}$  is the degree of node  $i$ . To make the notation compact, in the above expression we solely highlight the state components which are modified by the occurring transition [16–20]. The calculation is however exact: all components are accounted for, and no approximation is involved (see Appendix A).

A straightforward manipulation yields (see Appendix A and [14, 21]) the following equation for the time evolution of the mean-field node density  $\rho_i(t) = \lim_{M \rightarrow \infty} \langle m_i \rangle / M$ :

$$\frac{d\rho_i}{dt} = \sum_j \Delta_{ij} \left[ f(\rho_j)g(\rho_i) - \frac{k_j}{k_i} f(\rho_i)g(\rho_j) \right] = \mathcal{L}_i(\rho), \quad (1)$$

where  $\Delta_{ij} = a_{ij}/k_j - \delta_{ij}$  is the random walk Laplacian and the nonlinear operator  $\mathcal{L}_i(\rho)$  is defined by the right-most equality. Notice that the above mean field equation has been obtained by neglecting terms which are  $1/M$  smaller than the others. This is an approximation at  $M$  finite, but holds exactly in the limit  $M \rightarrow \infty$ , i.e., when  $1/M$  corrections vanish [16]. From Eq. (1) it is immediate to conclude that the mass, namely the total number

of walkers, is an invariant of the dynamics. The quantity  $\sum_{i=1}^N \rho_i(t)/N$  is hence conserved and equals to the average node congestion  $\beta$ .

### III. EQUILIBRIUM DISTRIBUTION

The stationary solution of Eq. (1) can be computed, for any choice of the nonlinear functions  $f$  and  $g$  (see Appendix B). We study here the case in which we set  $f(x) = x$  and  $g(x) = (1-x)^\sigma$ , with  $\sigma \geq 0$ . Modulating the exponent  $\sigma$ , means selecting different exploration strategies of the walkers. More specifically, the larger  $\sigma$  the more the walkers will try to avoid densely populated nodes. In other terms, the value of  $\sigma$  quantifies the level of anti-social behaviour of the walkers. Notice that the diluted limit of non-interacting walkers is recovered by letting  $\sigma \rightarrow 0$  (and also  $\beta \rightarrow 0$ ). For the case at hand, the stationary solution  $\rho_i^*$  should match the following implicit equation:

$$\frac{\rho_i^*}{c_\sigma k_i} = (1 - \rho_i^*)^\sigma \quad \forall i, \quad (2)$$

where  $c_\sigma$  is a normalisation factor which depends on the selected  $\sigma$ . Recalling the definition of  $\beta$  yields  $c_\sigma = \beta N / \sum_j k_j (1 - \rho_j^*)^\sigma$ , a condition which should complement Eq. (2) for a self-consistent determination of the stationary equilibrium. To interpret the above asymptotic solution we will draw a comparison with that obtained when assuming linear transition rates,  $\sigma = 1$ . When  $\sigma < 1$ , agents accumulate on the nodes characterised by a large degree, by consequently depleting those displaying modest connectivity (see Fig. 1). At variance, when  $\sigma > 1$ , hubs are progressively emptied and the walkers tend to preferentially fill peripheral nodes with respect to the linear case. It is instructive to compute the critical degree  $k^{\text{crit}}$  where such inversion takes place for a generic  $\sigma$  with respect to the reference case  $\sigma = 1$ . A direct computation (see Appendix B 1) returns:

$$k^{\text{crit}} = \left[ \left( \frac{c_1}{c_\sigma} \right)^{1/(1-\sigma)} - 1 \right] \frac{1}{c_1}. \quad (3)$$

In short, for all  $k_i > k^{\text{crit}}$ , we have  $\rho_i^*|_{\sigma < 1} > \rho_i^*|_{\sigma = 1}$ , while the opposite inequality holds true if  $k_i < k^{\text{crit}}$ . The sign of the inequalities reverse when  $\sigma > 1$  (see Fig. 1).

### IV. EXPLORATION UNDER CONGESTED CONDITIONS

The entropy rate of a random walk on a complex network characterises the walkers ability to explore the network, resulting in a non trivial indicator where topology and dynamical rules are mutually entangled [10, 22, 23]. We will hence evaluate the entropy rate of the process under study to quantify the performance of the walkers in exploring a given network under different level of

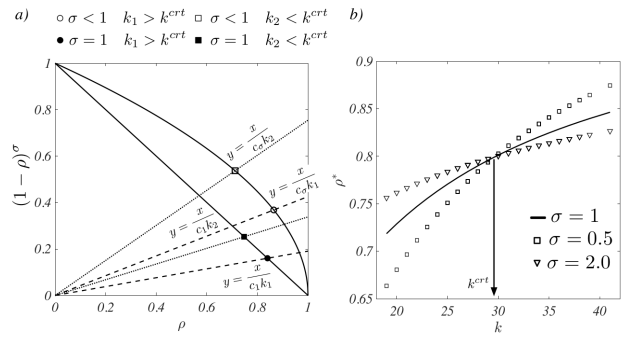


FIG. 1. **The stationary solution.** (a) The stationary distribution of the walkers at a node of degree  $k$  is determined as the intersection between the line  $\rho/(c_\sigma k)$  and the curve  $(1-\rho)^\sigma$ . The normalisation constant  $c_\sigma$  depends on  $\sigma$ , the level of crowding in the network. (b)  $\rho^*$  is plotted versus  $k$ . When  $\sigma < 1$  agents cluster on nodes with a large degree, while for  $\sigma > 1$  hubs are progressively depleted (with respect to the case  $\sigma = 1$ ). The solutions obtained for  $\sigma \neq 1$  intersect the curve relative to  $\sigma = 1$  at  $k^{\text{crit}}$ .

congestion. The entropy rate  $h$  of a stationary Markov chain with transition matrix  $\Pi = \{\pi_{ij}\}$  and stationary distribution  $\mathbf{w}^* = \{w_i^*\}$  can be written as  $h = -\sum_{ij} \pi_{ij} w_i^* \log \pi_{ij}$ . In the present case one gets:

$$h = -\sum_{ij} \rho_i^* a_{ij} \frac{\rho_i^* (1 - \rho_j^*)^\sigma}{k_i} \log \left[ a_{ij} \frac{\rho_i^* (1 - \rho_j^*)^\sigma}{k_i} \right]. \quad (4)$$

The entropy rate depends on the dynamics of the walkers, via the stationary probability  $\rho_i^*$ , the nonlinearity exponent  $\sigma$ , and the congestion parameter  $\beta$ , but also on the structure of the underlying network, via its adjacency matrix  $A = \{a_{ij}\}$ . The entropy rate in Eq. (4) (normalised to the system size  $N$ ) can be rewritten in the Heterogeneous Mean Field (HMF) approximation, by dividing the nodes in different degree classes, considering the asymptotic densities of nodes with the same degree and performing sums over degree classes (see Appendix C) [10, 24].

Fig. 2 a) and c) show the entropy rate per node,  $h/N$ , versus  $\beta$ , for synthetic networks with the same average degree  $\langle k \rangle$  and heterogeneous or homogeneous degree distributions respectively. Symbols refer to a direct (and exact) computation through Eq. (4). Solid lines are instead the results in the HMF approximation. We notice that, for any value of  $\sigma$ , there is an associated value of the crowding parameter  $\beta^{\text{opt}}$  which maximises the entropy rate  $h/N$ . An adequate and network-dependent amount of congestion seems therefore necessary to favour the network explorability for any given level of anti-social behaviour (as measured by the value of  $\sigma$ ). The maximum entropy  $h^{\text{opt}} = h(\beta^{\text{opt}})$  increases by decreasing  $\sigma$ , namely when the antisocial behaviour of the walkers is reduced. A trivial global optimum is eventually obtained when the constraint of a finite carrying capacity is completely re-

moved. Notice also that the entropy rate approaches zero when  $\beta \rightarrow 1$ , namely under extremely crowded conditions, i.e. when the agents are practically stuck in their positions. Interestingly, both  $\beta^{\text{opt}}$  and the value of  $h^{\text{opt}}$  depend on the topology of the network. As an example, when  $\sigma = 0.5$ ,  $\beta^{\text{opt}} \sim 0.68$  and  $h^{\text{opt}} \sim 0.48 \times N$  for Erdős-Rényi random graphs, while  $\beta^{\text{opt}} \sim 0.64$  and  $h^{\text{opt}} \sim 0.34 \times N$  for scale-free networks. Complementary insights can be obtained by looking at the iso-level lines of  $h/N$  in the plane  $(\beta, \sigma)$  reported in Fig. 2 b) and d). In order to maintain the same level of explorability, the walkers need to adjust the value of the dynamic bias  $\sigma$ , depending on the traffic load  $\beta$  in the network. Intriguingly enough,  $\sigma$  is a non-monotonic function of  $\beta$  on iso- $h$  curves. For small values of  $\beta$ , the walkers have to strengthen their antisocial behaviour (i.e. to increase  $\sigma$ ) to keep the same value of  $h$ . Above a critical value of the average node crowding  $\beta$ , the walkers need instead to weaken their antisocial bias (i.e. to decrease  $\sigma$ ).

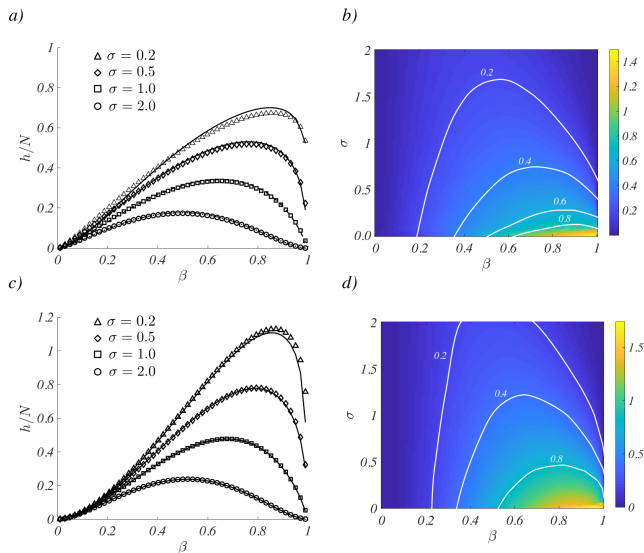


FIG. 2. **Entropy rate and iso-explorability on synthetic networks.** The asymptotic entropy rate per node  $h/N$  is shown as a function of the average node congestion  $\beta$  and for different choices of the parameter  $\sigma$ . Panel a) refers to scale-free networks with  $N = 1000$ ,  $\gamma = 2.5$  and average degree  $\langle k \rangle = 6.9$ . Panel c) is obtained for Erdős-Rényi networks, with  $N = 1000$  and  $\langle k \rangle = 6.9$ . Symbols refer to the exact computation performed from Eq. (4). Lines are the analytical predictions obtained in the heterogeneous mean field approximation. In panels b) (relative to SF networks) and c) (for the case of Erdős-Rényi graphs) the iso-level lines  $h/N$  are depicted in the reference plane  $(\beta, \sigma)$ . A constant level of explorability is obtained by modulating  $\sigma$  as a nonlinear function of  $\beta$ .

Further, we have analysed how the average node degree of a network, impacts the entropy rate of the walkers. To this end, we build different Erdős-Rényi networks with the same number of nodes but different average node degrees. Fig. 3 shows the entropy rate per node as a func-

tion of  $\beta$  and its maximum as a function of  $\langle k \rangle$ , for three values of the nonlinear bias parameter  $\sigma$  (0.5 top panels, 1.0 middle panels and 2.0 bottom panels). Increasing the network connectivity, yields a global enhancement of the entropy rate and of its associated maximum. The larger the connectivity, in fact, the richer the variety of routes available to the motion of the walkers. As a further point, we stress that random architectures return lower entropy values at peak, as compared to lattices, a counter-intuitive conclusion that is made quantitative in the Appendix B 2, where we also derive closed analytical formulae for the entropy rate on  $k$ -regular lattices.

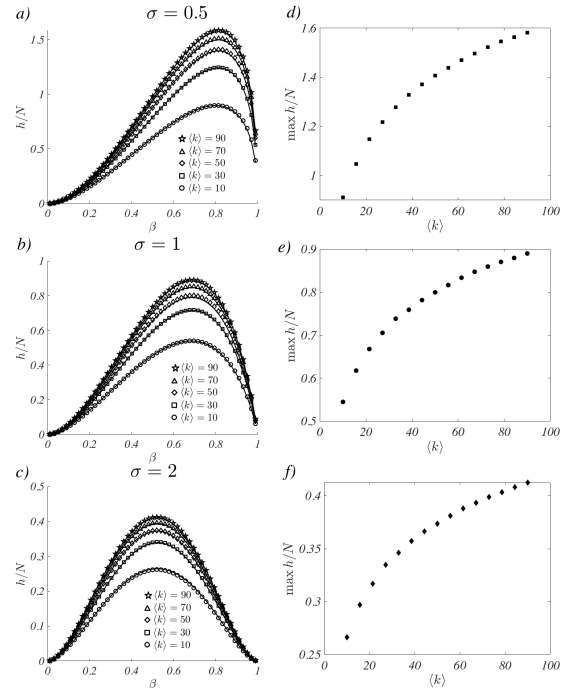


FIG. 3. **Entropy rate versus  $\langle k \rangle$ .** The asymptotic entropy rate  $h/N$  per node and its maximum value are reported as a function of the network load (left panels) and the average nodes degree (right panels), for several values of the nonlinear bias parameter  $\sigma$ . Erdős-Rényi networks with  $N = 100$  nodes have been used. Lines in the left panels are the analytical predictions.

Finally, we studied the properties of our model of nonlinear random walkers on several networks taken from the real world. We computed the entropy rate as a function of the crowding parameter, determining in each case the optimal values  $\beta^{\text{opt}}$  and  $h^{\text{opt}}$ , for several values of the nonlinear bias  $\sigma$ . Results are compared to those obtained on randomized versions of the networks. Two different types of randomization have been adopted: the first one preserves the degree of each node, while the second one maintains the network average degree only. As an example Fig. 4 shows the results obtained for: (a) a snapshot of the social network of Facebook [25], and (b) for the air transportation network among the 500 largest US airports [5, 26]. First, we confirm the non-monotonic be-

behaviour of the entropy rate: this latter vanishes for small and large values of  $\beta$ , and exhibits a maximum at an optimal value of the crowding parameter  $\beta^{\text{opt}}$ . In addition to this, we notice that, for intermediate and large values of  $\beta$ , the entropy rate of the walkers on both these two real-world networks is larger than that on the randomised versions of the networks preserving the degree distribution. In Appendix D and Table I we report on the results obtained for a large collection of real networks. Although some of them can also exhibit smaller values of entropy rate than their randomised versions, we have found that all the networks analysed, which describe urban street patterns, achieve a better explorability.

(CÉCI)” (<http://www.ceci-hpc.be>). V. L. acknowledges support from the Leverhulme Trust Research Fellowship CREATE: the network components of creativity and success.

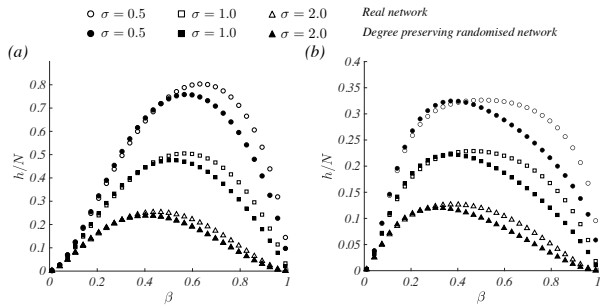


FIG. 4. **Entropy rate for real networks.** The entropy rate per node,  $h/N$ , for different values of  $\sigma$  is shown for two real networks: (a) the social network of Facebook [25] and (b) the transportation network of the 500 largest US airports [5, 26]. Filled symbols refer to the average entropy rate obtained for an ensemble of 50 randomizations which preserve the node degrees of the two real networks.

## V. CONCLUSIONS AND OUTLOOK

Summing up we have here discussed a general approach to the modelling of biased random walks, under crowded conditions. The formulation of the problem is not limited to the specific framework analysed here (see Appendix E for a generalisation in which also function  $f$  is a power law) and paves the way to devising novel algorithms for an efficient transport on networks, even in more complex adaptive settings where the dynamics of the walkers is coevolving with the underlying network [27].

## ACKNOWLEDGEMENTS

This research used resources of the ”Plateforme Technologique de Calcul Intensif (PTCI)” (<http://www.ptci.unamur.be>) located at the University of Namur, Belgium, which is supported by the FNRS-FRFC, the Walloon Region, and the University of Namur (Conventions No. 2.5020.11, GEQ U.G006.15, 1610468 et RW/GEQ2016). The PTCI is member of the ”Consortium des Équipements de Calcul Intensif

### Appendix A: From the master equation to the deterministic density evolution.

The goal of this section is derive the mean field equations for the densities, namely Eqs. (3) in the main body of the paper, from the Master Equation:

$$\frac{d}{dt}P(\mathbf{m}, t) = \sum_{\mathbf{m}'} [T(\mathbf{m}|\mathbf{m}')P(\mathbf{m}', t) - T(\mathbf{m}'|\mathbf{m})P(\mathbf{m}, t)], \quad (\text{A1})$$

where  $P(\mathbf{m}, t)$  denotes the probability to find the system at time  $t$  in the state  $\mathbf{m} = (m_1, \dots, m_N)$ . Recall that the above sum is restricted to states  $\mathbf{m}'$  compatible with  $\mathbf{m}$ . Because at a given time only a walker can hop from a given node to one of its neighbours, the states  $\mathbf{m}'$  take the form  $\mathbf{m}_j = (\dots, m_i \pm 1, \dots, m_j \mp 1, \dots)$ , for all  $j$  with  $a_{ij} = 1$ .

Let us introduce the average number of agents in node  $i$  at time  $t$ ,  $\langle m_i(t) \rangle \equiv \sum_{\mathbf{m}} m_i P(\mathbf{m}, t)$ , and the densities  $\rho_i(t) = \lim_{M \rightarrow \infty} \langle m_i(t) \rangle / M$ . Then, by taking the time-derivative of  $\langle m_i \rangle$ , recalling (A1) and accounting for the subsets of compatible states we get:

$$\begin{aligned} \frac{d}{dt} \langle m_i \rangle &= \sum_{j, \mathbf{m}_j} m_i \left[ -T(m_i - 1, m_j + 1 | m_i, m_j) P(m_i, m_j, t) + T(m_i, m_j | m_i + 1, m_j - 1) P(m_i + 1, m_j - 1, t) \right. \\ &\quad \left. + \sum_{j, \mathbf{m}_j} m_i \left[ -T(m_i + 1, m_j - 1 | m_i, m_j) P(m_i, m_j, t) + T(m_i, m_j | m_i - 1, m_j + 1) P(m_i - 1, m_j + 1, t) \right] \right], \end{aligned}$$

or equivalently

$$\frac{d}{dt} \langle m_i \rangle = \sum_j \left[ -\langle T(m_i - 1, m_j + 1 | m_i, m_j) \rangle + \langle T(m_i + 1, m_j - 1 | m_i, m_j) \rangle \right].$$

Consider now the transition probabilities. These latter are expressed in terms two functions,  $f$  node and  $g$  as:

$$T(m_i - 1, m_j + 1 | m_i, m_j) = \frac{a_{ij}}{k_i} f\left(\frac{m_i}{M}\right) \hat{g}\left(\frac{M - m_j}{M}\right),$$

then we eventually get:

$$\frac{d}{dt} \langle m_i \rangle = \sum_j \left[ -\frac{a_{ij}}{k_i} \langle f\left(\frac{m_i}{M}\right) \hat{g}\left(\frac{M - m_j}{M}\right) \rangle + \frac{a_{ji}}{k_j} \langle f\left(\frac{m_j}{M}\right) \hat{g}\left(\frac{M - m_i}{M}\right) \rangle \right].$$

By introducing the rescale time  $\tau = t/M$  and performing the limit  $M \rightarrow \infty$  (which in turn amounts to neglecting correlations, i.e.  $\langle f(\cdot) \rangle = f(\langle \cdot \rangle)$ , similarly for  $g$ ) yields:

$$\frac{d}{d\tau} \rho_i = \sum_j \left[ -\frac{a_{ij}}{k_i} f(\rho_i) \hat{g}(1 - \rho_j) + \frac{a_{ji}}{k_j} f(\rho_j) \hat{g}(1 - \rho_i) \right].$$

By introducing the random walk Laplacian  $\Delta_{ij} = a_{ij}/k_j - \delta_{ij}$  and making use of the symmetry of the adjacency matrix, we obtain the sought equation for the time evolution of the density  $\rho_i$ :

$$\frac{d}{d\tau} \rho_i = \sum_j \Delta_{ij} \left[ f(\rho_j) \hat{g}(1 - \rho_i) - \frac{k_j}{k_i} f(\rho_i) \hat{g}(1 - \rho_j) \right]. \quad (\text{A2})$$

### Appendix B: Asymptotic solution

We now set to calculate the asymptotic density  $\rho_i^*$  as displayed on each node of the network. To do this end we equate to 0 the right hand side of Eq. (A2), and rewrite the ensuing equation as follows:

$$\forall i = 1, \dots, N \quad 0 = \sum_j \Delta_{ij} \psi_j(i) \quad \text{where} \quad \psi_j(i) = \left[ f(\rho_j) \hat{g}(1 - \rho_i) - \frac{k_j}{k_i} f(\rho_i) \hat{g}(1 - \rho_j) \right],$$

that is for all  $i$ ,  $(\psi_1(i), \dots, \psi_N(i))^T$  should be the eigenvector of the random walk Laplace matrix  $\mathbf{\Delta}$  associated with the null eigenvalue. In other words, for some constant  $\mu(i)$ :

$$\psi_j(i) = k_j \mu(i).$$

Observe that  $\psi_i(i) = 0$  for all  $i$  and thus  $k_i\mu(i) = 0$ , which implies  $\mu(i) = 0$ . Indeed,  $k_i \neq 0$  for all  $i$  since the network is connected. In conclusion, the asymptotic solution  $\rho_i^*$  must satisfy:

$$f(\rho_j^*)\hat{g}(1 - \rho_i^*) - \frac{k_j}{k_i}f(\rho_i^*)\hat{g}(1 - \rho_j^*) = 0 \quad \forall i, j.$$

Reordering the terms one gets:

$$\frac{f(\rho_j^*)}{k_j\hat{g}(1 - \rho_j^*)} = \frac{f(\rho_i^*)}{k_i\hat{g}(1 - \rho_i^*)} \quad \forall i, j.$$

The above condition is met, for any  $i$  and  $j$ , only if the terms on the right and left hand-side equate to a constant  $c$  (namely if they do not bear a reflex of the associated index):

$$\frac{f(\rho_i^*)}{k_i\hat{g}(1 - \rho_i^*)} = c \quad \forall i. \quad (\text{B1})$$

For generic  $f$  and  $\hat{g}$  the previous equation can exhibit multiple solutions. To rule out such possibility, and eventually focus on the interesting setting where just one solution is allowed for, we can assume: (i)  $f$  to be non-decreasing function, vanishing at  $x = 0$ ; (ii)  $\hat{g}$  to be non-increasing function, vanishing at  $x = 1$ . In such a way, by continuity, the curves  $f(\rho)$  and  $ck\hat{g}(1 - \rho)$  intersect only once, for any choice of  $c > 0$  and  $k > 0$ .

### 1. About $k_{crit}$ .

In Fig.1 (main text) we have shown the non trivial behaviour of the stationary solution  $\rho_i^*$  as a function of  $\sigma$  and the node degree  $k_i$  responsible for the interesting phenomenon of accumulation / depletion of hubs and leaves with respect to the case  $\sigma = 1$ . For any given  $\sigma > 0$  there exists a unique critical values for the node degree,  $k_{crit}$ , where such inversion takes place which indirectly defines ‘‘large’’ versus ‘‘small’’ degrees.

To compute such critical value we need to impose the equality among the stationary solution  $\rho^*|_{\sigma}$ , for  $\sigma \neq 1$ , and the same quantity for  $\sigma = 1$ ,  $\rho^*|_{\sigma=1}$ , both associated to a node with degree  $k$ . From Eq. (4) (main text) with  $\sigma = 1$  we obtain  $\rho^*|_{\sigma=1} = c_1k/(1 + c_1k)$ ; assuming  $\rho^*|_{\sigma=1} = \rho^*|_{\sigma}$  and substituting this value again in (4) we get

$$\frac{c_1k_{crit}}{1 + c_1k_{crit}} = c_{\sigma}k_{crit} \left( 1 - \frac{c_1k_{crit}}{1 + c_1k_{crit}} \right)^{\sigma},$$

from which we straightforward obtain

$$\frac{c_1}{c_{\sigma}} = \frac{1}{(1 + c_1k_{crit})^{(\sigma-1)}},$$

which gives the Eq. (5) (main text).

### 2. The case of $k$ -regular networks

The asymptotic solution Eq. (B1) simplifies in the case of  $k$ -regular networks, for which i.e.  $k_i = k$  for all  $i$ ; in this case indeed, the dependence on the node index  $i$  disappears and thus all the nodes will display the same asymptotic density. The total mass conservation allows to determine the latter as

$$\rho_i^* = \beta \quad \forall i = 1, \dots, N, \quad (\text{B2})$$

independently of the nonlinear functions  $f$  and  $g$ . These latter are instead used in determining the normalising constant  $c$  entering in Eq. (B1)

$$c = \frac{1}{k} \frac{f(\beta)}{\hat{g}(1 - \beta)}. \quad (\text{B3})$$

Given the exact asymptotic solution one can explicitly compute the entropy rate given by Eq. (4) (in the main text with the choice  $f(x) = x$  and  $g(x) = (1 - x)^{\sigma}$  or the following Eq. (C1)). Indeed the sum over the index  $j$  allows to

simplify  $\sum_j a_{ij}$  with the degree  $k_i$  at the denominator, while the second sum returns the factor  $N$ , being the remaining part independent from  $i$ . One gets therefore:

$$h = -N\beta f(\beta)g(\beta) \log \left[ \frac{f(\beta)g(\beta)}{k} \right],$$

where we have used that  $g(x) = \hat{g}(1-x)$ .

Assuming  $f(x) = x$  and  $g(x) = (1-x)^\sigma$  one can compute the value of  $\beta$  which maximises  $h$  for a fixed  $\sigma$ , namely  $\beta^{\text{opt}}$ . Moreover one can calculate the parameter  $\sigma^{\text{opt}}$  which returns the maximum of  $h$  for a fixed  $\beta$ . To this end one needs to perform the partial derivative,  $\partial_\beta h$ , respectively  $\partial_\sigma h$ , and equating these latter to 0. In this way one can draw an interesting conclusion on  $\sigma^{\text{opt}}$ ; indeed one can obtain

$$\sigma^{\text{opt}} = \frac{\log[k/(e\beta)]}{\log(1-\beta)}, \quad (\text{B4})$$

and thus if  $k < e\beta$  one gets  $\sigma^{\text{opt}} > 0$  while on the contrary one obtains  $\sigma^{\text{opt}} < 0$ . The first constraint can be realised only with  $k = 2$ , that is for a 1D ring where each node is connected to its two neighbours, one on the left and one on the right, and for  $\beta$  sufficiently large, i.e.  $\beta > 2/e \sim 0.736$ . These facts can explain why in the case of the Erdős-Rényi and scale-free networks one always found  $\sigma^{\text{opt}} < 0$  and thus  $h$  is a decreasing function of  $\sigma$  (see Fig. 7).

The computation for  $\beta^{\text{opt}}$  follows the same reasoning. One can in particular obtain an implicit equation for the optimal value of  $\beta$  for a generic function  $g(x)$ :

$$\log \frac{\beta^{\text{opt}} g(\beta^{\text{opt}})}{k} = -1 + \frac{g(\beta^{\text{opt}})}{g(\beta^{\text{opt}}) + (\beta^{\text{opt}} g(\beta^{\text{opt}}))'}$$

For the particular choice  $g(x) = (1-x)^\sigma$  we obtain:

$$\log \frac{\beta^{\text{opt}}(1-\beta^{\text{opt}})^\sigma}{k} = -1 + \frac{1}{2 - \frac{\sigma\beta^{\text{opt}}}{1-\beta^{\text{opt}}}}$$

We conclude this section by observing that more regular topologies can be associated to larger entropy rates and thus to a stronger ergodic behaviour. In particular in Fig. 5 we compare the maximum of the entropy rate achieved for a  $k$ -regular 1D lattice against the same quantity computed for an Erdős-Rényi network with the same average degree and the same number of nodes. We can observe that for all the values of the average degree, the entropy rate is always larger in the case of the regular lattice than for the random network.

**Remark 1** (The  $k$ -Cayley trees). *A similar analysis can be performed in the case of  $k$ -Cayley trees, where each node has degree  $k$  (also called coordination number), but the leaves that by definition have degree 1. This implies that there will be two values for the asymptotic density, one associated to the leaves,  $\rho_{\text{out}}^*$ , and one for the remaining nodes, i.e. the inner ones,  $\rho_{\text{inn}}^*$ , determined by:*

$$\frac{\rho_{\text{inn}}^*}{k(1-\rho_{\text{inn}}^*)^\sigma} = c \quad \text{and} \quad \frac{\rho_{\text{out}}^*}{(1-\rho_{\text{out}}^*)^\sigma} = c. \quad (\text{B5})$$

*The constraint on the conservation of the total mass and the observation that in the limit of infinitely large Cayley tree, i.e. for a diverging number of shells, the number of inner nodes divided by the number of leaves converges to  $1/(k-2)$ , provide a third relation:*

$$\rho_{\text{inn}}^* \frac{1/(k-2)}{1+1/(k-2)} + \rho_{\text{out}}^* \frac{1}{1+1/(k-2)} = \beta. \quad (\text{B6})$$

*From Eqs. (B5) and (B6) one can determine the three variables  $\rho_{\text{inn}}^*$ ,  $\rho_{\text{out}}^*$  and  $c$ , and then again the entropy rate  $h$ . Let us observe that in this limiting case the average degree of the Cayley tree converges to 2 and thus we cannot fairly compare its entropy rate with the one obtained for the  $k$ -regular 1D lattice or the Erdős-Rényi network with the same average degree.*

### 3. Analytical approximation for the asymptotic solution, when $f(x) = x$ and $g(x) = (1-x)^\sigma$

Assuming  $f(x) = x$  and  $g(x) = (1-x)^\sigma$ ,  $\sigma > 0$ , for  $0 \leq x < 1$  and 0 otherwise, the asymptotic solution for the density Eq. (B1) is implicitly given by

$$\rho_i^* = k_i c (1 - \rho_i^*)^\sigma \quad \forall i. \quad (\text{B7})$$



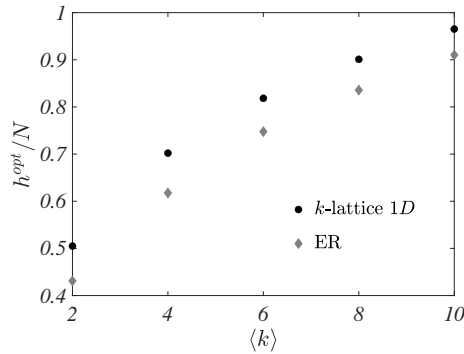


FIG. 5. **Entropy and topology regularity.** We compare, as a function of the network connectivity, the maximum of the entropy rate per node computed for a  $k$ -regular 1D-lattice and an Erdős-Rényi network with the same average degree and the same number of nodes ( $N = 50$ ). Results show that the regular topology always exhibits the highest maximum. For large network connectivities the two computed quantities converge to a shared value (indeed both networks converge to the same complete network).

In the following we shall write  $\rho_i^*(\sigma)$  to stress the dependence on the parameter  $\sigma$ . For  $\sigma = 1$  the solution to the latter problem takes the form [14]

$$\rho_i^*(1) = \frac{k_i c}{1 + k_i c} \quad \forall i.$$

Let us introduce  $y_i = 1 - \rho_i^*$  and rewrite the equation for the implicit solution as

$$1 - y = \kappa y^\sigma, \quad (\text{B8})$$

where for a sake of clarity we dropped the index  $i$  and we introduced  $\kappa = k_i c$ . One can thus look for a series expansion of  $y(\sigma)$  in terms of  $(\sigma - 1)$ , that should converge in a neighbourhood of  $\sigma = 1$ :

$$y(\sigma) = \sum_{n \geq 0} \frac{y_n}{n!} (\sigma - 1)^n.$$

Inserting this power series into Eq. (B8), recalling that

$$\frac{d}{d\sigma} [y(\sigma)]^\sigma = [y(\sigma)]^\sigma \log y + \sigma [y(\sigma)]^{\sigma-1} \frac{dy}{d\sigma},$$

and equating terms corresponding to the same powers of  $(\sigma - 1)$  on the left and the right hand sides of Eq. (B8), we can express  $y_n$  as a function of the terms  $y_l$ ,  $0 \leq l \leq n - 1$ . This recursive (infinite) system of equations can be explicitly solved. The first few terms are given by

$$\begin{aligned} y_0 &= \frac{1}{1 + \kappa} \\ y_1 &= -\frac{\kappa}{1 + \kappa} y_0 \log y_0 \\ y_2 &= -\frac{\kappa}{1 + \kappa} [(y_0 \log y_0 + y_1) \log y_0 + y_1] + \frac{\kappa^2}{(1 + \kappa)^2} y_0 \log y_0 (1 + \log y_0). \end{aligned}$$

Back to  $\rho_i^*(\sigma)$  we obtain

$$\rho_i^*(\sigma) = \frac{ck_i}{1+ck_i} + \frac{ck_i}{(1+ck_i)^2} \log(1+ck_i) \times (\sigma-1) + \rho_{i,2}^* \times \frac{(\sigma-1)^2}{2} + \mathcal{O}[(\sigma-1)^3], \quad (\text{B9})$$

with

$$\rho_{i,2}^* = \frac{ck_i \log(1+ck_i)}{(1+ck_i)^2} \left[ \left( \log(1+ck_i) - \frac{ck_i}{(1+ck_i)} \log(1+ck_i) \right) - \frac{ck_i}{(1+ck_i)} \right] - \frac{(ck_i)^2}{(1+ck_i)^2} \frac{\log(1+ck_i)}{1+ck_i} [1 - \log(1+ck_i)].$$

We can thus write, for  $\sigma \sim 1$ , the following approximate solution:

$$\rho_i^*(\sigma) = \rho_{i,0}^* + \rho_{i,1}^* \times (\sigma-1) + \rho_{i,2}^* \times \frac{(\sigma-1)^2}{2} + \mathcal{O}(\sigma-1)^3. \quad (\text{B10})$$

From the explicit form of the coefficients  $\rho_{i,n}^*$  one can analyse the dependence of the asymptotic density on the nodes degree and on the normalising parameter  $c$ , that, we recall, is a function of the crowding amount  $\beta$ , for fixed  $\sigma$ . Indeed we observe that for  $ck_i \gg 1$  the zeroth order correction is of the order of the unity,  $\rho_{i,0}^* \rightarrow 1$ , while the high order corrections,  $n \geq 1$ , do satisfy  $\rho_{i,n}^* = \mathcal{O}((\log ck_i)^n / ck_i)$ . Thus, they are negligible provided at least one among  $k_i$  and  $c$  is sufficiently large. The former condition implies that  $i$  is a hub, the latter amounts to operate under crowded conditions, namely  $\beta \rightarrow 1$ , which in turn implies  $c \rightarrow \infty$ . On the other hand,  $ck_i \ll 1$  (the network is connected and thus  $k_i \neq 0$  for all  $i$ ) yields  $\rho_{i,n}^* = \mathcal{O}((ck_i)^{n+1})$ ; hence, in very diluted conditions,  $\beta \rightarrow 0$ , high degree nodes can exhibit a very low density.

To check the accuracy of the approximation we quantify the discrepancy between the approximate formula Eq. (B9), up to a given order  $m$ ,  $\rho_i^{(m)}(\sigma)$ , and the exact numerical solution of Eq. (B7),  $\rho_i^*(\sigma)$ , both for the same fixed value of  $\sigma$ . The error is specifically defined as:

$$\delta_m(\sigma) = \max_{i=1,\dots,N} \left| 1 - \frac{\rho_i^{(m)}(\sigma)}{\rho_i^*(\sigma)} \right|.$$

Results reported in Fig. 6 testify on the accuracy of the proposed approximation for  $\sigma$  close to 1; observe that for over a significant window in  $\sigma$ , the error stays bounded to a few percents. The actual error depends also on the crowding parameter  $\beta$  and on the network topology. The top panels of Fig. 6 refer to a weakly crowded environment,  $\beta = 0.2$ , while bottom ones are obtained when considering a more pronounced degree of imposed crowding,  $\beta = 0.8$ . One can observe that the error deteriorates, as  $\beta$  increases. To test the impact of the topology of the underlying network, we created 10 Erdős-Rényi networks made by  $N = 100$  nodes and assuming a probability for the existence of a link  $p = 0.2$ . For each network, we computed  $\delta_7(\sigma)$ . The solid line in the right panels of Fig. 6 displays the average of the computed errors, while the boundaries of the grey shadow are set at one standard deviation by the mean. A similar behaviour (data not shown) is obtained when employing different schemes of network generation (as adopted in the main body of the paper).

### Appendix C: Entropy rate and the Heterogeneous Mean Field hypothesis.

The aim of this section is provide additional information on the application of the approximate Heterogeneous Mean Field hypothesis (HMF). Working under this assumption, we will characterise the entropy rate and then derive a simplified formula which holds when correlations among nodes degree can be neglected.

Consider again the entropy rate given by:

$$h = - \sum_{ij} \rho_i^* a_{ij} \frac{f(\rho_i^*) g(1-\rho_j^*)}{k_i} \log \left[ a_{ij} \frac{f(\rho_i^*) g(1-\rho_j^*)}{k_i} \right], \quad (\text{C1})$$

where  $a_{ij}$  is the adjacency matrix of the underlying network,  $(k_i)_{1 \leq i \leq N}$  the nodes degree and  $\rho_i^*$  the stationary probability. The first step consists in reorganising the sums as follows

$$h = - \sum_i \rho_i^* \frac{f(\rho_i^*)}{k_i} \log \frac{f(\rho_i^*)}{k_i} \sum_j a_{ij} g(1-\rho_j^*) - \sum_i \rho_i^* \frac{f(\rho_i^*)}{k_i} \sum_j a_{ij} g(1-\rho_j^*) \log (a_{ij} g(1-\rho_j^*)).$$

Then we invoke the Heterogeneous Mean Field hypothesis, namely we aggregate together nodes which share the same connectivity. Instead of summing on the node's index, we perform the sum on the degree [10]. Let thus denote by

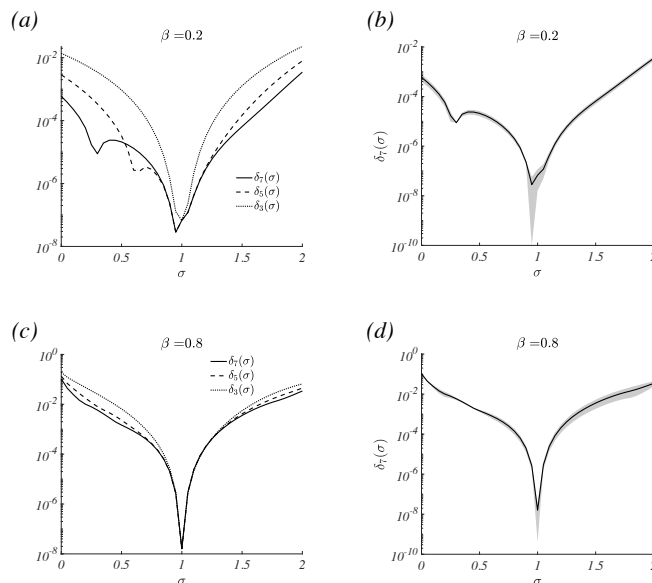


FIG. 6. **Analytical approximation for the asymptotic solution.** Left panels (a & c): We compare the approximate formula up to the  $m$ -th order ( $m = 3$  dotted line,  $m = 5$  dashed line and  $m = 7$  solid line) with the exact solution of Eq. (B7), as obtained via numerical methods. Right panels (b & d): for a fixed approximation order,  $m = 7$ , we compute the average error  $\delta_7(\sigma)$  over 10 realisations of the underlying network. The boundaries of the grey shadow are at one standard deviation from the mean. Top panels refer to  $\beta = 0.2$  while bottom ones to  $\beta = 0.8$ . The underlying network is generated according to the Erdős-Rényi recipe, with  $N = 100$  nodes and probability for the existence of a link  $p = 0.2$ .

$P(k)$  the probability for a generic node to have degree  $k$  and let  $P(k'|k)$  the conditional probability that a generic node with degree  $k$  is connected to a node with degree  $k'$ , then :

$$h_{\text{HMF}} = \frac{h}{N} = - \sum_k P(k) \hat{\rho}_k^* \frac{f(\hat{\rho}_k^*)}{k} \log \frac{f(\hat{\rho}_k^*)}{k} \sum_{k'} P(k'|k) k g(1 - \hat{\rho}_{k'}^*) - \sum_k P(k) \hat{\rho}_k^* \frac{f(\hat{\rho}_k^*)}{k} \sum_{k'} P(k'|k) k g(1 - \hat{\rho}_{k'}^*) \log (g(1 - \hat{\rho}_{k'}^*)),$$

where  $\hat{\rho}_k^*$  is the density of the nodes that share connectivity  $k$ .

Assuming an uncorrelated network,  $P(k'|k) = k'P(k')/\langle k \rangle$ , we get:

$$h_{\text{HMF}}^{\text{uncorr}} = - \sum_k P(k) \hat{\rho}_k^* f(\hat{\rho}_k^*) \log \frac{f(\hat{\rho}_k^*)}{k} \sum_{k'} \frac{k' P(k')}{\langle k \rangle} g(1 - \hat{\rho}_{k'}^*) - \sum_k P(k) \hat{\rho}_k^* f(\hat{\rho}_k^*) \sum_{k'} \frac{k' P(k')}{\langle k \rangle} g(1 - \hat{\rho}_{k'}^*) \log (g(1 - \hat{\rho}_{k'}^*)),$$

and using the equilibrium definition,  $g(1 - \hat{\rho}_{k'}^*) = f(\hat{\rho}_{k'}^*)/(ck')$ , we eventually get:

$$h_{\text{HMF}}^{\text{uncorr}} = - \sum_k P(k) \hat{\rho}_k^* f(\hat{\rho}_k^*) \log \frac{f(\hat{\rho}_k^*)}{k} \sum_{k'} \frac{P(k')}{\langle k \rangle} \frac{f(\hat{\rho}_{k'}^*)}{c} - \sum_k P(k) \hat{\rho}_k^* f(\hat{\rho}_k^*) \sum_{k'} \frac{P(k')}{\langle k \rangle} \frac{f(\hat{\rho}_{k'}^*)}{c} \log \frac{f(\hat{\rho}_{k'}^*)}{ck'},$$

and after some straightforward computations

$$\begin{aligned} h_{\text{HMF}}^{\text{uncorr}} &= - \frac{1}{c \langle k \rangle} \left[ \langle \hat{\rho}_k^* f(\hat{\rho}_k^*) \log \frac{f(\hat{\rho}_k^*)}{k} \rangle \langle f(\hat{\rho}_{k'}^*) \rangle + \langle \hat{\rho}_k f(\hat{\rho}_k) \rangle \langle f(\hat{\rho}_k^*) \log \frac{f(\hat{\rho}_k^*)}{ck} \rangle \right] = \\ &= - \frac{1}{c \langle k \rangle} \left[ \langle \hat{\rho}_k^* f(\hat{\rho}_k^*) \log f(\hat{\rho}_k^*) \rangle \langle f(\hat{\rho}_k^*) \rangle - \langle \hat{\rho}_k^* f(\hat{\rho}_k^*) \log k \rangle \langle f(\hat{\rho}_k^*) \rangle \right. \\ &\quad \left. + \langle \hat{\rho}_k^* f(\hat{\rho}_k^*) \rangle \langle f(\hat{\rho}_k^*) \log f(\hat{\rho}_k^*) \rangle - \langle \hat{\rho}_k^* f(\hat{\rho}_k^*) \rangle \langle f(\hat{\rho}_k^*) \log ck \rangle \right]. \end{aligned}$$

Recalling that  $f(x) = x$  (for the case analysed in the main body of the paper) we get:

$$h_{\text{HMF}}^{\text{uncorr}} = - \frac{1}{c \langle k \rangle} \left[ \langle (\hat{\rho}_k^*)^2 \log \hat{\rho}_k^* \rangle \langle \hat{\rho}_k^* \rangle - \langle (\hat{\rho}_k^*)^2 \log k \rangle \langle \hat{\rho}_k^* \rangle + \langle (\hat{\rho}_k^*)^2 \rangle \langle \hat{\rho}_k^* \log \hat{\rho}_k^* \rangle - \langle (\hat{\rho}_k^*)^2 \rangle \langle \hat{\rho}_k^* \log ck \rangle \right]. \quad (\text{C2})$$

In the main text we have shown that the entropy rate per node is a non-monotonic function of the crowding parameter  $\beta$  and thus the existence of an optimal value,  $\beta^{\text{opt}}$ , i.e. a value for which  $h$  attains its maximum. The

latter depends on the nonlinearity bias  $\sigma$ . In Fig. 7 we report results of some dedicated simulations proving that the optimal value of the crowding parameter is a decreasing function of the nonlinear bias  $\sigma$ , in the case of uncorrelated scale free networks, for several values of  $\gamma$ . A qualitatively similar result holds true also for other network topologies (data not shown).

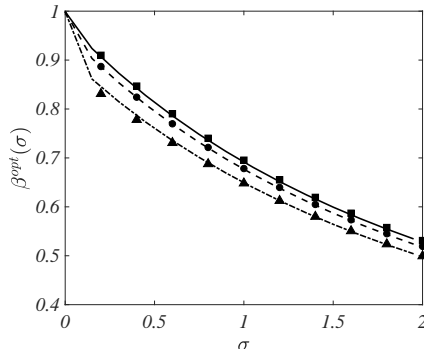


FIG. 7. **Optimal value for the crowding coefficient  $\beta^{\text{opt}}$  as a function of  $\sigma$ .** We report  $\beta^{\text{opt}}(\sigma)$  as numerically obtained by means of the HMF hypothesis (curves), under the assumption  $P(k) \sim 1/k^\gamma$  ( $\gamma = 3$  dashed line,  $\gamma = 2.5$  dash-dotted line and  $\gamma = 3.5$  solid line) and compare it with the corresponding quantity estimated for scale free networks which implement an identical scaling exponent ( $\gamma = 2.5$  triangles,  $\gamma = 3.0$  circles and  $\gamma = 3.5$  squares). Each symbol is the average over 10 different networks realisations.

To study the impact of the assortativity on the entropy rate, we build 400 random assortative and disassortative synthetic networks made of  $N = 1000$  nodes, each one characterised by its assortativity coefficient,  $r \in [-1, 1]$ . For each network we computed the entropy rate using Eq. (4) (in the main body of the paper), or equivalently Eq. (C2), that is taking into account the possible correlations among nodes degree; then we applied a degree-preserving randomisation process to the network, namely we build a null model by rewiring links without changing the nodes degree. Eventually we computed the entropy rate by using the nodes degree distribution  $P(k)$  common to both networks, via Eq. (C2). This amounts in turn to neglect the degree correlations. Let us observe that in this way both networks exhibit the same asymptotic nodes densities which depend only on the node degree for fixed  $\beta$  and  $\sigma$ , see Eq. (4) (in the main body of the paper). Hence, any possible differences as stemming from the usage of Eq. (C2) and Eq. (4) (in the main body of the paper) should be traced back to nodes degree correlations [34]. To measure the discrepancies as originated by the two aforementioned formulas, we define  $\Delta h = \max_\beta |h_{\text{HMF}} - h_{\text{HMF}}^{\text{uncorr}}|$ . The results reported in Fig. 8 show a dependence of the latter on  $r$ ;  $\Delta h$  vanishes for  $r \rightarrow 0$  coherently with the fact that, for non-assortative networks  $h_{\text{HMF}}$  and  $h_{\text{HMF}}^{\text{uncorr}}$  should eventually coincide. Then,  $\Delta h$  is maximal for  $|r| \rightarrow 1$  and in the worst case scenario, the difference is bounded by a few percents. We also stress that this behaviour does not depend on the value of the nonlinear bias  $\sigma$  imposed (data not shown).

#### Appendix D: Real and synthetic Networks

The synthetic uncorrelated scale-free networks used in our analyses have been created by means of the configuration model [5], i.e. by drawing a set of positive integer numbers according to the distribution  $\sim 1/k^\gamma$ , and then by using the latter as the degree sequence. In the Erdős-Rényi networks with  $N$  nodes, each of the  $N(N-1)/2$  edges is created independently with probability  $p_{\text{ER}}$ .

To study the behaviour of our model of nonlinear random walkers on real-world networks we have selected various networks from different domains and with a different number of nodes, links and level of assortativity. The networks considered and the results obtained are summarized in Table I. In particular, we have evaluated the maximal entropy rate  $h^{\text{opt}}$  for each network, and we have compared this value to that obtained in two different types of null models. The quantity  $\Delta h_{\text{rew}}^{\text{opt}}$  (resp.  $\Delta h_{\text{rnd}}^{\text{opt}}$ ) denotes the difference between the value of  $h^{\text{opt}}$  and that of a null model (both normalised to the system size) that consists in randomising the original network under the assumption of preserving

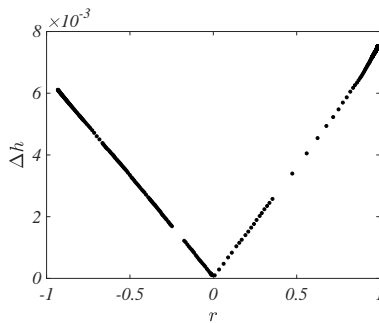


FIG. 8. **Entropy rate for assortative/disassortative synthetic networks.** We built several random assortative and disassortative synthetic networks, made of  $N = 1000$  nodes, each one characterised by its assortativity coefficient,  $r \in [-1, 1]$ . We then compute the deviation between the entropy rate  $h_{\text{HMF}}$  and the analogous quantity obtained when degree correlations are silenced using a null model,  $h_{\text{HMF}}^{\text{uncorr}}$ . We can observe a slight dependence of  $\Delta h = \max_{\beta} |h/N - h_{\text{HMF}}|$  on  $r$ , which in the worst case scenario reaches a few percents.

its node degrees (resp. the average degree) in the randomisation. In formulae:

$$\Delta h_{\text{rew}}^{\text{opt}} = \frac{h^{\text{opt}} - h_{\text{null,rew}}^{\text{opt}}}{N} \quad \text{and} \quad \Delta h_{\text{rnd}}^{\text{opt}} = \frac{h^{\text{opt}} - h_{\text{null,rnd}}^{\text{opt}}}{N}. \quad (\text{D1})$$

Let us observe that, by definition, networks in the first type of null model have the same degree sequence as the original real-world network. The asymptotic distribution in Eq. (B1) depends, for a fixed load  $\beta$ , only on the degree sequence of the network. We can hence conclude that both the real network and the rewired ones exhibit the same asymptotic density of the walkers  $\rho_i^*$  for all  $i = 1, \dots, N$ . From the expression of the entropy rate in Eq. (C1) we can thus obtain that  $h_{\text{null,rew}}$  differs from  $h$  only for the following contribution coming from differences in the adjacency matrices:

$$h - h_{\text{null,rew}} = - \sum_{ij} \rho_i^* (a_{ij} - a_{ij}^{(\text{null,rew})}) \frac{f(\rho_i^*)g(1 - \rho_j^*)}{k_i} \log \left[ \frac{f(\rho_i^*)g(1 - \rho_j^*)}{k_i} \right],$$

where  $a_{ij}$  is the adjacency matrix of the original network while  $a_{ij}^{(\text{null,rew})}$  the one obtained after the degree-preserving randomisation. Differences between the two matrices are limited by the constraint imposed by fixing the degree sequence, and so are the differences between the entropy rates. From the values of  $\Delta h_{\text{rew}}^{\text{opt}}$  we observe that in general real-world networks, with the exception of Internet at the autonomous systems level and some social networks perform better in terms of explorability especially in the case of walkers with large values of  $\sigma$ .

On the other hand, using a null model that only preserves the average degree of the original network, we will obtain larger variations of the entropy rates because now the asymptotic density will also differ (see Fig. 9). Randomised networks obtained in this way in general exhibit larger maximal entropy rates, with the notable exception of urban street patterns (see Table I and Fig. 10) that have not only a positive value of  $\Delta h_{\text{rew}}^{\text{opt}}$ , but also always a positive value of  $\Delta h_{\text{rnd}}^{\text{opt}}$ . In these systems, in fact, randomised networks with the same average degree performs worse in terms of explorability, namely their entropy rate is always smaller than the one of the original network. Based on this it is tempting to speculate that road network have been assembled so as to optimise their structure for transport under congested conditions: any randomised version, that disrupt the local organisation of crossroads, will perform worse.

Our results imply that the entropy rate per node, for the synthetic networks, the real ones or a randomised version of these latter, exhibits the same functional behaviour;  $h$  vanishes for very small and very large values of  $\beta$  and then achieves a single maximum at an intermediate value of the parameter. This value,  $\beta^{\text{opt}}$ , corresponds thus to an optimal network crowding (the total load on the network) that facilitates the exploration of the network, being a maximum of the entropy rate. In Table II we report, for three choices of the nonlinear parameter  $\sigma$ , the values of the optimal crowding computed for the real networks presented in Table I, together with the corresponding differences with respect to the same quantities computed for the null model networks obtained by randomising the original network under the

assumption of preserving its node degrees (resp. the average degree). In formulae:

$$\Delta\beta_{\text{rew}}^{\text{opt}} = \beta^{\text{opt}} - \beta_{\text{null,rew}}^{\text{opt}} \quad \text{and} \quad \Delta\beta_{\text{rnd}}^{\text{opt}} = \beta^{\text{opt}} - \beta_{\text{null,rnd}}^{\text{opt}}. \quad (\text{D2})$$

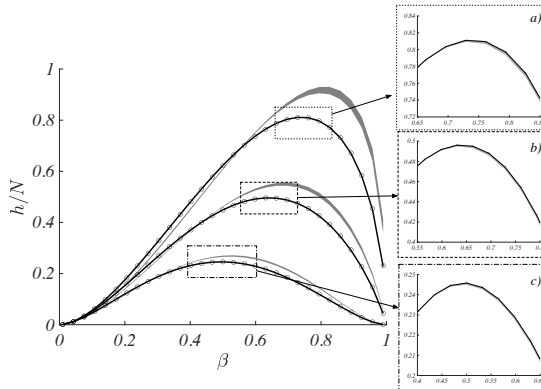


FIG. 9. **Comparison of the entropy rates for the different null models.** We report in the main panel the entropy rate for the real network “C. Elegans frontal” [30] (circles), the one computed using the HMF assumption (black line) and the one for the randomised network preserving the average degree (shaded grey areas corresponding to the interval  $[\min h/N, \max h/N]$  for the 50 replicas). In the panels a-b-c we zoom in the vicinity of the maximum of the entropy rate to appreciate the similarity with the entropy rate obtained for the degree-preserving null model.

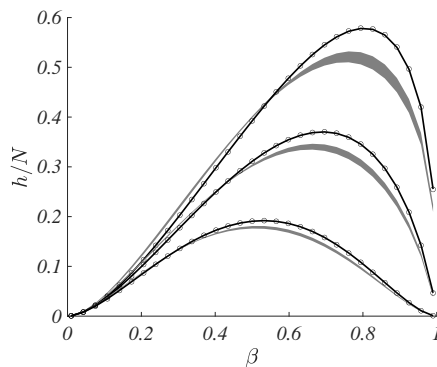


FIG. 10. **Entropy rates for the road networks.** We report the entropy rate for the road network of London [5, 31] (circles) together with the same quantity computed using the randomised network preserving the average degree (shaded grey areas corresponding to the interval  $[\min h/N, \max h/N]$  for the 50 replicas). The solid black line denotes the entropy rate computed using the HMF assumption, that also coincide with the entropy rate obtained using the randomised model that preserves the degree distribution (not visible at this scale).

To conclude, let us consider the role of the degree-degree correlations on the computation of the entropy rate of nonlinear random walkers. In few cases, e.g. the three Internet Autonomous Systems networks and the US Airplane network the randomisation process is not able to completely destroy the degree correlations and thus the entropy rate for the null model will deviated from the same quantity computed in the HMF approximation (see Fig. 11 top panels for the case of the US Airplane network). On the other hand, once the randomisation is able to wash out

the degree-degree correlations, we obtain a satisfying agreement between the null model entropy rate and the HMF approximation. We recall in fact that the HMF approximation implies neglecting degree-degree correlations (see Fig. 11 bottom panels for the case of the Facebook network).

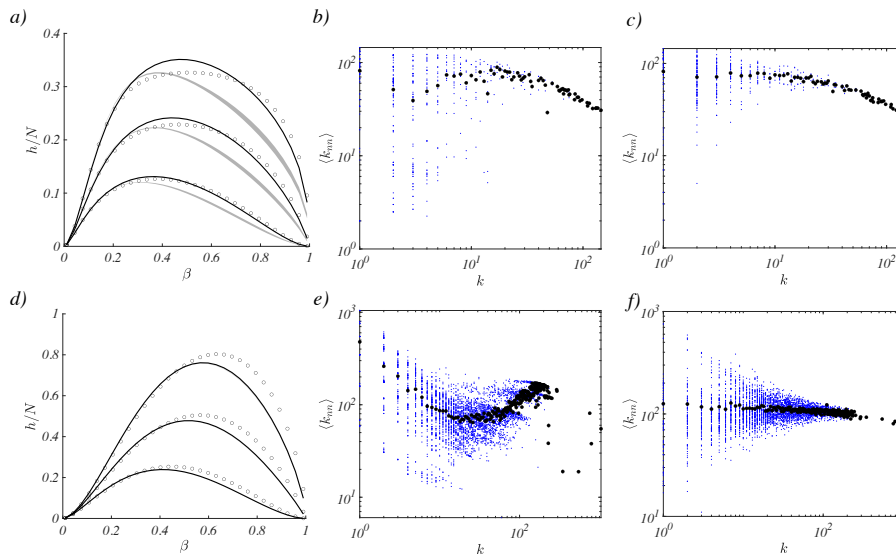


FIG. 11. **The role of degree-degree correlations.** Results reported for the US Airplane network [5, 26] (top panels) and Facebook [25] (bottom panels) indicate that, when the randomisation (preserving the degree distribution) is not able to completely wash out the degree-degree correlations, then the entropy rate of the randomised network (grey shaded area in panels a and d) differs from the one computed in the HMF approximation (black lines in panels a and d). In both cases, the circles denote the entropy rate for the original network. We report for the original network (panels b and e) and the randomised one (panels c and f) a measure of the presence of degree-degree correlations, namely the average degree  $\langle k_{nn} \rangle$  of neighbours of nodes of degree  $k$  as a function of  $k$ . The similarity of the scatter plots in panels b and c suggest the presence of degree-degree correlations in the randomised version of the US airplane network. Comparison of panels e and f suggests instead that correlations have been destroyed in the case of Facebook.

### Appendix E: Different choices of the bias functions $f$ and $g$

The aim of this section is to introduce and study some more general classes of biases functions  $f$  and  $g$ . The simplest straightforward generalisation is to consider

$$f(x) = x^\alpha \text{ and} \\ g(x) = (1-x)^\sigma \text{ for } 0 \leq x < 1 \text{ and } 0 \text{ otherwise,}$$

for some real positive  $\alpha$  and  $\sigma$ . Under such assumption the asymptotic solution given by Eq. (4) (in the main body of the paper) rewrites:

$$\frac{(\rho_i^*)^\alpha}{(1-\rho_i^*)^\sigma} = c_\sigma k_i \quad \forall i. \quad (\text{E1})$$

Observe that the function on the left hand side equals 0 for  $\rho_i^* = 0$  and monotonically diverge to  $+\infty$  for  $\rho_i^* \rightarrow 1^-$ . Hence, for any  $c_\sigma k_i$ , there exists one and only one value for  $\rho_i^*$  satisfying the equality. Stated differently looking for the asymptotic solution corresponds to finding the intersection of the two curves  $y = x^\alpha$  and  $y = \lambda(1-x)^\sigma$  (see Fig. 12 for two generic cases), and by continuity such intersection is always unique.

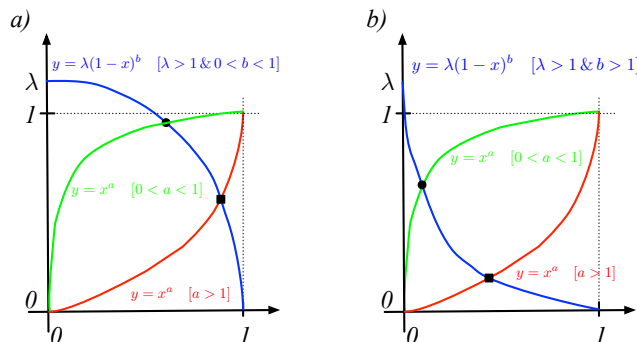


FIG. 12. Proof of the existence and uniqueness of the stationary solution via a qualitative approach. The solution of Eq. (E1) can be obtained as the intersection of the curves  $y = x^\alpha$  (red for  $\alpha > 1$  and green for  $\alpha < 1$ ) and  $y = \lambda(1-x)^\sigma$  (blue), where  $\lambda = c_\sigma k_i$  and  $x = \rho_i^*$ . For any choice of  $\alpha > 0$ ,  $\sigma > 0$  and  $\lambda > 0$  such two curves admit one and only one intersection in  $[0, 1]$ ; left panel  $0 < \sigma < 1$ , right panel:  $1 < \sigma$ .

- 
- [1] K. Pearson, *The problem of the random walk*, Nature, **72**, (1905), pp. 294
  - [2] Rayleigh, *The problem of the random walk*, Nature, **72**, (1905), pp. 318
  - [3] B.D. Hughes, *Random Walks and Random Environments, Vol. 1*, Clarendon Press, Oxford, (1995)
  - [4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, *Complex networks: Structure and dynamics*, Phys. Rep., **424**, (2006), pp. 424
  - [5] V. Latora, V. Nicosia, G. Russo, *Complex Networks: Principles, Methods and Applications*, Cambridge University Press, (2017)
  - [6] A. Barrat, M. Barthélemy, A. Vespignani, *Dynamical Processes on Complex Networks*, CUP, Cambridge, (2008)
  - [7] N. Masuda, M. A. Porter, and R. Lambiotte, *Random walks and diffusion on networks*, Phys. Rep., **716**, (2017)
  - [8] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, (1991)
  - [9] S. Meloni, J. Gómez-Gardenes, V. Latora and Y. Moreno, *Scaling breakdown in flow fluctuations on complex networks*, Phys. Rev. Lett., **100**, (2008), pp. 208701
  - [10] J. Gómez-Gardenes and V. Latora, *Entropy rate of diffusion processes on complex networks*, Phys. Rev. E, **78**, (2008), pp. 065102(R)
  - [11] S. Manfredi, E. Di Tucci, V. Latora, *Mobility and congestion in dynamical multilayer networks with finite storage capacity*, Phys. Rev. Lett., **120**, (2018), pp. 068301
  - [12] G. Cencetti, F. Battiston, D. Fanelli and V. Latora, *Reactive random walkers on complex networks*, Phys. Rev. E, **98**, (2018), pp. 052302
  - [13] V. Nicosia, F. Bagnoli, V. Latora, *Impact of network structure on competitive evolutionary processes*, Europhys. Lett., **94**, (2011), pp. 68009



Networks	$N$	$\langle k \rangle$	$r$	$\sigma = 0.5$ $\Delta h_{\text{rew}}^{\text{opt}}, \Delta h_{\text{rnd}}^{\text{opt}}$	$\sigma = 1$ $\Delta h_{\text{rew}}^{\text{opt}}, \Delta h_{\text{rnd}}^{\text{opt}}$	$\sigma = 2$ $\Delta h_{\text{rew}}^{\text{opt}}, \Delta h_{\text{rnd}}^{\text{opt}}$
Facebook [25]	4039	43.69	0.0636	$0.5 \cdot 10^{-1}, -6.0 \cdot 10^{-1}$	$0.3 \cdot 10^{-1}, -3.0 \cdot 10^{-1}$	$0.2 \cdot 10^{-1}, -1.2 \cdot 10^{-1}$
US Airplane [5, 26]	500	11.92	-0.2679	$0.3 \cdot 10^{-3*}, -6.0 \cdot 10^{-1}$	$0.6 \cdot 10^{-2}, -3.3 \cdot 10^{-1}$	$0.6 \cdot 10^{-2}, -1.5 \cdot 10^{-1}$
Internet Autonomous Systems network (AS-19971108) [5]	3015	3.42	-0.2289	$-1.8 \cdot 10^{-2}, -2.9 \cdot 10^{-1}$	$-1.3 \cdot 10^{-2}, -1.7 \cdot 10^{-1}$	$-7.8 \cdot 10^{-3}, -8.3 \cdot 10^{-2}$
Internet Autonomous Systems network (AS-19980402) [5]	3522	3.59	-0.210	$-1.9 \cdot 10^{-2}, -3.0 \cdot 10^{-1}$	$-1.4 \cdot 10^{-2}, -1.8 \cdot 10^{-1}$	$-8.0 \cdot 10^{-3}, -8.5 \cdot 10^{-2}$
Internet Autonomous Systems network (AS-19980703) [5]	3797	3.65	-0.207	$-1.8 \cdot 10^{-2}, -3.0 \cdot 10^{-1}$	$-1.3 \cdot 10^{-2}, -1.8 \cdot 10^{-1}$	$-7.8 \cdot 10^{-3}, -8.5 \cdot 10^{-2}$
C. Elegans frontal [30]	131	10.49	0.0218	$1.2 \cdot 10^{-5*}, -1.1 \cdot 10^{-1}$	$4.5 \cdot 10^{-4}, -5.4 \cdot 10^{-2}$	$4.6 \cdot 10^{-4}, -2.3 \cdot 10^{-2}$
Ahmedabad [5, 31]	2870	3.05	0.025	$0.3 \cdot 10^{-3}, 8.1 \cdot 10^{-2}$	$-0.2 \cdot 10^{-5*}, 4.2 \cdot 10^{-2}$	$0.3 \cdot 10^{-5*}, 1.9 \cdot 10^{-2}$
Barcelona [5, 31]	210	3.08	-0.037	$1.0 \cdot 10^{-3}, 3.2 \cdot 10^{-2}$	$0.5 \cdot 10^{-3}, 1.6 \cdot 10^{-2}$	$0.1 \cdot 10^{-3}, 6.7 \cdot 10^{-3}$
Bologna [5, 31]	541	2.85	0.024	$0.3 \cdot 10^{-3}, 5.2 \cdot 10^{-2}$	$0.2 \cdot 10^{-3}, 2.7 \cdot 10^{-2}$	$0.9 \cdot 10^{-4}, 1.2 \cdot 10^{-2}$
Cairo [5, 31]	1496	3.01	-0.0057	$0.1 \cdot 10^{-3}, 7.8 \cdot 10^{-2}$	$0.6 \cdot 10^{-4}, 4.1 \cdot 10^{-2}$	$0.1 \cdot 10^{-4}, 1.8 \cdot 10^{-2}$
London [5, 31]	488	2.99	-0.0301	$0.6 \cdot 10^{-3}, 5.7 \cdot 10^{-2}$	$0.3 \cdot 10^{-3}, 3.0 \cdot 10^{-2}$	$0.7 \cdot 10^{-4}, 1.3 \cdot 10^{-2}$
Venice [5, 31]	1840	2.61	-0.0921	$1.4 \cdot 10^{-3}, 3.8 \cdot 10^{-2}$	$0.6 \cdot 10^{-3}, 1.9 \cdot 10^{-2}$	$0.1 \cdot 10^{-3}, 8.6 \cdot 10^{-3}$
Medicis Family [5, 32]	15	2.67	-0.3748	$-0.8 \cdot 10^{-3*}, -8.2 \cdot 10^{-3}$	$-0.7 \cdot 10^{-3*}, -4.6 \cdot 10^{-3}$	$-0.4 \cdot 10^{-3*}, -1.9 \cdot 10^{-3}$
Zachary's Karate Club [5]	34	4.59	-0.4756	$-1.5 \cdot 10^{-3}, -1.2 \cdot 10^{-1}$	$-2.0 \cdot 10^{-3}, -6.3 \cdot 10^{-2}$	$-1.6 \cdot 10^{-3}, -2.7 \cdot 10^{-2}$
Kindergarten [5]	16	4.88	0.2234	$-1.2 \cdot 10^{-3}, -2.1 \cdot 10^{-2}$	$0.4 \cdot 10^{-3}, -1.2 \cdot 10^{-2}$	$0.7 \cdot 10^{-3}, -5.2 \cdot 10^{-3}$
Primates [5, 33]	14	4.43	-0.5046	$-0.6 \cdot 10^{-3}, -1.0 \cdot 10^{-1}$	$-0.1 \cdot 10^{-3*}, -5.2 \cdot 10^{-2}$	$0.7 \cdot 10^{-4*}, -2.3 \cdot 10^{-2}$

TABLE I. **Nonlinear random walkers on real-world networks.** Number of nodes  $N$ , average node degree  $\langle k \rangle$  and degree-degree correlation coefficient  $r$  [28, 29] for each of the networks considered. For three different choices of the nonlinear bias, namely  $\sigma = 0.5, 1$  and  $2$ , we report the difference between the maximum value of the entropy rate  $h^{\text{opt}}$  in the original network and in our two null models,  $\Delta h_{\text{rew}}^{\text{opt}}$  and  $\Delta h_{\text{rnd}}^{\text{opt}}$ . 50 different network realisations have been considered to evaluate averages and standard deviations for each null model. Values marked by \* are those smaller than one standard deviation.

Networks	$\sigma = 0.5$			$\sigma = 1$			$\sigma = 2$		
	$\beta^{\text{opt}}$	$\Delta\beta_{\text{rew}}^{\text{opt}}$	$\Delta\beta_{\text{rnd}}^{\text{opt}}$	$\beta^{\text{opt}}$	$\Delta\beta_{\text{rew}}^{\text{opt}}$	$\Delta\beta_{\text{rnd}}^{\text{opt}}$	$\beta^{\text{opt}}$	$\Delta\beta_{\text{rew}}^{\text{opt}}$	$\Delta\beta_{\text{rnd}}^{\text{opt}}$
Facebook [25]	$6.3 \cdot 10^{-1}$	$6.1 \cdot 10^{-2}$	$-2.3 \cdot 10^{-1}$	$5.7 \cdot 10^{-1}$	$5.1 \cdot 10^{-2}$	$-1.7 \cdot 10^{-1}$	$4.5 \cdot 10^{-1}$	$4.0 \cdot 10^{-2}$	$-1.0 \cdot 10^{-1}$
US Airplane [5, 26]	$5.1 \cdot 10^{-1}$	$1.2 \cdot 10^{-1}$	$-3.2 \cdot 10^{-1}$	$4.6 \cdot 10^{-1}$	$9.2 \cdot 10^{-2}$	$-2.4 \cdot 10^{-1}$	$3.8 \cdot 10^{-1}$	$6.7 \cdot 10^{-2}$	$-1.6 \cdot 10^{-1}$
Internet Autonomous Systems network (AS-19971108) [5]	$6.3 \cdot 10^{-1}$	$-7.6 \cdot 10^{-2}$	$-3.1 \cdot 10^{-2}$	$5.3 \cdot 10^{-1}$	$-6.7 \cdot 10^{-2}$	$-3.5 \cdot 10^{-2}$	$4.1 \cdot 10^{-1}$	$-5.0 \cdot 10^{-2}$	$-3.1 \cdot 10^{-2}$
Internet Autonomous Systems network (AS-19980402) [5]	$6.3 \cdot 10^{-1}$	$-7.2 \cdot 10^{-2}$	$-3.5 \cdot 10^{-2}$	$5.4 \cdot 10^{-1}$	$-6.4 \cdot 10^{-2}$	$-3.7 \cdot 10^{-2}$	$4.1 \cdot 10^{-1}$	$-4.8 \cdot 10^{-2}$	$-3.3 \cdot 10^{-2}$
Internet Autonomous Systems network (AS-19980703) [5]	$6.3 \cdot 10^{-1}$	$-7.0 \cdot 10^{-2}$	$-3.3 \cdot 10^{-2}$	$5.4 \cdot 10^{-1}$	$-6.2 \cdot 10^{-2}$	$-3.6 \cdot 10^{-2}$	$4.2 \cdot 10^{-1}$	$-4.7 \cdot 10^{-2}$	$-3.2 \cdot 10^{-2}$
C. Elegans frontal [30]	$7.3 \cdot 10^{-1}$	$2.2 \cdot 10^{-3}$	$-6.3 \cdot 10^{-2}$	$6.4 \cdot 10^{-1}$	$2.3 \cdot 10^{-3}$	$-4.4 \cdot 10^{-2}$	$5.0 \cdot 10^{-1}$	$2.1 \cdot 10^{-3}$	$-2.8 \cdot 10^{-2}$
Ahmedabad [5, 31]	$8.3 \cdot 10^{-1}$	$3.1 \cdot 10^{-3}$	$7.0 \cdot 10^{-2}$	$7.1 \cdot 10^{-1}$	$1.7 \cdot 10^{-3}$	$4.8 \cdot 10^{-2}$	$5.4 \cdot 10^{-1}$	$4.2 \cdot 10^{-4}$	$2.9 \cdot 10^{-2}$
Barcelona [5, 31]	$7.7 \cdot 10^{-1}$	$5.7 \cdot 10^{-4}$	$9.8 \cdot 10^{-3}$	$6.7 \cdot 10^{-1}$	$-5.4 \cdot 10^{-5}$	$7.7 \cdot 10^{-3}$	$5.2 \cdot 10^{-1}$	$-4.9 \cdot 10^{-4}$	$5.5 \cdot 10^{-3}$
Bologna [5, 31]	$7.9 \cdot 10^{-1}$	$4.0 \cdot 10^{-4}$	$3.6 \cdot 10^{-2}$	$6.9 \cdot 10^{-1}$	$2.0 \cdot 10^{-4}$	$2.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-1}$	$8.0 \cdot 10^{-6}$	$1.6 \cdot 10^{-2}$
Cairo [5, 31]	$8.2 \cdot 10^{-1}$	$4.6 \cdot 10^{-5}$	$6.6 \cdot 10^{-2}$	$7.0 \cdot 10^{-1}$	$-8.7 \cdot 10^{-6}$	$4.5 \cdot 10^{-2}$	$5.4 \cdot 10^{-1}$	$-5.1 \cdot 10^{-5}$	$2.8 \cdot 10^{-2}$
London [5, 31]	$8.0 \cdot 10^{-1}$	$2.7 \cdot 10^{-4}$	$4.3 \cdot 10^{-2}$	$6.9 \cdot 10^{-1}$	$-9.5 \cdot 10^{-5}$	$2.9 \cdot 10^{-2}$	$5.3 \cdot 10^{-1}$	$-3.3 \cdot 10^{-4}$	$1.8 \cdot 10^{-2}$
Venice [5, 31]	$7.8 \cdot 10^{-1}$	$5.3 \cdot 10^{-4}$	$2.0 \cdot 10^{-2}$	$6.8 \cdot 10^{-1}$	$-4.1 \cdot 10^{-4}$	$1.5 \cdot 10^{-2}$	$5.2 \cdot 10^{-1}$	$-9.1 \cdot 10^{-4}$	$9.9 \cdot 10^{-3}$
Medicis Family [5, 32]	$7.6 \cdot 10^{-1}$	$-2.6 \cdot 10^{-3}$	$-1.2 \cdot 10^{-2}$	$6.6 \cdot 10^{-1}$	$-1.8 \cdot 10^{-3}$	$-6.4 \cdot 10^{-3}$	$5.1 \cdot 10^{-1}$	$-1.3 \cdot 10^{-3}$	$-2.9 \cdot 10^{-3}$
Zachary's Karate Club [5]	$7.1 \cdot 10^{-1}$	$-9.8 \cdot 10^{-3}$	$-3.9 \cdot 10^{-2}$	$6.2 \cdot 10^{-1}$	$-9.5 \cdot 10^{-3}$	$-3.3 \cdot 10^{-2}$	$4.8 \cdot 10^{-1}$	$-8.1 \cdot 10^{-3}$	$-2.2 \cdot 10^{-2}$
Kindergarten [5]	$7.7 \cdot 10^{-1}$	$3.9 \cdot 10^{-3}$	$-2.6 \cdot 10^{-2}$	$6.7 \cdot 10^{-1}$	$4.1 \cdot 10^{-3}$	$-1.7 \cdot 10^{-2}$	$5.2 \cdot 10^{-1}$	$3.6 \cdot 10^{-3}$	$-1.0 \cdot 10^{-2}$
Primates [5, 33]	$7.3 \cdot 10^{-1}$	$8.6 \cdot 10^{-4}$	$-4.9 \cdot 10^{-2}$	$6.3 \cdot 10^{-1}$	$1.1 \cdot 10^{-3}$	$-3.7 \cdot 10^{-2}$	$4.9 \cdot 10^{-1}$	$1.1 \cdot 10^{-3}$	$-2.4 \cdot 10^{-2}$

TABLE II. **Nonlinear random walkers on real-world networks (II)**. For three different choices of the nonlinear bias, namely  $\sigma = 0.5, 1$  and  $2$ , we report the optimal value of the crowding parameter,  $\beta^{\text{opt}}$ , for the original network and in the two null models we used (degree distribution and average degree preserving)  $\Delta\beta_{\text{rew}}^{\text{opt}}$  and  $\Delta\beta_{\text{rnd}}^{\text{opt}}$ . 50 different network realisations have been considered to evaluate averages and standard deviations for each null model.

- [14] M. Asllani et al., *Hopping in the Crowd to Unveil Network Topology*, Phys. Rev. Lett., **120**, (2018), pp. 158301
- [15] D. Fanelli and A. J. McKane, *Diffusion in a crowded environment*, Phys. Rev. E, **82**, (2010), pp. 021113
- [16] A. McKane and T. J. Newman, Phys. Rev. Lett., **94**, (2005), pp. 218102
- [17] C. A. Lugo and A. J. McKane, Phys. Rev. E, **78**, (2008), pp. 051911
- [18] T. Biancalani, D. Fanelli, F. Di Patti, Phys. Rev. E, **81**, (2010), pp. 046215
- [19] M. Asllani, T. Biancalani, D. Fanelli, A. McKane Europ. Phys. J. B, **86**, (2013), pp. 476
- [20] P. de Anna et al., Phys. Rev. E, **81**, (2010), pp. 056110
- [21] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, Elsevier, Amsterdam, (2007)
- [22] Z. Burda, J. Duda, J. M. Luck, and B. Waclaw, *Localization of the Maximal Entropy Random Walk*, Phys. Rev. Lett., **102**, (2009), pp.160602
- [23] R. Sinatra, J. Gómez-Gardenes, R. Lambiotte, V. Nicosia and V. Latora, *Maximal-entropy random walks in complex networks with limited information*, Phys. Rev. E, **83**, (2011), pp. 030103
- [24] S.N. Dorogovtsev, A.V. Goltsev, J. F. F. Mendes, *Critical phenomena in complex networks*, Rev. Mod. Phys., **80**, (2008), pp. 1275
- [25] J. Leskovec and J. J. McAuley, *Learning to discover social circles in ego networks*, Advances in neural information processing systems, (2012), pp. 539
- [26] V. Colizza, R. Pastor-Satorras, and A. Vespignani. *Reaction-diffusion processes and metapopulation models in heterogeneous networks*, Nat. Phys., **3**, (2007), pp. 276.
- [27] I. Iacopini, S. Milojević and V. Latora, *Network Dynamics of Innovation Processes*, Phys. Rev. Lett., **120**, (2018), pp. 048301
- [28] M.E.J. Newman, *Phys. Rev. Lett.*, **89**, (2002), pp. 208701
- [29] M.E.J. Newman, *Phys. Rev. E*, **67**, (2003), pp. 026126
- [30] M. Kaiser and C. C. Hilgetag, *Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems*, PLoS Computational Biology, **2**, (7), (2006), p. e95.
- [31] P. Crucitti, V. Latora and S. Porta, *Centrality measures in spatial networks of urban streets*, *Phys. Rev. E*, **73**, (2006), pp. 036125.
- [32] S. Wasserman and K. Faust, *Social network analysis: Methods and applications*, **8**, Cambridge university press, (1994).
- [33] M. G. Everett and S. P. Borgatti, *The centrality of groups and classes*, J. Math. Sociol. **23**, (1999), pp. 181-201.
- [34] A randomisation algorithm preserving the total number of nodes and the average degree will generically induce different asymptotic nodes densities and thus deviation of  $h_{\text{HMF}}^{\text{uncorr}}$  from  $h_{\text{HMF}}$  could be imputed to both the nodes degree correlation and the different nodes densities.