Global mean first-passage times of random walks on complex networks

V. Tejedor,[1,2] O. Bénichou,[1] and R. Voituriez[1]

[1]*Laboratoire de Physique Théorique de la Matière Condensée (UMR 7600), Université Pierre et Marie Curie, 4 Place Jussieu, 75255 Paris Cedex, France*

[2]*Physics Department, Technical University of Munich, James Franck Strasse, 85747 Garching, Germany*

We present a general framework, applicable to a broad class of random walks on complex networks, which provides a rigorous lower bound for the mean first-passage time of a random walker to a target site averaged over its starting position, the so-called global mean first-passage time (GMFPT). This bound is simply expressed in terms of the equilibrium distribution at the target and implies a minimal scaling of the GMFPT with the network size. We show that this minimal scaling, which can be arbitrarily slow, is realized under the simple condition that the random walk is transient at the target site and independently of the small-world, scale-free, or fractal properties of the network. Last, we put forward that the GMFPT to a specific target is not a representative property of the network since the target averaged GMFPT satisfies much more restrictive bounds.

I. INTRODUCTION

Complex networks have appealed a lot of interest in the past few years [1–3], mainly because of the broad range of systems that they model, from biology to computer science or sociology. Despite this variety and their intrinsic topological complexity, many real networks have been shown to share some common features, such as the small-world property [2,4], the scale-free property [1,5], or even fractal scalings [6]. A crucial issue, still debated, is to understand the impact of the topological complexity of such systems on transport properties. As a paradigm of dynamical processes, random walks on complex networks have been intensely studied [7–11] and in particular first-passage times have been widely used as a quantitative indicator of transport efficiency [12,13]. The mean first-passage time (MFPT) to a target point was for instance calculated in the case of fractal networks [14,15].

Following the seminal work of Montroll [16], many papers have focused on the MFPT averaged over the starting point of the walker [8,17–25], sometimes called the global mean first-passage time (GMFPT). Recently, a sublinear dependence on the size N of the network of the GMFPT to the most connected node of network examples was shown [18,24] and was interpreted as favorable for an efficient trapping. This finding, in strong contrast with previously known results in the case of regular [16] or fractal lattices [8,19,21,22] lattices, has motivated an increasing number of works [20,23,25,26] that have tried to find examples of networks with high trapping efficiency, namely, displaying weaker and weaker dependence on N of the GMFPT. Relying on these specific examples, the heterogeneity, and more precisely the scale-free property, was put forward as advantageous [23,24], whereas the fractal property was suggested to be unfavorable [25].

Here, we propose a general framework, applicable to a broad class of networks, which deciphers the dependence of the GMFPT on the network size N and provides a global understanding of recent results obtained on specific examples

[19,23–26]. We first show on the example of a new set of networks that the GMFPT to the most connected node can scale as N^θ, with θ arbitrarily close to 0 despite the fractal property of the network. We then present an analytical approach which yields (i) rigorous bounds on the N dependence of the GMFPT and (ii) a simple criterion under which this bound is reached, which in particular provides a condition for a sublinear scaling with N, which is independent of the scale-free, small-world, or fractal nature of the network. Last, we show that a sublinear scaling is never representative of the network in the sense that the GMFPT averaged over the target site always scales faster than N.

II. DEFINITION OF THE PROBLEM AND NOTATIONS

We consider a set of graphs $\{\mathcal{G}_g\}_{g\in\mathbb{N}}$ where N_g denotes the number of sites of the graph $\mathcal{G}_g$ at generation g, such that $N_g \to \infty$ when $g \to \infty$. We consider a discrete time random walker on $\mathcal{G}_g$. We assume that the transition probabilities w_{ij} from site i to site j defining the walk are such that an equilibrium distribution P_{eq} satisfying detailed balance exists. We further assume that $\sup_{X\in\mathcal{G}_g} P_{\text{eq}}(X) \to 0$ when $g \to \infty$. We denote by $F_{S\to T}(\mathbf{T}=n)$ the probability that the walker reaches the target site T starting from site S for the first time after $\mathbf{T}=n$ steps and write $\overline{\mathbf{T}}_{S\to T}$ for the MFPT from S to T. Note that this first average $\overline{\cdot}$ is taken over the realizations of the random walk. Taking the average of the MFPT over the starting point, we define the GMFPT [denoted $G(T)$] by

$$G(T) = \langle \overline{\mathbf{T}}_{S\to T}\rangle_{S\neq T} = \frac{\underset{S\neq T}{\Sigma} P_{\text{eq}}(S)\overline{\mathbf{T}}_{S\to T}}{1 - P_{\text{eq}}(T)}. \tag{1}$$

Note that this quantity depends on the target point T. Here the space average $\langle\cdot\rangle$ is taken over the equilibrium distribution P_{eq} and slightly differs from the definition used in [20,23–25,27] where the average is taken over the flat distribution. It can be checked numerically on networks recently studied in the literature that both definitions lead to the same scaling with N_g.
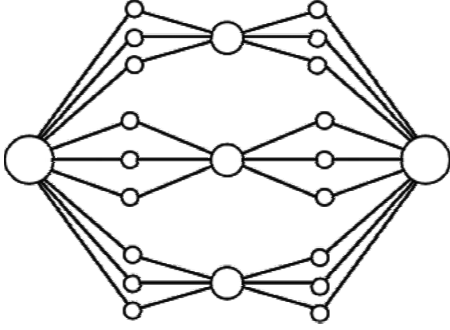
FIG. 1. A fractal network leading to efficient trapping by the hub T (on the most connected sites): case of $k=3$, $u=2$ at generation $g=3$.

III. EFFICIENT TRAPPING ON A FRACTAL NETWORK

We first exhibit a set of fractal networks which extends the so-called (u,u) flowers introduced in [28] and whose GMFPT to the most connected node scales as N^θ, with θ arbitrarily close to 0. Note that a similar scaling has been reported in [18] for an example of nonfractal network. The first generation of the graph consists in two nodes connected by one link; then, at each iteration, every link is broken and replaced by k paths of $u \geq 2$ links (cf. Fig. 1). It is clear that this network is fractal with a fractal dimension $d_f = \ln(ku)/\ln(u)$ since the diameter of the network at generation g is $L_g \sim u^g$ while the number of sites is $N_g \sim (ku)^g$ (the usual (u,u) flowers in [28] correspond to the special case $k=2$). Taking the target as one of the initial nodes, it is easily seen that the determination of the GMFPT is actually a one-dimensional (1D) problem since all the points $n(r)$ at the same distance r of the target are equivalent by symmetry and thus lead all to the same $\overline{T}(r)$. Noting next that for all $r \in [1, u^g - 1]$, $P_{\text{eq}}(r)n(r) = 2P_{\text{eq}}(T)$ and $P_{\text{eq}}(u^g)n(u^g) = P_{\text{eq}}(T)$, and using the classical 1D expression $\overline{T}(r) = r(2u^g - r)$ [29], we obtain the following exact expression: $G(T) = 2u^g(2u^g+1)/6 \propto N^{\ln(u^2)/\ln(ku)}$. Hence, for k large enough, the trapping at the hub is *arbitrarily efficient* on this network despite its fractal property.

IV. LOWER BOUND OF THE GMFPT

In order to gain understanding in the real parameters relevant to the scaling of the GMFPT with the size N, we now derive a general lower bound for the GMFPT. This derivation follows from the generalization of the Kac formula [30,31] which we briefly recall here for the sake of self-consistency. We start from the following backward equation satisfied by $F_{S \to T}$ for $n \geq 2$ (see [29]): $F_{S \to T}(n) = \Sigma_{j \neq T} w_{Sj}F_{j \to T}(n-1)$, which is completed by $F_{S \to T}(n=1) = w_{ST}$. Laplace transforming and averaging this equation over S [with a weight $P_{\text{eq}}(S)$ as in Eq. (1)] yield the generalized Kac formula,

$$\frac{P_{\text{eq}}(T)}{1-P_{\text{eq}}(T)}[\hat{F}_{T \to T}(s) - e^{-s}] = (e^{-s}-1)\langle \hat{F}_{S \to T}(s) \rangle_{S \neq T}, \quad (2)$$

where $\hat{F}_{S \to T}(s) \equiv \Sigma_{n=1}^{\infty} e^{-sn} F_{S \to T}(n)$. This very general equation, derived in a similar form in [31], relates the distribution

of the first return time to a site T to the distribution of the global first-passage time to T. Expanding Eq. (2) to first order in s yields the classical Kac formula $\overline{\mathbf{T}}_{T \to T} = 1/P_{\text{eq}}(T)$ [30,31]. In turn, the second order in s gives

$$G(T) = \frac{1}{2} \frac{P_{\text{eq}}(T)\overline{\mathbf{T}^2}_{T \to T} - 1}{1 - P_{\text{eq}}(T)}. \quad (3)$$

Using next $\overline{\mathbf{T}^2}_{T \to T} \geq \overline{\mathbf{T}}^2_{T \to T}$ and the classical Kac formula, the above exact expression gives a lower bound for the GMFPT,

$$G(T) \geq \frac{1}{2P_{\text{eq}}(T)}. \quad (4)$$

Note that this lower bound is in close analogy with the one obtained in [32] in the context of continuous space.

We now discuss under which conditions this lower bound is reached. Strictly speaking, this requires the very restrictive condition that the variance of $\mathbf{T}_{T \to T}$ is zero. More generally we can discuss under which conditions the right- and the left-hand sides of Eq. (4) share the same scaling in the large size limit. To do so, we consider a sequence of target sites $\{T_g \in \mathcal{G}_g\}_{g \in \mathbb{N}}$, which can be, for instance, hubs of the networks at each generation as in [20,23,24]. Using Eq. (4) and recalling that we have assumed $P_{\text{eq}}(T_g) \to 0$ for $g \to \infty$, we define the minimal scaling of the GMFPT for $g \to \infty$ by

$$G(T_g) = O[1/P_{\text{eq}}(T_g)]. \quad (5)$$

Equation (3) then shows that this minimal scaling is realized as soon as the reduced variance of the first return time is finite in the large size limit, namely, $(\overline{\mathbf{T}^2}_{T \to T} - \overline{\mathbf{T}}^2_{T \to T})/\overline{\mathbf{T}}^2_{T \to T} = O(1)$.

We now show that this condition for a minimal scaling with the network size N_g is actually equivalent to the transience property of the random walk at the target site T_g in the large size limit. We first derive an alternative exact expression for the GMFPT. Let us introduce the pseudo-Green's functions [14,31] defined as

$$H_{S \to T} = \sum_{n=1}^{\infty} [P_{S \to T}(n) - P_{\text{eq}}(T)], \quad (6)$$

where $P_{S \to T}(n)$ is the propagator, namely, the probability that the walker is at T at time n starting from S. It can be shown (see [9,14,31]) that the MFPT is then given by the exact expression $\overline{\mathbf{T}}_{S \to T} = [1/P_{\text{eq}}(T)](H_{T \to T} - H_{S \to T})$. Making use of the relation $P_{\text{eq}}(S)H_{S \to T} = P_{\text{eq}}(T)H_{T \to S}$, which follows from detailed balance (see also [9]), we obtain a second exact expression for the GMFPT,

$$G(T) = \frac{H_{T \to T}}{P_{\text{eq}}(T)[1 - P_{\text{eq}}(T)]}. \quad (7)$$

This equation provides an alternative condition under which the minimal scaling is realized, given by $H_{T \to T} = O(1)$ in the large g limit. From definition (6) of $H_{T \to T}$, this condition states that the random walk is transient at site T_g in the limit $g \to \infty$, i.e., that in this limit, a random walker returns on average only a *finite* number of times to T_g [33]. Conversely, Eq. (7) indicates that if the walk is recur-

rent at T_g for $g \to \infty$, that is if $H_{T_g \to T_g}$ diverges for $g \to \infty$, then $G(T_g)$ grows faster than $1/P_{eq}(T_g)$.

The lower bound [Eq. (4)] and minimal scaling [Eq. (5)] for the GMFPT obtained above call for comments. (i) First, our analysis puts forward a very general criterion to have a minimal scaling of the GMFPT with the network size, namely, the type (transient or recurrent) [33] of the random walk at the target site. We stress that this criterion is independent of the scale-free, small-world, or fractal properties of the network. Note that for a generic set of graphs $\{\mathcal{G}_g\}_{g \in \mathbb{N}}$, the type of the random walk for $g \to \infty$ is a site dependent property [8,28,33]. (ii) Second, the minimal scaling [Eq. (5)] is fully determined by the equilibrium distribution at the target site, which is generally much easier to obtain than dynamical quantities and which crucially depends on the connectivity of the target site. In particular for an isotropic random walk, for which $w_{ij} = 1/k_i$ if i and j are neighbors and else 0, where k_i denotes the connectivity of site i, the minimal scaling reads $G(T_g) \sim N_g \langle k \rangle / k_{T_g}$, where $\langle k \rangle$ is the connectivity averaged over all sites. (iii) Note finally that in the case of a recurrent random walk at the target the minimal scaling is not realized but the scaling of the GMFPT can however be sublinear if the growth of the connectivity at the target is fast enough.

Our analysis provides a comprehensive view of recent papers highlighting a sublinear dependence of the GMFPT to a hub on different examples of networks. (i) In the case of supercritical Erdös-Rényi and Molloy-Reed graphs studied in [18], both leading to transient random walks, the minimal scaling of Eq. (5) is indeed realized in agreement with our approach. (ii) The same remark applies to the example of deterministic scale-free graph proposed in [20], where the transience of the random walk at the target site (as defined above) is shown in the limit of large size since the probability to come back at the hub in a finite time is null. (iii) Zhang *et al.* [23,24] studied different examples of small-world scale-free networks (Apollonian networks [23] and (u,v) flowers [24]) where the GMFPT to the main hub displays a sublinear scaling. In these examples the scaling of $G(T_g)$ is strictly faster than our predicted minimal scaling $1/P_{eq}(T_g)$. Our criterion therefore implies that random walks on such structures are recurrent at the target site in the large size limit.

V. BOUNDS ON THE AVERAGE GMFPT

As demonstrated previously, the GMFPT highly depends on the target site, especially in the case of scale-free network where the connectivity can be very heterogeneous. Therefore the GMFPT to a specific target site cannot be taken as a general characteristic of the network. Actually, as we proceed to show the GMFPT averaged over the target site, defined by $\langle G \rangle = \Sigma_T P_{eq}(T) G(T)$, has scaling properties with N_g which can widely differ from the case of a fixed target site studied above. The inequality [Eq. (4)] gives straightforwardly the following lower bound for $\langle G \rangle$ (see also [30]): $\langle G \rangle \geq N_g/2$. Hence, the averaged GMFPT always scales faster than N_g and sublinear scalings discussed above are pointwise properties which are *never* representative of the network. Interest-

ingly, we can also propose an upper bound for $\langle G \rangle$ following [30]. First we define (see also [7,8]) the mean commute time as $\tau_{ij} = \bar{T}_{i \to j} + \bar{T}_{j \to i}$. The quantity τ_{ij} can actually be bounded using a very useful electrical analogy. Let us assign a unitary resistance to each link of the graph. Then it can be shown (see [34]) that the following general relation holds $\tau_{ij} = N_g \langle k \rangle r_{ij}$, where r_{ij} is the effective electrical resistance of the network between sites i and j. It is then straightforward to obtain that $r_{ij} \leq d_{ij}$, where d_{ij} is the distance between i and j. Indeed, d_{ij} is the resistance of a path of length d_{ij} between i and j and any parallel paths can only lower the resistance. We therefore finally obtain

$$\frac{N_g}{2} \leq \langle G \rangle \leq \frac{N_g \langle k \rangle \langle d \rangle}{2}, \qquad (8)$$

where $\langle d \rangle$ is the weighted average over pairs of the point to point distance d_{ij}.

Importantly, this shows that the scaling of $\langle G \rangle$ is much more constrained than the scaling of the GMFPT for a fixed target. This is particularly striking in the case of small-world networks for which $\langle d \rangle \sim \ln N_g$: hence in case of small-world networks with finite $\langle k \rangle$, widespread in nature [2], this shows that $\langle G \rangle$ always scales linearly with N_g (up to logarithmic corrections). Note also that these bounds [Eq. (8)] are compatible with the linear scaling of $\langle G \rangle$ with N_g reported in the case of Apollonian networks [13] and (1,2) flowers [8]. The conditions for which the scaling of each of the bounds in Eq. (8) is realized can also be discussed. As for the scaling of the lower bound, a sufficient condition for its realization is that for any sequence of targets $\{T_g \in \mathcal{G}_g\}_{g \in \mathbb{N}}$, the random walk is transient at T_g in the limit $g \to \infty$. Note however that this condition is not necessary, and the bound can be reached for networks having mixed type properties, as in the case of (1,2) flowers already mentioned [8]. As for the scaling of the upper bound, first notice that for any tree graph, r_{ij} is exactly the distance d_{ij} as discussed above using the electrical analogy. We conclude that for any tree the scaling of the upper bound is realized. In particular we find that $\langle G \rangle \sim N_g \ln N_g$ for *any* small-world tree.

Additional comments are in order. (i) First, Eq. (8) provides as a by-product an upper bound for the GMFPT itself, leading finally to $1/[2P_{eq}(T)] \leq G(T) \leq N_g \langle k \rangle \langle d \rangle / [2P_{eq}(T)]$. (ii) Second, this upper bound for GMFPT inductively gives an upper bound of the trapping time in the case of a moving target using the Pascal principle [35]. (iii) Last, we underline that in the case of fractal networks, characterized by a fractal dimension d_f and a walk dimension d_w [6] an explicit scaling of $\langle G \rangle$ can be obtained (see [8,19]). Indeed, using for instance the asymptotics of the MFPT between points separated by a distance r [14] and averaging over r, one gets the following scaling:

$$\langle G \rangle \sim \begin{cases} N_g & \text{if } d_w < d_f \\ N_g \ln(N_g) & \text{if } d_w = d_f \\ N_g^{d_w/d_f} & \text{if } d_w > d_f, \end{cases} \qquad (9)$$

which depends on the type of the random walk (transient if $d_f > d_w$).

VI. CONCLUSION

We have presented a general framework, applicable to a broad class of networks, which provides rigorous bounds on the size dependence of the GMFPT to a target site. We have shown that the GMFPT has the same scaling in the large size limit as this lower bound under the condition that the random walk is transient at the target site. This shows that the type of the random walk (transient or recurrent) is a crucial criterion to determine the scaling of the GMFPT, widely independent of its scale-free, small-world, or fractal properties. This study reconciles recent works on GMFPT for random walks on various network examples. Additionally, we have demonstrated that the scaling of the GMFPT to a specific target is not a representative property of the network since the target averaged GMFPT satisfies much more restrictive bounds, which in particular forbid any sublinear scaling with the network size.

[1] R. Albert and A.-L. Barabasi, Rev. Mod. Phys. **74**, 47 (2002).

[2] A. Barrat, M. Barthélémy, and A. Vespignani, *Dynamical Processes in Complex Networks* (Cambridge University Press, New York, 2008).

[3] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev. Mod. Phys. **80**, 1275 (2008).

[4] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[5] R. Albert, H. Jeong, and A.-L. Barabasi, Nature (London) **401**, 130 (1999).

[6] C. Song, S. Havlin, and H. A. Makse, Nature (London) **433**, 392 (2005); Nat. Phys. **2**, 275 (2006).

[7] A. Baronchelli, M. Catanzaro, and R. Pastor-Satorras, Phys. Rev. E **78**, 011114 (2008).

[8] E. M. Bollt and D. ben Avraham, New J. Phys. **7**, 26 (2005).

[9] J. D. Noh and H. Rieger, Phys. Rev. Lett. **92**, 118701 (2004).

[10] A. N. Samukhin, S. N. Dorogovtsev, and J. F. F. Mendes, Phys. Rev. E **77**, 036115 (2008).

[11] S. Nechaev and R. Voituriez, J. Phys. A **36**, 43 (2003); R. Voituriez, Nucl. Phys. B **621**, 675 (2002).

[12] L. K. Gallos, C. Song, S. Havlin, and H. A. Makse, Proc. Natl. Acad. Sci. U.S.A. **104**, 7746 (2007).

[13] Z. G. Huang, X. J. Xu, Z. X. Wu, and Y. H. Wang, Eur. Phys. J. B **51**, 549 (2006).

[14] S. Condamin *et al.*, Nature (London) **450**, 77 (2007).

[15] O. Bénichou, B. Meyer, V. Tejedor, and R. Voituriez, Phys. Rev. Lett. **101**, 130601 (2008).

[16] E. W. Montroll, J. Math. Phys. **10**, 753 (1969).

[17] V. Sood, S. Redner, and D. ben Avraham, J. Phys. A **38**, 109 (2005).

[18] A. Kittas *et al.*, EPL **84**, 40008 (2008).

[19] E. Agliari, Phys. Rev. E **77**, 011128 (2008).

[20] E. Agliari and R. Burioni, Phys. Rev. E **80**, 031125 (2009).

[21] C.P. Haynes and A. P. Roberts, Phys. Rev. E **78**, 041111 (2008).

[22] J. J. Kozak and V. Balakrishnan, Phys. Rev. E **65**, 021105 (2002).

[23] Z. Zhang *et al.*, EPL **86**, 10006 (2009).

[24] Z. Zhang, Y. Qi, S. Zhou, W. Xie, and J. Guan, Phys. Rev. E **79**, 021127 (2009).

[25] Z. Zhang *et al.*, e-print arXiv:0905.1521.

[26] Z. Zhang *et al.*, Phys. Rev. E **79**, 061113 (2009).

[27] Z. Zhang *et al.*, J. Math. Phys. **50**, 033514 (2009).

[28] H. D. Rozenfeld, S. Havlin, and D. ben Avraham, New J. Phys. **9**, 175 (2007).

[29] S. Redner, *A Guide to First Passage Time Processes* (Cambridge University Press, Cambridge, 2001).

[30] D. Aldous and J. Fill, Reversible Markov chains and random walks on graphs, 1999, www.stat.berkeley.edu /users/aldous/ RWG/book.html.

[31] S. Condamin, O. Benichou, and M. Moreau, Phys. Rev. Lett. **95**, 260601 (2005); Phys. Rev. E **75**, 021111 (2007).

[32] O. Bénichou, M. Coppey, M. Moreau, P. H. Suet, and R. Voituriez, Europhys. Lett. **70**, 42 (2005).

[33] R. Burioni and D. Cassi, J. Phys. A **38**, R45 (2005).

[34] A. K. Chandra, P. Raghavan, W. L. Ruzzo, and R. Smolensky, STOC '89: Proceedings of the 21st Annual ACM symposium on Theory of Computing, 1989 (unpublished).

[35] M. Moreau, G. Oshanin, O. Benichou, and M. Coppey, Phys. Rev. E **67**, 045104(R) (2003); **69**, 046101 (2004).