Journal of Physics A:
Mathematical and
Theoretical

Analytical results for the distribution of first return times of non-backtracking random walks on configuration model networks

To cite this article before publication: Dor Lev-Ari *et al* 2025 *J. Phys. A: Math. Theor.* in press https://doi.org/10.1088/1751-8121/ae257d

Analytical results for the distribution of first return times of non-backtracking random walks on configuration model networks

Dor Lev-Ari[1], Ido Tishby[1], Ofer Biham[1,*], Eytan Katzav[1] and Diego Krapf[2]

[1] Racah Institute of Physics, The Hebrew University, Jerusalem 9190401, Israel
[2] Department of Electrical and Computer Engineering and School of Biomedical Engineering, Colorado State University, Fort Collins, Colorado 80523, USA

E-mail: dor.lev-ari@mail.huji.ac.il, ido.tishby@mail.huji.ac.il, ofer.biham@mail.huji.ac.il, eytan.katzav@mail.huji.ac.il, diego.krapf@colostate.edu

Abstract.

We present analytical results for the distribution of first return (FR) times of non-backtracking random walks (NBWs) on undirected configuration model networks consisting of N nodes with degree distribution $P(k)$. We focus on the case in which the network consists of a single connected component. Starting from a random initial node i at time $t = 0$, an NBW hops into a random neighbor of i at time $t = 1$ and at each subsequent step it continues to hop into a random neighbor of its current node, excluding the previous node. We calculate the tail distribution $P(T_{\rm FR} > t)$ of first return times from a random initial node to itself. It is found that $P(T_{\rm FR} > t)$ is given by a discrete Laplace transform of the degree distribution $P(k)$. This result exemplifies the relation between structural properties of a network, captured by the degree distribution, and properties of dynamical processes taking place on the network. Using the tail-sum formula, we calculate the mean first return time $\mathbb{E}[T_{\rm FR}]$. Surprisingly, $\mathbb{E}[T_{\rm FR}]$ coincides with the result obtained from Kac's lemma that applies to simple random walks (RWs). We also calculate the variance ${\rm Var}(T_{\rm FR})$, which accounts for the variability of first return times between different NBW trajectories. We apply this formalism to Erdős-Rényi networks, random regular graphs and configuration model networks with exponential and power-law degree distributions and obtain closed-form expressions for $P(T_{\rm FR} > t)$ as well as its mean and variance. These results provide useful insight on the advantages of NBWs over simple RWs in network exploration, sampling and search processes.

Keywords: Random network, configuration model, random walk, first return time.

* Author to whom any correspondence should be addressed.

1. Introduction

Random walk (RW) models [1] provide useful tools for the analysis of dynamical processes on random networks [2–7]. Here we focus on the case of undirected networks. Starting at time $t = 0$ from a random initial node i, at each time step $t \geq 1$ an RW (also referred to as a simple RW) hops randomly to one of the neighbors of its current node. In some of the time steps the RW visits nodes that have not been visited before, while in other time steps it revisits nodes that have already been visited at an earlier time. The mean number $\langle S \rangle_t$ of distinct nodes visited by an RW on a random network up to time t was studied in Ref. [8]. It was found that in the infinite network limit, in which random networks exhibit a tree structure, $\langle S \rangle_t \simeq rt$, where the coefficient $r < 1$ depends on the network topology. In this case, the revisits are due to backtracking steps in which the RW hops back to the previous node and subsequent retroceding steps in which it keeps hopping backwards along its own path [9].

In order to perform systematic studies of random walks on random networks, it is useful to focus on configuration model networks. The configuration model is an ensemble of uncorrelated random networks consisting of N nodes, whose degree sequences are drawn from a given degree distribution $P(k)$. The admissible degrees are often restricted to a finite range $k_{\min} \leq k \leq k_{\max}$, where $k_{\min}$ is the minimal degree and $k_{\max}$ is the maximal degree, such that for any value of k outside this range $P(k) = 0$. The mean degree $\langle K \rangle$ is denoted by c. To ensure that an RW starting from any initial node i will be able to reach any other node j, we focus on the case in which the whole network consists of a single connected component. Using the terminology of percolation theory, these are networks in which the giant component encompasses the whole network. In the large network limit, a sufficient condition for a configuration model network to consist of a single connected component is that $k_{\min} \geq 3$ [6, 10]. In fact, a weaker condition of $k_{\min} \geq 2$ is sufficient in the large network limit as long as a finite fraction of the nodes satisfy $k \geq 3$ [11]. We thus avoid isolated nodes of degree $k = 0$ and leaf nodes of degree $k = 1$, which may form isolated tree structures.

The first return (FR) time $T_{\rm FR}$ of an RW is the first time at which it returns to the initial node i [12]. The first return time varies between different instances of the random walk trajectory and its properties can be captured by a suitable distribution. The distribution of first return times may depend on the specific realization of the random network and on the choice of the initial node i. The distribution of first return times from a random node to itself in a given ensemble of random networks is denoted by $P(T_{\rm FR} = t)$. A classical result regarding first return times is Kac's lemma, which states that the mean first return time of an RW from a given node i to itself is given by [13–15]

$$\mathbb{E}[T_{\rm FR}(i)] = \frac{1}{P_i(\infty)}, \tag{1}$$

where $P_i(\infty)$ is the probability that an RW will reside at node i at a given time step

under steady state conditions (which are achieved in the long time limit $t \to \infty$). In the case of undirected networks, Eq. (1) can be expressed in a more explicit form, namely [15]

$$\mathbb{E}[T_{\rm FR}(i)] = \frac{Nc}{k_i}, \tag{2}$$

where k_i is the degree of node i. This implies that the mean first return time from a random node to itself is given by

$$\mathbb{E}[T_{\rm FR}] = \left\langle \frac{Nc}{K} \right\rangle, \tag{3}$$

where $\langle X \rangle$ is the average of the random variable X over the degree distribution $P(k)$.

One can distinguish between two types of first return trajectories: first return trajectories in which the RW retrocedes its own steps backwards all the way back to the initial node i and first return trajectories in which the RW returns to i via a path that does not retrocede its own steps [9, 16]. In the retroceding trajectories, each edge that belongs to the RW trajectory is crossed the same number of times in the forward and backward directions. In the non-retroceding trajectories the RW path includes at least one cycle. In the infinite system limit, in which the network exhibits a tree structure, the only way to return to the initial node is via a retroceding trajectory [17, 18]. In finite networks both scenarios coexist, where the distribution $P(T_{\rm FR} = t)$ is dominated by retroceding trajectories at short times and by non-retroceding trajectories at long times [9, 19, 20].

A more general problem involves the calculation of the first passage (FP) time $T_{\rm FP}$, which is the first time at which a random walk starting from an initial node i at time $t = 0$ visits a specified target node j [12, 16, 21, 22]. The first return problem is a special case of the first passage problem, in which the initial node coincides with the target node. The distribution $P(T_{\rm FR} = t)$ of first return times of RWs was studied on the Bethe lattice, which exhibits a tree structure of an infinite size [17, 23–26] and on random regular graphs (RRGs) [9, 19].

An important variant of the RW model is the non-backtracking random walk (NBW), in which the move backwards to the previous node is excluded [27]. Since backtracking steps are excluded, in the infinite network limit in which the network exhibits a tree structure, an NBW never revisits a previously visited node. In particular, it never returns to the initial node. In a finite network, the first return process of NBWs takes place only via non-retroceding trajectories, which rely on the existence of cycles.

NBWs are important for several reasons, which are summarised below. They provide a more efficient way to explore and analyze complex networks, compared to standard random walks. This is due to the fact that by avoiding backtracking steps they can cover more of the network in less time [28]. NBWs are useful for identifying community structures within networks [29]. In certain types of networks, NBWs can help mitigate localization effects that might trap standard random walks in specific regions of the network.

In this paper we present analytical results for the distribution of first return times of NBWs on configuration model networks consisting of N nodes with degree distribution $P(k)$. An NBW starting from an initial node i forms a random trajectory in the network and eventually returns to i without backtracking its steps even once. In order to return to the initial node, the trajectory must include at least one cycle. The first return time may take any even or odd value that satisfies $T_{\rm FR} \geq 3$. Using probabilistic methods we calculate the tail-distribution of first return times $P(T_{\rm FR} > t|K = k)$ of NBWs starting from a random node of degree k. Averaging over the degree distribution, we obtain the overall tail distribution of first return times $P(T_{\rm FR} > t)$. We find that $P(T_{\rm FR} > t)$ is given by a discrete Laplace transform of the degree distribution $P(k)$. We calculate the mean first return time $\mathbb{E}[T_{\rm FR}]$ and show that it coincides with the result of Fasino et al. [30], which extends Kac's lemma to second order random walks. We also calculate the variance $\mathrm{Var}(T_{\rm FR})$ which accounts for the variability of first return times between different NBW trajectories. We apply this formalism to random regular graphs, Erdős-Rényi networks and configuration model networks with exponential and power-law degree distributions and obtain closed-form expressions for $P(T_{\rm FR} > t)$ and its first two moments. The analytical results are found to be in very good agreement with the results obtained from computer simulations.

The paper is organized as follows. In Sec. 2 we present the configuration model networks, their construction and essential properties. In Sec. 3 we present the non-backtracking random walk. In Sec. 4 we derive formulae for the distribution $P(T_{\rm FR} > t)$ of first return times of NBWs on configuration model networks and for its mean $\mathbb{E}[T_{\rm FR}]$ and variance $\mathrm{Var}(T_{\rm FR})$. In Sec. 5 we apply these results to RRGs, Erdős-Rényi (ER) networks and to configuration model networks with exponential and power-law distributions. The results are discussed in Sec. 6 and summarised in Sec. 7.

2. Configuration model networks

The configuration model is an ensemble of uncorrelated random networks whose degree sequences are drawn from a given degree distribution $P(k)$ [31–35]. These networks are simple graphs in the sense that each pair of nodes is connected by at most a single edge and there are no self-loops. The first moment (mean degree) and the second moment of $P(k)$ are denoted by $\langle K^n \rangle$, where $n = 1$ and 2, respectively, while the variance is given by $\mathbb{V}[K] = \langle K^2 \rangle - \langle K \rangle^2$. The support of the degree distribution of random networks is often bounded from below by $k_{\rm min} \geq 1$ such that $P(k) = 0$ for $0 \leq k \leq k_{\rm min} - 1$, with non-zero values of $P(k)$ only for $k \geq k_{\rm min}$. For example, the commonly used choice of $k_{\rm min} = 1$ eliminates the possibility of isolated nodes in the network. Choosing $k_{\rm min} = 2$ also eliminates the leaf nodes. One may also control the upper bound by imposing $k \leq k_{\rm max}$. This may be important in the case of finite networks with heavy-tail degree distributions such as power-law distributions. The configuration model network ensemble is a maximum entropy ensemble under the condition that the degree distribution $P(k)$ is imposed [32–34]. Here we focus on the case of undirected

networks.

To generate a network instance drawn from an ensemble of configuration model networks of N nodes, with a given degree distribution $P(k)$, one draws the degrees of the N nodes independently from $P(k)$. This gives rise to a degree sequence of the form $k_1, k_2, \ldots, k_N$. For the discussion below it is convenient to list the degree sequence in a decreasing order of the form $k_1 \geq k_2 \geq \cdots \geq k_N$. It turns out that not every possible degree sequence is graphical, namely admissible as a degree sequence of a network. Therefore, before trying to construct a network with a given degree sequence, one should first confirm the graphicality of the degree sequence. To be graphical, a degree sequence must satisfy two conditions. The first condition is that the sum of the degrees is an even number, namely $\sum_i k_i = 2L$, where L is an integer that represents the number of edges in the network. The second condition is expressed by the Erdős-Gallai theorem, which states that an ordered sequence of the form $k_1 \geq k_2 \geq \cdots \geq k_N$ that satisfies the first condition is graphical if and only if the condition

$$\sum_{i=1}^{n} k_i \leq n(n-1) + \sum_{i=n+1}^{N} \min(k_i, n) \tag{4}$$

holds for all values of n in the range $1 \leq n \leq N-1$ [36, 37].

To construct a network instance consisting of N nodes with a given degree sequence $k_1, k_2, \ldots, k_N$ (where $\sum_{i=1}^{N} k_i = 2L$ and L is the total number of undirected edges), we create a multiset of $2L$ stubs which includes k_i stubs for each node i. Pairs of stubs are then selected randomly and connected to each other to form edges between the corresponding nodes. To illustrate the process we represent the stubs by $2L$ balls, where the k_i balls associated with node i are marked by i. We then choose a random arrangement of the $2L$ balls in an array of L cells, such that each cell includes exactly two balls. In practice, a random arrangement of balls into cells can be obtained by generating a random permutation of the $2L$ balls and grouping them sequentially into L pairs, making the construction straightforward to implement. A cell containing balls i and j represents an edge between nodes i and j. The representation in terms of balls and cells is particularly convenient for implementation on the computer, since a single random permutation of the $2L$ balls produces a uniformly random pairing of stubs, from which the network can be constructed directly.

The network obtained from the procedure described above is a multigraph with the given degree sequence, which may include self-loops (edges connecting a node to itself) or multiple edges (two or more edges connecting the same pair of nodes). To eliminate the self-loops and multiple edges, we apply an edge switching process, which yields a simple graph while preserving the degree sequence. In this process, as long as the network has not yet become a simple graph, at each time step we select randomly one of the self-loops (i, i) or one of the multiple edges (i, j). In case that a self-loop (i, i) was selected, we select a random edge (i', j') and swap the two edges into (i, i') and (i, j'). Similarly, in case that a multiple edge (i, j) was selected, we select a random edge (i', j') and swap the two edges into (i, i') and (j, j'). In both cases, we complete

the move only after we make sure that the swapping does not create a new self-loop or a new multiple edge. This random edge-switching process continues until no self-loops or multiple edges remain. The procedure described above provides the random simple graph ensemble used in the simulations.

The elimination of multiple edges may introduce some degree-degree correlations in the resulting simple graph. To keep these degree-degree correlations negligible, the degree distribution must exhibit a structural cutoff such that the expected number of nodes of degree $k > \sqrt{Nc}$ is $o(1)$ [38, 39]. In the case of fat-tailed degree distributions such as the power-law degree distribution, one needs to impose an upper cutoff $k_{\max} < \sqrt{Nc}$. The degree distributions of all the network models considered here satisfy the above conditions, so degree-degree correlations are negligible.

Some commonly studied configuration model networks can be described in terms of single parameter families of degree distributions. These include the RRG, the ER network and configuration model networks with exponential and power-law degree distributions. A particularly convenient choice of the parameter is the mean degree $c = \langle K \rangle$. In this case, the degree distribution can be expressed by $P(k) = P_c(k)$, such that small values of c correspond to the dilute network limit while large values of c correspond to the dense network limit.

Configuration model networks in which the lower bound of the degree distribution satisfies $k_{\min} = 0$ or 1, may exhibit a percolation transition at some value c_0 of the mean degree, referred to as the percolation threshold. Below the transition the network consists of finite tree components, while above the transition a giant component emerges. The percolation transition is a second order phase transition, whose order parameter is the fraction g of nodes that reside on the giant component. Below the transition, where $c < c_0$, the order parameter is $g = 0$, while for $c > c_0$ the fraction $g = g(c)$ of nodes that reside on the giant component gradually increases. The giant component of a configuration model network consists of a 2-core which is decorated by tree branches [40]. The 2-core is a connected component, such that each node on the 2-core has links to at least two other nodes that reside on the 2-core. The nodes that reside on the tree branches have the property that their deletion would break the giant component into two or more components. Such nodes are referred to as articulation points [41, 42]. Similarly, the deletion of an edge that resides on one of the tree branches would break the giant component into two components. Such edges are referred to as bredges [43].

In this paper we focus on the case in which the whole network consists of a single connected component, for which $g = 1$. Below we discuss the conditions for $g = 1$ in RRGs, ER networks and configuration model networks with exponential and power-law distributions.

Consider an RRG that consists of N nodes of degree c (where Nc is even). For $c = 1$ the nodes form dimers. For $c = 2$ the network, which is referred to as a 2-random regular graph (2-RRG), consists of closed loops or cycles. In the large network limit, the expected number of cycles is $N_C \simeq \frac{1}{2} \ln N$ and the cumulative distribution of cycle lengths is given by $P(L \leq \ell) \simeq \ln \ell / \ln N$, where $\ell \leq N$ [44]. Here we focus

on RRGs with $c \geq 3$, which in the large network limit consist of a single connected component [45, 46].

In the case of ER networks, in the large network limit there is a phase transition at $c_1 = \ln N$, where $g \to 1$ [6, 10]. Above this point, the giant component encompasses the whole network, linking all the nodes into a single connected component.

In general, a sufficient condition for a configuration model network with degree distribution $P(k)$ to consist of a single connected component in the large network limit $N \to \infty$ is $k_{\min} \geq 3$ [47]. In fact, a weaker condition of $k_{\min} \geq 2$ is also sufficient, as long as a finite fraction of the nodes in the network are of degrees $k \geq 3$ [11]. In the analysis presented below of NBWs on configuration model networks with exponential and power-law distributions, we chose networks of size $N = 1000$ that satisfy $k_{\min} \geq 3$. We checked each network instance to confirm that it consists of a single connected component.

In a finite configuration model network, there is a non-zero probability that the network will consist of a single connected component even if it includes some nodes of degree $k = 1$. It was recently shown [48] that as the network size N is increased, it may still consist of a single connected component with high probability as long as the number n_1 of nodes of degree $k = 1$ grows more slowly than $\sqrt{N}$. However, in networks that include leaf nodes of degree $k = 1$, NBWs that enter these nodes will get stuck. Therefore, in the study of NBWs it is important not only to ensure that the network consists of a single connected component, but also that this component does not include any leaf nodes. This implies that the 2-core of the network (namely the largest subgraph in which all the nodes are of degree $k \geq 2$) encompasses the whole network. It also implies that the network does not include any articulation points [41,42] or bredges [43].

3. Non-backtracking random walks

NBWs are RWs for which the move backwards to the previous node is excluded. They belong to the class of second-order random walks, in which the transition probabilities depend not only on the current node but also on the previous node [27, 30]. This introduces memory into the process, which makes it no longer Markovian in the traditional sense. The challenge is to analyze such processes using methods that are typically applied to Markov chains, which rely on the memoryless property. Recently, Fasino et al. introduced a mapping of second order random walks into first order processes on a larger state space, referred to as the pullback process [30]. Instead of viewing the random walk as taking place between the nodes of the original graph, the pullback process considers a random walk on the directed-line graph associated with the original graph. Using this method they showed that the mean first return time $\mathbb{E}[T_{\rm FR}]$ of any second order random walk (including NBWs) on undirected networks satisfies Eq. (1), thus extending the validity of Kac's lemma to second order random walks on undirected networks [30]. Note that Kac's lemma deals with the mean first return time and has no implications on the overall shape of the distribution and its higher order

moments.

NBWs exhibit faster mixing times than standard random walks, meaning they converge to their stationary distribution more quickly [49, 50]. They thus inspire the design of more efficient algorithms for various graph-based problems, including link prediction and node centrality measures [51]. The non-backtracking (Hashimoto) matrix B associated with these walks has spectral properties that can reveal important information about the network structure, often more clearly than traditional adjacency matrices [15, 52, 53]. Actually, the mixing time is inversely proportional to the spectral gap of the matrix B [14]

$$t_{\rm mix} \propto \frac{1}{|\lambda_1 - \lambda_2|}, \tag{5}$$

where λ_1 and λ_2 are the largest and second largest eigenvalues of B, respectively.

For the special case of RRGs it was shown that the mixing time of NBWs (and RWs) scales like $t_{\rm mix} \propto \ln N$ [28, 54]. This result was later generalized to a broader class of configuration model networks with $k_{\rm min} \geq 3$ [55]. This result sits well with the fact that both the mean distance [56, 57] and the diameter [58–62] of RRGs are proportional to $\ln N$. It implies that an NBW starting from a random initial node i at time $t = 0$ may reach any other node in the network within $\ln N$ time steps. Moreover, using the shell structure around the initial node i as a spherical coordinate system, the radial component of the location of each node is given by its distance from i. Since RRGs are locally tree-like at distances in the range $\ell \ll \ln N$ [63], an NBW starting from i essentially moves deterministically to the next shell away from i as far as the tree-like structure persists. This is unlike the case of RWs which behave like biased random walks along the radial axis, moving outwards with probability $1 - 1/c$ and inwards with probability $1/c$ [9, 57].

4. The distribution of first return times

Consider an NBW on an undirected random network, starting from a random initial node i at time $t = 0$. At time $t = 1$ it hops into a random neighbor of i and at each subsequent step it hops randomly into one of the neighbors of its current node, excluding the previous node. Here we focus on the case of configuration model networks that consist of a single connected component, such that an NBW starting from any initial node can reach any other node in the network.

At each time step $t \geq 3$ an NBW may either step into a yet-unvisited node or into a node that has already been visited two or more time steps earlier. Similarly, at each time step $t \geq 4$ an NBW may go through an edge from node i to node i', that has been crossed before in the same direction, or through an edge that has not yet been crossed in that direction. We thus distinguish between the two possibilities of crossing an edge: from i to i' and from i' to i. In a network of size N and mean degree c, the expected number of such 'directed' edges is Nc. Below we consider the expected number of

distinct 'directed' edges $\langle L\rangle_t$ crossed by an NBW up to time t on a configuration model network. The initial condition is $\langle L\rangle_0 = 0$. The probability that at time step t an NBW will cross a yet uncrossed 'directed' edge is given by

$$\Delta L_t = \langle L\rangle_{t+1} - \langle L\rangle_t. \tag{6}$$

In the first three time steps the NBW crosses new 'directed' edges with probability 1, which implies that $\Delta L_t = 1$ for $t = 0, 1$ and 2. For $t \geq 3$ we use a mean-field approach, which essentially assumes that the 'directed' edges that have already been crossed and those that have not yet been crossed are distributed uniformly in the network and can thus be visited with equal probability at any time step. This approach applies under the condition that the network consists of a single connected component. A further condition is that the network will not be dominated by linear chains consisting of nodes of degree $k = 2$, namely that $P(k = 2)$ will be sufficiently small (networks that do not satisfy this condition are referred to as almost 2-RRGs [11]).

In configuration model networks that consist of a single connected component with $k_{\min} \geq 3$ and no leaf nodes, the mixing time scales like $t_{\text{mix}} \propto \ln N$ [54, 55], while the mean first return time scales like $\mathbb{E}[T_{\text{FR}}] \propto N$. Thus, for sufficiently large networks $t_{\text{mix}} \ll \mathbb{E}[T_{\text{FR}}]$. This separation of time scales implies that apart from the very early stages of the first return trajectories, NBWs sample the 'directed' edges in a unifom fashion. Under these conditions, the probability that at time $t+1$ the NBW will cross a 'directed' edge which has been crossed before is equal to the fraction of 'directed' edges that have already been crossed. This fraction is given by $(\langle L\rangle_t - 2)/(Nc - 2)$, where the subtraction of 2 from the numerator and the denominator accounts for the fact that the 'directed' edges crossed at times $t-1$ and t cannot be crossed again at time $t+1$. This implies that the probability ΔL_t is given by

$$\Delta L_t = 1 - \frac{\langle L\rangle_t - 2}{Nc - 2}. \tag{7}$$

To simplify the analysis, we reduce Eq. (7) to the form

$$\Delta L_t = 1 - \frac{\langle L\rangle_t}{Nc}. \tag{8}$$

The reduction from Eq. (7) to Eq. (8) relies on the assumption that the network is both sufficiently large and sufficiently dense, such that the product Nc satisfies $Nc \gg 2$. In addition, this reduction becomes accurate when the expected number of 'directed' edges $\langle L\rangle_t$ which have already been visited by the NBW satisfies $\langle L\rangle_t \gg 1$. This condition is indeed satisfied for sufficiently long times. Since for $t \ll Nc$ the number of distinct 'directed' edges visited by the NBW satisfies $\langle L\rangle_t \simeq t$, the condition $\langle L\rangle_t \gg 1$ can be replaced by $t \gg 1$.

Inserting ΔL_t from Eq. (6) into Eq. (8), we obtain the recursion equation

$$\langle L\rangle_{t+1} = \langle L\rangle_t \left(1 - \frac{1}{Nc}\right) + 1. \tag{9}$$

Solving Eq. (9), we obtain

$$\langle L\rangle_t = \begin{cases} t & 0 \le t \le 3 \\ 3e^{-\frac{t-3}{Nc}} + Nc\left(1 - e^{-\frac{t-3}{Nc}}\right) & t > 3. \end{cases} \tag{10}$$

While Eq. (8) is valid to a good approximation for $t \geq 3$, it becomes precise above the mixing time, where the random walker samples 'directed' edges in a uniform fashion. Thus, apart from the first few steps, Eq. (10) can be approximated by

$$\langle L\rangle_t = Nc\left(1 - e^{-\frac{t}{Nc}}\right). \tag{11}$$

The probability that an NBW will not visit a specific random node of degree k for the first time up to time t, can be expressed by $\left(1 - \frac{\langle L\rangle_{t-1}}{Nc}\right)^k$. This is due to the fact that in order to visit a node of degree k the NBW must enter via one of the k edges connected to i. Since an NBW quickly loses memory of its initial node, the probability of not returning to an initial node of degree k up to time t is the same as the probability not to visit any node of the same degree up to time t. Therefore, the tail distribution of first return times, under the condition that the initial node i is of degree k, is given by

$$P(T_{\rm FR} > t|K = k) = \left(1 - \frac{\langle L\rangle_{t-1}}{Nc}\right)^k. \tag{12}$$

Inserting $\langle L\rangle_{t-1}$ from Eq. (10) into Eq. (12), and using the fact that $k \ll Nc$, we obtain the tail distribution of first return times for initial nodes of degree k, which is given by

$$P(T_{\rm FR} > t|K = k) = e^{-\frac{k}{Nc}t}. \tag{13}$$

To obtain the tail distribution $P(T_{\rm FR} > t)$ of first return times of an NBW starting from a random node, we average over all possible initial nodes. This amounts to averaging over all possible degrees, with weights given by $P(k)$. We obtain

$$P(T_{\rm FR} > t) = \sum_{k=0}^{\infty} e^{-\frac{t}{Nc}k} P(k). \tag{14}$$

Interestingly, the right hand side of Eq. (14) is a discrete Laplace transform of the degree distribution $P(k)$. This transform is related to the one-sided Z-transform and to the starred transform [64]. To illustrate this point, we express Eq. (14) in the form

$$P(T_{\rm FR} > t) = \sum_{k=0}^{\infty} z^k P(k), \tag{15}$$

where

$$z = e^{-\frac{t}{Nc}}. \tag{16}$$

In fact, the right hand side of Eq. (15) is equal to the generating function $G_0(z)$ of the degree distribution $P(k)$. The generating function is known to play a central role

in the analysis of structural properties of random networks such as the percolation threshold [32] and the distribution of shortest path lengths [65]. Therefore, Eq. (15) provides a remarkable connection between structural properties of a network, captured by $G_0(z)$ and properties of dynamical processes taking place on the network.

From known properties of the (discrete) Laplace transform, we infer that the tail of $P(T_{\mathrm{FR}} > t)$ is determined by the abundances of the lowest degree nodes at the left end of $P(k)$. In contrast, the left end of $P(T_{\mathrm{FR}} > t)$ is determined by the highest degree nodes (or hubs) in the tail of $P(k)$.

The probability mass function of first return times is given by the difference

$$P(T_{\mathrm{FR}} = t) = P(T_{\mathrm{FR}} > t-1) - P(T_{\mathrm{FR}} > t). \tag{17}$$

The moments of the distribution of first return times can be obtained from the tail-sum formula [66]. In particular, the mean first return time is given by

$$\mathbb{E}[T_{\mathrm{FR}}] = \sum_{t=0}^{\infty} P(T_{\mathrm{FR}} > t), \tag{18}$$

and the second moment is given by

$$\mathbb{E}\left[T_{\mathrm{FR}}^2\right] = \sum_{t=0}^{\infty} (2t+1) P(T_{\mathrm{FR}} > t). \tag{19}$$

The variance is given by

$$\mathrm{Var}(T_{\mathrm{FR}}) = \mathbb{E}\left[T_{\mathrm{FR}}^2\right] - \mathbb{E}[T_{\mathrm{FR}}]^2. \tag{20}$$

To evaluate the mean first return time, we insert $P(T_{\mathrm{FR}} > t)$ from Eq. (14) into Eq. (18) and obtain

$$\mathbb{E}[T_{\mathrm{FR}}] = \sum_{t=0}^{\infty} \sum_{k=0}^{\infty} e^{-\frac{k}{Nc} t} P(k). \tag{21}$$

Exchanging the order of the summations and carrying out the sum over t, we obtain

$$\mathbb{E}[T_{\mathrm{FR}}] = \sum_{k=0}^{\infty} \frac{1}{1 - e^{-\frac{k}{Nc}}} P(k). \tag{22}$$

Expanding the exponent in the denominator in terms of $k/(Nc) \ll 1$ and taking the leading term, we obtain

$$\mathbb{E}[T_{\mathrm{FR}}] \simeq \sum_{k=0}^{\infty} \frac{Nc}{k} P(k) = \left\langle \frac{Nc}{K} \right\rangle. \tag{23}$$

This result coincides with Kac's lemma, which is obtained from general properties of discrete stochastic processes [13, 30]. Eq. (23) also implies that conditioning on initial nodes of a given degree k, the mean first return time is given by

$$\mathbb{E}[T_{\mathrm{FR}}|K=k] \simeq \frac{Nc}{k}. \tag{24}$$

Eq. (23) implies that the mean first return time is proportional to the mean inverse degree $\left\langle \frac{1}{K} \right\rangle$. In order to express this quantity in terms of $\langle K \rangle$ and $\mathrm{Var}(K)$, one can use a Taylor expansion of $1/K$ around $\langle K \rangle$ and obtain its expectation value

$$\left\langle \frac{1}{K} \right\rangle = \frac{1}{\langle K \rangle} + \frac{1}{\langle K \rangle^3}\mathrm{Var}(K) - \frac{1}{\langle K \rangle^4}\left\langle \left(K - \langle K \rangle\right)^3 \right\rangle + \dots. \tag{25}$$

This expansion is suitable for narrow distributions that are concentrated around their mean value. Moreover, in the case of symmetric distributions the third term on the right hand side of Eq. (25) vanishes and the first two terms are expected to provide accurate results for $\left\langle \frac{1}{K} \right\rangle$. Within the domain of validity of Eq. (25), we conclude that the mean first return time $\mathbb{E}[T_{\mathrm{FR}}]$ of an NBW on a configuration model network is proportional to the variance of the degree distribution of the network.

Using a similar derivation for the second moment, which is based on Eq. (19), we obtain

$$\mathbb{E}\left[T_{\mathrm{FR}}^2\right] = \sum_{k=0}^{\infty} \frac{1+e^{-\frac{k}{Nc}}}{\left(1-e^{-\frac{k}{Nc}}\right)^2} P(k) \simeq 2\sum_{k=0}^{\infty}\left(\frac{Nc}{k}\right)^2 P(k) = 2\left\langle \left(\frac{Nc}{K}\right)^2 \right\rangle. \tag{26}$$

This results goes beyond the generalization of Kac's lemma for second-order random walks [30] and is valid for the specific case of the NBW.

Conditioning on initial nodes of a given degree k, the second moment is given by

$$\mathbb{E}\left[T_{\mathrm{FR}}^2|K=k\right] \simeq 2\left(\frac{Nc}{k}\right)^2 \tag{27}$$

Inserting $\mathbb{E}\left[T_{\mathrm{FR}}^2\right]$ from Eq. (26) and $\mathbb{E}[T_{\mathrm{FR}}]$ from Eq. (23) into Eq. (20), we obtain

$$\mathrm{Var}(T_{\mathrm{FR}}) \simeq N^2c^2\left(2\left\langle \frac{1}{K^2} \right\rangle - \left\langle \frac{1}{K} \right\rangle^2\right). \tag{28}$$

Thus, the variance of the distribution of first return times, conditioned on initial nodes of a given degree k, is given by

$$\mathrm{Var}(T_{\mathrm{FR}}|K=k) = \frac{N^2c^2}{k^2}. \tag{29}$$

This indicates that the variance $\mathrm{Var}(T_{\mathrm{FR}})$ of the distribution of first return times can be divided into two parts, according to

$$\mathrm{Var}(T_{\mathrm{FR}}) = N^2c^2\ \mathrm{Var}\left(\frac{1}{K}\right) + N^2c^2\left\langle \frac{1}{K^2} \right\rangle, \tag{30}$$

where the first term on the right hand side of Eq. (30) can be attributed to the variations in the degrees between different initial nodes, while the second term can be attributed to the variation in the first return times between NBW trajectories originating from nodes of the same degree.

5. Application to specific random networks

In this section we apply the general results derived above to NBWs on specific random networks that belong to the class of configuration model networks [32]. In particular we study the first return process on RRGs, ER networks and configuration model networks with exponential and power-law degree distributions. For each type of network we calculate the tail distribution of first return times as well as the mean and variance.

5.1. Random regular graphs

Random regular graphs are random networks of a finite size in which all the nodes are of the same degree, but the connectivity is random [10]. They thus belong to the class of configuration model networks. Consider an RRG that consists of N nodes of degree $c \geq 3$. In such network, in the large N limit, all the nodes reside on a single connected component. As a result, an RW (or NBW) starting from any initial node i may reach any other node j.

The degree distribution of an RRG is a degenerate distribution of the form

$$P(k) = \delta_{k,c}, \tag{31}$$

where the mean degree $\langle K \rangle = c$ is an integer and the variance $\mathrm{Var}(K) = 0$.

Inserting $P(k)$ from Eq. (31) into Eq. (14), we obtain the tail distribution of first return times, which is given by

$$P(T_{\mathrm{FR}} > t) = e^{-\frac{t}{N}}. \tag{32}$$

It would be useful to compare the distribution of first return times of NBWs on RRGs to the corresponding distribution of simple RWs on RRGs. The latter distribution consists of a contribution from retroceding trajectories, which are dominant at short times and non-retroceding trajectories, which are dominant at long times. The distribution $P(T_{\mathrm{FR}} > t)$, given by Eq. (32), is analogous to the contribution of the non-retroceding RW trajectories in simple RWs, which for sufficiently long times is given by [9]

$$P(T_{\mathrm{FR}} > t | \neg \mathrm{RETRO}) = \exp\left[-\left(\frac{c-2}{c-1} \right) \frac{t}{N} \right]. \tag{33}$$

This implies that the backtracking and retroceding steps slow down the first return process of RWs by a factor of $\frac{c-2}{c-1}$ compared to NBWs.

In Fig. 1 we present analytical results, obtained from Eq. (32), for the tail distribution $P(T_{\rm FR} > t)$ (solid line) of first return times of an NBW on an RRG of size $N = 1000$. Note that the right hand side of Eq. (32) does not depend on the degree c, which implies that the results are valid for RRGs with any degree $c \geq 3$. Indeed, the analytical results are found to be in very good agreement with the results obtained from computer simulations for RRGs with $c = 3$ ($\times$) and $c = 10$ ($\circ$).

The mixing time of an NBW on an RRG of size N and degree c is given by [54]

$$t_{\rm mix}(N, c) = \frac{\ln N}{\ln(c-1)} + \mathcal{O}(1). \tag{34}$$

Applying this result to NBWs on RRGs of size $N = 1000$ with degrees of $c = 3$ and $c = 10$, it is found that $t_{\rm mix}(1000, 3) \simeq 10$ and $t_{\rm mix}(1000, 10) \simeq 3$, which are clearly much smaller than the time scales that are relevant to the first return process.

For the simulations we generated 20 independent instances of the network. On each network instance, we generated $100,000$ NBW trajectories, where each trajectory starts from a random initial node i at time $t = 0$. Each NBW trajectory was terminated upon its first return to the initial node i. The first return time t is thus equal to the length of the trajectory. The simulation results were obtained by averaging the results over all these trajectories.

Inserting Eq. (32) into Eq. (23), we obtain the mean first return time, which is given by

$$\mathbb{E}[T_{\rm FR}] \simeq N, \tag{35}$$

thus the mean first return time does not depend on the degree c. This result is in agreement with Kac's lemma, expressed by Eq. (1). Since all the nodes in an RRG are of the same degree, the probability that an RW (or an NBW) will reside at any given node at time t is $P_i(\infty) = 1/N$. Inserting $P_i(\infty)$ into Eq. (1), we obtain $\mathbb{E}[T_{\rm FR}] = N$. This result is also in agreement with the mean first return time of a simple RW on an RRG, calculated in Ref. [9].

Similarly, one can calculate the second moment, which is given by

$$\mathbb{E}\left[T_{\rm FR}^2\right] \simeq 2N^2. \tag{36}$$

Therefore, the variance is

$$\mathrm{Var}(T_{\rm FR}) \simeq N^2. \tag{37}$$

Since in an RRG all the nodes are of the same degree, this variance reflects the variability between first return trajectories originated from nodes of the same degree. Going back to Eq. (30), we conclude that Eq. (37) represents the lowest possible variance in the distribution of first return times for random networks consisting of N nodes.

Interestingly, for a simple RW on an RRG, it was found that the variance of the distribution of first return times is given by [9]

$$\mathrm{Var}(T_{\mathrm{FR}}) \simeq \frac{c}{c-2}N^2. \tag{38}$$

This result is larger than the variance for NBWs by a multiplicative factor of $\frac{c}{c-2}$. This factor is significant for sparse RRGs and approaches 1 as c is increased. It is due to the fact that in simple RWs the distribution of first return times is bimodal, consisting of two different types of first return trajectories. At short times it is dominated by retroceding trajectories while at long times it is dominated by non-retroceding trajectories. This separation of time scales broadens the distribution and increases the variance. The difference in the variance between NBWs and simple RWs reflects the fact that Kac's lemma applies only to the mean first return time and does not provide any prediction for the variance.

5.2. Erdős-Rényi networks

Consider an Erdős-Rényi network that consists of N nodes. In such network, each pair of nodes is connected by an edge with probability p [67–69]. As a result, the degree distribution is a Poisson distribution of the form [3]

$$P(k) = \frac{e^{-c}c^k}{k!}, \tag{39}$$

for $k = 0, 1, 2, \ldots$, where $c = (N-1)p$ is the mean degree $\langle K \rangle$ and the variance is given by $\mathrm{Var}(K) = c$.

In general, for $c > 1$ an ER network consists of a giant component and finite tree components. Since we focus in this paper on networks that consist of a single connected component, we restrict ourselves to the case in which $c > \ln N$, where in the large network limit the giant component encompasses the whole network [6, 10]. In the case that $c > \ln N$, the probability that a random node will be isolated is $P(K = 0) < 1/N$, which implies that in a typical network instance the expected number of isolated nodes will be smaller than 1. Since we study NBWs we would like to ensure that the network instances we consider will also not include leaf nodes of degree $k = 1$. Therefore, in the analysis we focus on the limit of sufficiently dense networks that satisfy $c > -W(-1/N)$, where $W(x)$ is the Lambert W function [70]. In this limit the probability that a random node will be a leaf node satisfies $P(K = 1) < 1/N$. In practice, when we generate network instances for the computer simulations, we discard network instances that include isolated nodes or leaf nodes.

For an NBW starting from a random node i on an ER network, the tail distribution of first return times is obtained by inserting Eq. (39) into Eq. (14), which yields

$$P(T_{\mathrm{FR}} > t) = \exp\left[c\left(e^{-\frac{t}{Nc}} - 1\right)\right]. \tag{40}$$

Note that in the long time limit of $t \to \infty$, $P(T_{\mathrm{FR}} > t) \to e^{-c}$, which is bounded by $1/N$ for $c > \ln N$ and hence vanishes in the large system limit. However, for finite networks the fact that $P(T_{\mathrm{FR}} > t)$ does not vanish in the limit of $t \to \infty$ and therefore

the moments diverge. In order to deal with this issue, we adjust the degree distribution by eliminating the possibility of isolated nodes of degree $k = 0$ and leaf nodes of degree $k = 1$. The adjusted degree distribution is given by

$$P(k|K > 1) = \frac{1}{1 - e^{-c} - ce^{-c}} \frac{e^{-c}c^k}{k!}, \tag{41}$$

for $k \geq 2$. Inserting the adjusted degree distribution from Eq. (41) into Eq. (14), we obtain

$$P(T_{\rm FR} > t|K > 1) = \frac{e^{-c}}{1 - e^{-c} - ce^{-c}} \left[\exp\left(ce^{-\frac{t}{Nc}}\right) - 1 - ce^{-\frac{t}{Nc}}\right]. \tag{42}$$

Taking the long time limit of Eq. (42), we obtain the leading order asymptotic behavior, which exhibits an exponential tail of the form

$$P(T_{\rm FR} > t|K > 1) \simeq \frac{c^2 e^{-c}}{2(1 - e^{-c} - ce^{-c})} e^{-\frac{2}{Nc}t}. \tag{43}$$

This tail is dominated by the lowest degree nodes in the network, whose degree is $k = 2$.

In Fig. 2 we present analytical results for the tail distribution $P(T_{\rm FR} > t|K > 1)$ (solid lines) of first return times of an NBW on an Erdős-Rényi network of size $N = 1000$ and mean degree $c = 10$. The analytical results, obtained from Eq. (42), are in very good agreement with the results obtained from computer simulations (circles).

Inserting $P(T_{\rm FR} > t|K > 1)$ from Eq. (42) into Eq. (23), we obtain

$$\mathbb{E}[T_{\rm FR}|K > 1] = \mathbb{E}\left[\frac{Nc}{K}\middle| K > 1\right]. \tag{44}$$

Evaluating the mean on the right hand side of Eq. (44), we obtain

$$\mathbb{E}[T_{\rm FR}|K > 1] = Nc \frac{e^{-c}}{1 - e^{-c} - ce^{-c}} \left[\mathrm{Ei}(c) - c - \ln c - \gamma\right], \tag{45}$$

where $\mathrm{Ei}(x)$ is the exponential integral [70]

$$Ei(x) = \int_{-\infty}^{x} \frac{e^t}{t} dt, \tag{46}$$

and γ is the Euler-Mascheroni constant [70]. In the limit of large mean degree c, Eq. (45) can be simplified to

$$\mathbb{E}[T_{\rm FR}|K > 1] = N\left[1 + \frac{1}{c} + \mathcal{O}\left(\frac{1}{c^2}\right)\right], \tag{47}$$

where $\mathcal{O}(1/c^2)$ means that the terms of order $1/c^2$ and higher are ignored in the expansion. This is in agreement with the first two terms on the right hand side of Eq. (25), confirming the validity of the expansion to the Poisson distribution, for sufficiently large values of the mean degree c. Eq. (47) shows that the mean first return time in an

ER network is larger than in an RRG of the same size, and is a decreasing function of c.

In Fig. 3 we present analytical results for the mean first return time $\mathbb{E}[T_{\rm FR}|K>1]$ (solid line) of an NBW on an Erdős-Rényi network of size $N = 1000$, as a function of the mean degree c. The analytical results, obtained from Eq. (45), are in very good agreement with the results obtained from computer simulations (circles).

Similarly, we can calculate the second moment, by plugging Eq. (41) into Eq. (26). We obtain

$$\mathbb{E}\left[T_{\rm FR}^2|K>1\right] \simeq 2N^2c^2\frac{e^{-c}}{1-e^{-c}-ce^{-c}}\sum_{k=2}^{\infty}\frac{c^k}{k!}\frac{1}{k^2}. \tag{48}$$

Carrying out the summation on the right hand size of Eq. (48), we obtain

$$\mathbb{E}\left[T_{\rm FR}^2|K>1\right] \simeq 2N^2c^3\frac{e^{-c}}{1-e^{-c}-ce^{-c}}\left[{}_3F_3\left(\begin{matrix}1,1,1\\2,2,2\end{matrix}\middle|c\right)-1\right]. \tag{49}$$

where ${}_3F_3\left(\begin{matrix}a_1,a_2,a_3\\b_1,b_2,b_3\end{matrix}\middle|z\right)$ is the generalized hypergeometric function [70]. Thus, the variance of $P(T_{\rm FR}>t|K>1)$ is give by

$$\begin{aligned}\mathrm{Var}(T_{\rm FR}|K>1) = {} & 2N^2c^3\frac{e^{-c}}{1-e^{-c}-ce^{-c}}\left[{}_3F_3\left(\begin{matrix}1,1,1\\2,2,2\end{matrix}\middle|c\right)-1\right]\\ & - N^2c^2\left(\frac{e^{-c}}{1-e^{-c}-ce^{-c}}\right)^2\left[\mathrm{Ei}(c)-c-\ln c-\gamma\right]^2. \end{aligned}\tag{50}$$

In the limit of large mean degree c, one can simplify Eq. (50), which takes the form

$$\mathrm{Var}(T_{\rm FR}|K>1) = N^2\left[1+\frac{4}{c}+\mathcal{O}\left(\frac{1}{c^2}\right)\right]. \tag{51}$$

In Fig. 4 we present analytical results for the variance $\mathrm{Var}(T_{\rm FR}|K>1)$ (solid line) of the distribution of first return times of an NBW on an Erdős-Rényi network of size $N = 1000$, as a function of the mean degree c. The analytical results, obtained from Eq. (50), are in very good agreement with the results obtained from computer simulations (circles).

5.3. Configuration model networks with an exponential degree distribution

Consider an ensemble of configuration model networks with an exponential degree distribution of the form

$$P(k) = Ae^{-\alpha k}, \tag{52}$$

where $\alpha > 0$ is the rate parameter and the degree k takes values in the range $k_{\rm min} \le k \le \infty$ [$P(k) = 0$ for $0 \le k \le k_{\rm min}-1$]. The parameter A is a normalization

factor and it is given by $A = (1 - e^{-\alpha})e^{\alpha k_{\min}}$. In order to obtain a network that consists of a single connected component, one needs to choose $k_{\min} \geq 2$.

Imposing the normalization condition and parameterising the distribution in terms of the mean degree $c = \langle K \rangle$, one can rewrite the degree distribution in the form [42]

$$P(k) = \frac{1}{c - k_{\min} + 1} \left(\frac{c - k_{\min}}{c - k_{\min} + 1} \right)^{k - k_{\min}}, \tag{53}$$

for $k \geq k_{\min}$. The parameter α from Eq. (52) can be expressed in the form

$$\alpha = \ln \left(\frac{c - k_{\min} + 1}{c - k_{\min}} \right). \tag{54}$$

The variance of the exponential degree distribution can be expressed in the form

$$\mathrm{Var}(K) = (c - k_{\min} + 1)(c - k_{\min}), \tag{55}$$

such that in the limit of a broad degree distribution, where $c \gg k_{\min}$ it can be approximated by

$$\mathrm{Var}(K) \simeq c^2. \tag{56}$$

Inserting $P(k)$ from Eq. (53) into Eq. (14) and carrying out the summation, we obtain the distribution of first return times

$$P(T_{\mathrm{FR}} > t) = \frac{1}{c + 1 - k_{\min}} \left(\frac{e^{-\frac{k_{\min}}{Nc} t}}{1 - \frac{c - k_{\min}}{c + 1 - k_{\min}} e^{-\frac{1}{Nc} t}} \right). \tag{57}$$

To explore the asymptotic long time tail of $P(T_{\mathrm{FR}} > t)$ we expand the right hand side of Eq. (57) in powers of $\exp\left(-\frac{t}{Nc}\right) \ll 1$, we obtain

$$P(T_{\mathrm{FR}} > t) \simeq \frac{1}{c + 1 - k_{\min}} e^{-\frac{k_{\min}}{Nc} t}. \tag{58}$$

As can be seen, this tail is dominated by the lowest degree nodes, whose degree is $k_{\min}$.

In Fig. 5 we present analytical results for the tail distribution $P(T_{\mathrm{FR}} > t)$ (solid lines) of first return times of an NBW on a configuration model network of size $N = 1000$ that exhibits an exponential degree distribution with $k_{\min} = 3$ and mean degree $c = 10$. The analytical results, obtained from Eq. (57), are in very good agreement with the results obtained from computer simulations (circles).

To calculate the mean of the distribution of first return times, we use Eq. (23), and obtain

$$\mathbb{E}[T_{\mathrm{FR}}] = N \frac{c}{c + 1 - k_{\min}} \Phi \left(\frac{c - k_{\min}}{c + 1 - k_{\min}}, 1, k_{\min} \right), \tag{59}$$

where $\Phi(z, s, \alpha)$ is the Lerch transcendent [70]. In the limit of large mean degree c, we obtain

$$\mathbb{E}[T_{\rm FR}] = N\left[\ln c - H_{k_{\rm min}-1} + \mathcal{O}\left(\frac{\ln c}{c}\right)\right], \tag{60}$$

where H_m is the Harmonic number [70]. Note that the leading term in $\mathbb{E}[T_{\rm FR}]$ is given by $N \ln c$, unlike the the RRG and ER network, in which $\mathbb{E}[T_{\rm FR}] \simeq N$. This reflects the fact that the variance of the degrees in the exponential case is much larger than in the Poisson distribution and that it increases as c is increased. Since the exponential distribution is broad and highly asymmetric, the expansion presented in Eq. (25) cannot be used to reproduce the results of Eq. (60). Eq. (60) shows that the mean first return time in configuration model networks with an exponential degree distribution increases logarithmically with the mean degree c.

In Fig. 6 we present analytical results for the mean first return time $\mathbb{E}[T_{\rm FR}]$ (solid line) of an NBW on a configuration model network of size $N = 1000$ which exhibits an exponential degree distribution with $k_{\rm min} = 3$, as a function of the mean degree c. The analytical results, obtained from Eq. (59), are in very good agreement with the results obtained from computer simulations (circles).

Similarly, we calculate the second moment, using Eq. (26). We obtain

$$\mathbb{E}\left[T_{\rm FR}^2\right] = 2N^2 \frac{c^2}{c+1-k_{\rm min}} \Phi\left(\frac{c-k_{\rm min}}{c+1-k_{\rm min}}, 2, k_{\rm min}\right). \tag{61}$$

In the limit of large c, we obtain

$$\mathbb{E}\left[T_{\rm FR}^2\right] = 2N^2\left[(c+2k_{\rm min}-1)\zeta(2,k_{\rm min}) - \ln c + H_{k_{\rm min}-1} - 1 + \mathcal{O}\left(\frac{\ln c}{c}\right)\right], \tag{62}$$

where $\zeta(s,a)$ is the Hurwitz zeta function [70].

The variance is given by

$$\mathrm{Var}(T_{\rm FR}) = (Nc)^2\left\{\frac{2\Phi\left(\frac{c-k_{\rm min}}{c+1-k_{\rm min}}, 2, k_{\rm min}\right)}{c+1-k_{\rm min}} - \left[\frac{\Phi\left(\frac{c-k_{\rm min}}{c+1-k_{\rm min}}, 1, k_{\rm min}\right)}{c+1-k_{\rm min}}\right]^2\right\}. \tag{63}$$

In the limit of large c, one can express the variance in a simpler form, namely

$$\begin{aligned}\mathrm{Var}(T_{\rm FR}) \simeq N^2\Big[&2c\,\zeta(2,k_{\rm min}) - (\ln c)^2 + 2\left(H_{k_{\rm min}-1}-1\right)\ln c \\ &+ 2\left(2k_{\rm min}-1\right)\zeta(2,k_{\rm min}) + 2H_{k_{\rm min}-1} \\ &- H_{k_{\rm min}-1}^2 - 2 + \mathcal{O}\left(\frac{\ln c}{c}\right)\Big].\end{aligned} \tag{64}$$

Note that the leading term is proportional to the mean degree c.

In Fig. 7 we present analytical results for the variance $\mathrm{Var}(T_{\rm FR})$ (solid line) of the distribution of first return times of an NBW on a configuration model network of

size $N = 1000$, which exhibits an exponential degree distribution with $k_{\min} = 3$, as a function of the mean degree c. The analytical results, obtained from Eq. (63), are in very good agreement with the results obtained from computer simulations (circles).

5.4. Configuration model networks with a power-law degree distribution

Consider a configuration model network with a power-law degree distribution of the form

$$P(k) = Ak^{-\gamma}, \tag{65}$$

where the degree k takes values in the range $k_{\min} \leq k \leq k_{\max}$. The parameter $A = [\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)]^{-1}$ is a normalization constant. Here we focus on the case that $k_{\min} \geq 2$, in which the network consists of a single connected component.

Since a power-law distribution may allow nodes of high degree, it is important to note that in order to enable the construction of a configuration model network in which degree-degree correlations are negligible, on must impose an upper cutoff on the degree distribution, which satisfies $k_{\max} < \sqrt{Nc}$ [38, 39].

The mean degree is given by [42]

$$c = \langle K \rangle = \frac{\zeta(\gamma - 1, k_{\min}) - \zeta(\gamma - 1, k_{\max} + 1)}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}, \tag{66}$$

and the second moment of the degree distribution is

$$\langle K^2 \rangle = \frac{\zeta(\gamma - 2, k_{\min}) - \zeta(\gamma - 2, k_{\max} + 1)}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}. \tag{67}$$

The variance of the degree distribution is given by $\mathrm{Var}(K) = \langle K^2 \rangle - \langle K \rangle^2$.

For $\gamma \leq 2$ the mean degree (and the variance) diverge when $k_{\max} \to \infty$. For $\gamma > 3$ both the mean degree and the variance are bounded. In the intermediate range of $2 < \gamma < 3$ the mean degree $\langle K \rangle$ is bounded while the variance $\mathrm{Var}(K)$ diverges. In this regime, as $k_{\max}$ is increased, the variance diverges like

$$\mathrm{Var}(K) \simeq \frac{1}{(3 - \gamma)[\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)]} (k_{\max})^{3-\gamma}. \tag{68}$$

Inserting $P(k)$ from Eq. (65) into Eq. (14) and carrying out the summation, we obtain

$$P(T_{\mathrm{FR}} > t) = \frac{e^{-\frac{t}{Nc} k_{\min}} \Phi\left(e^{-\frac{t}{Nc}}, \gamma, k_{\min}\right) - e^{-\frac{t}{Nc}(k_{\max}+1)} \Phi\left(e^{-\frac{t}{Nc}}, \gamma, k_{\max} + 1\right)}{\zeta(\gamma, k_{\min}) - \zeta(\gamma, k_{\max} + 1)}. \tag{69}$$

To explore the asymptotic long time tail of $P(T_{\mathrm{FR}} > t)$ we expand the right hand side of Eq. (69) in powers of $\exp\left(-\frac{t}{Nc}\right) \ll 1$, we obtain

$$P(T_{\rm FR} > t) \simeq \frac{(k_{\rm min})^{-\gamma}}{\zeta(\gamma, k_{\rm min}) - \zeta(\gamma, k_{\rm max}+1)} e^{-\frac{k_{\rm min}}{Nc}t}. \quad (70)$$

As can be seen, this tail is dominated by the lowest degree nodes, whose degree is $k_{\rm min}$.

In Fig. 8 we present analytical results for the tail distribution $P(T_{\rm FR} > t)$ (solid lines) of first return times of an NBW on a configuration model network of size $N = 1000$ that exhibits a power-law degree distribution with $k_{\rm min} = 3$, $k_{\rm max} = 30$ and $\gamma = 2.5$, which yields a mean degree of $\langle K \rangle \simeq 5.58$. The analytical results, obtained from Eq. (69), are in very good agreement with the results obtained from computer simulations (circles).

Inserting $P(k)$ from Eq. (65) into Eq. (23) and carrying out the summation, we obtain the mean first return time, which is given by

$$\mathbb{E}[T_{\rm FR}] = N\langle K \rangle \frac{\zeta(\gamma+1, k_{\rm min}) - \zeta(\gamma+1, k_{\rm max}+1)}{\zeta(\gamma, k_{\rm min}) - \zeta(\gamma, k_{\rm max}+1)}. \quad (71)$$

Since the power-law distribution is broad and highly asymmetric, the expansion presented in Eq. (25) for $\langle \frac{1}{K} \rangle$ cannot be used to reproduce the results of Eq. (71).

In the limit of $k_{\rm max} \to \infty$, inserting $\langle K \rangle$ from Eq. (66), Eq. (71) is reduced to

$$\mathbb{E}[T_{\rm FR}] = N \frac{\zeta(\gamma-1, k_{\rm min})\zeta(\gamma+1, k_{\rm min})}{[\zeta(\gamma, k_{\rm min})]^2}. \quad (72)$$

For $\gamma \neq 1$ and $k > 0$, the Hurwitz zeta function can be expressed in the form

$$\zeta(\gamma, k) = \frac{k^{-\gamma}}{2} + \frac{k^{1-\gamma}}{\gamma - 1} + \frac{1}{\Gamma(\gamma)} \int_0^\infty \left(\frac{1}{e^x - 1} - \frac{1}{x} + \frac{1}{2} \right) x^{\gamma-1} e^{-kx} dx, \quad (73)$$

where $\Gamma(y)$ is the Gamma function [70]. In the context of this paper the Hurwitz zeta function $\zeta(\gamma, k)$ is evaluated in the range of $\gamma > 1$ and $k \geq 3$. Exploring the terms on the right hand size of Eq. (73) in this range of values, it was found that the contribution of the integral is negligible. Thus, Eq. (73) can be simplified to

$$\zeta(\gamma, k) \simeq \frac{1}{2} k^{-\gamma} \left(1 + \frac{2k}{\gamma - 1} \right). \quad (74)$$

Inserting $\zeta(\gamma, k)$ from Eq. (74) into Eq. (72), we obtain

$$\mathbb{E}[T_{\rm FR}] \simeq N \frac{\left(1 + \frac{2}{\gamma-2} k_{\rm min}\right)\left(1 + \frac{2}{\gamma} k_{\rm min}\right)}{\left(1 + \frac{2}{\gamma-1} k_{\rm min}\right)^2}. \quad (75)$$

For sufficiently large values of $k_{\rm min}$, Eq. (72) can be approximated by

$$\mathbb{E}[T_{\rm FR}] \simeq N \left[\frac{(\gamma-1)^2}{\gamma(\gamma-2)} + \mathcal{O}\left(\frac{1}{k_{\rm min}} \right) \right]. \quad (76)$$

In practice, for $k_{\rm min} = 3$ there is a slight deviation between the right hand sides of Eqs. (72) and (76), which becomes negligible for $k_{\rm min} \geq 5$. It is found that $\mathbb{E}[T_{\rm FR}]$ is a

monotonically decreasing function of the exponent γ. In the limit of $\gamma \gg 1$ it converges towards $\mathbb{E}[T_{\rm FR}] \simeq N$, where it coincides with the result for RRGs. In the opposite limit, when $\gamma \to 2^+$ the mean first return time diverges (given that $k_{\max} \to \infty$).

Going back to Eq. (71), taking the limit of $k_{\max} \gg k_{\min}$ and using Eq. (74) to approximate the ratio $\zeta(\gamma+1,k_{\min})/\zeta(\gamma,k_{\min})$ while leaving $\langle K\rangle$ unchanged, we obtain

$$\mathbb{E}[T_{\rm FR}] \simeq N\frac{\gamma-1}{\gamma}\frac{\langle K\rangle}{k_{\min}}. \tag{77}$$

While the results obtained from Eq. (77) are not as accurate as those obtained from Eq. (76), it provides useful insight on the relation between the mean first return time and the mean degree $\langle K\rangle$. Comparing Eq. (77) to Eq. (24) shows that the mean first return time is dominated by the lowest degree nodes.

The second moment, obtained from Eq. (26), is given by

$$\mathbb{E}\left[T_{\rm FR}^2\right] = 2(N\langle K\rangle)^2\frac{\zeta(\gamma+2,k_{\min})-\zeta(\gamma+2,k_{\max}+1)}{\zeta(\gamma,k_{\min})-\zeta(\gamma,k_{\max}+1)}, \tag{78}$$

and the variance is given by

$$\begin{aligned}\mathrm{Var}(T_{\rm FR}) = {} & 2(N\langle K\rangle)^2\frac{\zeta(\gamma+2,k_{\min})-\zeta(\gamma+2,k_{\max}+1)}{\zeta(\gamma,k_{\min})-\zeta(\gamma,k_{\max}+1)} \\ & - (N\langle K\rangle)^2\left[\frac{\zeta(\gamma+1,k_{\min})-\zeta(\gamma+1,k_{\max}+1)}{\zeta(\gamma,k_{\min})-\zeta(\gamma,k_{\max}+1)}\right]^2.\end{aligned} \tag{79}$$

In the limit of $k_{\max} \to \infty$, and for values of γ which are sufficiently far above $\gamma = 2$, the variance of the distribution of first return times can be roughly approximated by

$$\begin{aligned}\mathrm{Var}(T_{\rm FR}) \simeq {} & 2N^2\frac{\left(1+\frac{2}{\gamma-1}k_{\min}\right)^2\left(1+\frac{2}{\gamma+2}k_{\min}\right)}{\left(1+\frac{2}{\gamma}k_{\min}\right)^3} \\ & - N^2\frac{\left(1+\frac{2}{\gamma-1}k_{\min}\right)^2\left(1+\frac{2}{\gamma+1}k_{\min}\right)^2}{\left(1+\frac{2}{\gamma}k_{\min}\right)^4}.\end{aligned} \tag{80}$$

For sufficiently large values of $k_{\min}$, Eq. (80) can be approximated by

$$\mathrm{Var}(T_{\rm FR}) \simeq N^2\frac{\gamma^3(\gamma^2+2\gamma+2)}{(\gamma-1)^2(\gamma+1)^2(\gamma+2)}. \tag{81}$$

In the limit of large γ, the variance converges towards N^2, in agreement with the result for RRGs. Interestingly, in the opposite limit of $\gamma \to 2^+$ the variance remains finite, unlike the mean first return time that tends to diverge.

6. Discussion

A key observation, expressed by Eq. (13) is that the distribution of first return times for an initial node of a given degree k depends only on the degree k and on the total number of 'directed' edges in the network, given by Nc. It does not depend on the degree distribution $P(k)$, which accounts for the way in which the $Nc-2k$ remaining 'directed' edges are divided among the $N-1$ remaining nodes. This implies that the distribution of first return times is determined by local properties of the network and is not sensitive to the global structure.

In Fig. 9 we present analytical results for the distribution of first return times of NBWs starting from a random initial node of degree $k=8$ in a configuration model network of size $N=1000$ and mean degree $c=8$, given by Eq. (13). The analytical results are found to be in very good agreement with simulation results for RRGs ($\circ$), ER networks ($\times$) and configuration model networks with exponential ($+$) and power-law ($\square$) distributions. As can be seen, the distribution $P(T_{\rm FR}>t|K=8)$ does not depend on the degree distribution $P(k)$ but only on the mean degree c and on the degree of the initial node.

Comparing the results obtained for the four random network models considered above, we conclude that the mean first return time strongly depends on the variability of the degrees of nodes in the network. More specifically, as the degree distribution $P(k)$ becomes broader the mean first return time $\mathbb{E}[T_{\rm FR}]$ increases. This is illustrated by the fact that for an NBW on an RRG $\mathbb{E}[T_{\rm FR}]\simeq N$, for an NBW on an ER network $\mathbb{E}[T_{\rm FR}|K>1]\simeq N\left(1+\frac{1}{c}\right)$, for an NBW on a configuration model network with an exponential degree distribution $\mathbb{E}[T_{\rm FR}]\simeq N\ln c$ and for an NBW on a configuration model network with a power-law degree distribution $\mathbb{E}[T_{\rm FR}]\sim Nc/k_{\rm min}$. In light of these results, it is interesting to note that the dependence of the mean first return time on the mean degree c is a non-trivial issue, which depends on the details of the degree distribution. In the examples studied here we observe three different behaviors: in RRGs the mean first return time is independent of c, in ER networks it decreases with c and in configuration model networks with an exponential degree distribution it increases with c.

In all the network ensembles considered above, the long-time tail of $P(T_{\rm FR}>t)$ exhibits a decaying exponential form, which is determined by the lowest-degree nodes in the network. From a broader perspective, it implies that the distribution of first return times is mostly characterised by low-degree nodes that reside in the periphery of the network. This is unlike the outburst dynamics of other processes such as the spreading of information and infections, which are dominated by the highest degree nodes (or hubs) that reside in the core of the network [5].

Apart from the first return process, there are other significant events that take place over the lifetime of an NBW (and other RWs) on a random network. One of them is the first hitting (FH) process, which is the first time at which an NBW steps into a previously visited node. Starting from a random initial node i, in the early stages of its

trajectory, an NBW visits a new node at each time step. During this time, the statistical properties of the NBW trajectory are identical to those of a self avoiding walk [71]. After the first hitting event, in some of the time steps the NBW visits yet-unvisited nodes and in other time steps it revisits nodes that it has already visited before. The distribution of first hitting times of RWs and NBWs on ER networks were studied in Refs. [72] and [73], respectively. It was found that in both cases, for sufficiently dense ER networks in which there are no leaf nodes of degree $k = 1$, the distribution $P(T_{\rm FH} > t)$ of first hitting times is given by a product of an exponential distribution and a Rayleigh distribution, which is a special case of the Weibull distribution. In this limit, the mean first hitting time of NBWs on ER networks is given by

$$\mathbb{E}[T_{\rm FH}] = \sqrt{\frac{\pi}{2}}\sqrt{N}. \tag{82}$$

Similar results were also obtained for first hitting processes on RRGs [74].

The results presented in this paper shed light on the more general class of first passage processes. Consider an NBW starting from a random initial node i, seeking a target node j, where $j \neq i$. Unlike the first return event of an NBW which may take place only at $t \geq 3$, a first passage event may take place even at $t = 1$ (in case that i and j are connected by an edge). We thus conclude that to a very good approximation, the distribution of first passage times of NBWs on configuration model networks can be expressed in the form

$$P(T_{\rm FP} > t) \simeq P(T_{\rm FR} > t + 2). \tag{83}$$

Another important event, which occurs at much longer time scales, is the step at which an NBW (or RW) completes visiting all the nodes in the network. The time at which this happens is called the cover-time. For RWs on RRGs it was shown that the mean cover time scales like

$$\mathbb{E}[T_{\rm C}] \propto N \ln N. \tag{84}$$

This means that on average an RW visits each node $\ln N$ times before it completes visiting all the nodes in the network at least once. The distribution of cover times of RWs on RRGs was studied in Refs. [4, 75, 76]. Since they do not backtrack their steps, NBWs scan the network more efficiently than RWs. This is expected to affect the pre-factor of the scaling relation on the right hand side of Eq. (84) but is not expected to change the way the cover time scales with N.

The results presented in this paper were derived in the context of configuration model networks. However, we expect them to apply within a good approximation to a somewhat broader class of small-world networks which are sufficiently strongly connected without bottlenecks and exhibit short mixing times of the NBW, determined by the spectral gap of the non-backtracking (Hashimoto) matrix. In contrast, these results are not expected to apply in the case of modular networks, which consist of

several modules with weak connections between them and for networks that exhibit long mixing times. However, in the case of modular networks, at short times, the distribution of first return times is likely to behave as if the module on which the initial node resides is isolated from other modules. This behavior persists until the probability that the NBW will hop into some other module becomes significant.

In essence, the derivation presented above requires that the mixing time will be much shorter than the mean first return time, such that the assumption that the NBW samples uniformly the 'directed' edges is justified. In light of this it is interesting to discuss the effect of degree-degree correlations. In general, negative or disassortative correlations tend to enhance the connectivity of the network [42] and hence shortens the mixing times [77] by increasing the spectral gap. On the other hand, positive or assortative correlations are known to decrease the spectral gap [78], thus increasing the mixing time. This is particularly relevant in networks that have many high degree nodes, such as scale-free networks where high assortativity may break the network into disconnected components [42]. However, for low correlations the overall impact of degree-degree correlations on the spectrum is not large, especially on short range correlations between eigenvalues that follow the predictions of random matrix theory [79]. In summary, in the case of disassortative networks we expect our results to hold. Regarding networks that exhibit low to mild positive assortativity and to the extent that they do not break the network into disconnected components, we expect the results to hold to a good approximation.

Another key factor influencing the mixing times is the clustering coefficient, primarily mediated through its effect on the spectral gap. The clustering coefficient measures how often a node's neighbors form triangles indicating the degree of local inter-connectedness. Networks with higher clustering coefficients typically have smaller spectral gaps. This occurs because increased clustering introduces more local structure. A smaller spectral gap leads to longer mixing times, since random walks become trapped in tightly connected neighborhood before fully exploring the network [80]. Consequently, higher clustering increases the mixing times. We thus expect our results to be valid as long as the clustering is not too strong.

While directed networks are of significant theoretical interest, they introduce complexities that fall outside the scope of this paper. In directed networks the asymmetry of edges creates distinct behavior as the random walk dynamics are heavily influenced by both the in-degrees and out-degrees. This asymmetry complicates the analysis of return times and of the mixing behavior, often leading to nodes with low in-degrees being visited only rarely or potentially not at all. Additionally, in weakly connected networks random walkers may become trapped in certain domains rendering the analysis of return times more intricate. As the focus of this paper is on undirected networks, we leave these issues to future work.

7. Summary

We presented analytical results for the distribution of first return times of NBWs on configuration model networks consisting of N nodes with degree distribution $P(k)$, focusing on the case in which the network consists of a single connected component. It was found that the tail distribution $P(T_{\mathrm{FR}} > t)$ of first return times is given by a discrete Laplace transform of the degree distribution $P(k)$. This result demonstrates the relation between structural properties of a network, captured by the degree distribution, and the properties of dynamical processes taking place on the network. It was found that $P(T_{\mathrm{FR}} > t)$ exhibits an exponential tail, which is determined by the properties of the low-degree nodes that reside in the periphery of the network. We calculated the mean first return time and found that $\mathbb{E}[T_{\mathrm{FR}}] = \langle \frac{Nc}{K} \rangle$. Surprisingly, this result coincides with the result of Kac's lemma that applies to simple RWs, in agreement with recent rigorous results by Fasino et al. [30]. We also calculated the variance $\mathrm{Var}(T_{\mathrm{FR}})$, which accounts for the variability of the first return times between different NBW trajectories. We applied this formalism to random regular graphs, Erdős-Rényi networks and configuration model networks with exponential and power-law degree distributions and obtained closed-form expressions for $P(T_{\mathrm{FR}} > t)$ and its first two moments. These results provide useful insight on the advantages of NBWs over simple RWs in network exploration, sampling and search processes. Our results are expected to hold for a broader class of networks, in which the mixing time is much shorter than the mean first return times.

This work was supported by Grant no. 2020720 from the United States-Israel Binational Science Foundation (BSF) and grant no. 2102832 from the National Science Foundation (NSF).

References

[1] Lawler G F and Limic V 2010 *Random Walk: A Modern Introduction* (Cambridge: Cambridge University Press)
[2] Havlin S and Cohen R 2010 *Complex Networks: Structure, Robustness and Function* (New York: Cambridge University Press)
[3] Newman M E J 2018 *Networks: an Introduction, Second Edition* (Oxford: Oxford University Press)
[4] Masuda N, Porter M A and Lambiotte R 2017 Random walks and diffusion on networks *Physics Reports* **716** 1
[5] Barrat A, Barthélemy M and Vespignani A 2012 *Dynamical Processes on Complex Networks* (Boston: Cambridge University Press)
[6] van der Hofstad 2017 *Random Graphs and Complex Networks, Volume One* (Cambridge: Cambridge University Press)
[7] van der Hofstad 2024 *Random Graphs and Complex Networks, Volume Two* (Cambridge: Cambridge University Press)
[8] De Bacco C, Majumdar S N and Sollich P 2015 The average number of distinct sites visited by a random walker on random graphs *J. Phys. A* **48** 205004
[9] Tishby I, Biham O and Katzav E 2021 Analytical results for the distribution of first return times of random walks on random regular graphs *J. Phys. A* **54** 325001
[10] Bollobas B 2001 *Random Graphs, Second Edition* (London: Academic Press)

[11] Federico L 2023 Almost-2-regular random graphs *Australasian Journal of Combinatorics* **86** 76

[12] Redner S 2001 *A Guide to First Passage Processes* (Cambridge: Cambridge University Press)

[13] Kac M 1947 On the notion of recurrence in discrete stochastic processes, *Bull. Amer. Math. Soc.* **53** 1002

[14] Levin D A, Peres Y and Wilmer E L 2017 *Markov chains and mixing times, Second Edition* (Providence: American Mathematical Society)

[15] Dorogovtsev S N and Mendes J F F 2022 *The Nature of Complex Networks* (Oxford: Oxford University Press)

[16] Tishby I, Biham O and Katzav E 2022 Analytical results for the distribution of first passage times of random walks on random regular graphs *J. Stat. Mech.* P113403

[17] Masuda N and Konno N 2004 Return times of random walk on generalized random graphs *Phys. Rev. E* **69** 066113

[18] Evnin O and Horinouchi W 2025 First return times on sparse random graphs *J. Phys. A* **58** 075002

[19] Hormann E, Lambiotte R and Cantwell G 2025 An approximation for return time distributions of random walks on sparse networks *Phys. Rev. E* **111** 064306

[20] Albert J 2024 Exact analytic expressions for discrete first-passage time probability distributions in Markov networks arXiv:2403.14149

[21] Sood V, Redner S and ben-Avraham D 2005 First-passage properties of the Erdős–Rényi random graph *J. Phys. A* **38** 109

[22] Peng J, Sandev T and Kocarev L 2021 First encounters on Bethe lattices and Cayley trees *Communications in Nonlinear Science and Numerical Simulation* **95** 105594

[23] Hughes B D and Sahimi M 1982 Random walks on the Bethe lattice *J. Stat. Phys.* **29** 781

[24] Cassi D 1989 Random walks on Bethe lattices *Europhys. Lett.* **9** 627

[25] Giacometti A 1995 Exact closed form of the return probability on the Bethe lattice *J. Phys. A* **28** L13

[26] Martin O and Šulc P 2010 Return probabilities and hitting times of random walks on sparse Erdős-Rényi graphs *Phys. Rev. E* **81** 031111

[27] Fitzner R and van der Hofstad R 2013 Non-backtracking random walks *J. Stat. Phys.* **150** 264

[28] Alon N, Benjamini I, Lubetzky E and Sodin S 2007 Non-backtracking random walks mix faster *Commun. Contemp. Math.* **9** 585

[29] Arrigo F, Grindrod P, Higham D J and Noferini V 2018 Non-backtracking walk centrality for directed networks *J. Complex Networks* **6** 54

[30] Fasino D, Tonetto A and Tudisco F 2023 Hitting times for second-order random walks *Euro. Jnl of Applied Mathematics* **34** 642

[31] Bollobas B 1980 A probabilistic proof of an asymptotic formula for the number of labelled regular graphs *European J. Combin.* **1** 311

[32] Newman M E J, Strogatz S H and Watts D J 2001 Random graphs with arbitrary degree distributions and their applications *Phys. Rev. E* **64** 026118

[33] Molloy M and Reed A 1995 A critical point for random graphs with a given degree sequence *Rand. Struct. Alg.* **6** 161

[34] Molloy M and Reed A 1998 The size of the giant component of a random graph with a given degree sequence *Combinatorics, Probability and Computing* **7** 295

[35] Fosdick B K, Larremore D B, Nishimura J and Ugander J 2018 Configuring random graph models with fixed degree sequences it SIAM Rev. **60** 315

[36] Erdős P and Gallai T 1960 Gráfok előírt fokszámú pontokkal *Matematikai Lapok* **11** 264

[37] Choudum S A 1986 A simple proof of the Erdős-Gallai theorem on graph sequences *Bulletin of the Australian Mathematical Society* **33** 67

[38] Boguñá M, Pastor-Satorras R and Vespignani A (2004) Cut-offs and finite size effects in scale-free networks *Eur. Phys. J. B* **38** 205

[39] Catanzaro M, Boguñá M and Pastor-Satorras R (2005) Generation of uncorrelated random scale-free networks *Phys. Rev. E* **71** 027103

[40] Tishby I, Biham O, Katzav E and Kühn R 2018 Revealing the micro-structure of the giant cluster in random graph ensembles *Phys. Rev. E* **97** 042318
[41] Tian L, Bashan A, Shi D N and Liu Y Y 2017 Articulation points in complex networks, *Nature Communications* **8** 14223
[42] Tishby I, Biham O, Kühn R and Katzav E 2018 Statistical analysis of articulation points in configuration model networks *Physical Review E* **98** 062301
[43] Bonneau H, Biham O, Kühn R and Katzav E 2020 Statistical analysis of edges and bredges in configuration model networks *Phys. Rev. E* **102** 012314
[44] Tishby I, Biham O, Katzav E and Kühn R 2023 Distribution of the number of cycles in directed and undirected random regular graphs of degree 2 *Phys. Rev. E* **107** 024308
[45] Bender E and Canfield E R 1978 The asymptotic number of labeled graphs with given degree sequences *Journal of Combinatorial Theory A* **24** 296301
[46] Wormald N C 1981 The asymptotic connectivity of labelled regular graphs *Journal of Combinatorial Theory B* **31** 156
[47] Wormald N C 1999 Models of random regular graphs *Surveys in Combinatorics, LMS Lecture Note Series* **267** 239
[48] Federico L and van der Hofstad R 2017 Critical window for connectivity in the configuration model *Combinatorics, Probability and Computing* **26** 660
[49] Cioba S M and Xu P 2015 Mixing rates of random walks with little backtracking *Contemp. Math.* **655** 27
[50] Ben-Hamou A, Lubetzky E and Peres Y 2019 Comparing mixing times on sparse random graphs *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* **55** 1116
[51] Wu Y, Bian Y and Zhang X 2016 Remember where you came from: on the second-order random walk based proximity measures *Proc. VLDB Endow.* **10** 13
[52] Karrer B, Newman M E J and Zdeborova L 2014 Percolation on sparse networks *Phys. Rev. Lett.* **113** 208702
[53] Krzakala F, Moore C, Mossel E, Neeman J, Sly A, Zdeborova L and Zhang P 2013 Spectral redemption in clustering sparse networks *Proc. Nat. Acad. Sci. USA* **110** 20935
[54] Lubetzky E and Sly A 2010 Cutoff phenomena for random walks on random regular graphs *Duke Math. J.* **153** 475
[55] Ben-Hamou A and Salez J 2017 Cutoff for nonbacktracking random walks on sparse random graphs *Ann. Probab.* **45** 1752
[56] van der Hofstad R, Hooghiemstra G and Van Mieghem P 2005 Distances in random graphs with finite variance degrees *Random Structures & Algorithms* **27** 76
[57] Tishby I, Biham O, Kühn R and Katzav E 2022 The mean and variance of the distribution of shortest path lengths of random regular graphs *J. Phys. A: Math. Theor.* **55** 265005
[58] Fernholz D and Ramachandran V 2007 The diameter of sparse random graphs *Random Structures & Algorithms* **31** 482
[59] Riordan O and Wormald N 2010 The diameter of sparse random graphs *Combinatorics, Probability & Computing* **19** 835
[60] Chung F and Lu L 2001 The diameter of sparse random graphs *Advances in Applied Mathematics* **26** 257
[61] Bollobás B and de la Vega W F 1982 The diameter of random regular graphs *Combinatorica* **2** 125
[62] Van Mieghem P 2023 *Graph Spectra for Complex Networks, Second Edition* (Cambridge: Cambridge University Press)
[63] Bonneau H, Hassid A, Biham O, Kühn R and Katzav E 2017 Distribution of shortest cycle lengths in random networks *Phys. Rev. E* **96** 062307
[64] Phillips C L, Nagle H T, and Chakrabortty A 2015 *Digital Control System: Analysis and Design, Fourth Edition* (Harlow: Pearson Education)
[65] Nitzan M, Katzav E, Kühn R and Biham O 2016, Distance distribution in configuration-model networks *Phys. Rev. E***93** 062309

[66] Pitman J 1993 *Probability* (New York: Springer-Verlag)
[67] Erdős P and Rényi A 1959 On random graphs I, *Publicationes Mathematicae* **6** 290
[68] Erdős P and Rényi A 1960 On the evolution of random graphs *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17
[69] Erdős P and Rényi A 1961 On the evolution of random graphs II *Bull. Inst. Int. Stat.* **38** 343
[70] Olver F W J, Lozier D M, Boisvert R R and Clark C W 2010 *NIST Handbook of Mathematical Functions* (Cambridge: Cambridge University Press)
[71] Tishby I, Biham O and Katzav E 2016 The distribution of path lengths of self avoiding walks on Erdős–Rényi networks *J. Phys. A* **49** 285002
[72] Tishby I, Biham O and Katzav E 2017 The distribution of first hitting times of random walks on Erdős–Rényi networks *J. Phys. A* **50** 115001
[73] Tishby I, Biham O and Katzav E 2017 The distribution of first hitting times of nonbacktracking random walks on Erdős–Rényi networks *J. Phys. A* **50** 205003
[74] Tishby I, Biham O and Katzav E 2021 Analytical results for the distribution of first hitting times of random walks on random regular graphs *J. Phys. A* **54** 145002
[75] Cooper C and Frieze A M 2005 The cover time of random regular graphs *SIAM J. Discrete Math.* **18** 728
[76] Tishby I, Biham O and Katzav E 2022 Analytical results for the distribution of cover times of random walks on random regular graphs *J. Phys. A* **55** 015003
[77] Noldus R and Van Mieghem P 2015 Assortativity in complex networks *J. Comp. Networks* **3** 507
[78] Van Mieghem P, Wang H, Ge X, Tang S and Kuipers F A 2010 Influence of assortativity and degree preserving rewiring on the spectra of networks *Eur. Phys. J. B* **76** 643
[79] Jalan S and Yadav A 2015 Assortative and disassortative mixing investigated using the spectra of graphs *Phys. Rev. E* **91** 012813
[80] Minh Pham T, Peron T and Metz L F 2024 Effects of clustering heterogeneity on the spectral density of sparse networks *Phys. Rev. E* **110** 054307

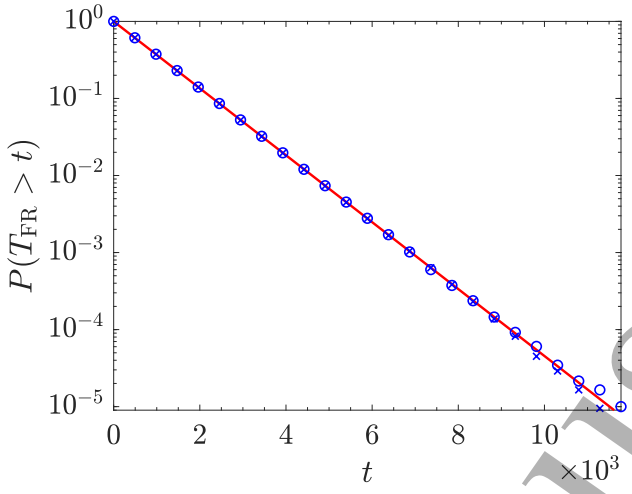

Figure 1. Analytical results (solid line), obtained from Eq. (32), for the tail distribution $P(T_{\rm FR} > t)$ of first return times of an NBW on an RRG of size $N = 1000$. The right hand side of Eq. (32) does not depend on the degree c, which implies that these results are valid for RRGs with any degree $c \geq 3$. Indeed, the analytical results are in very good agreement with the results obtained from computer simulations for $c = 3$ ($\times$) and for $c = 10$ ($\circ$). Each data point of the simulation results was obtained by averaging the results obtained for 20 network instances and 100, 000 NBW trajectories for each network instance.

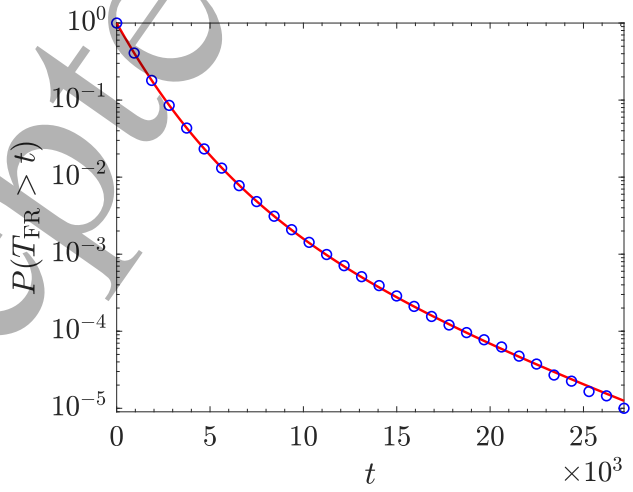

Figure 2. Analytical results for the tail distribution $P(T_{\rm FR} > t|K > 1)$ (solid line) of first return times of an NBW on an Erdős-Rényi network of size $N = 1000$ and mean degree $c = 10$. The analytical results, obtained from Eq. (42), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 1.

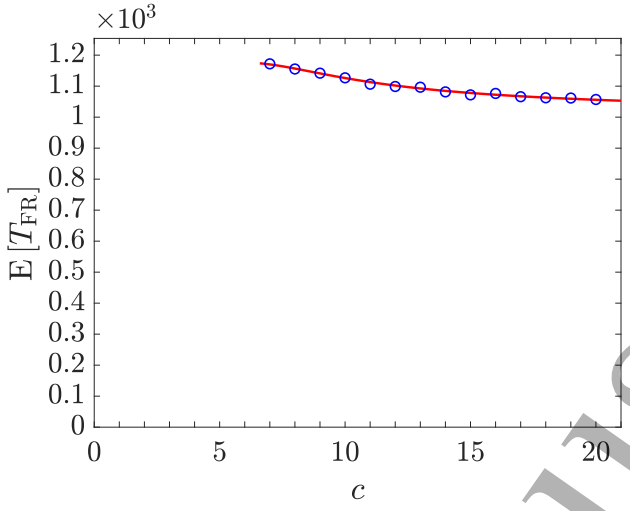

Figure 3. Analytical results for the mean first return time $\mathbb{E}[T_{\rm FR}|K > 1]$ (solid line) of an NBW on an Erdős-Rényi network of size $N = 1000$, as a function of the mean degree c, for $c > \ln N$, where the whole network consists of a single connected component and network instances that include leaf nodes are discarded. The analytical results, obtained from Eq. (45), are in very good agreement with the results obtained from computer simulations (circles). Each data point of the simulation results was obtained by averaging the results obtained for 20 network instances and 10,000 NBW trajectories for each network instance.

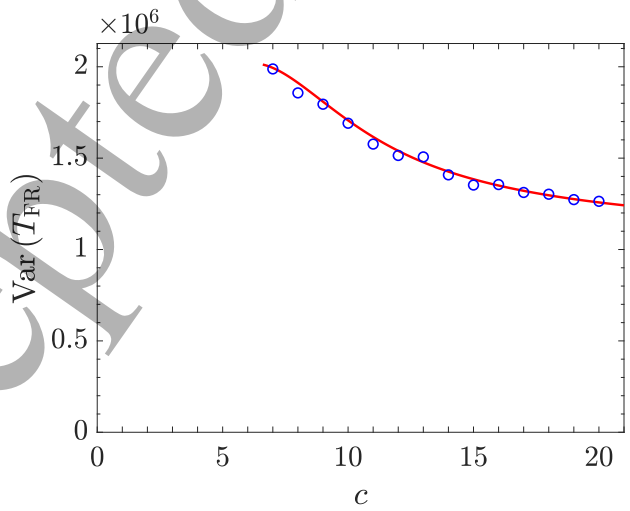

Figure 4. Analytical results for the variance $\mathrm{Var}(T_{\rm FR}|K > 1)$ (solid line) of the distribution of first return times of an NBW on an Erdős-Rényi network of size $N = 1000$, as a function of the mean degree c. The analytical results, obtained from Eq. (50), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 3.

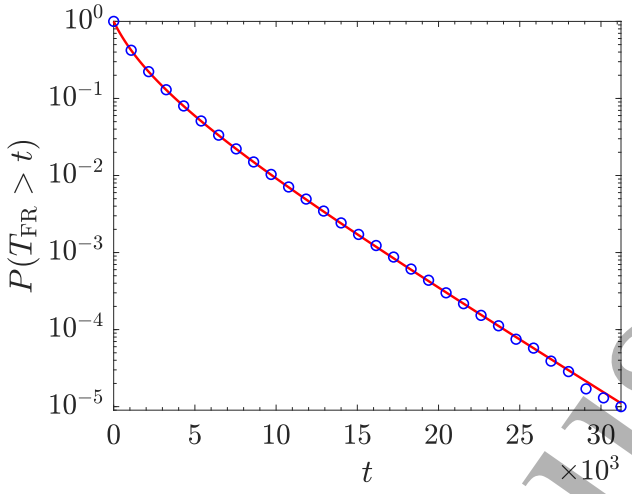

Figure 5. Analytical results for the tail distribution $P(T_{\rm FR} > t)$ (solid lines) of first return times of an NBW on a configuration model network of size $N = 1000$ which exhibits an exponential degree distribution with $k_{\rm min} = 3$ and mean degree $c = 10$. The analytical results, obtained from Eq. (57), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 1.

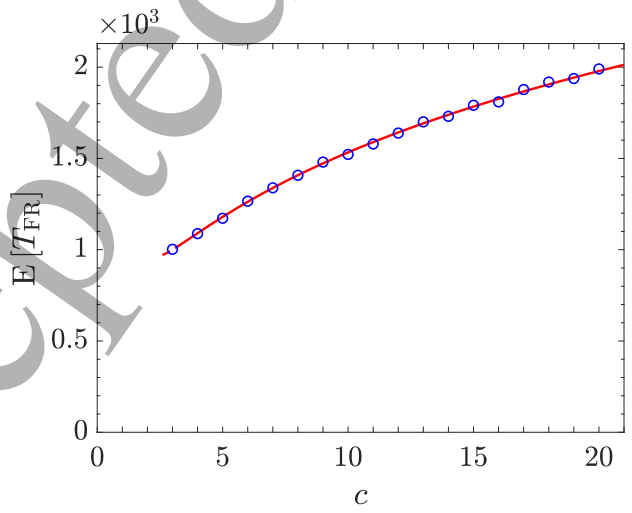

Figure 6. Analytical results for the mean first return time $\mathbb{E}[T_{\rm FR}]$ (solid line) of an NBW on a configuration model network of size $N = 1000$, which exhibits an exponential degree distribution with $k_{\rm min} = 3$, as a function of the mean degree c. The analytical results, obtained from Eq. (59), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 3.

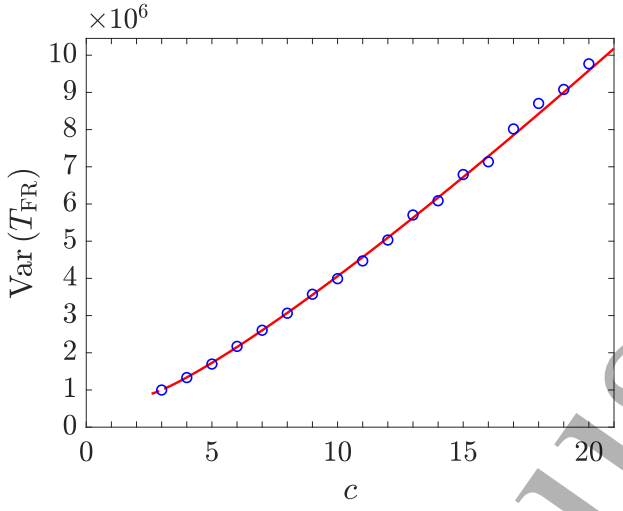

Figure 7. Analytical results for the variance $\mathrm{Var}(T_{\mathrm{FR}})$ (solid line) of the distribution of first return times of an NBW on a configuration model network of size $N = 1000$, which exhibits an exponential degree distribution with $k_{\mathrm{min}} = 3$, as a function of the mean degree c. The analytical results, obtained from Eq. (63), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 3.

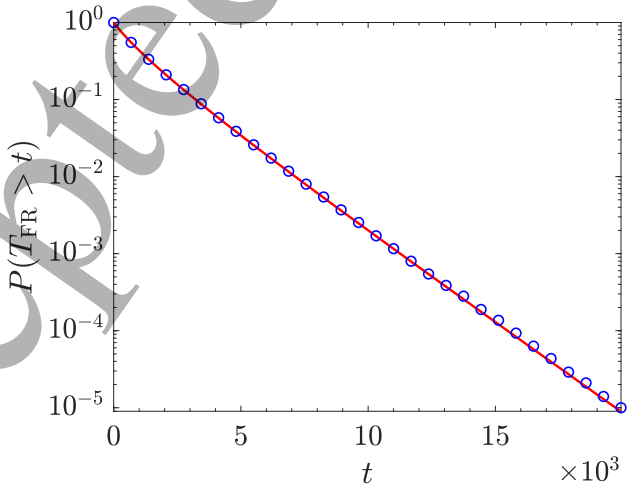

Figure 8. Analytical results for the tail distribution $P(T_{\mathrm{FR}} > t)$ (solid lines) of first return times of an NBW on a configuration model network of size $N = 1000$ which exhibits a power-law distribution with $k_{\mathrm{min}} = 3$, $k_{\mathrm{max}} = 30$ and $\gamma = 2.5$. The analytical results, obtained from Eq. (69), are in very good agreement with the results obtained from computer simulations (circles). The simulation results were obtained using the same averaging procedure as in Fig. 1.

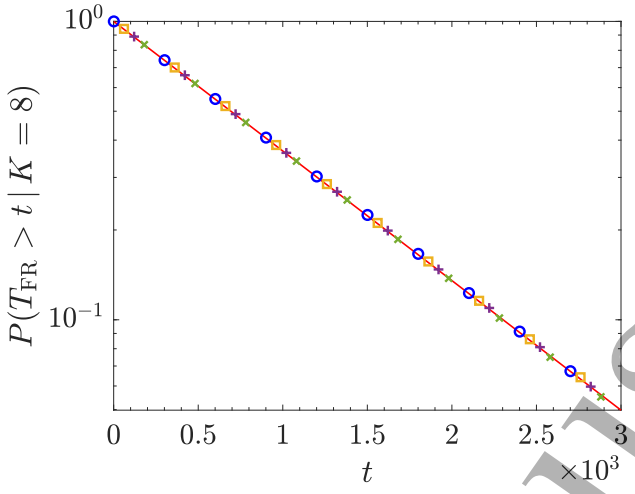

Figure 9. Analytical results for the distribution of first return times for NBWs starting from an initial node of degree $k = 8$ in configuration model networks of size $N = 1000$ and mean degree $c = 8$, given by Eq. (13). The analytical results are found to be in very good agreement with simulation results for RRGs ($\circ$), ER networks ($\times$) and configuration model networks with exponential ($+$) and power-law ($\square$) distributions. Each data point in the simulation results was obtained by averaging over 20 independent network instances and 10,000 NBW trajectories in each network instance, starting from the same initial node of degree $k = 8$.