

# Continuous-Time Random Walks and Temporal Networks

Renaud Lambiotte<sup>1</sup>

<sup>1</sup>*Mathematical Institute, University of Oxford, Oxford (UK)*

(Dated: June 17, 2019)

Real-world networks often exhibit complex temporal patterns that affect their dynamics and function. In this chapter, we focus on the mathematical modelling of diffusion on temporal networks, and on its connection with continuous-time random walks. In that case, it is important to distinguish active walkers, whose motion triggers the activity of the network, from passive walkers, whose motion is restricted by the activity of the network. One can then develop renewal processes for the dynamics of the walker and for the dynamics of the network respectively, and identify how the shape of the temporal distribution affects spreading. As we show, the system exhibits non-Markovian features when the renewal process departs from a Poisson process, and different mechanisms tend to slow down the exploration of the network when the temporal distribution presents a fat tail. We further highlight how some of these ideas could be generalised, for instance in the case of more general spreading processes.

PACS numbers: Random walks; Renewal processes; Diffusion; Stochastic Temporal Networks

## I. INTRODUCTION

[?] Random walks are a paradigmatic model for stochastic processes, finding applications in a variety of scientific domains [1, 2], and helping to understand how the random motion of particles leads to diffusive processes at the macroscopical scale. Classically defined on infinitely large regular lattices or on continuous media, random walks have long been studied on non-trivial topologies, as different parts of the system, adjacent or not, are connected through a predefined transition probability. In a finite and discrete setting, random walks are equivalent to Markov chains [3], whose behaviour is entirely encoded in their transition matrix. The matrix allows to characterise the succession of states visited by the walker, whose dynamics is seen as a discrete-time process and where time is measured by the number of jumps experienced by the walker. However, the complete description of a trajectory requires an additional input about the statistical properties of the timings at which the jumps take place, usually under the form of a waiting-time distribution for the walker. Taken together, the modelling of where to and when the next step will be form the core of the theory of continuous-time random walks.

Random walks also play a central role within the field of network science, and provide a simple framework where to understand the relation between their structure and dynamics. Random walks on networks have been used to model diffusion of ideas in social networks, or diffusion of people in location networks, for instance [4]. In their dual form, they are also used for the modelling of decentralised consensus [5]. In addition, random walks have been exploited to extract non-local information from the underlying network structure. Take Pagerank for instance, defined as the density of walkers on a node at stationarity [6]; or random walk-based kernels defining a similarity measure between nodes, and embedding nodes in low-dimensional space [7]; or community detection where clusters are defined in terms of

their tendency to trap a walker long times [8–10]. In each of these examples, properties of the process at a slow time scale are essentially used in order to capture large-scale information in the system. Importantly, these works usually rely on discrete-time random walks or on basic Poisson processes in their continuous counterpart.

The structure of networks has been the subject of intense investigation since the early works on small-world or scale-free networks [11]. This activity has originally been driven by the availability of large relational datasets in a variety of disciplines, leading to the design of new methods to uncover their properties and of models to reproduce the empirical findings. Yet, a vast majority of datasets only provided static snapshots of networks, or information about their growth in certain conditions. It is only more recently that the availability of fine-grained longitudinal data motivated the study of the temporal properties of networks [12–14]. Several works have shown that the dynamics of real-world networks are non-trivial and exhibit a combination of temporal correlations and non-stationarity.

In this chapter, we will focus on a particular aspect that has attracted much attention in the literature, the presence of burstiness in the temporal series of network activity [15]. Take a specific node, or a specific edge, and look at the sequence of events associated to that object. The resulting distribution of inter-event times has been shown to differ significantly from an exponential, even after discarding confounding factors [16]. After a short introduction on relevant concepts, and a clarification of the differences between active and passive diffusion, we will use the language of continuous-time random walks to identify how the shape of temporal distributions affects diffusion. In analogy with static networks, where deviations from a binomial degree distribution play a central role, we will focus on renewal processes whose inter-event distribution differs from an exponential. Finally, we will widen the scope and discuss possible generalisations of the models, for instance in the case of non-conservative

spreading processes.

## II. MODELS OF GRAPHS AND OF TEMPORAL SEQUENCES

Random models play an important role when analysing real-world data. The main purpose of this section is to introduce simple random models for graphs and for temporal sequences, and to highlight similarities between them.

### A. Random Graphs

The most fundamental model of random graph is the Erdős-Rényi model. Usually denoted by  $\mathcal{G}(N, q)$ , it takes as parameters the number of nodes  $N$ , and the probability  $q$  that two distinct nodes are connected by a link. By construction, each pair of node is a Bernoulli process, whose realisation is independent from that of other pairs in the graph. The Erdős-Rényi model, as any random graph model, has to be considered as an ensemble of graphs, whose probability of having been realised depends on the model parameters. Due to the independence between the processes defined on each edge, several properties of the model can be computed exactly. This includes the degree distribution, expected to take the form of the binomial distribution

$$p(k) = \binom{N-1}{k} q^k (1-q)^{N-1-k}, \quad (1)$$

but also the number of cliques of any size, or the percolation threshold. Underlying assumptions of a Erdős-Rényi model are often violated in empirical data. Take connections in a social network, where triadic closure induces correlations between neighbouring edges for instance. Yet, the model's simplicity and analytical tractability naturally make it a baseline model, and more realistic network models can be developed systematically by relaxing its assumptions. Well-known examples include:

- *The configuration model.* Real-life networks tend to present a strong heterogeneity in their degrees, associated to fat tailed distributions very different from a binomial. The configuration model is defined as a random graph in which all possible configurations appear with the same probability, with the constraint that each node  $i$  has a prescribed degree  $k_i$ , that is with a tuneable degree distribution.
- *Stochastic block models.* Real-life networks are not homogeneous and their nodes tend to be organised in groups revealing their function in the system. These groups may take the form of assortative or disassortative communities and may be reproduced by stochastic block models where the nodes are divided into  $k$  classes, and the probability  $p_{ij}$  for two nodes  $i, j$ , belonging to class  $c_i$  and  $c_j$ , to be connected is encoded in an affinity matrix  $\Omega_{c_i c_j}$ .

Note that both models can be combined to form so-called degree-corrected stochastic block models [17]. In essence, both models assume that the processes on each edge are independent but break the assumption that they are identical. Certain pairs of nodes are more or less likely to be connected within this framework. Models questioning the independence of different edges include

- *The preferential attachment model.* In this mechanistic model for growing networks, nodes are added one at a time and tend to connect with a higher probability to high degree nodes. The resulting networks naturally produce fat-tailed degree distributions and exhibit correlations building in the course of the process. Variations of the model include divergence-duplication models [18], or copying models [19], whose correlations produce a high density of triangles, and cliques of all size, that are negligible in the afore-mentioned models.

### B. Poisson and renewal processes

Let us now turn our attention to the modelling of temporal sequences of events. Examples include the sequences of retweets of an original tweet, of meeting a cat in the street, or of nuclei to disintegrate in a radioactive material. As a first order approximation, these systems can be modelled by a Poisson process, assuming that events are independent of each other and that their rate is constant over time. As in case of the Erdős-Rényi model, these assumptions are often unrealistic, but the resulting simplicity allows us to derive analytically its statistical properties. For Poisson processes, the inter-event time between two consecutive events, usually denoted by  $\tau$ , is exponentially distributed according to

$$\psi(\tau) = \lambda e^{-\lambda\tau}, \quad (2)$$

where  $\lambda$  is the rate at which events occur. Likewise, the distribution of the number of events observed within a given time window is readily found to be

$$p(n, t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} \quad (3)$$

for any  $n \geq 0$ . Deviations from these distributions in empirical data indicate that the assumptions of a Poisson process are not verified and that a more complicated process is at play. Generalised models relaxing some of these assumptions include:

- *Renewal processes.* Empirical data often show fat-tailed inter-event time distributions, which can be captured by renewal processes. In a renewal process, inter-event times are independent of each other and drawn from the same distribution  $\psi(\tau)$ . When  $\psi(\tau) = \lambda e^{-\lambda\tau}$ , we recover a Poisson process. The properties of renewal processes are usually best

analysed in the frequency domain. After defining the Laplace transform

$$\hat{\psi}(s) = \int_0^\infty \psi(\tau) e^{-s\tau} d\tau. \quad (4)$$

and noting that a convolution in time translates into a product in the Laplace domain, one readily finds the probability of having performed  $n$  steps at time  $t$

$$\hat{p}(n, s) = [\hat{\psi}(s)]^n \frac{1 - \hat{\psi}(s)}{s}. \quad (5)$$

As we will see below, this quantity is critical, as it relates two ways to count time: one in terms of the number of events,  $n$ , and the other in terms of the physical time,  $t$ .

- *Non-homogeneous Poisson processes.* Real-life time series are often non-stationary, which can be incorporated in a Poisson process with a time-dependent event rate  $\lambda(t)$ . For a non-homogeneous Poisson process, (3) is extended as

$$p(n, t) = \frac{\Lambda(t)^n}{n!} e^{-\Lambda(t)}, \quad (6)$$

where

$$\Lambda(t) = \int_0^t \lambda(t') dt'. \quad (7)$$

Similarly, the distribution of inter-event times is given by

$$\psi(\tau) = \lambda(\tau) e^{-\Lambda(\tau)}, \quad (8)$$

and leads to time-dependent, non-exponential distributions in general.

The previous two generalisations assume that successive inter-event times are independent processes, which is expected to be invalid in many situations. Take discussion threads between individuals for instance, or cascades of events in social media. This type of situation can instead be modelled by

- *Self-exciting processes.* The underlying idea is that an event induces an additional event rate for future events. This property is at the core of Hawkes processes [20, 21] where, in their simplest form, the event rate at time  $t$  is given by

$$\lambda(t) = \lambda_0 + \sum_{\ell, t_\ell \leq t} \chi(t - t_\ell), \quad (9)$$

where  $t_\ell$  is the time of the  $\ell$ th event. The model incorporates a baseline rate  $\lambda_0$  independent of self excitation and a memory kernel  $\chi(t)$  describing the additional rate incurred by an event. It is generally assumed that  $\chi(t)$  peaks at  $t = 0$  and monotonically decay towards zero as  $t$  increases.

As in the case of random graph models, different generalisations can be combined to form more realistic models. This is the case of TideH, for instance, a model for retweet dynamics combining ingredients from Hawkes processes and non-homogeneous Poisson processes [22].

### III. TRAJECTORIES ON NETWORKS

#### A. Discrete-time dynamics

Let us now turn to the case of static networks and the description of trajectories on their nodes. A canonical example could be one-dollar bill transiting between individuals forming a large social network [23]. For the sake of simplicity, we will assume that social interactions are undirected and unweighted, and that the whole network forms a connected component. The structure of the network is encoded through its adjacency matrix  $A$ , whose element  $A_{ij}$  determines the presence of an edge between nodes  $i$  and  $j$ . As a first step, we consider the case when the random walk process takes place at the discrete times  $n$ . A trajectory on the network is thus characterised by a sequence  $X_0, X_1, \dots, X_n, \dots$ , where  $X_n$  is a random variable denoting the node visited by the walker at time  $n$ . In general, the state  $X_{n+1}$  may depend on all preceding states of the dynamics and the probability of visiting a certain node  $i$  requires information about the full history of the process

$$p(X_{n+1} = i_{n+1} | X_n = i_n, \dots, X_1 = i_1, X_0 = i_0). \quad (10)$$

The process simplifies drastically in situations when the system is stationary and the conditional probability depends only on the state at time  $n$ . The process then takes the form of a Markov chain, and is fully described by its initial state and the  $N \times N$  transition matrix

$$p(X_{n+1} = j | X_n = i) \equiv T_{ij} = \frac{A_{ij}}{k_i}, \quad (11)$$

where  $k_i$  is, as before, the degree of node  $i$ . For instance, the probability that state  $i$  is visited at time  $n$ , denoted by  $P_i(n)$ , evolves according to

$$P_j(n+1) = \sum_{i=1}^N P_i(n) T_{ij} \quad (12)$$

or, in matrix notations,

$$P(n+1) = P(n)T, \quad (13)$$

yielding the formal solution

$$P(n) = P(0)T^n. \quad (14)$$

When the underlying network is connected and the corresponding Markov chain is ergodic, it can be shown that any initial condition converges to the stationary density  $P_i^* = k_i/2m$  solution of

$$P^* = P^*T. \quad (15)$$

## B. Fourier modes

The solution, Eq. (14), involves products of matrices, which can be simplified by rewriting the system in the basis formed by the eigenvectors of the transition matrix. This operation is sometimes called graph Fourier transform [24], and allows to replace the matrix products by algebraic products for amplitudes associated to the different dynamical modes. To show so, we consider the symmetric matrix

$$\tilde{A}_{ij} = \frac{A_{ij}}{\sqrt{k_i k_j}}, \quad (16)$$

and its spectral decomposition

$$\tilde{A}_{ij} = \sum_{\ell=1}^N \lambda_{\ell} u_{\ell} u_{\ell}^{\top}, \quad (17)$$

where  $u_{\ell}$  is the normalised eigenvector of eigenvalue  $\lambda_{\ell}$ , and where we have assumed that all eigenvalues are distinct to avoid unnecessary complications. By construction, the eigenvectors verify  $\langle u_{\ell}, u_{\ell'} \rangle = \delta_{\ell\ell'}$  and form a proper basis for signals defined on nodes of the network. One should also note that the eigenvalues are in the interval  $[-1, 1]$ , that the multiplicity of the dominant eigenvalue 1 gives the number of connected components in the graph, and that an eigenvalue equal to  $-1$  indicates that the graph is bipartite. It is straightforward to show that the left and right eigenvectors of  $T$  are given by

$$u_{\ell}^L = \left( (u_{\ell})_1 \sqrt{k_1} \cdots (u_{\ell})_N \sqrt{k_N} \right), \quad (18)$$

$$u_{\ell}^R = \left( (u_{\ell})_1 / \sqrt{k_1} \cdots (u_{\ell})_N / \sqrt{k_N} \right)^{\top} \quad (19)$$

respectively which implies, after some algebra, that the state of the random walk after  $n$  steps is given by a linear combination of the eigenmodes

$$P_i(n) = \sum_{\ell=1}^N a_{\ell}(n) (u_{\ell}^L)_i, \quad (20)$$

where the amplitude of the modes evolve as

$$a_{\ell}(n) = \lambda_{\ell}^n a_{\ell}(0). \quad (21)$$

and  $a_{\ell}(0) \equiv \langle P(0), u_{\ell}^R \rangle$  is given by the initial condition.

The spectral decomposition (20) helps to understand how the structure of a network affects the diffusion of a random walker. The stationary density corresponds to the mode with  $\lambda_{\ell} = 1$ , assumed to be unique as the network is connected. In addition, in situations when the network is non-bipartite, and no eigenvalue is equal to  $-1$ , all the other modes asymptotically decay to 0, each one with a time-scale associated to its eigenvalue. The long-time relaxation to the stationary density is determined by the eigenmode associated to  $\lambda_{2\text{nd}}$ , the second largest eigenvalue, which is related to the spectral gap  $1 - \lambda_{2\text{nd}}$ . A small spectral gap entails slow relaxation and the presence of a bottleneck between communities in the network, as shown by the Cheeger inequality [25].

## C. Continuous-time dynamics

We have described the trajectory of the walker in terms of the number of jumps so far. We now turn to the situation when the jumps take place in continuous time, which motivates the use of continuous-time random walk processes. The passage from discrete to continuous time is usually done by modelling the sequence of jumps of the walker as a renewal process: the walker waits between two jumps for a duration  $\tau$  given by the probability density function  $\psi(\tau)$  before performing a transition according to the Markov chain. We have assumed here that the waiting time distribution is identical on each node. The position of the walker at time  $t$  is given

$$P_i(t) = \sum_{n=0}^{\infty} P_i(n) p(n, t), \quad (22)$$

where we have used the fact that arriving on node  $i$  in  $n$  steps, and performing  $n$  steps in time  $t$  are independent processes.

By going into the Laplace domain, and combining (5) and (14), we find

$$\hat{P}(s) = \frac{1 - \psi(s)}{s} P(0) \left[ I - T \hat{\psi}(s) \right]^{-1}, \quad (23)$$

whose inverse Laplace transform provides the probability  $P_i(t)$  for the walker to be on node  $i$  at time  $t$ . The latter usually involves convolutions in time, reflecting the lack of Markovianity of the process for general renewal process. (23) also takes the equivalent form

$$\left( \frac{1}{\psi(s)} - 1 \right) \hat{P}(s) = \left( \frac{1}{\psi(s)} - 1 \right) \frac{1}{s} P(0) + \hat{P}(s) L \quad (24)$$

where  $T - I = L$  is the normalised Laplacian of the network. This expression simplifies drastically in the case when the renewal process is a Poisson process, and  $\psi(\tau) = \lambda e^{-\lambda\tau}$ , leading to the standard rate equation

$$\frac{dP(t)}{dt} = P(t) L. \quad (25)$$

It is important to emphasise that (24) provides an exact solution to the problem and that it departs from (25) through its causal operator  $\left( \frac{1}{\psi(s)} - 1 \right)$ , which translates the input from the neighbours of a node into a change of its state, different from a usual  $\frac{d}{dt}$ . Note that this operator asymptotically behaves like a fractional derivative in situations when the waiting time distribution has a power-law tail [26].

This solution also helps to understand us to clarify the impact of the shape of the waiting time distribution on the speed of diffusion. In the basis of eigenvectors of the transition matrix, it is straightforward to generalise (21) to obtain

$$\hat{P}(s) = \frac{1 - \hat{\psi}(s)}{s} \sum_{\ell=1}^N \frac{a_{\ell}(0)}{1 - \lambda_{\ell} \hat{\psi}(s)} u_{\ell}^L, \quad (26)$$

In other words, the time evolution of the amplitude of each mode is given by

$$\hat{a}_\ell(s) = \frac{1 - \hat{\psi}(s)}{s [1 - \lambda_\ell \hat{\psi}(s)]} a_\ell(0), \quad (27)$$

which is, in general, different from an exponential decay. The asymptotic behaviour can be obtained by performing a small  $s$  expansion

$$\hat{\psi}(s) = 1 - \langle \tau \rangle s + \frac{1}{2} \langle \tau^2 \rangle s^2 + o(s^2), \quad (28)$$

and assuming a finite mean and variance, yielding the dominant terms

$$a_\ell(s) = \frac{\langle \tau \rangle}{1 - \lambda} \left[ 1 - s \left( \frac{\lambda_\ell \langle \tau \rangle}{1 - \lambda_\ell} + \frac{\langle \tau^2 \rangle}{2 \langle \tau \rangle} \right) \right]. \quad (29)$$

and thus a characteristic time  $t_\ell$ ,

$$t_\ell = \langle \tau \rangle \left( \frac{1}{\epsilon_\ell} + \beta \right), \quad (30)$$

where  $\epsilon_\ell = 1 - \lambda_\ell$  is an eigenvalue of the normalised Laplacian and

$$\beta = \frac{\sigma_\tau^2 - \langle \tau \rangle^2}{2 \langle \tau \rangle^2}, \quad (31)$$

is the variance of  $\tau$ . Poisson processes yield  $\beta = 0$  and this expression clearly shows that negative values of  $\beta$  tend to accelerate the relaxation of each mode, while larger values slow them down. The former happens in the case of discrete-time dynamics for instance, when  $\psi(\tau)$  is a delta distribution. The latter is when the distribution has a fat-tail. Importantly, (30) shows that a combination of structural and temporal information determine the temporal properties of the process.

## IV. DIFFUSION ON TEMPORAL NETWORKS

### A. Active versus passive walks

What about temporal networks? The results derived in the previous section focus on random walks on static networks. They are nonetheless instructive to model and understand diffusion on temporal networks. As a first step, we should emphasise that the distinction between the dynamics on the network and the dynamics of the network is not always clear-cut [27, 28]. The temporal nature of a network usually comes from time series of events taking place on nodes or edges. There are situations when it is the diffusive process itself that defines the temporal network. Take the action of sending an email or an SMS to a friend, and the modelling of information diffusion in the resulting network. In this case of **active** diffusion, the motion of the random walker is defining the

temporal patterns of activity on links existing, as transition trigger the activation of a potential edge. The model of section III C is then a good candidate to explore the interplay between structure and dynamics in the resulting temporal network. Note that even the Poisson model described by (25) can then be seen as generating a temporal network, even if it is not a very interesting one.

There are other situations, however, when the motion of the walker does not trigger the connections between nodes, but it is instead constrained by their temporal patterns. A good example would that of a person random walking a public transportation network, or of a disease spreading in a contact network. In that case, the temporal sequence of edges restrains the time-respecting paths that are available for the walker and one talks of **passive** diffusion. As we will see, passive random walks can also be mapped to continuous-time random walks, to some extent, but this operation must be performed more carefully. As a simple model of temporal networks supporting diffusion, let us consider a stochastic temporal network. The network is modelled as a set of potential edges between nodes, each one evolving as an independent renewal process, with a distribution of inter-event times  $\phi(\tau)$ . A random walker located on a node  $i$  performs a jump as soon as an edge appears, for instance to node  $j$ , where it waits until the next available edge.

### B. Bus paradox and backtracking transitions

When considering passive random walks, it is critical to clearly distinguish the waiting-time distribution  $\psi(\tau)$  from the inter-event time distribution  $\phi(\tau)$ . The former characterises the times that a walker has to wait on a node before its next move. The latter gives the time between edge activations in the renewal process defining the stochastic temporal network. The inter-event time distribution is a parameter of the network model but the motion of the walker is directly affected by the waiting-time distribution, and it should thus be estimated. To do so, let us first focus on the case of a walker arriving at a node  $j$  from  $i$  and calculating the time before an edge to another edge  $k$  appears. Assuming the independence between the act of arriving on node  $j$  and the appearance of the edge to  $k$ , one finds that both distributions are related as

$$\psi(\tau) = \frac{1}{\langle \tau \rangle_\phi} \int_\tau^\infty \phi(\tau') d\tau', \quad (32)$$

where  $\langle \tau \rangle_\phi$  is the average inter-event time distribution. Most strikingly, the average waiting time is

$$\langle \tau \rangle_\psi = \frac{1}{2} \frac{\langle \tau^2 \rangle_\phi}{\langle \tau \rangle_\phi} \quad (33)$$

and depends on the variance of the inter-activation time. At a fixed value of  $\langle \tau \rangle_\phi$ , the waiting time can thus be arbitrarily large if the variance of  $\phi(\tau)$  is sufficiently large.

The waiting-time paradox is a standard result in queuing theory [29] and is an example of length-biased sampling. It is often called bus paradox, after the observation that the average waiting time at a bus stop tends to be longer than half of the average interval between two buses expected from the timetable.

As a second step, it is important to note that the approximation behind the derivation of (32) is, in general, not respected if the walker passes several times by the same edge, as information about the previous passage time may help to predict the next activation time. This effect is most apparent in (but is not limited to) the case of undirected networks. Consider a walker taking an edge from node  $i$  to node  $j$ . The time before the next activation of the edge back to  $i$  is clearly not given by  $\psi(\tau)$  in (32), but simply by  $\phi(\tau)$ . The waiting time for a backtracking edge is therefore typically different, and shorter for fat-tailed distributions, than the waiting time of a non-backtracking edge. This difference is particularly critical because edges are in competition with each other [27]. When a walker waits on a node, the model specifies that it takes the first edge to appear. For this reason, the prevalence of an edge over another may bias the trajectory of the walker, and make the process non-Markovian [28].

Let us quantify this effect when the walker arrives on a node  $j$  with degree  $k_j$ , and consider the probability of the time  $t$  at which the walker takes a specific edge. As before, the inter-event times of the links are identically and independently distributed according to  $\phi(\tau)$ . Starting from the time when the walker arrived on  $j$ , the time of the next activation for a backtracking step is simply  $\phi(\tau)$ . The time for another edge to activate is instead determined by the waiting-time distribution  $\psi(\tau)$ , where we assume that the approximation (32) is valid. For the walker to take an edge at time  $t$ , no other edge can have appeared in  $[0, t]$ . Therefore, we obtain

$$f_{\text{back}}(t) \approx \phi(t) \left[ \int_t^\infty \psi(\tau) d\tau \right]^{k_i-1} \quad (34)$$

$$f_{\text{non-back}}(t) \approx \psi(t) \left[ \int_t^\infty \psi(\tau) d\tau \right]^{k_i-2} \int_t^\infty \phi(\tau) d\tau \quad (35)$$

As we discussed, if  $\psi$  has a fat tail, the waiting-time is larger than the inter-event time on average and the walker preferentially backtracks, thereby leading to non-Markovian trajectories. Here, we should clarify the distinction between two types of non-Markovianity that can be induced on temporal networks. In (24), the sequence of nodes visited by the walker is described by a Markov chain, but the statistical properties of the timings induce long-term memory effects. In (34), instead, the sequence of nodes visited by the walker can not be produced by a first-order Markov process.

To summarise, when considering passive diffusion on a stochastic temporal network, the rate at which the random walker explores the network is slowed down in three ways when the inter-event time distribution has a fat tail,

namely through:

- *the bus paradox*, because the waiting time of the walker on the nodes tends to be longer on average. As the speed of diffusion is controlled by the sum of the waiting times of the walker, this effect leads to a slow down of diffusion.
- *the backtracking bias*. The random walker tends to backtrack, which hinders its exploration of the network and can be shown to increase the mixing time of the process [30].
- *the variance of the waiting-time distribution*. On top of the slow down due to the bus paradox, and an increase of the average waiting time, the variance of the waiting time also slows down diffusion through the same mechanisms as for active random walks, in (30).

## V. PERSPECTIVES

The main purpose of this chapter was to provide an overview on theoretical results for diffusion on temporal networks. As we discussed, the problem may be understood through the lens of continuous-time random walks, after carefully distinguishing between waiting time and inter-event time, on the one hand, and active versus passive walks, on the other hand. An analytical approach allows us to identify unambiguously the mechanisms accelerating or slowing down the diffusion, and also helps to warn against caveats that could be met with numerical simulations. A good example concerns the use of null models to determine how the temporal nature of a real-world network affects diffusion. A popular solution consists in comparing numerical simulations of diffusion on the original data and on different versions of randomised data [31]. The results from section IV B show that randomised null models in which temporal correlations are removed yet have the tendency towards backtracking, and thus to slow down the exploration of the network.

The results presented in this chapter also open different perspectives for future research. As we discussed in section II A, different types of random graph models have been proposed for static networks. A model like the stochastic temporal network incorporates temporal activity on a given network structure, which opens the question of how to properly define generalisations of the configuration model or stochastic block model in this context. Answers may be found by clarifying connections with activity-driven models, where dynamics can also be generated by general renewal processes [32, 33]. The temporal networks presented here also suffer from limitations that may hinder their applicability. Those include their stationarity, the absence of correlations between edges and the instantaneity of the interactions. To address the last two limitations, we point the reader to the possibility to use higher-order Markov models for the data [34, 35],

and recent generalisations based on continuous-time random walks allowing for interactions with a finite duration [36]. The latter emphasises a critical aspect of temporal networks, which are characterised by different processes and their corresponding time scales. Those include one associated to the motion of the random walker, one to the time between successive activity periods of the edges and another to the duration of the activity periods. Depending on the model parameter, one process may dominate the others and lead to mathematical simplifications for the dynamics.

As a final comment, we would also like to come back on our observation that temporal networks may be generated by active diffusive processes. This chapter was limited to conservative spreading processes, where the number of diffusing entities is preserved in time. There are many practical situations, however, when this is not the case. Take the spreading of viruses in human populations or of hashtags in online social networks for instance. In that case, other spreading models should be considered to generate more realistic temporal networks. A promising candidate is multivariate Hawkes processes, generalising (9) to interacting entities, and whose equation of evolution takes a form very similar to (24)

$$\lambda(s) = \frac{\lambda_0}{s} + \chi(s)\lambda(s)A, \quad (36)$$

where the vector  $\lambda(s)$  is the Laplace transform of the average rate of activity on each node,  $\chi(s)$  of the memory kernel and  $A$  is the adjacency matrix of the network. The presence of the adjacency matrix instead of the Laplacian is a clear sign of the epidemic nature of the spreading. Note also that a heterogeneous mean-field, *à la configuration model*, version of the model has been considered for the modelling of retweet cascades [37]. Alternatively, the related family of Bellman-Harris branching processes could be used, where a node  $i$  remains infected for a duration  $t$ , determined by a distribution  $\rho(t)$ , before infecting its neighbours, leading to

$$\lambda(s) = \frac{1 - \rho(s)}{s} + \rho(s)\lambda(s)A. \quad (37)$$

In each case, the active process allows for the formation of a growing cascade of infections, and includes a non-trivial causal operator.

- 
- [1] Jari Saramäki and Petter Holme. Exploring temporal networks with greedy walks. *The European Physical Journal B*, 88(12):334, 2015.
  - [2] Radu Balescu. *Statistical dynamics: matter out of equilibrium*. Imperial College London, 1997.
  - [3] Joseph Klafter and Igor M Sokolov. *First steps in random walks: from tools to applications*. Oxford University Press, New York, 2011.
  - [4] László Lovász et al. Random walks on graphs: A survey. *Combinatorics, Paul erdos is eighty*, 2(1):1–46, 1993.
  - [5] Naoki Masuda, Mason A Porter, and Renaud Lambiotte. Random walks and diffusion on networks. *Physics reports*, 716:1–58, 2017.
  - [6] Vincent D Blondel, Julien M Hendrickx, Alex Olshevsky, and John N Tsitsiklis. Convergence in multiagent coordination, consensus, and flocking. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2996–3000. IEEE, 2005.
  - [7] S. Brin and L. Page. Anatomy of a large-scale hyper-textual web search engine. *Proceedings of the Seventh International World Wide Web Conference*, pages 107–117, 1998.
  - [8] François Fouss, Marco Saerens, and Masashi Shimbo. *Algorithms and models for network data and link analysis*. Cambridge University Press, 2016.
  - [9] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA*, 105:1118–1123, 2008.
  - [10] J. C. Delvenne, S. N. Yaliraki, and M. Barahona. Stability of graph communities across time scales. *Proc. Natl. Acad. Sci. USA*, 107:12755–12760, 2010.
  - [11] R. Lambiotte, J. C. Delvenne, and M. Barahona. Random walks, Markov processes and the multiscale modular organization of complex networks. *IEEE Trans. Netw. Sci. Eng.*, 1:76–90, 2014.
  - [12] Mark Newman. *Networks: an introduction*. Oxford university press, 2010.
  - [13] Petter Holme and Jari Saramäki. Temporal networks. *Physics reports*, 519(3):97–125, 2012.
  - [14] Petter Holme. Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9):1–30, 2015.
  - [15] Naoki Masuda and Renaud Lambiotte. *A guide to temporal networks*. World Scientific, London, 1996.
  - [16] Albert-Laszlo Barabasi. The origin of bursts and heavy tails in human dynamics. *Nature*, 435(7039):207, 2005.
  - [17] R Dean Malmgren, Daniel B Stouffer, Adilson E Motter, and Luís AN Amaral. A poissonian explanation for heavy tails in e-mail communication. *Proceedings of the National Academy of Sciences*, 105(47):18153–18158, 2008.
  - [18] Brian Karrer and Mark EJ Newman. Stochastic block-models and community structure in networks. *Physical review E*, 83(1):016107, 2011.
  - [19] Iaroslav Ispolatov, PL Krapivsky, and A Yuryev. Duplication-divergence model of protein interaction network. *Physical review E*, 71(6):061911, 2005.
  - [20] R Lambiotte, PL Krapivsky, U Bhat, and S Redner. Structural transitions in densifying networks. *Physical review letters*, 117(21):218301, 2016.
  - [21] A. G. Hawkes. Point spectra of some mutually exciting point processes. *J. R. Stat. Soc. B*, 33:438–443, 1971.
  - [22] Naoki Masuda, Taro Takaguchi, Nobuo Sato, and Kazuo

- Yano. Self-exciting point process modeling of conversation event sequences. In *Temporal Networks*, pages 245–264. Springer, 2013.
- [23] Ryota Kobayashi and Renaud Lambiotte. Tideh: Time-dependent hawkes process for predicting retweet dynamics. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [24] Dirk Brockmann, Lars Hufnagel, and Theo Geisel. The scaling laws of human travel. *Nature*, 439(7075):462, 2006.
- [25] Nathanaël Perraudin and Pierre Vandergheynst. Stationary signal processing on graphs. *IEEE Transactions on Signal Processing*, 65(13):3462–3477, 2017.
- [26] Fan RK Chung and Fan Chung Graham. *Spectral graph theory*. Number 92. American Mathematical Soc., 1997.
- [27] Sarah De Nigris, Anthony Hastir, and Renaud Lambiotte. Burstiness and fractional diffusion on complex networks. *The European Physical Journal B*, 89(5):114, 2016.
- [28] T. Hoffmann, M. A. Porter, and R. Lambiotte. Generalized master equations for non-Poisson dynamics on networks. *Phys. Rev. E*, 86:046102, 2012.
- [29] L. Speidel, R. Lambiotte, K. Aihara, and N. Masuda. Steady state and mean recurrence time for random walks on stochastic temporal networks. *Phys. Rev. E*, 91:012806, 2015.
- [30] A. O. Allen. *Probability, Statistics, and Queueing Theory: With Computer Science Applications*. Academic Press, Boston, MA, USA, second edition, 1990.
- [31] Martin Gueuning, Renaud Lambiotte, and Jean-Charles Delvenne. Backtracking and mixing rate of diffusion on uncorrelated temporal networks. *Entropy*, 19(10):542, 2017.
- [32] Márton Karsai, Mikko Kivelä, Raj Kumar Pan, Kimmo Kaski, János Kertész, A-L Barabási, and Jari Saramäki. Small but slow world: How network topology and burstiness slow down spreading. *Physical Review E*, 83(2):025102, 2011.
- [33] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. Random walks in non-poissonian activity driven temporal networks. *arXiv preprint arXiv:1904.10749*, 2019.
- [34] Antoine Moinet, Michele Starnini, and Romualdo Pastor-Satorras. Burstiness and aging in social temporal networks. *Physical review letters*, 114(10):108701, 2015.
- [35] Ingo Scholtes, Nicolas Wider, René Pfitzner, Antonios Garas, Claudio J Tessone, and Frank Schweitzer. Causality-driven slow-down and speed-up of diffusion in non-markovian temporal networks. *Nature communications*, 5:5024, 2014.
- [36] Renaud Lambiotte, Martin Rosvall, and Ingo Scholtes. From networks to optimal higher-order models of complex systems. *Nature physics*, page 1, 2019.
- [37] Julien Petit, Martin Gueuning, Timoteo Carletti, Ben Lauwens, and Renaud Lambiotte. Random walk on temporal networks with lasting edges. *Physical Review E*, 98(5):052307, 2018.
- [38] Qingyuan Zhao, Murat A Erdogdu, Hera Y He, Anand Rajaraman, and Jure Leskovec. Seismic: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1513–1522. ACM, 2015.