



RESEARCH ARTICLE

# Temporal configuration model: Statistical inference and spreading processes

Thien Minh Le<sup>1</sup> , Hali Hambridge<sup>2</sup> and Jukka-Pekka Onnela<sup>2</sup> 

<sup>1</sup>Department of Mathematics, University of Tennessee at Chattanooga, Chattanooga, TN, USA and <sup>2</sup>Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

**Corresponding author:** Jukka-Pekka Onnela; Email: [onnela@hsph.harvard.edu](mailto:onnela@hsph.harvard.edu)

## Abstract

We introduce a family of parsimonious network models that are intended to generalize the configuration model to temporal settings. We present consistent estimators for the model parameters and perform numerical simulations to illustrate the properties of the estimators on finite samples. We also derive analytical solutions for the basic and effective reproduction numbers for the early stage of the discrete-time SIR spreading process for our temporal configuration model (TCM). We apply three distinct TCMs to empirical student proximity networks and compare their performance.

**Keywords:** time-varying networks; estimation; epidemics; reproduction number; generating functions

## 1. Introduction

In its simplest form, a network is a static collection of nodes joined together by edges. Ties between nodes can be directed or undirected and can be dichotomous or valued, and different kinds of ties may function differently. Networks can be interpersonal, micro-level networks where nodes represent individual actors, or inter-organizational, macro-level networks where nodes represent organizations (Chen et al., 2022). Broadly speaking, networks capture the pattern of interactions between elements of a system, reducing the system to a basic topological structure that facilitates analysis. Network variables can serve as both dependent (outcome) and independent (explanatory) variables (Borgatti and Foster, 2003). Furthermore, research can focus on the impact of individual dyads, or take a whole-of-network approach (Provan et al., 2007).

Networks that change over time, or temporal networks, can model both changes in relational states (i.e., the formation, transformation, or dissolution of a tie) and the occurrence of relational events (i.e., transactional events occurring between two actors) (Chen et al., 2022). In this paper, we present the temporal configuration model (TCM), a simple yet flexible way to model temporal networks. This approach can be used to model networks where edges represent relational states or where ties denote relational events. Furthermore, it can be applied to a diverse set of open research questions, including studying the mechanisms of network dynamics, the outcomes of network dynamics, and the interplay of relational states and relational events (Chen et al., 2022).

Over the last decade, temporal networks have emerged as a fascinating topic that has received substantial research interest from a variety of disciplines. The majority of research focuses on developing various temporal models that describe real-world data. Popular approaches include dynamic link modeling, activity-driven modeling, link-node memory modeling, and community dynamics modeling (Holme, 2013; Perra et al., 2012; Vestergaard et al., 2014; Peixoto and Rosvall, 2017). These models provide key insights into the mechanisms of network dynamics, that is, how network ties form, dissolve, and evolve. A related line of research, relational event models (REMs),

focuses on the intensity and sequencing of relational events (Butts, 2008; Butts et al., 2023). While REMs are useful for studying specific event-based processes and the timing of interactions, they do not directly capture tie persistence or larger-scale structural change. Stochastic actor-oriented models (SAOMs) provide another influential framework, modeling micro-mechanisms of tie formation and dissolution, but their estimation relies on extensive simulations and repeated full-network observations (Snijders, 2001; Snijders et al., 2010). Dynamic random graph approaches, such as Zhang et al., propose generative rules where edges form and dissolve at constant rates and provide estimators for those rates (Zhang et al., 2017). Together, these approaches illustrate the progress of temporal network modeling, but a common limitation is that their theoretical guarantees may fail when complete sequences of network snapshots are unavailable, a situation often encountered in empirical applications where networks are only partially observed or aggregated over time.

Our proposed TCM addresses this limitation by allowing more freedom in how edges are formed and dissolved through the flexibility of CM degree distributions with an explicit mechanism for tie persistence. At the same time, it provides explicit, consistent estimators that clarify how much temporal network information is required to reliably estimate model parameters from data. Beyond its theoretical contribution, the TCM is also relevant to applications where tie persistence is essential, such as biological networks, human mobility and communication networks, and financial systems (Barabasi and Oltvai, 2004; Barbosa et al., 2018; Karsai et al., 2018; Barucca and Lillo, 2016).

We also present an application of our model to infectious disease modeling on interpersonal networks. When a network evolves slowly relative to a spreading process, the dynamics can be approximated using a static network. When a network evolves very rapidly relative to the spreading process, the dynamics can be approximated well by a time-averaged version of the network. However, when the network and spreading process evolve at comparable time scales, the interplay between the two becomes important (Holme and Saramäki, 2012; Holme, 2015; Holme and Saramäki, 2019; Jordan et al., 2020; Hosseinzadeh et al., 2022). Thus, regardless of whether one studies the spread of infectious disease, health behaviors, or information, it is essential that the model capture important temporal changes in the underlying network topology.

Recent technological advances have made high-fidelity individual-level proximity data easier to capture, thereby allowing us to study empirical time-varying networks more precisely. Wearable sensors, like RFID tags and Bluetooth-enabled devices, such as smartphones, allow researchers to collect accurate, granular data in a passive manner (Vanhems et al., 2013; Sapiezynski et al., 2019; Neto et al., 2021). These technologies also allow investigators to capture proximity information with high precision, measuring the exact time and duration that two individuals are in close contact, along with their approximate distance from one another. This high-fidelity data allows us to quantify exposure, making it ideal for studying the spread of contagion. However, this begs the question: how much temporal granularity is needed to accurately capture spreading processes over networks? Very granular temporal resolution poses its own challenges, as it may capture unimportant interactions and erode privacy-protecting measures. Coarser collection can help preserve privacy and be less burdensome to store and analyze. However, if the network information is too coarse, we may miss key temporal dynamics, thereby inhibiting our ability to accurately model spreading processes over the network.

The remainder of the paper is organized as follows. In Section 2, we present a generative model for temporal networks which employs an edge persistence rate, which can be a simple constant, generated from a distribution, or some function of the empirical data. We then establish consistent estimators for the model parameters. Under the proposed modeling framework, the temporal granularity required for a consistent estimator depends on the number of edges that persist up to  $k$  time steps from the original network, where  $k$  corresponds to the highest moment of the generating distribution we seek to estimate. For instance, if the generating distribution is Normal or Beta, then only  $k = 2$  is required, since the first and second moments alone suffice to characterize the

distribution. We find that constructing an estimator using as much temporal network information as possible does not necessarily result in a superior estimator. Section 3 presents a simulation that illustrates some of the key statistical properties of our proposed estimators. In Section 4, we derive some theoretical results for a spreading process on temporal configuration networks, with a specific focus on the reproduction number, a fundamental concept in infectious disease epidemiology. In Section 5, we demonstrate how to fit the proposed model to an empirical dataset in the context of an infectious disease spreading through student contact on a college campus. Finally, Section 6 summarizes the main contributions of the study and suggests avenues for future research.

## 2. Temporal network model

### 2.1 Model specification

The configuration model (CM) is a widely used network model as it balances both realism and simplicity (Newman, 2018; Holme and Saramäki, 2012; Fosdick et al., 2018). Unlike many other network models, the CM incorporates arbitrary degree distributions. The exact degree of each individual node is fixed, which in turn fixes the total number of edges. To construct a network realization from a CM, one first specifies the total number of nodes in the network, denoted  $N$ . Then, for a given node  $i$ , one specifies the degree of that node,  $k_i$ , repeating this process for all nodes in the network. Each node is then given “stubs” equal to its degree, where a stub is simply an edge that is only connected to the node in question with its remaining end free. Pairs of stubs are then selected uniformly at random and connected to form edges until no stubs remain. The CM has several attractive properties. First, each possible matching of stubs is generated with equal probability. Specifically, the probability of nodes  $i$  and  $j$  being connected is given by  $\frac{k_i k_j}{2m-1}$ , where  $m$  is the total number of edges. Because any stub is equally likely to be connected to any other, many of the properties of the CM can be solved exactly. While the CM does allow for multi-edges and self-edges, their probability tends to zero as  $N$  approaches infinity. The CM is recognized as “one of the most important theoretical models in the study of networks” and many consider it one of the canonical network models (Newman, 2018). When studying a new question or process, the CM is often the first model network scientists employ, making it an ideal foundation for our temporal network model (Newman, 2018; Fosdick et al., 2018).

We present a time-varying version of the CM that we call the TCM. Like the CM, the TCM allows for an arbitrary degree distribution. Additionally, it encodes persistent ties, a key motivation for using and studying network models, by assigning each dyad a latent persistence probability.

To construct a TCM, one creates an initial network,  $G_0$ , via the standard CM algorithm. At each subsequent discrete time point  $t$ , a new network,  $G_t$ , is generated as follows. For each edge in  $G_{t-1}$ , a Bernoulli trial is conducted with success probability equal to that edge’s latent persistence probability. That is, for an edge between nodes  $i$  and  $j$ , we perform a Bernoulli trial with success probability  $p_{ij}$ . If the trial results in a success, the edge remains. If the trial is unsuccessful, the edge is broken, creating two stubs. After Bernoulli trials have been completed for all edges in  $G_{t-1}$ , stubs are matched uniformly at random with one another, forming edges in the new network,  $G_t$ . This process is repeated for each time step until one obtains the desired sequence of graphs,  $G_0, G_1, \dots, G_T$ .

The TCM can take on several intuitive forms by adjusting the latent edge-wise persistence probability. We begin by considering the two extremes. First, we can recover the standard CM simply by setting  $p_{ij} = 1$  for all  $i, j$ . This creates a static network such that  $G_t = G_0$  for all times  $t$ . Second, if we set  $p_{ij} = 0$  for all  $i, j$ , we simply generate independent and identically distributed random draws of graphs from a CM ensemble. At each  $t$ , a CM with the specified number of nodes and degree sequence will be generated, but this instance will be independent of the previous network realizations.

A more interesting scenario is obtained by setting  $p_{ij}$  equal to some fixed  $p \in (0, 1)$  for all pairs  $i, j$ . That is, we specify a single or homogeneous persistence probability for all edges in the network. This parameter dictates how quickly the network changes over time. If  $p$  is large, the network will change slowly as edges are highly likely to persist forward in time. If  $p$  is small, the network will change rapidly.

If we wanted the persistence probability to vary across the network, creating a more heterogeneous population, we can draw edge-level probabilities from some distribution for all node pairs. We could also imagine circumstances where it may make sense to encode some functional form for  $p_{ij}$ . For instance, edge-level persistence probabilities could be a function of node-level attributes if such information is available. We could envision a scenario where some individuals are more likely to stay within a particular social group while others may frequently move between groups, be it due to age, location, or some other factor or combination of factors. In this setting, we might construct  $p_{ij}$  to be some function of node  $i$ 's attributes,  $a_i$ , and node  $j$ 's attributes,  $a_j$ . As such, every edge would have a distinct probability that would relate back to the two individuals in question. We could take this one step further and have edge-level persistence probabilities be a function of the relationship between the two nodes. For example, edges between close friends may have a higher persistence probability than those between acquaintances.

## 2.2 Inference

We focus on three of the intuitive forms outlined above, namely a single persistence probability for the entire network, edge-level persistence parameters drawn from a probability distribution, and edge-level persistence parameters that are product of node-level persistence parameters drawn from a probability distribution. In the first model, the entire network has a single persistence probability,  $p$ . We then consider the case where the persistence rate  $p_{ij}$  of edge  $(i, j)$  is generated from a given distribution  $W$ , that is,  $p_{ij} \sim W$ , for all  $i, j$ . Finally, we consider the case where  $p_{ij} = p_i p_j$ , where  $p_i \sim W$ , for  $i = 1, \dots, N$  are independently drawn from a given distribution  $W$ . Notice that the first model is a special case of the second model as the distribution  $W$  shrinks to a constant  $p$ .

We observe that the future status of each edge after one step follows a Bernoulli distribution with mean equal to the persistence rate of that edge, and the status of each edge after two time steps follows a Bernoulli distribution with mean equal to the square of the persistence rate. The number of edges remaining after one or two time steps in the network is the sum of independent Bernoulli distributions. Thus, to estimate the first moment of the distribution, we can simply use the proportion of edges persisted one time step. To estimate the second moment of the distribution, we can use the proportion of edges that persisted two time steps. Next, we show that these estimators are consistent for each scenario.

**Remark:** As long as  $W$  is uniquely determined from its first  $k$  moments, the proposed approach can specify the distribution of  $W$ .

**Model 1:** All edges of the network have a single persistence probability  $p$  for some  $0 \leq p \leq 1$ .

Let  $X_0 = x_0$  denote the number of edges in the initial graph,  $G_0$ . Notice that  $x_0$  is a fixed constant for a given initial graph  $G_0$ . Let  $X_1$  denote the number of edges that persist to time  $t = 1$ . That is,  $X_1 = \sum_{(i,j) \in G_0} \text{Bernoulli}(p)$ . We can then use the ratio  $Z_1 = \frac{X_1}{X_0}$  to estimate the persistence probability  $p$ . Lemma 2.1 gives the properties of estimator  $Z_1$ ; the proof can be found in Appendix A.

**Lemma 2.1.**  $\frac{Z_1 - p}{\sqrt{(p(1-p))/X_0}}$  converges to a standard normal distribution, as  $X_0 \rightarrow \infty$ .

Since the network continues to evolve over time, we can incorporate this information to refine our estimate of  $p$ . Following the same logic as in Lemma 2.1, we can show that the ratio of edges remaining after each time step is also an unbiased and consistent estimator of  $p$ . To obtain a more

precise estimator for the persistence probability  $p$ , we can use the average of ratios over all time steps, denoted  $\bar{Z}$ .

Before discussing the proposed estimator  $\bar{Z}$ , we first walk through the evolution of the temporal network and set up the necessary notation. During the first time step, the broken edges of the original network  $G_0$  form stubs and are rematched to create new edges. Let  $Y_1$  denote the number of newly formed edges (excluding self-loops and multi-edges) at  $t = 1$ . For large networks, the number of self-loops and multi-edges is negligible relative to the total number of edges; thus,  $Y_1 = X_0 - X_1 + o(1)$ , where  $o(1)$  captures the number of self-loops and multi-edges. Let  $X_1^+$  denote the total number of edges at  $t = 1$  after adding the new edges to the remaining edges. That is,  $X_1^+ = X_1 + Y_1 = X_0 + o(1)$ . Then, at  $t = 2$ , the number of edges that persist is  $X_2 \sim \text{Binomial}(X_1^+, p)$  and the total number of edges after the rematching is  $X_2^+ = X_1^+ + o(1)$ . Similarly, for time  $t = T$ ,  $X_T$  is the total number of edges persisting from  $X_{T-1}^+$  and  $X_T^+$  is the number of edges after the rematching.

Let  $Z_t$  denote the fraction of edges that persist to time  $t$ , that is,  $Z_t = \frac{X_t}{X_{t-1}^+}$ , for  $t = 1, \dots, T$ .

Then the proposed estimator  $\bar{Z}$  is defined as  $\bar{Z} = \frac{1}{T} (Z_1 + Z_2 + \dots + Z_T)$ . Next, we present some important properties of the proposed estimator  $\bar{Z}$ , with a proof provided in Appendix B.

**Theorem 2.2.** *The estimator  $\bar{Z}$  is an unbiased and consistent estimator of the persistence probability  $p$ . In addition, when  $T = o(X_0)$  and  $T \rightarrow \infty$ ,  $\bar{Z}$  converges to  $p$  at the rate of  $O_p\left(\sqrt{\frac{1}{X_0}}\left(\frac{1}{T}\right)^{1/2-\delta}\right)$ , for some small  $\delta > 0$ .*

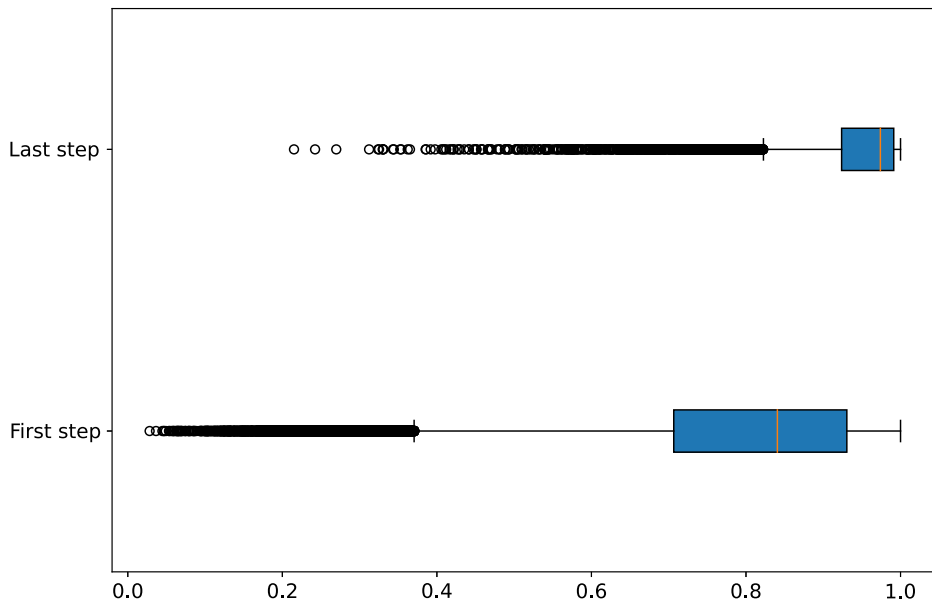
**Remark:** From Theorem 2.1 and Lemma 2.1, we see that by incorporating information about network evolution over time, the estimator  $\bar{Z}$  converges to the persistence probability  $p$  faster than  $Z_1$  by a factor of  $(\frac{1}{T})^{1/2-\delta}$  for some small positive  $\delta$ .

**Model 2:** Now, consider drawing  $p_{ij}$  from a given distribution  $W$ , that is,  $p_{ij} \sim W$  for all  $i, j$ . The simplest way to structure this model is to set the persistence probability for edge  $(i, j)$  as  $p_{ij}$ , fixing it over time. Alternatively, one could fix the persistence probability of an edge over a time window of  $T_0$  time steps for some  $T_0 > 1$ . Under this model, edge persistence probabilities are regenerated from  $W$  anew for each time window. This latter approach may be more realistic in some settings.

We will show that in either case the ratio of edges remaining after the first time step from the original network  $G_0$  and the ratio of edges remaining after the first two steps from the original network  $G_0$  can serve as good estimators for the first and second moment of the distribution  $W$ , respectively. For simplicity, we only provide the proofs for the estimator of the first moment; the proof for the estimator of the second moment is similar. Recycling the notation from Model 1, we will use  $Z_1 = \frac{X_1}{X_0}$  to estimate the first moment  $E(W)$ . Lemma 2.3 gives us the convergence property of the estimator  $Z_1$ ; the proof is provided in Appendix C.

**Lemma 2.3.**  *$Z_1$  is an unbiased and consistent estimator of the first moment  $E(W)$  of distribution  $W$ , as  $X_0 \rightarrow \infty$ .*

Next, we consider how to utilize temporal network information to refine the estimator  $Z_1$ . Consider the case where we generate  $p_{ij}$  from distribution  $W$  and assign fixed persistence probabilities  $p_{ij}$  for the  $(i, j)$  dyad. In this case,  $\bar{Z}$ , constructed under Model 1, will not be a good estimator for the first moment  $E(W)$  of  $W$ . As the network evolves, edges with higher persistence probabilities are more likely to be retained in the network. Therefore, over time, the persistence probabilities of edges remaining in the network will represent a shifted version of the original distribution  $W$ . As a result, using the estimator  $\bar{Z}$  will result in a biased estimate of  $E(W)$ . Figure 1 shows this phenomenon, where the initial network has 1000 nodes and the edge persistence probabilities are generated from  $W = \text{Beta}(4, 1)$  for the first time steps. After 100 time steps, edge persistence probabilities of edges retained in the network are shifted compared with the original distribution significantly.



**Figure 1.** Box plots of edge persistence probabilities at the first step and the last step (step 100) for persistence edges of a network of size 1000 over 100 time steps. Edge persistent probabilities are generated from a Beta(4, 1) distribution.

Next, consider the variant of Model 2 where the edge persistence probabilities are generated anew from  $W$  for each time window  $T_0$  and kept fixed for each edge throughout each time window. More specifically, starting with  $X_0$  edges from the set of edges  $A_0$  in the initial graph  $G_0$  at time  $t = 0$ , we generate persistence probabilities  $p_{ij}^{(0)}$ ,  $i, j = 1, \dots, N$  from a distribution  $W$  and assign the probability  $p_{ij}^{(0)}$  to the edge  $(i, j)$  if the edge appears at time  $t = 0$  and persists to time  $T_0 - 1$ . At time step  $T_0$ , persistence probabilities  $p_{ij}^{(1)}$ ,  $i, j = 1, \dots, N$  are generated anew from  $W$  and assigned to the edge  $(i, j)$  if the edge appears at time  $t = T_0$  and persists to time  $2T_0 - 1$ . This process continues until we reach the desired number of time steps  $T$ . If additional information is available regarding the time window  $T_0$ , we can use information from the first  $k$  time steps ( $k < T_0$ ) from each window to obtain a better estimator for the first  $k$  moments of distribution  $W$ .

Denote the set of edges of  $A_0$  that persist to  $t = 1$  as  $A_1$ . After the rematching process at  $t = 1$ , the set of edges now becomes  $A_1^+$ . Let  $X_1$  denote the number of edges in  $A_1$  and  $X_1^+$  the number of edges in  $A_1^+$  and so on for  $t = 2, \dots, T$ . Let  $\mathcal{F}_k^+$  denote the  $\sigma$ -algebra generated by all sets of edges up to time  $k$ , that is,  $\mathcal{F}_k^+ = \sigma\{A_1^+, \dots, A_k^+\}$ . Finally, let  $Z_k$  represent the fraction of edges remaining after one time step starting from each time window, that is,  $Z_1 = \frac{X_1}{X_0}$ ,  $Z_2 = \frac{X_{T_0+1}}{X_{T_0}^+}$ ,  $Z_3 = \frac{X_{2T_0+1}}{X_{2T_0}^+}$ ,  $\dots$ . Then we can use  $\bar{Z} = \sum_{k=1}^m Z_k/T$  to estimate the first moment  $E(W)$ , where  $m = \lceil T/T_0 \rceil$ .

**Theorem 2.4.** *If the periodic time interval  $T_0$  is correctly specified, the proposed estimator  $\bar{Z}$  is an unbiased and consistent estimator for  $E(W)$ . In addition, when  $T = o(X_0)$  and  $T \rightarrow \infty$ ,  $\bar{Z}$  converges to  $E(W)$  at the rate of  $O_p\left(\sqrt{\frac{1}{X_0}}\left(\frac{1}{m}\right)^{1/2-\delta}\right)$  for some small  $\delta > 0$ ,  $m = \lceil T/T_0 \rceil$ .*

The proof is provided in Appendix D.

**Model 3:** In Model 3, we consider the case where  $p_{ij} = p_i p_j$ , where  $p_i$ ,  $i = 1, \dots, N$  are independently drawn from a distribution  $W$ . As with Model 2, we consider two model variants. Since the persistence probability  $p_{ij} = p_i p_j$ ,  $E(p_{ij}) = E(p_i p_j) = E(p_i)E(p_j) = E(W)^2$  and  $E(p_{ij}^2) = E(p_i^2 p_j^2) = E(p_i^2)E(p_j^2) = E(W^2)^2$ . Therefore, we can use the same estimation strategy as above to estimate



moments of the distribution  $W$ . The estimator  $\bar{Z}$  will only be beneficial for the estimation process if we can correctly specify the time window  $T_0$ . Using the same arguments as in Lemma 2.3 and Theorem 2.4, we obtain the following:

**Lemma 2.5.** *The proposed estimator  $Z_1$  is an unbiased and consistent estimator for  $E(W)^2$ .*

**Theorem 2.6.** *If the periodic time interval  $T_0$  is correctly specified, the proposed estimator  $\bar{Z}$  is an unbiased and consistent estimator for  $E(W)^2$ . In addition, when  $T = o(X_0)$  and  $T \rightarrow \infty$ ,  $\bar{Z}$  converges to  $E(W)^2$  at the rate of  $O_p\left(\sqrt{\frac{1}{N}}\left(\frac{1}{m}\right)^{1/2-\delta}\right)$  for some small  $\delta > 0$ ,  $m = \lceil T/T_0 \rceil$ .*

Proofs are provided in Appendices E and F.

### 3. Simulation studies of proposed estimators

In this section, we perform some numerical simulations to illustrate the properties of proposed estimators on finite samples. We consider three different sets of simulation studies for each of the three models. The first set of simulations corresponds Model 1, where all edge persistence probabilities are fixed at  $p = 0.8$ . The second set of simulations examines Model 2, where edge persistence probabilities are generated from a Beta(1, 4) distribution, kept fixed for a time window of  $T_0 = 2$ , and then resampled at the start of a new window. More specifically, persistence probabilities for any given edge  $(i, j)$  of the network at any two consecutive time steps  $2(k-1)$  and  $2(k-1) + 1$  are the same and drawn from Beta(1, 4), for  $k = 1, \dots, \lceil T/2 \rceil$ . Finally, we consider Model 3, where the persistence probability  $p_{ij}$  of a given edge  $(i, j)$  is the product of  $p_i$  and  $p_j$ , where node-level persistence is drawn from Beta(1, 4), kept fixed for any two consecutive time steps  $2(k-1)$  and  $2(k-1) + 1$  and resampled at the start of a new window. To understand the effect of network size and number of time steps, we consider three different network sizes  $N = 10, 100$ , and  $1000$ . For each  $N$ , we allow the network to evolve through  $T = 30$  and  $T = 100$  time steps. The original graph  $G_0$  is first generated via the standard CM, where the degree sequence is generated from a Poisson distribution with mean 6. The network then evolves through  $T$  time steps with the persistence probabilities corresponding to each of the settings outlined above.

We use the proposed estimators  $Z_1$  and  $\bar{Z}$  to estimate the persistence probability  $p$  in the first set of simulations and the first moment of the underlying distribution in the second and third simulations. To estimate the second moment of the underlying distribution in the last two simulations, we use estimators  $V_1$  and  $\bar{V}$ , where  $V_1$  denotes the proportion of edges remaining after the first two time steps and  $\bar{V}$  utilizes the average of  $V_k$  when the time window  $T_0$  is specified correctly. To evaluate the proposed estimators' accuracy, we compute the absolute relative bias and the standard deviation of each estimator based on 100 replications. The absolute relative bias of estimators  $Z_1$  and  $\bar{Z}$  are defined as  $\frac{1}{100} \sum_{k=1}^{100} \left| \frac{Z_1^{(i)} - p}{p} \right|$  and  $\frac{1}{100} \sum_{k=1}^{100} \left| \frac{\bar{Z}^{(i)} - p}{p} \right|$ , respectively, for the first set of simulations where,  $Z_1^{(i)}$  and  $\bar{Z}^{(i)}$  correspond to the  $i$ th replication. For the second and third simulation, the absolute relative biases are  $\frac{1}{200} \left( \sum_{k=1}^{100} \left| \frac{Z_1 - E(W)}{E(W)} \right| + \sum_{k=1}^{100} \left| \frac{V_1 - E(W^2)}{E(W^2)} \right| \right)$  and  $\frac{1}{200} \left( \sum_{k=1}^{100} \left| \frac{\bar{Z} - E(W)}{E(W)} \right| + \sum_{k=1}^{100} \left| \frac{\bar{V} - E(W^2)}{E(W^2)} \right| \right)$ . We denote absolute relative bias and its standard deviation as AbsRelBias and SdAbsRelBias, respectively.

Table 1 shows that both  $Z_1$  and  $\bar{Z}$  are good estimators of  $p$  with the estimator  $\bar{Z}$  outperforming the estimator  $Z_1$ . The estimator  $Z_1$  seems to be a reasonable estimator for  $p$  when the network is greater than or equal to  $N = 100$ , but does not perform well when the network size is 10.  $\bar{Z}$ , on

**Table 1.** Absolute relative bias and standard deviations of the proposed estimators  $Z_1$  and  $\bar{Z}$  for Model 1 when edge persistence probability is fixed at  $p = 0.8$

<i>N</i>	<i>T</i>	Estimator $\bar{Z}$		Estimator $Z_1$	
		AbsRelBias	SdAbsRelBias	AbsRelBias	SdAbsRelBias
10	30	0.0022	0.0297	0.0097	0.1052
	100	0.0017	0.0205	0.0023	0.1139
100	30	0.0007	0.0056	0.0002	0.0323
	100	0.0005	0.0037	0.0035	0.0257
1000	30	0.0000	0.0017	0.0007	0.0093
	100	0.0001	0.0008	0.0006	0.0095

**Table 2.** Absolute relative bias and standard deviations of the proposed estimators  $Z_1$  and  $\bar{Z}$  for Model 2 when edge persistence probabilities  $p_{ij}$  are drawn from Beta(1, 4)

<i>N</i>	<i>T</i>	Estimator $\bar{Z}$		Estimator $Z_1$	
		AbsRelBias	SdAbsRelBias	AbsRelBias	SdAbsRelBias
10	30	0.0315	0.3649	0.1351	0.6581
	100	0.0333	0.1949	0.0382	0.5996
100	30	0.0047	0.0503	0.0193	0.1704
	100	0.0057	0.0347	0.0054	0.1653
1000	30	0.0020	0.0122	0.0013	0.0514
	100	0.0011	0.0070	0.0029	0.0481

the other hand, is a good estimator of  $p$  in all cases. These results suggest that  $\bar{Z}$  is a more reliable estimator for  $p$  when working with Model 1, where all edges have a fixed probability  $p$ .

Tables 2 and 3 also demonstrate the consistency of the proposed estimators for large networks. As expected, with more information in hand,  $\bar{Z}$  also outperforms to the estimator  $Z_1$  as the number of time steps  $T$  increases. For a small network size of  $N = 10$ ,  $Z_1$  does a poor job while the estimator  $\bar{Z}$  improves as the number of time steps increases from  $T = 30$  to  $T = 100$ . As the network size increases, both estimators become more reliable and tend to concentrate at the underlying true value of the generating distribution.

4. Theoretical results from an infectious disease application

In this section, we demonstrate how the proposed model, despite its simplicity, can offer interesting insights, potentially opening new avenues for exploring theoretical properties in the study of spreading processes on temporal networks. While the TCM can be leveraged for any spreading process over a network, here we examine the spread of an infectious disease. With infectious diseases, the contagion propagates through the population via contact between an infected individual and a susceptible individual. Typically, infectious disease modeling is done via a fully mixed model, which assumes that every individual in the population interacts with every other individual with equal probability or rate. However, in the real world, individuals typically have a small circle of contacts with whom they interact repeatedly. This was emphasized during the COVID-19 pandemic when many individuals practiced social distancing, choosing to interact with only a small



**Table 3.** Absolute relative bias and standard deviations of the proposed estimators  $Z_1$  and  $\bar{Z}$  for Model 3 when edge persistence probabilities are modeled as  $p_{ij} = p_i p_j$  with  $p_i$  drawn from Beta(1, 4), for  $i = 1, \dots, N$

$N$	$T$	Estimator $\bar{Z}$		Estimator $Z_1$	
		AbsRelBias	SdAbsRelBias	AbsRelBias	SdAbsRelBias
10	30	0.3208	0.5165	0.4897	0.8497
	100	0.1125	0.4485	0.4202	0.9954
100	30	0.0028	0.0987	0.0811	0.3657
	100	0.0134	0.0668	0.0849	0.3683
1000	30	0.0025	0.0250	0.0013	0.1029
	100	0.0020	0.0140	0.0190	0.1067

group of people dubbed their “social bubble” or “pod” (Hambridge et al., 2021; Wu et al., 2022). These repeated contacts, or persistent ties, between individuals are well represented by a network structure and often lead to isolation of the contagion within a particular part of the network. However, as the network topology changes, paths to new, uninfected portions of the network can open up. As such, it is critical to account for both persistent ties and changes in network structure over time. An approach that captures both of these phenomena, like our TCM, may offer a more realistic look at how diseases spread through populations.

To model disease spread, we employ compartmental models, a class of models that divides the population into different groups with respect to disease state. These models assign transition rules, allowing individuals to move between different states. Here, we use the susceptible-infectious-recovered (SIR) model, which assumes that individuals obtain perfect immunity once they recover from the disease. Under this model, individual nodes can be in any of three states: susceptible, infectious, and recovered. In the susceptible state, the node has not been infected but could become infected if it came into contact with an infectious node. That is, the node has no immunity to the disease. In the infected state, the node is infectious and can infect others it comes into contact with. Finally, in the recovered state, the node has recovered from the illness and can no longer infect others or be reinfected. With a stochastic model, nodes move through states probabilistically. Due to the discrete nature of our data, we utilize a discrete-time approach wherein events are defined by transition probabilities per unit time, as opposed to the transition rates used in a continuous time framework. A contact between an infectious and a susceptible node will result in a transmission with probability  $\beta$ . Similarly, at any given time step a node will recover with probability  $\gamma$ . For simplicity, we assume that at each time step, each edge has persistence probability  $p$ .

Let  $R_0$  denote the average number of transmissions from the initially infected node and  $R_*$  denote the average number of transmissions in the early stages of the epidemic excluding those from the initial infection. For both a static and dynamic network,  $R_0 = \tau \sum_k k p_k$ , where  $\tau$  is the transmission probability for a link between an infected node and a susceptible node and  $p_k$  is the probability a node has degree  $k$  in the CM. For a static network,  $R_* = \tau \sum_k q_k (k - 1)$ , where  $q_k$  is the excess degree distribution of a node with degree  $k$ , that is,  $q_k = k p_k / \sum_k k p_k$  (Volz and Meyers, 2009).

Before further discussion of the reproduction number, we recall an important tool in studying disease spread on a network: the probability-generating function (PGF). Suppose the PGF for the degree distribution of the initial CM network is  $g(x) = \sum_k p_k x^k$ , where  $p_k$  is the probability that a randomly chosen edge has degree  $k$ . The PGF of the excess degree distribution

is  $g_1(x) = \sum_k q_k x^{k-1}$ , where  $q_k = kp_k / \sum_k kp_k$ . Thus, we have the following relations:

$$g'(1) = \sum_k kp_k \quad (1)$$

$$g''(1) = \sum_k k(k-1)p_k \quad (2)$$

$$g'_1(1) = \sum_k (k-1)q_k = \sum_k (k-1)kp_k / \sum_k kp_k = g''(1)/g'(1). \quad (3)$$

We now examine the reproduction number at the early stage of the SIR spreading process on the TCM; a proof is provided in Appendix G.

**Proposition 4.1.** *The reproduction number at the early stage of a discrete-time SIR spreading process on the temporal network is given by  $R_* = \tau[(1-\gamma)(1-p)/\gamma + (1-p + \gamma p)g''(1)/(\gamma g'(1))]$ , where  $\tau = \frac{\beta}{1-p(1-\beta)(1-\gamma)}$ .*

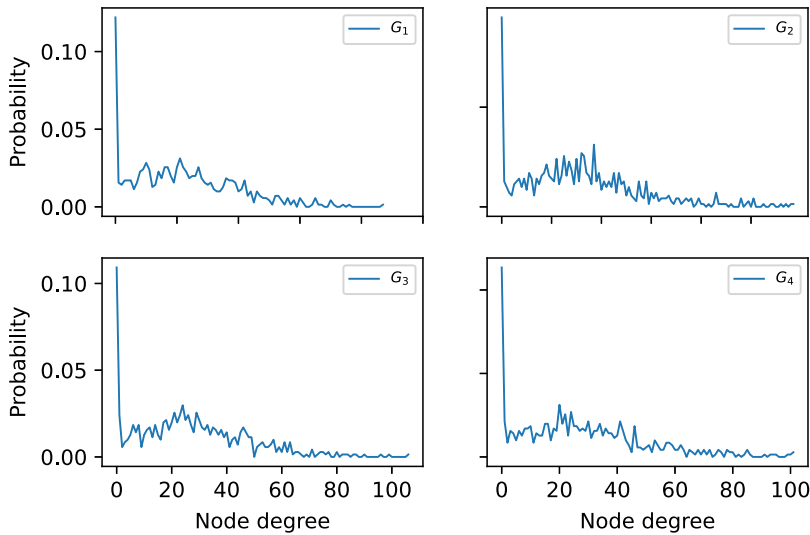
Next, we examine these results under specific degree distributions.

- When the degree distribution follows a Poisson distribution with mean  $\lambda$ ,  $p_k = \lambda^k \exp(-\lambda)/k!$ , its corresponding PGF is  $g(x) = \exp(\lambda(x-1))$ . Then  $g'(1) = \lambda$  and  $g''(1) = \lambda^2$ . Therefore,  $R_* = \tau[(1-\gamma)(1-p)/\gamma + (1-p + \gamma p)\lambda/\gamma]$ . For  $p = 1$ , this gives us  $R_* = \tau\lambda$ , which agrees with the results derived in Volz and Meyers (2009).
- When the degree distribution satisfies  $p_k = 1$ , if  $k = M$  for some  $1 \leq M \leq N$  where  $N$  is the network size, then the PGF is given by  $g(x) = x^M$ . So  $g''(1)/g'(1) = M-1$ , and therefore  $R_* = \tau[(1-\gamma)(1-p)/\gamma + (1-p + \gamma p)(M-1)/\gamma]$ . Notice that if  $k = N$ , we have a fully connected network. Under this scenario, the reproduction number is given by  $R_* = \tau[(1-\gamma)(1-p)/\gamma + (1-p + \gamma p)(N-1)/\gamma]$ . For  $p = 1$ , this reduces to  $R_* = \tau(N-1)$ .

Additionally, when the temporal network is static over time,  $p = 1$ . The reproduction number at the early stages is then  $R_* = \tau g''(1)/g'(1)$ , which agrees with the  $R_*$  value for a static network introduced above. When the temporal network evolves as independent draws from a CM, then  $p = 0$ . Under this scenario, the reproduction number at the early stage is  $R_* = \tau[(1-\gamma)/\gamma + g''(1)/(\gamma g'(1))] = \beta[(1-\gamma)/\gamma + g''(1)/(\gamma g'(1))]$ .

## 5. Fitting the model to empirical network data

In this section, we show how to use the three proposed TCMs to fit empirical network data and examine the fit of the generated temporal networks. The empirical data in this study was collected by the Copenhagen Network Study data and is publicly available (Sapiezynski et al., 2019). This data set contains information about the connectivity patterns of 706 students at the Technical University of Denmark over 28 days in February 2014. During the study period, participants agreed to use loaner cell phones from researchers as their primary phones. The proximity patterns were collected using Bluetooth where the approximate pairwise distances of phones were obtained using the received signal strength indicator (RSSI) every five minutes. We assigned a connection between two persons if there was at least one strong Bluetooth ping with  $\text{RSSI} \geq -75\text{dBm}$  (Hambridge et al., 2021). Because the empirical data reflects student proximity patterns, they fluctuate significantly on a daily basis. Students were more likely to connect during the week if they were in the same class, but on weekends, their proximity patterns were likely driven by their personal contact networks. Our models cannot capture the weekday-weekend variability because they require a comparable number of edges at each time instance. Therefore, we further processed the daily network data by combining each of the seven daily networks into a single weekly network. In particular, we established a weekly network by taking the union of the daily networks of



**Figure 2.** Degree distributions of weekly empirical networks  $G_1, G_2, G_3, G_4$ .

each weekly batch of 7 daily networks. As a result, the period of 28 days yielded 4 weekly networks, denoted as  $\{G_1, G_2, G_3, G_4\}$ . Figure 2 shows their degree distributions.

To fit the TCM models, we used  $G_1$  as the starting point and counted the number of edges in  $G_1$  that were retained in  $G_2$  and  $G_3$ . For Model 1, we calculated the fixed edge persistence rate as the ratio of the number of edges in  $G_1$  that persisted into  $G_2$ . For Models 2 and 3, we assumed that the edge-level and node-level persistence rates were generated from Beta distributions. The first and second moments of these Beta distributions were estimated by the ratios of the number of edges in  $G_1$  that persisted into  $G_2$  and  $G_3$ , respectively. Using the estimated first two moments, we determined the shape and scale parameters of the Beta distributions. The empirical weekly data gave us the estimate  $\hat{p} = 0.476$  for Model 1; the Beta distributions for Models 2 and 3 were estimated as  $\text{Beta}(0.975, 1.074)$  and  $\text{Beta}(1.171, 1.562)$ , respectively. With the estimated model parameters, we used  $G_1$  as the initial network and used the TCM models to obtain predicted networks  $\hat{G}_2^{(M_k)}, \hat{G}_3^{(M_k)}, \hat{G}_4^{(M_k)}$  for the empirical networks  $G_2, G_3, G_4$ , respectively. Here  $k = 1, 2, 3$  represents Models 1–3, respectively. As a benchmark, we employed the naive TCM, in which all edges were broken and rewired randomly at each time step. We refer to this CM as Model A. Furthermore, we also compare the proposed models with the dynamic random graphs with arbitrary expected degrees in Zhang et al. (2017). In this model, at each time step, new edges between each node pair  $i, j$  at is added with probability  $1 - e^{-\lambda_{ij}}$  and removing existing edges at with probability  $1 - e^{-\mu}$ , where  $\lambda_{ij} = \mu d_i d_j / 2m$ ,  $m = \frac{1}{2} \sum_i d_i$ ,  $d_i$  is degree of node  $i$  (Refer to Zhang et al., (2017) for more details). We denote this model as Model B.

To assess the performance of the TCM models, we compared the distances between the degree distributions of networks generated with Model A, Model B, Models 1–3 with the empirical networks  $G_2, G_3, G_4$ . We denote the distances between the model-based networks  $\hat{G}_2^{(M_k)}, \hat{G}_3^{(M_k)}, \hat{G}_4^{(M_k)}$  and the corresponding empirical networks  $G_2, G_3, G_4$  as  $D_2^{(M_k)}, D_3^{(M_k)}, D_4^{(M_k)}$ . We computed the average distances  $\bar{D}^{(M_k)} = \frac{1}{3} (D_2^{(M_k)} + D_3^{(M_k)} + D_4^{(M_k)})$  to assess the performance of the models. We calculated the means and standard deviations of  $\bar{D}^{(M_k)}$  from 100 simulation runs and used them as final goodness-of-fit metrics to assess the performance of the models. In particular, for run  $i$ , we calculated the average distances  $\bar{D}^{(M_k, i)}$ ; the final metrics for comparing the four models are the means and standard deviations of these metrics. The distances considered include the

**Table 4.** Assessment of model goodness-of-fit using means and standard deviations of total variation distance between degree distributions, Jaccard edge similarity, and absolute difference in average clustering coefficients over 100 runs

Model	Total variation distance	Jaccard edge similarity	Average clustering difference
Model A	0.2406± 0.0050	0.0359 ± 0.0007	0.2142 ± 0.0010
Model B	0.2079 ± 0.0073	0.0634 ± 0.0010	0.1860 ± 0.0013
Proposed model 1	0.2345 ± 0.0060	0.0829 ± 0.0012	0.2043 ± 0.0011
Proposed model 2	0.2248 ± 0.0051	0.1517 ± 0.0012	0.1576 ± 0.0014
Proposed model 3	0.2251 ± 0.0056	0.1521 ± 0.0033	0.1587 ± 0.0039

total variation distance,  $D_T = \frac{1}{2} \sum_{i=1}^k |p_i - q_i|$ ; the Jaccard edge similarity,  $D_J$  which is equal to the ratio of shared edges between  $G$  and  $\hat{G}$  to the total number of unique edges in  $G$  and  $\hat{G}$ ; and the absolute difference in average clustering coefficient,  $D_{cl} = |C_G - C_{\hat{G}}|$ . Here,  $P = (p_1, \dots, p_k)$  and  $Q = (q_1, \dots, q_k)$  denote the discrete degree distributions of the empirical network  $G$  and the model-based network  $\hat{G}$ , respectively, while  $C_G$  and  $C_{\hat{G}}$  are their corresponding average clustering coefficients. For total variation and average clustering difference, smaller distance values indicate that the model-based networks more closely resemble the empirical networks, suggesting a better model fit. In contrast, a higher Jaccard edge similarity value indicates a better fit.

Table 4 shows that Models 2, 3, and B yield the best fit to the empirical data. In general, the performance of Models 2 and 3 is comparable to that of Model B. While Model B outperforms Models 2 and 3 in terms of total variation distance, Models 2 and 3 achieve better performance than Model B with respect to the average clustering difference and Jaccard edge similarity. It is important to note that estimating Model B requires more detailed information than Models 2 and 3. Specifically, while Models 2 and 3 only require knowledge of the number of edges remaining after one and two time steps, Model B requires tracking edges status that emerge, dissolve, or remain unchanged during the first two time steps. Model 1 performed worse than Models 2 and 3, and Model A had the worst performance. The results demonstrate that model fit generally improves from Model 1 to Model 2 to Model 3. The improvement in model fit was to be expected because Model A uses data about the initial network only; Model 1 requires additional knowledge of the number of edges that remain in the initial network  $G_1$  after one time step, and Models 2 and 3 require knowledge of the number of edges that remain after one and two time steps.

To further evaluate the robustness of the proposed model, we repeated the procedure using student contact patterns from five weekdays only in the Copenhagen Network Study. The results, shown in Table 5, are consistent with those in Table 4, which was based on data from all seven days of the week. This consistency of the results suggests that the model is able to capture temporal contact patterns in social interaction networks.

6. Discussion and conclusion

We introduced the TCM, a family of generative models for a temporal network, as well as approaches for estimating model parameters. The proposed generative model is simple and flexible. The modeling framework allows the edge persistence probability to be fixed, generated from a distribution, or constructed from the corresponding node-level persistence probabilities. Our simulation studies showed that our estimators have good statistical properties. We then presented a theoretical application of the proposed model to the spread of an infectious disease and demonstrated how to fit the model to empirical data from the Copenhagen Network Study.

**Table 5.** Assessment of model goodness-of-fit using means and standard deviations of total variation distance between degree distributions, Jaccard edge similarity, and absolute difference in average clustering coefficients over 100 runs for five weekday networks

Model	Total variation distance	Jaccard edge similarity	Average clustering difference
Model A	$0.2429 \pm 0.0053$	$0.0338 \pm 0.0007$	$0.2203 \pm 0.0009$
Model B	$0.2063 \pm 0.0081$	$0.0606 \pm 0.0009$	$0.1927 \pm 0.0011$
Proposed model 1	$0.2383 \pm 0.0053$	$0.0812 \pm 0.0013$	$0.2102 \pm 0.0012$
Proposed model 2	$0.2228 \pm 0.0052$	$0.1549 \pm 0.0011$	$0.1593 \pm 0.0015$
Proposed model 3	$0.2235 \pm 0.0052$	$0.1554 \pm 0.0038$	$0.1603 \pm 0.0040$

We proposed consistent estimators for the model parameters and provided the convergence rate of the proposed estimators. We additionally showed that, when the persistence probability is constant across edges, using the network evolution process can give us a more precise estimator with a faster convergence rate. However, when persistence probabilities are generated at random or are the product of the node-level persistence probabilities, care must be taken when using an estimator that relies on averages over the entire course of the network's evolution. Instead, it might be advantageous to consider alternative modeling strategies, such as those in which persistence probabilities are periodically updated. If additional information on this time interval is available, then a design that incorporates the network evolution process will give us a finer estimator with an accelerated convergence rate. This finding addresses a key question regarding the level of temporal resolution needed for reliable estimation in dynamic network settings: the resolution depends strongly on the nature of the underlying temporal process. For the proposed class of temporal networks, the only data required for consistent estimation are the counts of edges that persist up to  $k$  time steps from the original network, where  $k$  corresponds to the highest-order moment of the generating distribution. This insight reduces the effort involved in data collection and storage for temporal networks governed by the TCM.

Additionally, we highlighted the merits of TCM through both a theoretical application to the study of spreading processes on temporal networks and an empirical data analysis. In the theoretical setting, we examined an SIR spreading model on a temporal network under the assumption of constant edge persistence probability and derived an explicit formula for the reproduction number during the early stage of the epidemic. In the empirical analysis, TCM demonstrated good fit with real-world temporal network data.

Further research directions might involve examining a good estimator for the periodic information in Models 2 and 3, extending the temporal model to include network size growth information, and investigating various epidemic characteristics of spreading processes on the temporal network of the three models. All of these potential developments, as well as the broad range of applications for temporal networks, offer an exciting future for the study of time-varying networks (Holme, 2015; Jordan et al., 2020; Hosseinzadeh et al., 2022). With technological advances, particularly Bluetooth technology, network data is becoming more accessible than ever. How we can appropriately employ these new datasets for public health insights is an interesting subject that requires further research. We believe that statistical techniques with robust theoretical underpinnings are the foundation for leveraging the vast amounts of data available for the common good.

An important contribution of the TCM is that it provides a principled solution to the problem of temporal granularity in network analysis. Prior work has shown that the resolution at which interactions are observed can substantially alter conclusions about spreading dynamics, robustness, and network structure (Holme and Saramäki, 2012; Holme, 2015). If data are aggregated too coarsely, essential ordering and persistence effects are lost; if collected at overly fine

resolution, data can be noisy, costly, and difficult to obtain (Vanhems et al., 2013; Sapiezynski et al., 2019). Despite this recognition, a rigorous statistical characterization of how much temporal information is sufficient has been lacking. Our results show that for the class of TCMs, consistent inference requires only edge persistence counts up to a finite horizon of length  $k$ , thereby providing a tractable statistical solution to this problem for this model class.

Our modeling framework is relevant in application domains where persistence and temporal resolution are essential. For example, within network biology, distinguishing stable from transient protein interactions (Barabasi and Oltvai, 2004) can be framed through persistence-based inference. In human mobility, separating recurrent commuting flows from irregular trips (Barbosa et al., 2018) can benefit from a principled treatment of temporal resolution. In studies of bursty human dynamics, persistence parameters can provide a way to capture stability amid temporal clustering (Karsai et al., 2018). In financial networks, where aggregation can obscure underlying creditor–debtor structures (Barucca and Lillo, 2016), the TCM can help track the evolving backbone of interbank relationships. The framework can also be applied to study changes in resource access in interpersonal networks, assessing growth in interorganizational partnerships, comparing persistence across communities, and analyzing the interplay between relational states and relational events. By formalizing persistence within a statistical framework, the TCM provides theoretical clarity and practical insights for temporal network research.

**Funding statement.** H.H. was supported by a Harvard University Department of Biostatistics scholarship and a U.S. Government scholarship. T.M.L. and J.P.O. were supported by a National Institutes of Health award (NIAID R01 AI138901). The funding sources had no role in study design, data analysis, data interpretation, or the writing of the paper.

**Competing interests.** None.

**Data availability statement.** The empirical network data is publicly available at The Copenhagen Networks Study interaction data Sapiezynski et al. (2019). Python code used in this study is publicly available at <https://github.com/onnella-lab/temporal-configuration-model>.

## References

- Barabasi, A.-L., & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*, 5(2), 101–113.
- Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T. . . . Tomasini, M. (2018). Human mobility: models and applications. *Physics Reports*, 734, 1–74.
- Barucca, P., & Lillo, F. (2016). Disentangling bipartite and core-periphery structure in financial networks. *Chaos, Solitons & Fractals*, 88, 244–253.
- Borgatti, S. P., & Foster, P. C. (2003). The network paradigm in organizational research: a review and typology. *Journal of Management*, 29(6), 991–1013.
- Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, 38(1), 155–200.
- Butts, C. T., Lomi, A., Snijders, T. A. B., & Stadtfeld, C. (2023). Relational event models in network science. *Network Science*, 11(2), 175–183.
- Chen, H., Mehra, A., Tasselli, S., & Borgatti, S. P. (2022). Network dynamics and organizations: a review and research agenda. *Journal of Management*, 48(6), 1602–1660.
- Dürre, A., & Paindaveine, D. (2023). On the consistency of incomplete u-statistics under infinite second-order moments. *Statistics & Probability Letters*, 193, 109714.
- Fosdick, B. K., Larremore, D. B., Nishimura, J., & Ugander, J. (2018). Configuring random graph models with fixed degree sequences. *Siam Review*, 60(2), 315–355.
- Hambridge, H. L., Kahn, R., & Onnela, J.-P. (2021). Examining sars-cov-2 interventions in residential colleges using an empirical network. *International Journal of Infectious Diseases*, 113, 325–330.
- Holme, P. (2013). Epidemiologically optimal static networks from temporal network data. *PLoS Computational Biology*, 9(7), e1003142.
- Holme, P. (2015). Modern temporal network theory: a colloquium. *The European Physical Journal B*, 88(9), 234.
- Holme, P., & Saramäki, J. (2012). Temporal networks. *Physics Reports*, 519(3), 97–125.
- Holme, P., & Saramäki, J. (2019). A map of approaches to temporal networks. In *Temporal network theory*, 1–24. Springer.



- Hosseinzadeh, M. M., Cannataro, M., Guzzi, P. H., & Dondi, R. (2022). Temporal networks in biology and medicine: a survey on models, algorithms, and tools. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 12(1), 10.
- Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: considerations as the paradigm shifts? *Journal of Clinical Psychology*, 76(9), 1591–1612.
- Karsai, M., Jo, H.-H., Kimmo, K. et al., (2018). *Bursty human dynamics*, vol. 1. Springer.
- Neto, O. L., Haenni, S., Phuka, J., Ozella, L., Paolotti, D., Cattuto, C., Robles, D., Lichand, G., et al. (2021). Combining wearable devices and mobile surveys to study child and youth development in malawi: implementation study of a multimodal approach. *JMIR Public Health and Surveillance*, 7(3), e23154.
- Newman, M. (2018). *Networks*. Oxford university press.
- Peixoto, T. P., & Rosvall, M. (2017). Modelling sequences and temporal networks with dynamic community structures. *Nature communications*, 8(1), 582.
- Perra, N., Gonçalves, B., Pastor-Satorras, R., & Vespignani, A. (2012). *Activity driven modeling of time varying networks*. Scientific Reports, 2(1), 469.
- Provan, K. G., Fish, A., & Sydow, J. (2007). Interorganizational networks at the network level: a review of the empirical literature on whole networks. *Journal of Management*, 33(3), 479–516.
- Sapiezynski, P., Stopczynski, A., Lassen, D. D., & Lehmann, S. (2019). Interaction data from the copenhagen networks study. *Scientific Data*, 6(1), 1–10.
- Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, 31(1), 361–395.
- Snijders, T. A. B., Van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), 44–60.
- Vanhems, P., Barrat, A., Cattuto, C., Pinton, J.-F., Khanafer, N., Régis, C., Kim, B., Comte, B., & Voirin, V. (2013). Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. *PLoS One*, 8(9), e73970.
- Vestergaard, C. L., Génois, M., & Barrat, A. (2014). How memory generates heterogeneous dynamics in temporal networks. *Physical Review*, 90(4), 042805.
- Volz, E., & Meyers, L. A. (2009). Epidemic thresholds in dynamic contact networks. *Journal of the Royal Society Interface*, 6(32), 233–241.
- Wu, M., Li, C., Shen, Z., He, S., Tang, L., Zheng, J., Fang, Y., Li, K., Cheng, Y., Shi, Z., et al. (2022). Use of temporal contact graphs to understand the evolution of covid-19 through contact tracing data. *Communications Physics*, 5(1), 270.
- Zhang, X., Moore, C., & Newman, M. E. J. (2017). Random graph models for dynamic networks. *The European Physical Journal B*, 90(10), 200.

## Appendix A. Proof of Lemma 2.1

**Proof.** We have  $Z_1 = X_1/X_0 = \sum_{(i,j) \in G_0} \text{Bernoulli}(p)/X_0$ . Since  $\sum_{(i,j) \in G_0} \text{Bernoulli}(p)$  is the sum of independent Bernoulli random variables, applying the central limit theorem we have  $\frac{Z_1 - p}{\sqrt{(p(1-p))/X_0}}$  converges to a standard normal distribution as  $X_0 \rightarrow \infty$ .  $\square$

Thus,  $Z_1$  is an unbiased and consistent estimator for  $p$ .

## Appendix B. Proof of Theorem 2.2

**Proof.** It follows from Lemma 2.1 that each estimator  $Z_k$  is an unbiased and consistent estimator for the persistence probability  $p$ , for all  $k = 1, \dots, T$ . Therefore,  $\bar{Z} = \frac{1}{T} \sum_{i=1}^T Z_i$  is also an unbiased and consistent estimator of  $p$ .

Denote  $S_T = \sum_{i=1}^T Z_i$  and  $\gamma_{st} = E[(Z_s - p)(Z_t - p)]$  for all  $s, t \in \{1, \dots, T\}$ . We now evaluate the variance of  $S_T$ .

$$\text{Var}(S_T) = E[(S_T - pT)^2] = E\left[\left(\sum_{i=1}^T Z_i - pT\right)^2\right] = \sum_{s=1}^T \sum_{t=1}^T E[(Z_s - p)(Z_t - p)] = \sum_{s=1}^T \sum_{t=1}^T \gamma_{st}.$$

For any  $s > t$ , we have

$$\begin{aligned}
 \gamma_{st} &= E[(Z_s - p)(Z_t - p)] = E(Z_s Z_t) - p E(Z_s) - p E(Z_t) + p^2 = E(Z_s Z_t) - p^2 - p^2 + p^2 \\
 &= E\left(\frac{X_s}{X_{s-1}^+} \frac{X_t}{X_{t-1}^+}\right) - p^2 = E\left[E\left(\frac{X_s}{X_{s-1}^+} \frac{X_t}{X_{t-1}^+} \middle| X_{s-1}^+\right)\right] - p^2 \\
 &= E\left[E\left(\frac{X_s}{X_{s-1}^+} \middle| X_{s-1}^+\right)\right] E\left(\frac{X_t}{X_{t-1}^+}\right) - p^2 \text{ by independence of } \frac{X_t}{X_{t-1}^+} \text{ and } E\left(\frac{X_s}{X_{s-1}^+} \middle| X_{s-1}^+\right) \\
 &= p E\left(\frac{X_t}{X_{t-1}^+}\right) - p^2 = p^2 - p^2 = 0.
 \end{aligned}$$

Thus,  $\gamma_{st} = 0$  for any  $s > t$ . Similarly,  $\gamma_{st} = 0$  for any  $s < t$ . Therefore,  $\text{Var}(S_T) = \sum_{s=1}^T \sum_{t=1}^T \gamma_{st} = \sum_{s=1}^T \gamma_{ss}$ .

For  $T = o(X_0)$ , we have  $X_{s-1}^+ = X_0 + o(1)$ , for all  $s \in 1, \dots, T$ . Therefore,

$$\begin{aligned}
 \gamma_{ss} &= E((Z_s - p)(Z_s - p)) = E(Z_s^2) - p^2 = E\left[\left(\frac{X_s}{X_{s-1}^+}\right)^2\right] - p^2 \\
 &= E\left[E\left(\left(\frac{X_s}{X_{s-1}^+}\right)^2 \middle| X_{s-1}^+\right)\right] - p^2 = E\left[\left(\frac{1}{X_{s-1}^+}\right)^2 E(X_s^2 | X_{s-1}^+)\right] - p^2 \\
 &= E\left[\left(\frac{1}{X_{s-1}^+}\right)^2 [X_{s-1}^+ p(1-p) + (p X_{s-1}^+)^2]\right] - p^2 \\
 &= E\left[\frac{p(1-p)}{X_{s-1}^+}\right] \leq \frac{C_1}{X_0}, \quad \text{for some } C_1 > 0.
 \end{aligned}$$

This gives us  $\text{Var}(S_T) = \sum_{s=1}^T \gamma_{ss} \leq T \frac{C_1}{X_0}$ . Applying the Chebyshev's, for some  $C_2 > 0$  and  $0 < \delta < 1/2$ , we have

$$\begin{aligned}
 P\left(|\bar{Z} - p| \geq C_2 \sqrt{\frac{1}{X_0}} \left(\frac{1}{T}\right)^{1/2-\delta}\right) &\leq \frac{E[(\bar{Z} - p)^2]}{C_2^2 \frac{1}{X_0} \left(\frac{1}{T}\right)^{1-2\delta}} = \frac{E[(S_T - T p)^2]}{T^2 C_2^2 \frac{1}{X_0} \left(\frac{1}{T}\right)^{1-2\delta}} \\
 &\leq \frac{T \frac{C_1}{X_0}}{T^2 C_2^2 \frac{1}{X_0} \left(\frac{1}{T}\right)^{1-2\delta}} = \frac{C_1}{C_2^2 T^{2\delta}} \rightarrow 0.
 \end{aligned}$$

Thus,  $\bar{Z}$  converges to  $p$  at the rate of  $O_p\left(\sqrt{\frac{1}{X_0}} \left(\frac{1}{T}\right)^{1/2-\delta}\right)$ . □

### Appendix C. Proof of Lemma 2.3

**Proof.** We first prove that, for any given combination of edge persistence probabilities drawn from  $W$ ,  $Z_1$  is an unbiased and consistent estimator for the average of the edge persistence probabilities. We have  $Z_1 = X_1/X_0 = \sum_{(i,j) \in G_0} \text{Bernoulli}(p_{ij})/X_0$ , where  $\text{Bernoulli}(p_{ij})$  are independent variables. For simplicity, let us re-enumerate the sequence of independent  $\text{Bernoulli}(p_{ij})$  for all  $(i, j) \in G_0$  and denote them as  $U_1, U_2, \dots, U_{X_0}$ , where  $U_i \sim \text{Bernoulli}(p_i)$  for  $i = 1, \dots, X_0$ . Then  $Z_1 = \frac{1}{X_0} \sum_{i=1}^{X_0} U_i$ . Denote  $\mu_{X_0} = \frac{1}{X_0} \sum_{i=1}^{X_0} p_i$  and  $\sigma_{X_0} = \frac{1}{X_0} \sqrt{\sum_{i=1}^{X_0} p_i(1-p_i)}$ .

Using the Lyapunov central limit theorem for independent non-identical distributions, we see that  $(Z_1 - \mu_{X_0})/\sigma_{X_0}$  will converge to a standard normal distribution if we can verify the Lyapunov condition (C1) below:

$$\lim_{X_0 \rightarrow \infty} \frac{1}{\sigma_{X_0}^{2+\delta}} \sum_{i=1}^{X_0} \left( \frac{1}{X_0} \right)^{2+\delta} E |U_i - p_i|^{2+\delta} \rightarrow 0, \text{ for some } \delta > 0. \quad (\text{C1})$$

Since  $0 \leq p_i \leq 1$  for all  $i$ ,  $\sum_{i=1}^{X_0} p_i(1-p_i)$  is of the same order as  $\sum_{i=1}^{X_0} p_i$ . Therefore,  $\sum_{i=1}^{X_0} p_i(1-p_i) \rightarrow \infty$  as  $X_0 \rightarrow \infty$ . In addition, we also have

$$\begin{aligned} E|U_i - p_i|^{2+\delta} &= p_i(1-p_i)^{2+\delta} + (1-p_i)p_i^{2+\delta} \\ &\leq p_i(1-p_i)^2 + (1-p_i)p_i^2 = p_i(1-p_i). \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{X_0 \rightarrow \infty} \frac{1}{\sigma_{X_0}^{2+\delta}} \sum_{i=1}^{X_0} \left( \frac{1}{X_0} \right)^{2+\delta} E |U_i - p_i|^{2+\delta} &= \lim_{X_0 \rightarrow \infty} \frac{1}{\left( \sum_{i=1}^{X_0} p_i(1-p_i) \right)^{1+\delta/2}} \sum_{i=1}^{X_0} E |U_i - p_i|^{2+\delta} \\ &\leq \lim_{X_0 \rightarrow \infty} \frac{\sum_{i=1}^{X_0} p_i(1-p_i)}{\left( \sum_{i=1}^{X_0} p_i(1-p_i) \right)^{1+\delta/2}} \\ &= \lim_{X_0 \rightarrow \infty} \frac{1}{\left( \sum_{i=1}^{X_0} p_i(1-p_i) \right)^{\delta/2}} \rightarrow 0. \end{aligned}$$

So,  $\frac{Z_1 - \mu_{X_0}}{\sigma_{X_0}}$  converges to a standard normal distribution, as  $X_0 \rightarrow \infty$ .

The central limit theorem for random observations generated from distribution  $W$  gives us  $\mu_{X_0} = E(W) + O_p\left(\sqrt{\frac{1}{X_0}}\right)$ . Furthermore, we also have that  $E(Z_1) = E\left(\frac{1}{X_0} \sum_{i=1}^{X_0} p_i\right) = E(W)$ . Therefore,  $Z_1$  is an unbiased and consistent estimator of the first moment  $E(W)$  of distribution  $W$  as  $X_0 \rightarrow \infty$ .  $\square$

## Appendix D. Proof of Theorem 2.4

**Proof.** It follows from Lemma 2.3 that each  $Z_k$  is an unbiased and consistent estimator for the first moment  $E(W)$ , for  $k = 1, \dots, m$ . Then  $E(\bar{Z}) = E\left(\sum_{k=1}^m Z_k/m\right) = E(W)$ . In other words,  $\bar{Z}$  is also an unbiased and consistent estimator for  $E(W)$ , as  $X_0, X_{T_0}, \dots, X_{mT_0} \rightarrow \infty$ .

To obtain the convergence rate of  $\bar{Z}$ , we first provide the upper bound for  $E\left(\sum_{s=1}^m Z_s - mE(W)\right)^2$ :

$$\begin{aligned} E\left(\sum_{s=1}^m Z_s - mE(W)\right)^2 &= \text{Var}\left(\sum_{s=1}^m Z_s - mE(W)\right) = \text{Var}\left(\sum_{s=1}^m Z_s\right) \\ &= \sum_{s=1}^m \text{Var}(Z_s) + 2 \sum_{s>t, s, t=1}^m \text{Cov}(Z_s, Z_t). \end{aligned}$$

For any  $s > t$ , we have

$$\begin{aligned}\text{Cov}(Z_s, Z_t) &= E \left( \frac{X_{(s-1)T_0+1}}{X_{(s-1)T_0}^+} \cdot \frac{X_{(t-1)T_0+1}}{X_{(t-1)T_0}^+} \right) - (E(W))^2 \\ &= E \left[ E \left( \frac{X_{(s-1)T_0+1}}{X_{(s-1)T_0}^+} \cdot \frac{X_{(t-1)T_0+1}}{X_{(t-1)T_0}^+} \middle| \mathcal{F}_{(s-1)T_0}^+ \right) \right] - (E(W))^2 \\ &= E \left[ E \left( \frac{X_{(s-1)T_0+1}}{X_{(s-1)T_0}^+} \middle| \mathcal{F}_{(s-1)T_0}^+ \right) \right] E \left( \frac{X_{(t-1)T_0+1}}{X_{(t-1)T_0}^+} \right) - (E(W))^2 \\ &= E(W)E(W) - (E(W))^2 = 0.\end{aligned}$$

In addition,

$$\begin{aligned}\text{Var}(Z_s) &= E \left[ \left( \frac{X_{(s-1)T_0+1}}{X_{(s-1)T_0}^+} \right)^2 \right] - (E(W))^2 = E \left[ E \left( \left( \frac{X_{(s-1)T_0+1}}{X_{(s-1)T_0}^+} \right)^2 \middle| \mathcal{F}_{(s-1)T_0}^+ \right) \right] - (E(W))^2 \\ &= E \left[ \left( \frac{1}{X_{(s-1)T_0}^+} \right)^2 E \left( X_{(s-1)T_0+1}^2 \middle| \mathcal{F}_{(s-1)T_0}^+ \right) \right] - (E(W))^2.\end{aligned}$$

Since  $X_{(s-1)T_0+1}$  is the summation of independent Bernoulli variables Bernoulli  $(p_{ij}^{(s)})$ , where  $(i, j) \in A_{(s-1)T_0}^+$ , we have

$$\begin{aligned}E \left( X_{(s-1)T_0+1}^2 \middle| \mathcal{F}_{(s-1)T_0}^+ \right) &= \text{Var} \left( X_{(s-1)T_0+1} \middle| \mathcal{F}_{(s-1)T_0}^+ \right) + \left[ E \left( X_{(s-1)T_0+1} \middle| \mathcal{F}_{(s-1)T_0}^+ \right) \right]^2 \\ &= \sum_{(i,j) \in A_{(s-1)T_0}^+} p_{ij}^{(s)} (1 - p_{ij}^{(s)}) + \left( \sum_{(i,j) \in A_{(s-1)T_0}^+} p_{ij}^{(s)} \right)^2.\end{aligned}$$

For  $T = o(X_0)$ , we have  $X_{(s-1)T_0}^+ = X_0 + o(1)$ , for all  $s = 1, \dots, m$ . Therefore,

$$\begin{aligned}\text{Var}(Z_s) &= E \left( \frac{\sum_{(i,j) \in A_{(s-1)T_0}^+} p_{ij}^{(s)} (1 - p_{ij}^{(s)})}{X_{(s-1)T_0}^{+2}} \right) + E \left( \frac{\left( \sum_{(i,j) \in A_{(s-1)T_0}^+} p_{ij}^{(s)} \right)^2}{X_{(s-1)T_0}^{+2}} \right) - (E(W))^2 \\ &\leq E \left( \frac{\sum_{(i,j) \in A_{(s-1)T_0}^+} p_{ij}^{(s)}}{X_{(s-1)T_0}^{+2}} \right) + E \left( \frac{\text{Var}(W)}{X_{(s-1)T_0}^+} \right) \leq \frac{C}{X_0}, \text{ for some } C > 0.\end{aligned}$$

This gives us  $E \left( \sum_{s=1}^m Z_s - m E(W) \right)^2 \leq \frac{mC}{X_0}$ . Using the same argument as in the proof of Theorem 2.2,  $\bar{Z}$  converges to  $E(W)$  at the rate of  $O_p \left( \sqrt{\frac{1}{X_0}} \left( \frac{1}{m} \right)^{1/2-\delta} \right)$ , for some small  $\delta > 0$ .  $\square$

## Appendix E. Proof of Lemma 2.5

**Proof.** As in Lemma 2.3, for any given combination of edge persistence probabilities generated from distribution  $W$ , the standardized version of  $Z_1$  converges to a standard normal distribution. Applying the Central Limit Theorem to the degenerate U-statistics (Theorem 2.1 in Dürre and Paindaveine (2023)), we have  $\mu_{X_0} = \frac{1}{X_0} \sum_{(i,j) \in G_0} p_i p_j = E(W)^2 + O_p\left(\sqrt{\frac{1}{N}}\right)$ . Furthermore, we also have that  $E(Z_1) = \frac{1}{X_0} \sum_{(i,j) \in G_0} p_i p_j = E(W)^2$ . Therefore,  $Z_1$  is an unbiased and consistent estimator for  $E(W)^2$ , as  $X_0 \rightarrow \infty$ .  $\square$

## Appendix F. Proof of Theorem 2.6

**Proof.** The proof follows by using the same arguments as in Theorem 2.4.  $\square$

## Appendix G. Proof of Proposition 4.1

**Proof.** Following the idea of Volz and Meyers (2009), we derive the reproduction number for the process through four steps as follows:

1. Find the probability of transmission  $\tau$  for each connection.
2. Find the PGF of the number of contacts  $M$  for a node with degree  $k$ .
3. Find the PGF  $\tilde{H}_1(x)$  of the number of contacts of the selected node proportional to the concurrent degree, but subtracting one transitory contact.
4. Find the PGF  $H_1(x)$  of the number of transmissions caused by the selected node proportional to the concurrent degree, but subtracting one transitory contact.

Then the reproduction number  $R_*$  at the early stage of the epidemic is the derivative of  $H_1(x)$  evaluated at  $x = 1$ .

We first find the probability of transmission  $\tau$  for each connection. Notice that for a given edge, the time  $X$  until the edge has broken follows a geometric distribution, where  $X \sim \text{Geo}(1-p)$ . Similarly, the time  $Y$  until an infected node transmits to a susceptible neighbor satisfies  $Y \sim \text{Geo}(\beta)$ , and the time  $Z$  until an infected node recovers satisfies  $Z \sim \text{Geo}(\gamma)$ . For an infected-susceptible connection, the overlap time  $U$  while the infected node is still infectious and the edge is still connected is  $U \sim \min(X, Z)$ . Notice that  $X$  and  $Z$  are independent, therefore  $U \sim \text{Geo}(1-p(1-\gamma))$  by the properties of a geometric distribution.

So, the probability that the infected node transmits the disease to the susceptible in at most  $k$  steps is  $P(Y \leq k-1) = 1 - (1-\beta)^k$ . Therefore, the transmission probability  $\tau$  during the overlap time is

$$\begin{aligned} \tau &= E\left(1 - (1-\beta)^U\right) = \sum_{k \geq 1} P(U=k) \left[1 - (1-\beta)^k\right] \\ &= \sum_{k \geq 1} \left[1 - (1-\beta)^k\right] [p(1-\gamma)]^{k-1} [1 - p(1-\gamma)] \\ &= [1 - p(1-\gamma)] \sum_{k \geq 1} [p(1-\gamma)]^{k-1} - [1 - p(1-\gamma)](1-\beta) \sum_{k \geq 1} [p(1-\beta)(1-\gamma)]^{k-1} \\ &= [1 - p(1-\gamma)] \frac{1}{1 - p(1-\gamma)} - [1 - p(1-\gamma)](1-\beta) \frac{1}{1 - p(1-\beta)(1-\gamma)} \\ &= 1 - \frac{[1 - p(1-\gamma)](1-\beta)}{1 - p(1-\beta)(1-\gamma)} = \frac{\beta}{1 - p(1-\beta)(1-\gamma)}. \end{aligned} \tag{G1}$$

We then find the PGF of number of contacts  $M$  for an infected node with degree  $k$  during time period  $\ell$ . For simplicity, we first construct the PGF for  $M$  corresponding to nodes with degree 1. For a node of degree 1, the number of edges swapped during  $\ell$  time steps follows  $\text{Binomial}(\ell, (1-p))$ . Therefore, the probability that the number of contacts for a node of degree 1 is equal to  $m$  during the period of  $\ell$  steps is  $\binom{\ell}{m-1}(1-p)^{m-1}p^{\ell-m+1}$ . Thus, for a node with degree 1, the PGF of the number of contacts  $M$  is  $\sum_{m \geq 1} \binom{\ell}{m-1}(1-p)^{m-1}p^{\ell-m+1}x^m = x \sum_{m \geq 1} \binom{\ell}{m-1}[x(1-p)]^{m-1}p^{\ell-m+1} = x[x(1-p) + p]^\ell$ . Therefore, for a node with degree  $k$ , the PGF for the number of contacts  $M$  is  $\{x[x(1-p) + p]^\ell\}^k$ .

Next, we provide the PGF  $\tilde{H}_1(x)$  of the number of transitory contacts of a selected node proportional to the concurrent degree, but subtracting one transitory contact. Construction of  $\tilde{H}_1(x)$  includes the following three steps:

- Sum over the infectious period  $\ell$ .
- Sum over all PGFs of all possible degree  $k$  while weighting for the excess degree,  $q_k = kp_k / (\sum kp_k)$ .
- Add an extra term, denoted as  $f_\ell(x)$ , corresponding to the PGF that accounts for the concurrent partnership of the infected node being considered, but that excludes the contact that infected the node being considered.

Following the same argument as for the PGF of number of contacts for a node with degree 1, we have  $f_\ell(x) = \sum_{m \geq 1} \binom{\ell}{m-1}(1-p)^{m-1}p^{\ell-m+1}x^{m-1} = \sum_{m \geq 1} \binom{\ell}{m-1}[x(1-p)]^{m-1}p^{\ell-m+1} = [x(1-p) + p]^\ell$ . Notice that the probability of the infectious period  $\ell$  is equal to the probability that the infected node recovers at time  $\ell + 1$ . Therefore, we have

$$\begin{aligned} \tilde{H}_1(x) &= \sum_{\ell \geq 0} \text{Pr}(Z = \ell + 1) \left[ f_\ell(x) + \sum_{k \geq 1} q_k (x f_\ell(x))^{k-1} \right] \\ &= \sum_{\ell \geq 0} (1-\gamma)^\ell \gamma \left\{ (x(1-p) + p)^\ell + \sum_{k \geq 1} q_k [x(x(1-p) + p)^\ell]^{k-1} \right\} \\ &= \sum_{\ell \geq 0} (1-\gamma)^\ell \gamma (x(1-p) + p)^\ell + \sum_{\ell \geq 0} (1-\gamma)^\ell \gamma \sum_{k \geq 1} q_k [x(x(1-p) + p)^\ell]^{k-1} \\ &= \gamma \sum_{\ell \geq 0} [(1-\gamma)(x(1-p) + p)]^\ell + \gamma \sum_{k \geq 1} q_k x^{k-1} \sum_{\ell \geq 0} [(1-\gamma)(x(1-p) + p)^{k-1}]^\ell \\ &= \gamma / [1 - (1-\gamma)(x(1-p) + p)] + \gamma \sum_{k \geq 1} q_k x^{k-1} / [1 - (1-\gamma)(x(1-p) + p)^{k-1}] \end{aligned}$$

where the summation exists as long as  $x < (1/(1-\gamma)^{1/(k-1)} - p)/(1-p)$  for all  $k \geq 2$ . Therefore,

$$\begin{aligned} \tilde{H}'_1(1) &= \gamma(1-\gamma)(1-p)/\gamma^2 + \gamma \sum_{k \geq 1} (\gamma q_k(k-1) + q_k(k-1)(1-\gamma)(1-p))/\gamma^2 \\ &= (1-\gamma)(1-p)/\gamma + (1-p + \gamma p) \sum_{k \geq 1} q_k(k-1)/\gamma, \end{aligned} \quad (\text{G2})$$

or

$$\tilde{H}'_1(1) = (1-\gamma)(1-p)/\gamma + (1-p + \gamma p)g''(1)/(\gamma g'(1)). \quad (\text{G3})$$

Finally, we find the PGF  $H_1(x)$  of the number of transmissions caused by the selected node proportional to the concurrent degree, but subtracting one transitory contact. The PGF of number



of transmissions corresponding to  $k$  transitory contacts is  $(1 - \tau + \tau x)^k$ . Since the PGF of the number of transmissions is the summation of all possible transitory contacts,

$$H_1(x) = \sum P(\text{\#transitory} = k)(1 - \tau + \tau x)^k = \tilde{H}_1(1 - \tau + \tau x). \quad (\text{G4})$$

Then we have,

$$R_* = (dH_1(x)/dx)_{x=1} = (d\tilde{H}_1(1 - \tau + \tau x)/dx)_{x=1} = \tau \tilde{H}'(1). \quad (\text{G5})$$

Applying (G5) and (G6), the reproduction number  $R_*$  is

$$R_* = \tau[(1 - \gamma)(1 - p)/\gamma + (1 - p + \gamma p)g''(1)/(\gamma g'(1))], \quad (\text{G6})$$

where  $\tau = \frac{\beta}{1 - p(1 - \beta)(1 - \gamma)}$ . Proposition 4.1 is thus proved.  $\square$