Temporal Graph Reproduction with RWIG

Shvydun, Sergey; Almasan, Anton David; Mieghem, Piet Van

DOI
10.1109/TNSE.2025.3555974

Publication date
2025

Document Version
Final published version

Published in
IEEE Transactions on Network Science and Engineering

Citation (APA)
Shvydun, S., Almasan, A. D., & Mieghem, P. V. (2025). Temporal Graph Reproduction with RWIG. *IEEE Transactions on Network Science and Engineering*, *12*(4), 3015-3024.
https://doi.org/10.1109/TNSE.2025.3555974

Temporal Graph Reproduction With RWIG

Sergey Shvydun, Anton-David Almasan, and Piet Van Mieghem, *Fellow, IEEE*

***Abstract*—We examine the Random Walkers Induced temporal Graph (RWIG) model, which generates temporal graphs based on the co-location principle of M independent walkers that traverse the underlying Markov graph with different transition probabilities. Given the assumption that each random walker is in the steady state, we determine the steady-state vector $\tilde{s}$ and the Markov transition matrix P_i of each walker w_i that can reproduce the observed temporal network $G_0, \ldots, G_{K\text{-}1}$ with the lowest mean squared error. We also examine the performance of RWIG for periodic temporal graph sequences.**

***Index Terms*—Generative models, Markov process, network dynamics, random walks, RWIG, temporal networks.**

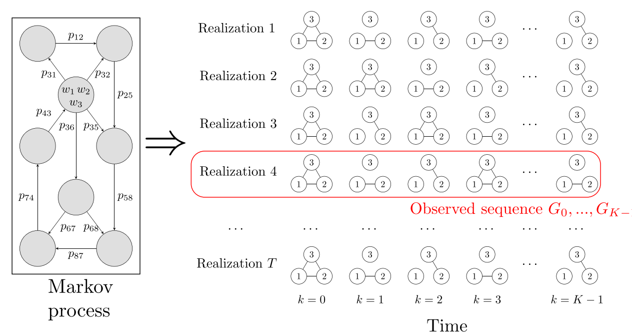


Fig. 1. Inverse problem for RWIG: non-identifiability of the initial process.

I. Introduction

Temporal networks have attracted a lot of attention in the last decades [1], [2], [3]. Many complex systems, including infrastructural, biological, social, and financial networks, evolve over time which, in turn, affects the topology and the processes that propagate over these networks. Modeling these systems enhances our understanding of their dynamics and the processes that propagate through them, such as diffusion and contagion, while also providing valuable predictions [4], [5].

Modeling temporal networks is more difficult and challenging than modeling static graphs. While the static network generation is aimed to reproduce some emergent network structure (e.g. power-law degree distribution [6], community structure [7], [8], motifs [9] or small-world [10] properties), the temporal network generation should take into account a highly non-trivial interplay between the topology and the evolutionary process of the network, which exhibits temporal patterns, memory effects and dependencies, including burstiness of links, their quasi-periodicity (e.g. day-night or weekly rhythms) and temporal correlation. There exist temporal network models that mimic real-world networks in terms of certain topological features such as the number of links, clustering coefficient, degree distribution, connected components or motifs [11], [12], [13]. A system theoretical approach towards emulating temporal graphs is presented in [14]. Various approaches to human mobility are discussed in [15], [16], [17], [18], [19]. However, most existing models do not provide precise knowledge about the underlying process that generates temporal graphs. If the underlying evolutionary process is stochastic, then an observed graph sequence $G_0, \ldots, G_{K\text{-}1}$ represents just one possible realization. Therefore, replicating such sequences without capturing the underlying process may result in overfitting and inaccurate predictions of future network dynamics and propagating processes.

This paper aims a deeper understanding of network evolutionary processes. Specifically, we examine the *Random Walkers Induced temporal Graph* (*RWIG*) model [19] that generates temporal graphs based on the underlying Markov process. The essential components of RWIG are

a) M random walkers on a fixed underlying graph,
b) the fixed underlying graph is a Markov graph on N states (nodes). The random walk on a Markov graph specifies a stochastic process that steers the walker, which we interpret as the walker policy. In principle, each independent walker w_i possesses its own Markov graph on N states) specified by an $N \times N$ probability transition matrix P_i.
c) the co-location principle creates the links in the contact graph G_k between walkers that are in the same state in the underlying graph at time k.

Fig. 1 exemplifies $M = 3$ random walkers, who traverse the same Markov graph (left) with $N = 8$ states (shaded) in discrete-time steps according to the transition probabilities p_{ij}. Fig. 1 (right) shows T possible contact sequences between M walkers. Initially, all walkers form a clique in the contact graph G_0, because they are in the same state at discrete time $k = 0$. At each discrete time k, RWIG generates the contact graph G_k by creating links between all walkers found in the same state in the Markov graph. Therefore, any sequence of graphs $G_0, \ldots, G_{K\text{-}1}$ consists of the union of disconnected cliques, which is in complete accord with the topology of the empirical

Received 7 October 2024; revised 25 March 2025; accepted 26 March 2025. Date of publication 31 March 2025; date of current version 27 June 2025. This work was supported by the European Research Council (ERC) through European Union's Horizon 2020 research and innovation program under Grant Agreement No 101019718. Recommended for acceptance by Dr. Yuan Shen. *(Corresponding author: Sergey Shvydun.)*

The authors are with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: S.Shvydun@tudelft.nl; A.D.Almasan@tudelft.nl; P.F.A.VanMieghem@tudelft.nl).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSE.2025.3555974, provided by the authors.

Digital Object Identifier 10.1109/TNSE.2025.3555974

co-location contact graphs in [20]. However, RWIG cannot reproduce real-world networks with a non-clique structure.

A physical interpretation of RWIG is a collection of individuals moving through space (e.g. city map). Each node of the underlying graph is a physical location (workplaces, homes, hospitals, schools, public transport stations, etc.) while links are physical paths between locations. RWIG assumes that the Markov process generates human motion over a set of places (states). Almasan et al. [19] derived closed-form solutions for the probability distribution of contact graphs at discrete time k. Here, we focus on the inverse problem and discuss how well RWIG can reproduce a temporal graph sequence. Our goal is to infer the initial state and the Markov graph of each walker that can generate a given K-length graph sequence $G_0, \ldots, G_{K-1}$.

Solving the inverse problem for RWIG is challenging. If the evolutionary process of the temporal network is not deterministic, the observed sequence $G_0, \ldots, G_{K-1}$ is only one possible realization of the stochastic process and does not replicate the entire process. Fig. 1 illustrates an example where only realization 4 is observed. Hence, identifying the RWIG's parameters from a single realization may not be feasible.

Our major contributions can be summarized:

- We determine the steady-state vectors of the walkers in RWIG that generate a temporal graph sequence with the same average graph density and clique distribution as in the observed graph sequence $G_0, \ldots, G_{K-1}$.
- If $G_0, \ldots, G_{K-1}$ is produced by RWIG in *the steady state*, then we identify the underlying common policy matrix (the Markov graph) of the walkers that generates the observed graph sequence $G_0, \ldots, G_{K-1}$.
- We demonstrate that RWIG with different policies of the walkers is able to reproduce any periodic graph sequence. We propose the algorithm that defines the Markov graph and initial state of each walker.

The paper is organized as follows. Section II describes RWIG and discusses the performance metrics for the inverse problem. Sections III and IV discuss RWIG in the steady state. Section III provides the solution for the steady-state vectors of the walkers. Section IV determines the Markov graph of the walkers. Section V discusses periodic graph sequences. We summarize the notation in the Appendix A.

II. Random Walkers Induced Temporal Graph (RWIG)

A. Temporal Networks

We consider a temporal graph $G = \{G_k(\mathcal{M}, \mathcal{L}_k)\}_{k=0}^{K-1}$, consisting of a set $\mathcal{M}$ of $|\mathcal{M}| = M$ walkers connected by a set $\mathcal{L}_k$ of L_k links at discrete time $k \in \{0, 1, \ldots, K-1\}$. The graph G_k is described by an $M \times M$ adjacency matrix $A[k]$ whose elements $a_{ij}[k]$ are either one or zero depending on whether there is a link between walkers w_i and w_j or not at discrete time k. For simplicity, we assume that the set of nodes is fixed and the graph is undirected without self-loops. Then all adjacency matrices $A[k] = (A[k])^T$ are real symmetric matrices. Since our focus is on the inverse problem of RWIG, we also assume that the graph G_k is composed of a set of disconnected cliques.

B. Description of RWIG

The random variable $X_i[k]$ denotes the state in the Markov graph of walker w_i at discrete time k and $\Pr[X_i[k] = l]$ is the probability that walker w_i is in state l in the Markov graph at discrete time k. The probability transition matrix P_i encodes the time-independent *policy* of walker w_i

$$(P_i)_{lj} = \Pr[X_i[k+1] = j | X_i[k] = l]$$

at any discrete time k. Given the initial $1 \times N$ state vector $s_i[0]$ of walker w_i (its initial position), the corresponding $1 \times N$ state vector $s_i[k]$ at discrete time k is defined in [21] as

$$s_i[k] = s_i[0] P_i^k,$$

where the l-th element of the probability state vector $s_i[k]$ for walker w_i at discrete time k is $(s_i[k])_l = \Pr[X_i[k] = l]$.

RWIG generates a sequence of contact graphs $G_0, \ldots, G_{K-1}$ where the contact graph G_k consists of the union of disconnected cliques. Hence, RWIG is not able reproduce sequences of non-clique graph structures.

Almasan et al. [19] derived the probability of an m-clique contact graph $g_k = \{\mathcal{A}_1[k], \ldots, \mathcal{A}_m[k]\}$ at discrete time k

$$\Pr[G_k = g_k] = \sum_{i_1=1}^{N} \cdots \sum_{\substack{i_m=1 \\ i_m \notin \{i_l\}_{l=1}^{m-1}}}^{N} \prod_{j=1}^{m} \prod_{w_u \in \mathcal{A}_j[k]} (s_u[k])_{i_j},$$

or, equivalently,

$$\Pr[G_k = g_k] = \sum_{\pi \in \mathcal{P}_g} \left(\prod_{\mathcal{C} \in \pi} (-1)^{|\mathcal{C}|-1} (|\mathcal{C}|-1)! \right) \prod_{\mathcal{A} \in g(\pi)} \sigma_{\mathcal{A}}[k],$$

where $\mathcal{A}_i[k]$ for all $i \in \{1, \ldots, m\}$ represent the cliques formed at discrete time k, $\mathcal{P}_g$ is the set of all partitions on g, $|\mathcal{C}|$ denotes the number of cliques $\mathcal{A}$ in cell $\mathcal{C}$ of partition π on g, $g(\pi)$ is the contact graph in which the cliques are formed by the cells $\mathcal{C}$ and $\sigma_{\mathcal{A}}[k]$ is the probability that walkers of a subset $\mathcal{A} \subseteq \mathcal{M}$ are in the same state at discrete time k

$$\sigma_{\mathcal{A}}[k] = \left(\bigodot_{w_j \in \mathcal{A}} s_j[0] P_j^k \right) u^T, \quad (1)$$

where $\bigodot$ denotes the Hadamard product [22] and u is the $1 \times N$ all-one vector.

Since the Markov process generates the temporal graph sequence, the probability of the observed contact graph sequence $g_0, \ldots, g_{K-1}$ is given by

$$\Pr[G_0 = g_0, \ldots, G_{K-1} = g_{K-1}] = \Pr[G_0 = g_0] \prod_{k=1}^{K-1} \Pr[G_k = g_k | G_{k-1} = g_{k-1}], \quad (2)$$

where $\Pr[G_k = g_k | G_{k-1} = g_{k-1}]$ is the conditional probability of an m-clique contact graph g_k at discrete time k given that an p-clique contact graph g_{k-1} occurs at time k–1 with

$$\Pr[G_k = g_k | G_{k-1} = g_{k-1}] = \sum_{\substack{c_1, \ldots, c_p = 1 \\ c_1 \neq \ldots \neq c_p}}^{N} \sum_{\substack{i_1, \ldots, i_m = 1 \\ i_1 \neq \ldots \neq i_m}}^{N} \prod_{r=1}^{p} \prod_{j=1}^{m} \prod_{w_l \in \mathcal{A}_j[k] \cap \mathcal{A}_r[k-1]} (P_l)_{c_r i_j}. \quad (3)$$

Intuitively, (3) considers possible positions of M walkers in the Markov graph at time k–1 and k and then evaluates the probability of walker transitions between these states. Identification of the transition probability matrix P_i and the initial state $s_i[0]$ of each walker w_i may not be analytically tractable for (2) because the conditional probability $\Pr[G_k = g_k | G_{k-1} = g_{k-1}]$ has $N(N-1)M$ variables and each term in (3) is the product of M variables[1]. Hence, we examine another performance metric for RWIG.

Theorem 1 defines the joint probability of a clique in a contact sequence, which is generated by RWIG.

Theorem 1: Consider RWIG with M walkers $w_1, \ldots, w_M$ where each walker w_i has the $N \times N$ transition probability matrix P_i and the initial state $s_i[0]$ with $s_i[k] = s_i[0]P_i^k$. Denote by $\mathbb{I}_C[k]$ the indicator variable whose values are either one or zero depending on whether a clique C among walkers exists in graph G_k. Then

1) The joint probability of a link (i, j) at discrete time $k + \kappa$ and k between walkers w_i and w_j is
$$\Pr\left[a_{ij}[k+\kappa]=1, a_{ij}[k]=1\right] = s_i[k]\mathrm{diag}\left(P_i^\kappa \left(P_j^\kappa\right)^T\right) s_j^T[k].$$
2) The joint probability that a clique $C \subseteq \mathcal{M}$ occurs at discrete time $k + \kappa$ and k is
$$\Pr[\mathbb{I}_C[k+\kappa] = \mathbb{I}_C[k] = 1] = \left(\bigodot_{w_i \in C} s_i[k]\right)\left(\bigodot_{w_i \in C} P_i^\kappa\right) u^T.$$
3) The joint probability that a clique $C \subseteq \mathcal{M}$ occurs from discrete time k to discrete time $k + \kappa$ is
$$\Pr[\mathbb{I}_C[l] = 1 \ k \le l \le k+\kappa] = \left(\bigodot_{w_i \in C} s_i[k]\right)\left(\bigodot_{w_i \in C} P_i\right)^\kappa u^T.$$

The proof of Theorem 1 is provided in Appendix B.

C. Performance Evaluation of RWIG

Suppose that RWIG generates the sequence $\widehat{A} = (\widehat{A}[0], \ldots, \widehat{A}[K-1])$ of adjacency matrices over K time slots. RWIG's performance can be assessed by measuring the *mean squared error (MSE)* between links in A and $\widehat{A}$

$$MSE_{link}\left(A, \widehat{A}\right) = \frac{1}{K\binom{M}{2}} \sum_{k=0}^{K-1} \sum_{i=1}^{M} \sum_{j=i+1}^{M} \left(a_{ij}[k] - \widehat{a}_{ij}[k]\right)^2, \quad (4)$$

where $a_{ij}[k]$ and $\widehat{a}_{ij}[k]$ are real and estimated entries corresponding to the link between w_i and w_j in graph G_k.

Since all random walkers move independently of each other in the Markov graph, the probability of a contact $\widehat{a}_{ij}[k]$ between two walkers w_i and w_j at discrete time k depends on the corresponding probability state vectors $s_i[k]$ and $s_j[k]$

$$\widehat{a}_{ij}[k] = \sum_{l=1}^{N} \Pr\left[X_i[k] = l, X_j[k] = l\right]$$
$$= \sum_{l=1}^{N} \prod_{u \in \{i,j\}} \Pr\left[X_u[k] = l\right] = s_i[k] \cdot s_j^T[k]. \quad (5)$$

Eq. (5) imposes constraints on vectors $s_1[k], \ldots, s_M[k]$ of the walkers at discrete time k to reproduce graph G_k.

Lemma 1: RWIG reproduces the sequence $\widehat{A} = (\widehat{A}[0], \ldots, \widehat{A}[K-1])$ of adjacency matrices with $MSE_{link}(A, \widehat{A}) = 0$ if and only if for any discrete time k

1) $s_i[k] = s_j[k] = e_l$ for $a_{ij}[k] = 1$ where e_l is the l-th standard basis vector, $l \in \{1, \ldots, N\}$;
2) $s_i[k] \perp s_j[k]$ for $a_{ij}[k] = 0$.

Substituting (5) into (4) yields

$$MSE_{link}\left(A, \widehat{A}\right) = \frac{1}{K\binom{M}{2}} \sum_{k=0}^{K-1} \sum_{i<j} \left(a_{ij}[k] - s_i[k]s_j^T[k]\right)^2. \quad (6)$$

Given the sequence $A = (A[0], \ldots, A[K-1])$, we need to identify the initial state vectors $s_{\mathcal{M}}[0] = (s_1[0], \ldots, s_M[0])$ of M walkers and their Markov matrices $P_{\mathcal{M}} = (P_1, \ldots, P_M)$ that minimize (6). However, identifying $s_{\mathcal{M}}[0]$ and $P_{\mathcal{M}}$ is not analytically tractable because $s_i[k]$ and $s_j[k]$ in (6) involve powers of unknown stochastic matrices $P_1, \ldots, P_M$ as well as unknown initial state vectors $s_1[0], \ldots, s_M[0]$. For instance, the term $s_i[k]s_j^T[k] = s_i[0]P_i^k(P_j^k)^T s_j^T[0]$ is a polynomial of degree $2(k+1)$ with $2(N^2-1)$ variables[2].

Eq. (4) compares only the links in A and $\widehat{A}$. Alternatively, we can extend $MSE_{link}(A, \widehat{A})$ and compare higher-order structures, such as cliques of size r. Denote by $\mathcal{M}_r$ a set of $\binom{M}{r}$ possible combinations of r walkers from a set $\mathcal{M}$. The MSE$_{clique}$ measure compares all cliques of size r ($2 \le r \le M$) in sequences A and $\widehat{A}$

$$MSE_{clique}\left(A, \widehat{A}\right) = \sum_{r=2}^{M} \frac{\sum_{k=0}^{K-1} \sum_{C \in \mathcal{M}_r} \left(\mathbb{I}_C[k] - \sigma_C[k]\right)^2}{K\binom{M}{r}}, \quad (7)$$

where $\sigma_C[k]$ is defined in (1) as the probability of clique C at discrete time k, $\binom{M}{r}$ is the number of cliques of size r in a complete graph and $\mathbb{I}_C[k]$ defines the existence of clique C in graph G_k. We analyze MSE$_{link}$ and MSE$_{clique}$ of RWIG in the steady state in Section III.

III. Generating Graphs in the Steady State

A. RWIG With a Single Stochastic Matrix P in the Steady State

Suppose that all walkers have the same $N \times N$ Markov transition matrix P (i.e., $P_i = P$ for all $w_i \in \mathcal{M}$), which admits a $1 \times N$ steady-state vector $\tilde{s} = \tilde{s}P$ with $\tilde{s}u = 1$. Then the steady-state probability vector of each walker $w_i \in \mathcal{M}$ is

$$\lim_{k \to \infty} s_i[k] = \tilde{s}.$$

Assume that all M walkers start in the same steady state $\tilde{s}$. Then $s_1[k] = \ldots = s_M[k] = \tilde{s}$ and the probability of contact $\widehat{a}_{ij}[k]$ between walkers w_i and w_j at discrete time k is

$$\widehat{a}_{ij}[k] = s_i[k] \cdot s_j^T[k] = \tilde{s} \cdot \tilde{s}^T = \sum_{l=1}^{N} \tilde{s}_l^2 = p. \quad (8)$$

[1]Each walker w_i has the $N \times N$ transition probability matrix P_i with $N(N-1)$ variables. The total number of walkers is M. The term $\prod_{j=1}^{m} \prod_{r=1}^{p} \prod_{w_l \in \mathcal{A}_{i_j} \cap \mathcal{B}_{c_r}} (P_l)_{i_j c_r}$ in (3) has polynomial degree M, because it contains one element from each matrix $P_1, \ldots, P_M$.

[2]Walker w_i has $N-1$ variables for the initial state vector $s_i[0]$ and N^2-N variables for the $N \times N$ transition probability matrix P_i. The total number of variables for each walker is N^2-1. Each entry of P_i^k has polynomial degree k and $s_i[0]P_i^k$ has degree $k+1$.

Applying the Cauchy-Schwarz inequality [22] to the $N\times 1$ vector $\tilde{s}$ and the $N\times 1$ all-one vector u gives

$$\left(\sum_{l=1}^{N} \tilde{s}_l\right)^2 \leq \sum_{l=1}^{N} \tilde{s}_l^2 \sum_{l=1}^{N} u_l^2,$$

or, equivalently,

$$\frac{1}{N} \leq \sum_{l=1}^{N} \tilde{s}_l^2 = p, \tag{9}$$

because $\sum_{l=1}^{N} \tilde{s}_l = 1$ and $0\leq\tilde{s}_l\leq 1$. Since $\sum_{l=1}^{N} \tilde{s}_l^2 \leq \sum_{l=1}^{N} \tilde{s}_l$, we conclude that $p\in[\frac{1}{N}, 1]$. Hence, there is always a non-zero probability that walkers w_i and w_j form a contact between each other in the steady-state.

Relation (8) shows some limitations of RWIG in the steady-state (RWIG$_{ss}$) for temporal network generation. First, the probability p of a contact between walkers w_i and w_j is invariant of the discrete time k, because RWIG is in the steady-state. Thus, an m-clique contact graph G is equally likely to emerge across K time slots. Second, the probability of a contact between any two walkers has a fixed probability $p = \tilde{s}\tilde{s}^T$, which is not realistic for many real networks. Third, relation (9) shows that the lower bound of a contact probability p is rather high for small N (e.g. $p \geq \frac{1}{4}$ for $N = 4$). Hence, it is unlikely for RWIG$_{ss}$ to accurately reproduce any specific labeled contact sequence. Nevertheless, although all links have the same probability p to appear, RWIG$_{ss}$ is different from any existing random graph model for static networks. Indeed, RWIG always produces a set of cliques. Consequently, RWIG$_{ss}$ can also be viewed as a new random graph model for fixed graphs. Lemma 2 defines the properties of RWIG$_{ss}$.

Lemma 2: Consider RWIG where all M walkers have the same Markov matrix P with N states, which admits a steady-state vector $\tilde{s} = \tilde{s}P$, and $s_i[0] = \tilde{s}$ for any $w_i \in \mathcal{M}$. Then,

1) the expected number of links $\overline{L}_k$ in graph G_k at any discrete time $k \in \{0,\ldots,K\text{-}1\}$ is, with $p = \tilde{s}\tilde{s}^T \in [\frac{1}{N}, 1]$,
$$\overline{L}_k = \frac{M(M-1)}{2}p \in \left[\frac{M(M-1)}{2N}, \frac{M(M-1)}{2}\right];$$
2) the joint probability of a link (i, j) at discrete time $k+\kappa$ and k in the steady state is
$$\Pr\left[a_{ij}[k+\kappa]=a_{ij}[k]=1\right]$$
$$= \tilde{s}\text{diag}\left(P^{\kappa}\left(P^{\kappa}\right)^T\right)\tilde{s}^T \in \left[\frac{1}{N^2}, 1\right].$$

The proof of Lemma 2 is provided in Appendix C.1.

B. Performance of RWIG$_{ss}$ With Respect to MSE$_{link}$

We analyze MSE$_{link}$ of RWIG, where all M walkers start in the steady state $\tilde{s}$. Substitution of (8) into (6) gives

$$MSE_{link}\left(A, \widehat{A}\right) = \left(\tilde{s}\tilde{s}^T - \overline{a}\right)^2 + b, \tag{10}$$

where b is a non-negative constant and $\overline{a}$ is the average density of a temporal graph with $\overline{a} = \frac{2\sum_{k=0}^{K-1}\sum_{i=1}^{M}\sum_{j=i+1}^{M} a_{ij}[k]}{KM(M-1)}$. Appendix C provides additional information on the simplification of MSE_{link}. The global minimum of (10) occurs at $\tilde{s}\tilde{s}^T = \overline{a}$. Indeed, (10) demonstrates the variational principle of variance $Var[X]$ [21] stating that the best least-square approximation of the random variable X is its mean $E[X]$.

Since the average density $\overline{a}\in[0, 1]$ for any arbitrary graph sequence and $\tilde{s}$ is the steady-state vector with $\tilde{s}\tilde{s}^T \in [\frac{1}{N}, 1]$, $\tilde{s}\tilde{s}^T = \overline{a}$ does not necessarily have a solution for a fixed N. Theorem 2 defines the accuracy of RWIG with respect to MSE$_{link}$ as well as the conditions on steady-state vector $\tilde{s}$.

Theorem 2: Consider RWIG where all M walkers have the same Markov matrix P with N states, which admits a steady-state vector $\tilde{s} = \tilde{s}P$, and $s_i[0] = \tilde{s}$ for any $w_i \in \mathcal{M}$. Any sequence $A = (A[0], \ldots, A[K-1])$ of $M\times M$ adjacency matrices with an average density $\overline{a}$ can be reproduced by RWIG with

1) $MSE_{link}(A, \widehat{A}) = \frac{\sum_{k=0}^{K-1}\sum_{i=1}^{M}\sum_{j=i+1}^{M}(a_{ij}[k]-\frac{1}{N})^2}{K\binom{M}{2}} > 0$ if $\overline{a} < \frac{1}{N}$. The steady-state vector is $\tilde{s} = \frac{u}{N}$.
2) $MSE_{link}(A, \widehat{A}) = \frac{Var[A]}{K\binom{M}{2}} \geq 0$ if $\overline{a} \geq \frac{1}{N}$. The steady-state vector $\tilde{s}$ satisfies $\tilde{s}\tilde{s}^T = \overline{a}$ and has finite solutions if and only if $\overline{a} \in \{\frac{1}{N}, 1\}$.

The minimal number of Markov states N to achieve the lowest MSE_{link} is $N = \lceil 1/\overline{a}\rceil$.

The proof of Theorem 2 is provided in Appendix C.3. Theorem 2 demonstrates that RWIG$_{ss}$, where $s_i[0] = \tilde{s}$ for any $i \in \mathcal{M}$, is able to reproduce the sequence of graphs $G_0, \ldots, G_{K-1}$ accurately ($MSE_{link}(A, \widehat{A}) = 0$) if and only if $Var[A] = 0$ and, consequently, $G_0, \ldots, G_{K-1}$ is a sequence of complete graphs or a sequence of null graphs (there are no links between walkers). Moreover, Theorem 2 shows that for $\overline{a} \notin \{\frac{1}{N}, 1\}$, it is impossible to identify the initial steady-state vector using MSE$_{link}$ because equation $\tilde{s}\tilde{s}^T = \overline{a}$ has infinitely many solutions. Hence, we defined a class of the steady-state vectors $\tilde{s}$ that generate a temporal graph sequence with the same average graph density $\overline{a}$ as in $G_0, \ldots, G_{K-1}$.

C. Performance of RWIG$_{ss}$ With Respect to MSE$_{clique}$

Section III-B demonstrates that, in most cases, the initial steady-state vector $\tilde{s}$ cannot be recovered based on the MSE$_{link}$ criterion. In this section, we show that the vector $\tilde{s}$ can be uniquely defined using the MSE$_{clique}$ criterion defined in (7).

We analyze MSE$_{clique}$ of RWIG, where M walkers start in the steady-state $\tilde{s}$ and follow the same policy P. Assume that all components of the steady-state vector $\tilde{s}$ are not zero. From (1), the probability that r walkers of set $\mathcal{A} = \{w_{i_1}, \ldots, w_{i_r}\}$ form a clique at discrete time k becomes

$$\sigma_{\mathcal{A}}[k] = \left(\bigodot_{w_j\in\mathcal{A}} s_j[0]P_j^k\right)u^T = \left(\bigodot_{w_j\in\mathcal{A}} \tilde{s}\right)u^T = \sum_{i=1}^{N} \tilde{s}_i^r, \tag{11}$$

Eq. (11) is invariant to the discrete time k and provides the same value for any set of r walkers, because they all have the same steady-state vector $\tilde{s}$. Hence, for simplicity, we denote $\sigma_{\mathcal{A}}[k]$ by σ_r. Substitution of (10) into (7) gives

$$MSE_{clique}\left(A, \widehat{A}\right) = \sum_{r=2}^{M}\left(\sigma_r - q_r\right)^2 + \sum_{r=2}^{M} \frac{b_r}{K\binom{M}{r}}, \tag{12}$$

where $b_r \geq 0$ is a constant, $q_r = \frac{\sum_{k=0}^{K-1}\sum_{C\in\mathcal{M}_r}\mathbb{I}_C[k]}{K\binom{M}{r}}$ denotes the probability of cliques of size r in the observed graph sequence

$G_0, \ldots, G_{K\text{-}1}$. Appendix C provides additional information on the simplification of MSE_{clique}.

The global minimum of (12) occurs when $\sigma_r = q_r$ for any $2 \le r \le M$. Explicitly, by (11)

$$\begin{cases} \sum_{i=1}^{N} (\tilde{s})_i = 1, \\ \sum_{i=1}^{N} (\tilde{s})_i^2 = q_2, \\ \cdots \\ \sum_{i=1}^{N} (\tilde{s})_i^M = q_M. \end{cases} \tag{13}$$

The set of equations in (13) can be solved using the Newton identities for polynomials [21], [23]. Consider a polynomial $p_N(z)$ of degree N in the complex variable z

$$p_N(z) = \sum_{r=0}^{N} a_r z^r = a_N \prod_{r=1}^{N} \left(z - (\tilde{s})_r\right), \tag{14}$$

where $a_0, \ldots, a_N$ are the coefficients of a polynomial $p_N(z)$ with $a_N = 1$ and $(\tilde{s})_1, \ldots, (\tilde{s})_N$ are the roots of $p_N(z)$. For each integer $r \ge 1$, the rth power sum σ_r is

$$\sigma_r = \sum_{i=1}^{N} \tilde{s}_i^r.$$

The relation between the coefficients $a_0, \ldots, a_N$ and the power sums $\sigma_1, \ldots, \sigma_N$ is derived in [21] as

$$a_r = -\frac{1}{N - r} \sum_{l=r+1}^{N} a_l \sigma_l. \tag{15}$$

Lemma 3 shows that RWIG$_{ss}$, where walkers start from $\tilde{s}$, is able to generate a temporal graph with the same probability of cliques as in $G_0, \ldots, G_{K\text{-}1}$.

Lemma 3: Let q_r be the empirical probability of cliques of size r with $1 \le r \le M$. The minimum of $MSE_{clique}(A, \widehat{A})$ satisfies (12) and occurs when the components of the $1 \times N$ vector $\tilde{s}$ are the non-zero roots of a polynomial of order M

$$p_M(z) = \sum_{r=0}^{M} a_r z^r, \tag{16}$$

where $a_r = -\frac{1}{M-r} \sum_{l=r+1}^{M} a_l q_l$ with $a_M = 1$. The minimum number of states in the Markov graph is

$$N = M - \min\{r | a_r \neq 0 \text{ and } a_l = 0 \text{ for all } l < r\}.$$

System (13) can be transformed into (16). Given the assumption that the observed contact sequence $G_0, \ldots, G_{K\text{-}1}$ between M walkers is produced by RWIG$_{ss}$ with $N \le M$ states, the rth power sum σ_r can be estimated from $G_0, \ldots, G_{K\text{-}1}$ for any $1 \le r \le M$ as $\sigma_r = q_r$. Given the power sums $\sigma_1, \ldots, \sigma_M$, we can evaluate the coefficients $a_0, \ldots, a_M$ using (15). The number of states N in the Markov graph can be defined based on $a_0, \ldots, a_M$, because the number of zero coefficients equals the number of zero roots of the polynomial $p_M(z)$. The non-zero roots of the polynomial in (16) provides[3] the values of the steady-state vector components $(\tilde{s})_1, \ldots, (\tilde{s})_N$. Any permutation of $(\tilde{s})_1, \ldots, (\tilde{s})_N$ forms a steady-state vector $\tilde{s}$. However, if $G_0, \ldots, G_{K\text{-}1}$ is generated by RWIG with $N > M$ states, we cannot identify the steady-state vector $\tilde{s}$, because the probabilities $q_{M+1}, \ldots, q_N$ are not observed and, consequently, coefficients $a_{M+1}, \ldots, a_N$ of the N-order polynomial $p_N(z)$ cannot be defined.

Lemma 3 requires that the empirical contact probabilities $q_1, \ldots, q_M$ are equal to the steady-state contact probabilities $\sigma_1, \ldots, \sigma_M$. The recoverability of the initial steady-state vector $\tilde{s}$ from a single realization of RWIG is discussed in Appendix D.

[3]For polynomials of degree $N = 5$ or higher, we can obtain the roots using the companion matrix in [22].

D. MSE_{link} of RWIG$_{ss}$ With Stochastic Matrices $P_1, \ldots, P_M$

Suppose that each random walker w_i has an $N \times N$ stochastic matrix P_i, which admits a steady-state distribution $\tilde{s}_i = \tilde{s}_i P_i$ and the initial state $s_i[0]$ of walker w_i is equal to the steady-state probability vector $\tilde{s}_i$, i.e. $s_i[0] = \tilde{s}_i$. The probability of contact $\widehat{a}_{ij}[k]$ between walkers w_i and w_j at discrete time k is

$$\widehat{a}_{ij}[k] = s_i[k] \cdot s_j^T[k] = \tilde{s}_i \tilde{s}_j^T = p_{ij} \in [0, 1]. \tag{17}$$

Relation (17) demonstrates that the probability of a contact between any two walkers at discrete time k is invariant to time in the steady state. Hence, any given m-clique contact graph G is equally likely to emerge across K time slots. This version of the simplified RWIG model is more flexible than the simplified RWIG model ($P_i = P$) from Section III-A, because it incorporates varying probabilities of contacts between walkers. RWIG$_{ss}$ can be also viewed as another random graph model for static graphs that generates a set of cliques from steady-state vectors $\tilde{s}_1, \ldots, \tilde{s}_M$.

Substitution of (17) into (6) gives

$$MSE_{link}\left(A, \widehat{A}\right) = \frac{1}{\binom{M}{2}} \sum_{i=1}^{M} \sum_{j=i+1}^{M} \left(\tilde{s}_i \tilde{s}_j^T - \overline{a}_{ij}\right)^2 + b, \tag{18}$$

where b is a non-negative constant and $\overline{a}_{ij}$ is the average number of links between w_i and w_j in $G_0, \ldots, G_{K\text{-}1}$ with $\overline{a}_{ij} = \frac{\sum_{k=0}^{K\text{-}1} a_{ij}[k]}{K}$. Appendix C provides additional information on the simplification of MSE_{link}.

Relation (18) shows that the global minimum of MSE_{link} occurs when $\tilde{s}_i \tilde{s}_j^T = \overline{a}_{ij}$ for $\forall i \neq j$. Therefore, we need to examine whether the system of equations

$$\begin{cases} \tilde{s}_i \tilde{s}_j^T = \overline{a}_{ij}, \quad \forall i, j \in \mathcal{M}, \\ \tilde{s}_i u = 1, \quad \forall i \in \mathcal{M}, \\ (\tilde{s}_i)_l \ge 0, \quad \forall i \in \mathcal{M}, \ \forall l \in \{1, \ldots, N\} \end{cases} \tag{19}$$

has a solution. The solution set of (19) depends on the average number of links $\overline{a}_{ij}$. For instance, if $\overline{a}_{ij} = 0$ for any two walkers w_i and w_j (the graph sequence has no links), then all walkers traverse different states of the Markov graph at any discrete time k. One possible solution for the walker w_i is the $M \times 1$ steady-state vector $\tilde{s}_i = \delta_{il}$ where δ is the Kronecker delta ($\delta_{il} = 1$ if $l = i$, and 0 otherwise). Another example, if $\overline{a}_{ij} = 1$ for any two walkers w_i and w_j (walkers form a complete graph at any discrete time k), the solution of (19) is the 1×1 steady-state vector $\tilde{s}_i = 1$ for any $w_i \in \mathcal{M}$ while the Markov graph has only $N = 1$ state. However, if $\overline{a}_{12} = 1$, $\overline{a}_{13} = 1$ and $\overline{a}_{23} = 0$, the solution of (19) does not exist[4]. Lemma 4 provides an example of static graph sequences for which system (19) always has a solution. The proof of Lemma 4 is provided in Appendix E.1.

Lemma 4: Consider RWIG where M walkers have $N\times N$ Markov transition matrices $P_1,\ldots,P_M$ with an arbitrary number N of states, which admit $1\times N$ steady-state vectors $\tilde{s}_1,\ldots,\tilde{s}_M$ and $s_i[0]=\tilde{s}_i$ for any $w_i\in\mathcal{M}$. Then any sequence of graphs $G_0,\ldots,G_{K\text{-}1}$ that do not change over time, where G_0 consists of the union of m disconnected cliques, can be accurately modelled by RWIG in the steady state where $N\geq m$.

The system (19) can be represented by an $M\times M$ Gram matrix $\mathcal{G}$ of all inner products of steady-state vectors $\tilde{s}_1,\ldots,\tilde{s}_M$

$$\mathcal{G}=\begin{bmatrix} g_{11} & \overline{a}_{12} & \cdots & \overline{a}_{1M} \\ \overline{a}_{12} & g_{22} & \cdots & \overline{a}_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{1M} & \overline{a}_{2M} & \cdots & g_{MM} \end{bmatrix}=SS^T,$$

where $S=[\tilde{s}_1\ \tilde{s}_2\ \ldots\ \tilde{s}_M]^T$ is an $M\times N$ matrix of the steady-state vectors and $g_{ii}=\tilde{s}_i\tilde{s}_i^T$. Geometrically, S represents the constellation of possible latent positions of M vectors on an (N-1)-dimensional simplex. Unfortunately, given $\mathcal{G}$, it is impossible to recover S. If R is an $N\times N$ orthogonal matrix ($R^TR=RR^T=I$), then an $M\times N$ matrix $\tilde{S}=SR$ has the same Gram matrix[5] [22]. Hence, the solution of (19) is identifiable only up to an orthogonal transformation.

The diagonal element $g_{ii}=\tilde{s}_i\tilde{s}_i^T$ of $\mathcal{G}$ can be interpreted as the probability that walker w_i remains at the same state of the Markov graph in the next time slot. Unfortunately, $g_{11},\ldots,g_{MM}$ are not known as we only observe contacts between walkers and we do not have access to the underlying Markov graph. However, the diagonal elements of $\mathcal{G}$ should satisfy certain conditions in order for system (19) to have at least one solution.

Theorem 3: Let $\mathcal{G}$ be an $M\times M$ symmetric matrix with $g_{ij}\in[0,1]$ for any $i\neq j$ and $\mathcal{G}=U\Lambda U^T$ be the eigenvalue decomposition of $\mathcal{G}$ where $U=[u_1\,u_2\ldots u_M]$ is an $M\times M$ orthogonal matrix formed by the scaled, real eigenvalues u_k belonging to eigenvalue $\lambda_k(\mathcal{G})$, $\Lambda=\text{diag}(\lambda_k(\mathcal{G}))$ is an $M\times M$ diagonal matrix of eigenvalues and $\lambda_1(\mathcal{G})\geq\ldots\geq\lambda_M(\mathcal{G})$. Then $\mathcal{G}$ is the Gram matrix of $1\times N$ steady-state vectors $\tilde{s}_1,\ldots,\tilde{s}_M$ if and only if

1) $\mathcal{G}$ is positive semidefinite;
2) $g_{ii}\in[\frac{1}{N},1]$;
3) $N>rank(\tilde{X})$, where $\tilde{X}$ is an $(M\text{–}1)\times M$ matrix with $\tilde{X}=[x_2\text{–}x_1\ldots x_M\text{–}x_1]^T$ and $X=U\Lambda^{1/2}=[x_1\ldots x_M]^T$;
4) there exists an $N\times 1$ non-zero vector n such that $$\begin{cases}\sum_{j=1}^{rank(\mathcal{G})} n_j(\tilde{X})_{ij}=0,\ \forall i\in\{1,\ldots,M\text{–}1\},\\ \frac{|\sum_{j=1}^{rank(\mathcal{G})} n_j(X)_{1j}|}{||n||}=\frac{1}{\sqrt{N}}.\end{cases}$$
5) there exists an $N\times N$ orthogonal matrix R such that $XR=[\tilde{s}_1\ \tilde{s}_2\ \ldots\ \tilde{s}_M]$ and $\frac{1}{\sqrt{N}||n||}Rn=\frac{u}{N}$.

The proof of Theorem 3 is provided in Appendix E. Theorem 3 defines the conditions on the diagonal of $\mathcal{G}$ and defines a class of the steady-state vectors $\tilde{s}_1,\ldots,\tilde{s}_M$ that generate a temporal graph sequence with the same probability of a contact between any pair of walkers as in $G_0,\ldots,G_{K\text{-}1}$. We provide the solution of (19) for $M=2$ and $M=3$ (for $N=2$) in Appendix E.3.

IV. Recovering the Markov Graph of RWIG

A. Recovering the Matrix P_i From the Steady-State Vector $\tilde{s}_i$

Section III discusses how to find the $1\times N$ steady-state vector $\tilde{s}_i$ of each walker w_i given the temporal graph sequence $G_0,\ldots,G_{K\text{-}1}$ and the initial condition $s_i[0]=\tilde{s}_i$. To recover an $N\times N$ stochastic matrix P_i from a given steady-state vector $\tilde{s}_i$, we need to utilize properties of stochastic matrices [21]. First, P_i is an $N\times N$ square matrix of nonnegative real numbers with each row summing to 1. Second, the steady-state vector $\tilde{s}_i$ is the left eigenvector of P_i that corresponds to the eigenvalue 1. Hence, the stochastic matrix P_i can be obtained from the set of linear equations

$$\begin{cases}P_iu^T=u^T,\\ \tilde{s}_iP_i=\tilde{s}_i,\\ (P_i)_{jk}\geq 0\quad\forall j,k\in\{1,\ldots,N\}.\end{cases}\tag{20}$$

The solution of (20) is always not unique. For instance, $P_i=I$ and $P_i=[\tilde{s}_i\ \tilde{s}_i\ \ldots\ \tilde{s}_i]^T$ are solutions of (20) for any $\tilde{s}_i$. Thus, the set of linear equations in (20) defines the class of stochastic matrices P_i having the same steady-state vector $\tilde{s}_i$.

Example 1: Let walker w_1 be in the steady state $\tilde{s}_1=[\frac{1}{4}\ \frac{3}{4}]$. Then the 2×2 stochastic matrix P_1 of walker w_1 is

$$P_1=\begin{bmatrix}-2+3p_{22} & 3-3p_{22}\\ 1-p_{22} & p_{22}\end{bmatrix},$$

where $\frac{2}{3}\leq p_{22}\leq 1$.

As a remark, a set of linear equations in (20) are sufficient to define the stochastic matrix P_i if and only if $s_i[0]=\tilde{s}_i$, because it is possible that $\lim_{k\to\infty}s_i[k]=\lim_{k\to\infty}s_i[0]P_i^k\neq\tilde{s}_i$ for some $s_i[0]$ and P_i from system (20). However, if the Markov chain is irreducible, aperiodic, and positive recurrent [21], then any initial vector $s_i[0]$ converges to the unique steady-state vector $\tilde{s}_i$.

Example 2: Let walker w_1 be in the steady state $\tilde{s}_1=[\frac{1}{2}\ \frac{1}{2}]^T$. Then the 2×2 stochastic matrix P_1 of walker w_1 is

$$P_1=\begin{bmatrix}p_{22} & 1-p_{22}\\ 1-p_{22} & p_{22}\end{bmatrix},$$

where $0\leq p_{22}\leq 1$. Suppose $p_{22}=0$ and $P_1=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$. If $s_1[0]=[\frac{1}{2}\ \frac{1}{2}]^T$ then $\tilde{s}_1=\lim_{k\to\infty}s_i[0]P_i^k=[\frac{1}{2}\ \frac{1}{2}]^T$. However, if $s_1[0]=[1\ 0]^T$, the steady-state vector $\tilde{s}_1$ does not exist.

B. Recovering the Matrix P From the Contact Sequence

Assume that the contact sequence $G_0,\ldots,G_{K\text{-}1}$ is generated by RWIG$_{ss}$ where all walkers have the same Markov policy P and steady-state vector $\tilde{s}$. We demonstrate that, under this assumption, we can recover the Markov graph using the time correlation between cliques. Since $s_i[k]=\tilde{s}$ for any walker w_i in the steady state, we rewrite the joint probability that a clique

[4]By Lemma 1, $\overline{a}_{12}=1$ and $\overline{a}_{13}=1$ implies $\tilde{s}_1=\tilde{s}_2=\tilde{s}_3=(1,0,0,\ldots)$, which in turn contradicts with $\overline{a}_{23}=0$.

[5]$\tilde{S}\tilde{S}^T=(SR)(SR)^T=S(RR^T)S^T=SS^T=\mathcal{G}$.

$C \subseteq \mathcal{M}$ occurs at discrete time k and $k+1$ from Theorem 1 as

$$\Pr\left[\mathbb{I}_C[k]=1, \mathbb{I}_C[k+1]=1\right] = \left(\bigodot_{w_i \in C} \tilde{s}\right)\left(\bigodot_{w_i \in C} P\right) u^T$$

$$= \sum_{c_0=1}^{N}\sum_{c_1=1}^{N} \left(\tilde{s}_{c_0} P_{c_0 c_1}\right)^{|C|}. \quad (21)$$

Eq. (21) is independent of the walker set C and the discrete time k and is determined solely by the size of C. For simplicity, we denote the joint probability in (21) by $\sigma_{r,1}$ where $r = |C|$.

Denote by W the $N \times N$ matrix obtained from matrix P and steady-state vector $\tilde{s}$

$$W = \begin{bmatrix} \tilde{s}_1 P_{11} & \tilde{s}_1 P_{12} & \cdots & \tilde{s}_1 P_{1N} \\ \tilde{s}_2 P_{21} & \tilde{s}_2 P_{22} & \cdots & \tilde{s}_2 P_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \tilde{s}_N P_{11} & \tilde{s}_N P_{12} & \cdots & \tilde{s}_N P_{1N} \end{bmatrix} = P \odot \tilde{s}^T.$$

The sum of elements in i-th row of W is $\sum_{j=1}^{N} \tilde{s}_i P_{ij} = \tilde{s}_i$. The sum of elements in j-th column of matrix W is $\sum_{i=1}^{N} \tilde{s}_i P_{ij} = \tilde{s}_j$, because $\tilde{s}$ is the steady-state vector of P. Thus, $Wu^T = \tilde{s}^T$ and $uW = \tilde{s}$. The sum of all elements of W is $\sum_{i=1}^{N}\sum_{i=j}^{N} w_{ij} = uWu^T = 1$. Furthermore, for each integer $r > 1$ the rth power sum of elements of W defines the joint probability that graphs G_k and G_{k+1} contain the same clique of size r

$$\sum_{i=1}^{N}\sum_{i=j}^{N} w_{ij}^r = \sigma_{r,1}.$$

Theorem 4 shows that the elements of W can be identified using the Newton identities for polynomials [21], [23]. The proof of Theorem 4 is provided in Appendix F.

Theorem 4: Consider the $N \times N$ matrix $W = P \odot \tilde{s}^T$. Given the steady-state vector $\tilde{s}$ and the steady-state joint probability $\sigma_{r,1}$ for each integer $2 \le r \le N^2$, the elements of matrix W are the roots of a polynomial of order N^2

$$p_{N^2}(z) = \sum_{r=0}^{N^2} a_r z^r, \quad (22)$$

where $a_r = -\frac{1}{N^2-r}\sum_{l=r+1}^{N^2} a_l \sigma_{l,1}$ with $a_{N^2} = 1$. The position of the rth root z_r of (22) in the matrix W is defined by the solution of the placement problem

$$\begin{cases} \sum_{r=1}^{N^2} y_{ir} z_r = \tilde{s}_i, & \text{for all } i \in \{1,\ldots,N\}, \\ \sum_{r=1}^{N^2} y_{rj} z_r = \tilde{s}_j, & \text{for all } j \in \{1,\ldots,N\}, \\ \sum_{j=1}^{N} y_{rj} = 1, & \text{for all } r \in \{1,\ldots,N^2\}, \\ \sum_{i=1}^{N} y_{ir} = 1, & \text{for all } r \in \{1,\ldots,N^2\}, \\ \sum_{r=1}^{N^2} y_{ir} = N, & \text{for all } i \in \{1,\ldots,N\}, \\ \sum_{r=1}^{N^2} y_{rj} = N, & \text{for all } j \in \{1,\ldots,N\} \end{cases} \quad (23)$$

where $y_{ir}, y_{rj} \in \{0,1\}$ for all $i,j \in \{1,\ldots,N\}$ are the binary variables that define if the rth root z_r, $r \in \{1,\ldots,N^2\}$ has position w_{ij} in matrix W.

Given the assumption that $G_0,\ldots,G_{K-1}$ is produced by RWIG$_{ss}$, we identify the steady-state vector $\tilde{s}$ using Lemma 3 and evaluate the joint probability $\sigma_{r,1}$ for any $1 < r \le M^2$. We define the coefficients $a_0,\ldots,a_{N^2}$ and identify the roots $\{z_r\}_{r=1}^{N^2}$ of (22). The solution of (23) is not unique because W and W^T satisfy $Wu^T = \tilde{s}^T$ and $uW = \tilde{s}$. The set of stochastic matrices P is derived from $\tilde{s}$ and W.

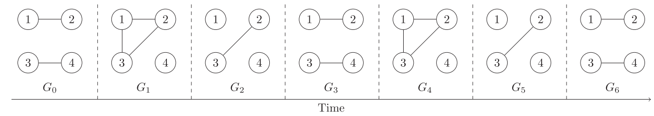


Fig. 2. 3-periodic sequence for a temporal graph with 4 walkers.

Example 3: Consider the Markov graph with $N = 3$ states and $M = 9$ walkers that traverse the Markov graph according to the 3×3 Markov transition probability matrix

$$P = \begin{bmatrix} 0.567 & 0.157 & 0.276 \\ 0.373 & 0.276 & 0.351 \\ 0.327 & 0.502 & 0.171 \end{bmatrix}.$$

Given the steady-state probability q_r of the clique of size r for $r \le 9$, Lemma 3 defines the steady-state vector $\tilde{s} = [0.447, 0.284, 0.269]$. Given the steady-state joint probability $\sigma_{r,1}$ that the same clique of size r occurs in two adjacent time slots for $r \le 9$, the roots of (22) are $z = [0.254, 0.135, 0.123, 0.106, 0.1, 0.088, 0.078, 0.07, 0.046]$. From (23), there are only two possible arrangements or roots $\{z_r\}_{r=1}^{9}$ in the W matrix

$$W_1 = \begin{bmatrix} 0.254 & 0.07 & 0.123 \\ 0.106 & 0.078 & 0.1 \\ 0.088 & 0.135 & 0.046 \end{bmatrix}, \quad W_2 = W_1^T.$$

Dividing each row i of W_1 and W_2 by the corresponding component $\tilde{s}_i$ of the steady-state vector produces transition probability matrices P_1 (the initial matrix) and P_2

$$P_1 = \begin{bmatrix} 0.567 & 0.157 & 0.276 \\ 0.373 & 0.276 & 0.351 \\ 0.327 & 0.502 & 0.171 \end{bmatrix}, P_2 = \begin{bmatrix} 0.567 & 0.236 & 0.197 \\ 0.248 & 0.276 & 0.476 \\ 0.459 & 0.37 & 0.171 \end{bmatrix}.$$

Example 3 illustrates how the initial Markov graph can be inferred from Theorem 4. We show in Appendix D how Theorem 4 is applied to identify the underlying topology of the real transportation system (PATH rail system) from contacts of random walkers.

V. RWIG for Periodic Sequences

Section III demonstrates that RWIG$_{ss}$ can accurately reproduce only static graph sequences $G_0,\ldots,G_{K-1}$. We call such graph sequences 1-periodic because $\forall k \in \{0,..,K-2\}$ $G_k = G_{k+p}$ for $p = 1$. In this section, we examine whether RWIG can accurately reproduce p-periodic graph sequences where $p > 1$. Our motivation for studying periodic sequences is that many real-world systems possess a quasi-periodic dynamic that repeats during a certain period of time [14], [24].

Definition 1: The graph sequence $G_0,\ldots,G_{K-1}$ is called p-periodic if p is the smallest positive integer $p < K-1$ such that $G_k = G_{k+p}$ for any $k \in \{0,..,K-p-1\}$.

An example of 3-periodic sequence for a temporal graph with $M = 4$ walkers is shown in Fig. 2.

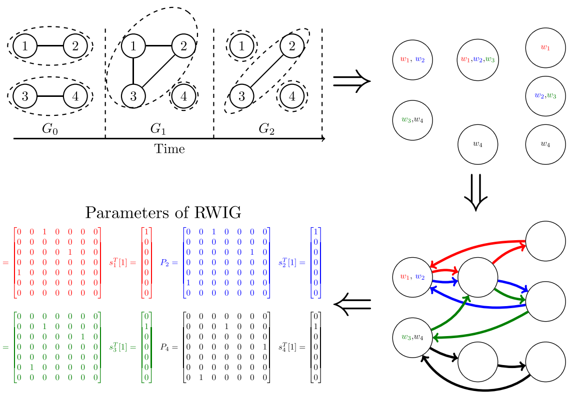


Fig. 3. RWIG with $N = 7$ states for 3-periodic graph sequence from Fig. 2.

Lemma 5: The RWIG model where the walkers have the same $N \times N$ Markov transition matrix P can accurately reproduce only 1-periodic graph sequences $G_0, \ldots, G_{K-1}$.

The proof of Lemma 5 is provided in Appendix G.1.

Consider RWIG where M walkers have $N \times N$ transition matrices $P_1, \ldots, P_M$. Since $a_{ij}[k] = a_{ij}[k+p]$ for any walkers w_i and w_j at any discrete time $k \in \{0, .., K{-}p{-}1\}$, RWIG reproduces p-periodic graph sequence $G_0, \ldots, G_{K-1}$ accurately if and only if $a_{ij}[k] = \widehat{a}_{ij}[k] = \widehat{a}_{ij}[k+p]$ for any $k \in \{0, .., K{-}p{-}1\}$ and any $w_i, w_j \in \mathcal{M}$ where $\widehat{a}_{ij}[k]$ is defined in (5). We assume that each walker w_i traverses the Markov graph with period p (i.e. $s_i[k] = s_i[k+p]$ for any $k \in \{0, .., K{-}p{-}1\}$) and prove that any p-periodic graph sequence $G_0, \ldots, G_{K-1}$ can be reproduced by RWIG if any graph G_k contains a set of cliques.

First, we demonstrate in Fig. 3 the solution for the 3-periodic graph sequence from Fig. 2. We introduce $N = 7$ states in the Markov graph because there are 2 cliques in G_0, 2 cliques in G_1 and 3 cliques in G_2. Each state of the Markov graph corresponds to one of the cliques in the temporal graph at discrete time $k \in \{0, 1, 2\}$. Initially, the walkers w_1, w_2 and w_3, w_4 are placed in the same state as they have a contact in G_0 (see Fig. 3, bottom right). At each discrete time k, walkers move to the same state if they have a contact between each other. Otherwise, a walker w_i moves to its own state in discrete time k (see Fig. 3, top right). The transitions of the walkers are shown in Fig. 3 (red for walker w_1, blue for walker w_2, green for walker w_3 and black for walker w_4). The total number of walker transitions in the Markov graph, or, equivalently, the total number of non-zero elements in the transition matrices $P_1, \ldots, P_4$ is sum of walkers periods, i.e., $\sum_{j=1}^{4} p_i = 12$.

Lemma 6: Any p-periodic graph sequences $G_0, \ldots, G_{K-1}$ can be accurately reproduced by RWIG where M walkers have different $N \times N$ Markov transition matrices $P_1, \ldots, P_M$ and $N = \sum_{k=0}^{p-1} n_c[k]$, $n_c[k]$ is the number of cliques in $G_0, \ldots, G_{p-1}$.

The proof of Lemma 6 is provided in Appendix G.2.

Lemma 6 shows that any periodic graph sequence can be represented by RWIG where each walker w_i follows p-periodic walk. We formulate two research questions:

1) What is the minimal period of each walker to reproduce p-periodic graph sequence?

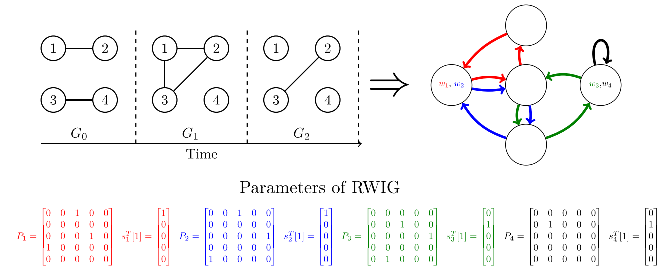


Fig. 4. RWIG with $N = 5$ states for 3-periodic graph sequence from Fig. 2. The transitions of the walkers are shown in red for w_1, blue for w_2, green for w_3 and black for w_4.

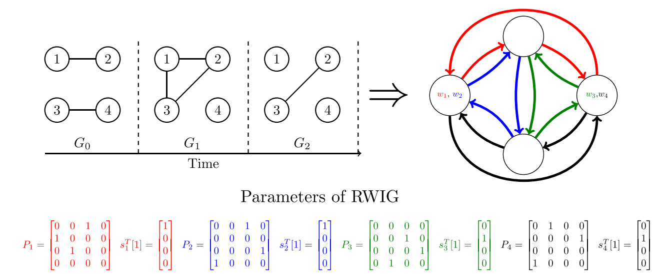


Fig. 5. RWIG with N=4 states for 3-periodic graph sequence from Fig. 2. The transitions of the walkers are shown in red for w_1, blue for w_2, green for w_3 and black for w_4.

2) What is the minimal number of states N in the Markov graph to reproduce p-periodic graph sequence?

Intuitively, the shorter the period p_i of each walker w_i, the fewer states are needed to describe the walk of w_i and, consequently, the fewer number of states N should be in the Markov graph. However, we demonstrate that the minimal period of the walkers does not imply the minimal number of states.

Lemma 7: Any non-zero p-periodic contact sequence between walkers w_i and w_j can be reproduced by RWIG where w_i and w_j traverse the Markov graph with periods p_i and p_j such that

- $LCD(p_1, p_2) \bmod p = 0$, where $LCD(p_1, p_2)$ is the least common denominator of p_1 and p_2.
- $\min(p_1, p_2) \geq \frac{LCD(p_1,p_2)}{p} \sum_{k=0}^{p-1} a_{ij}[k]$
- $\nexists k_1, k_2 < LCD(p_1, p_2)$ with $a_{ij}[k_1] = a_{ij}[k_2] = 1$ and

$$\begin{cases} k_1 \bmod p_1 = k_2 \bmod p_1, \\ k_1 \bmod p_2 = k_2 \bmod p_2. \end{cases}$$

Lemma 7 defines the set of periods (p_i, p_j) for walkers w_i and w_j to reproduce the observed contacts between them. For instance, the contact sequence [1 0 0] between w_3 and w_4 can be reproduced by RWIG if $(p_3, p_4) \in \{(1, 3), (2, 3), (3, 3), (3, 1), (3, 2)\}$. The proof of Lemma 7 is provided in Appendix G.3.

The minimal period p_i of walker w_i in p-periodic contact sequence is the period, which is present in all sets between walker w_i and other walkers. For instance, the 3-periodic graph sequence from Fig. 2 can be modelled by RWIG where walkers w_1, w_2, w_3, w_4 have periods $p_1 = p_2 = p_3 = 3$ and $p_4 = 1$ (see Fig. 4). Hence, the minimal number of states in the Markov graph is $N = 5$ because walkers w_1 and w_2 require 4 states w_1

Algorithm 1: Markov Graph Generator.

Input: periods $p_1, \ldots, p_M$ of M walkers, contact graph sequence $G_0, \ldots, G_{K-1}$.
Output: periodic states sequence $X_1, \ldots, X_M$ of walkers, the size N of the Markov graph.

$N \leftarrow 0$
for $i \leftarrow 1$ **to** M **do**
  $X_i \leftarrow 0_{p_i \times 1}$
  for $p \leftarrow 1$ **to** p_i **do**
    for $k \leftarrow p$ **to** K **step** p_i **do**
      $j \leftarrow$ GetFirstContact(G_k), i
      if $j \neq 0$ **then** $X_i[p] \leftarrow X_j[k \mod p_j]$ **break**
      end if
    end for
    if $X_i[p] = 0$ **then**
      $s \leftarrow$ GetFreeState$(N, [X_1, \ldots, X_i])$
      if $s = 0$ **then**
        $X_i[p] \leftarrow N + 1$, $N \leftarrow N + 1$
      else
        $X_i[p] \leftarrow s$
      end if
    end if
  end for
end for
return $[X_1 \ldots X_M], N$

and w_2 traverse three states, sharing two of them and walker w_4 requires one additional state w_4 remains in the same state. The total number of transitions in the Markov graph with $N = 5$ states is 10.

However, $N = 5$ states states are not minimal. Suppose that all walkers traverse the Markov graph with period $p = 3$. We propose Algorithm 1, which is a heuristic method that constructs the Markov graph with a small number of states for p-periodic graph sequence $G_0, \ldots, G_{K-1}$. First, Algorithm 1 selects walker w_1 and generates $N = p_1$ states in the Markov graph where the states of w_1 are $X_1[k] = k$ for $k \in \{1, \ldots, p_1\}$. Then Algorithm 1 chooses the next walker w_2 that traverses p_2 states. If walker w_2 has a contact with walker w_1 at discrete time $k \leq K$ (function "GetFirstContact"), then walker w_2 shares one of the existing states with w_1, i.e., $X_2[k \mod p_2] = X_1[k \mod p_1]$. However, if w_2 has no contact with w_1 at time k, w_2 visits the first available state l in the Markov graph (function "GetFreeState")[6]. If such states are not available, we add an additional state to the Markov graph for walker w_2. Thus, Algorithm 1 iteratively processes each walker w_i until all M walkers have been considered.

Fig. 5, obtained by Algorithm 1, shows that the 3-periodic graph sequence from Fig. 2 can be modelled by RWIG with $N = 4$ states. The total number of transitions in the Markov graph with $N = 4$ states is 12, which is more compared to Fig. 4. Hence, Algorithm 1 demonstrates that the minimal periods of the walkers do not imply the minimal number of states.

[6] State $l \in \{1, \ldots, N\}$ is available for walker w_2 if $X_1[k] \neq l$ for any $(k \mod p_1) = l$ (w_2 never meets w_1 at state l) and $X_2[t] \neq l$ for $t < k$ (w_2 have not visited state l).

VI. Discussion

We examined the inverse problem for RWIG in the steady state (RWIG$_{ss}$). If all walkers have the same transition probability matrix P and start from the steady-state vector $\tilde{s}$, RWIG$_{ss}$ is able to accurately reproduce only a sequence of complete or null graphs. Any other temporal graph $G_0, \ldots, G_{K-1}$ with M walkers can be approximately modelled by the graph sequence that has the same probability of a clique of size r and the joint probability $\sigma_{r,1}$ for each $r \leq M$. Furthermore, we demonstrate that inferring the initial ergodic Markov process is possible: given $G_0, \ldots, G_{K-1}$, we derive an exact analytical solution that defines the transition probability matrix P of the walkers. Our findings are based on fundamental results of Newton in polynomial theory and functional analysis. The scalability and computational complexity of our methodology are driven by the $O(n^2)$ complexity of existing root-finding methods for n-order polynomials.

If walkers have different transition probability matrices $P_1, \ldots, P_M$ that have steady-state vectors $\tilde{s}_1, .., \tilde{s}_M$, RWIG$_{ss}$ can accurately reproduce only a sequence of m-clique graphs that do not change over time. The lowest possible MSE occurs when $\tilde{s}_i \tilde{s}_j^T = \overline{a}_{ij}$ for all pairs of walkers w_i and w_j, where $\overline{a}_{ij}$ is the average number of links between w_i and w_j in $G_0, \ldots, G_{K-1}$. We imposed several constraints on $\tilde{s}_1, \ldots, \tilde{s}_M$ to achieve this MSE value. However, the general solution for $\tilde{s}_1, .., \tilde{s}_M$ for arbitrary N and M remains unknown.

For periodic sequences, we have proven that, if the walkers follow the same policy P in the Markov graph, RWIG can accurately reproduce only 1-periodic sequences. However, any p-periodic sequence can be reproduced by RWIG where M walkers traverse the Markov graph with different transition probability matrices $P_1, \ldots, P_M$. We provide the lemma that defines the minimal period p_i of each walker w_i.

There are several future directions for this research. First, we have identified the constraints on the temporal graph sequence, which can be reproduced by RWIG with the lowest MSE. However, if a given temporal graph sequence $G_0, \ldots, G_{K-1}$ does not satisfy these constraints, what are the parameters of RWIG$_{ss}$ to approximate $G_0, \ldots, G_{K-1}$? Second, how to infer the parameters of RWIG when the contact sequence is neither periodic nor generated in the steady state? Finally, our findings are valid only for graph sequences that can be generated by RWIG. Therefore, we emphasize the importance of solving the inverse problem for other generative models that produce non-clique structures.

References

[1] P. Holme and J. Saramäki, "Temporal networks," *Phys. Rep.*, vol. 519, no. 3, pp. 97–125, 2012.
[2] P. Holme, "Modern temporal network theory: A colloquium," *Eur. Phys. J. B*, vol. 88, pp. 1–30, 2015.
[3] P. Holme and J. Saramäki, *Temporal Network Theory*. New York, NY, USA: Springer, 2019.
[4] J. L. Iribarren and E. Moro, "Impact of human activity patterns on the dynamics of information diffusion," *Phys. Rev. Lett.*, vol. 103, 2009, Art. no. 038702.

[5] M. Karsai et al., “Small but slow world: How network topology and burstiness slow down spreading,” *Phys. Rev. E*, vol. 83, 2013, Art. no. 025102.

[6] A.-L Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, 1999.

[7] P. W. Holland, K. B. Laskey, and S. Leinhardt, “Stochastic blockmodels: First steps,” *Social Netw.*, vol. 5, no. 2, pp. 109–137, 1983.

[8] A. Grönlund and P. Holme, “Networking the seceder model: Group formation in social and economic systems,” *Phys. Rev. E*, vol. 70, 2004, Art. no. 036108.

[9] F. Bois and G. Gayraud, “Probabilistic generation of random networks taking into account information on motifs occurrence,” *J. Comput. Biol.*, vol. 22, no. 1, pp. 25–36, 2015.

[10] D. J. Watts and S. H. Strogatz, “Collective dynamics of small-world networks,” *Nature*, vol. 393, pp. 440–442, 1998.

[11] D. Zhou, L. Zheng, J. Han, and J. He, “A data-driven graph generative model for temporal interaction networks,” in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2020, pp. 401–411.

[12] G. Zeno, T. LA Fond, and J. Neville, “DYMOND: Dynamic motif-nodes network generative model,” in *Proc. Web Conf.*, 2021, pp. 718–729.

[13] A. Longa, G. Cencetti, S. Lehmann, A. Passerini, and B. Lepri, “Generating fine-grained surrogate temporal networks,” *Commun. Phys.*, vol. 7, 2024, Art. no. 22. [Online]. Available: https://www.nature.com/articles/s42005-023-01517-1

[14] S. Shvydun and P. Van Mieghem, “System identification for temporal networks,” *IEEE Trans. Netw. Sci. Eng.*, vol. 11, no. 2, pp. 1885–1895, 2024.

[15] H. Barbosa et al., “Human mobility: Models and applications,” *Phys. Rep.*, vol. 734, pp. 1–74, 2018.

[16] B. Chang et al., “Markov modulated process to model human mobility,” *Stud. Comput. Intell.*, vol. 1072, pp. 607–618, 2022.

[17] A. Panisson, A. Barrat, C. Cattuto, W. Van den Broeck, G. Ruffo, and R. Schifanella, “On the dynamics of human proximity for data diffusion in ad-hoc networks,” *Ad Hoc Netw.*, vol. 10, no. 8, pp. 1532–1543, 2012.

[18] G. Mauro, M. Luca, A. Longa, B. Lepri, and L. Pappalardo, “Generating mobility networks with generative adversarial networks,” *EPJ Data Sci.*, vol. 11, p. 58, 2022. [Online]. Available: https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-022-00372-4

[19] A.-D. Almasan, S. Shvydun, I. Scholtes, and P. Van Mieghem, “Generating temporal contact graphs using random walkers,” *IEEE Trans. Netw. Sci. Eng.*, vol. 12, no. 3, pp. 1649–1659, May-Jun. 2025, doi: 10.1109/TNSE.2025.3537162.

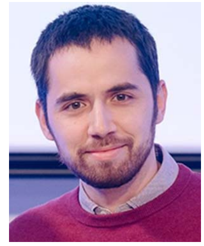
[20] M. Génois and A. Barrat, “Can co-location be used as a proxy for face-to-face contacts?,” *EPJ Data Sci.*, vol. 7, no. 11, pp. 1–18, 2018.

[21] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems*. Cambridge, U.K.: Cambridge University Press, 2014.

[22] P. Van Mieghem, *Graph Spectra for Complex Networks*. 2ed, Cambridge, U.K.: Cambridge University Press, 2023.

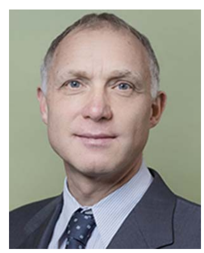
[23] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*. London, U.K.: Oxford University Press, 1998.

[24] E. Andres, A. Barrat, and M. Karsai, “Detecting periodic time scales of changes in temporal networks,” *J. Complex Netw.*, vol. 12, no. 2, 2024, Art. no. cnae004.



Sergey Shvydun received the master’s degree in business informatics with distinction in 2014, and the Ph.D. degree in applied mathematics cum laude,in 2020 from HSE University, Moscow, Russia. He was an Associate Professor with HSE University and as a Senior Research Fellow, the Institute of Control Sciences of the Russian Academy of Science, Moscow. Since 2023, he has been a Postdoctoral Researcher with the Delft University of Technology, Delft, The Netherlands. He is author of more than 30 papers in peer-reviewed journals and edited volumes and one book on centrality in networks. His research interests include network science, machine learning, operations research, and social choice theory.

Anton-David Almasan received the Master of Engineering degree with specialisation in information and computer engineering from the University of Cambridge, Cambridge, U.K. Since 2024, he has been the Ph.D. degree student with the Delft University of Technology, Delft, The Netherlands. He was an AI Consultant with Deepsea Technologies and as a Data Scientist, Thales U.K. His research interests include span machine learning, network science, and stochastic processes.



Piet Van Mieghem (Fellow, IEEE) received the master’s and Ph.D. degrees in electrical engineering from KU Leuven, Leuven, Belgium, in 1987 and 1991, respectively. He was with the Interuniversity Micro Electronic Center (IMEC) from 1987 to 1991. He was a Visiting Scientist with MIT from 1992 to 1993 and a Visiting Professor with UCLA in 2005; Cornell University in 2009; Stanford University in 2015, and Princeton University in 2022. Since 1998, he has been a Professor with the Delft University of Technology and the Chairman of the section Network Architectures and Services. He is the Author of four books: *Performance Analysis of Communications Networks and Systems*, *Data Communications Networking*, *Graph Spectra for Complex Networks* and *Performance Analysis of Complex Networks and Systems*. He was awarded an Advanced ERC grant 2020 for ViSiON, Virus Spread in Networks. During 1993 to 1998, he was a Member of the Alcatel Corporate Research Center in Antwerp. He is also a Board Member of the Netherlands Platform of Complex Systems, a Steering Committee Member of the Dutch Network Science Society, an external Faculty Member with the Institute for Advanced Study of the University of Amsterdam, Amsterdam, The Netherlands. He is currently on the editorial board of the OUP Journal of Complex Networks. He was a Member of the editorial board of Computer Networks 2005-2006, the IEEE/ACM Transactions on Networking 2008-2012, the Journal of Discrete Mathematics 2012-2014 and Computer Communications 2012-2015.