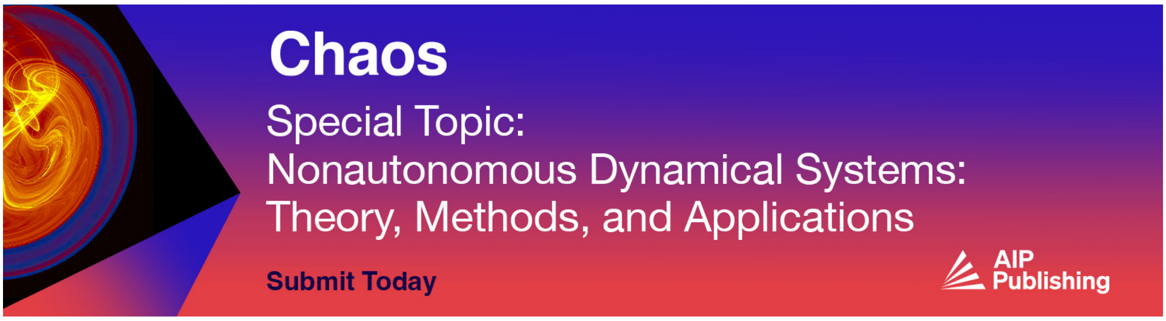
Chaos
An Interdisciplinary Journal of Nonlinear Science

RESEARCH ARTICLE | OCTOBER 31 2018

Inference of time-varying networks through transfer entropy, the case of a Boolean network model FREE

Maurizio Porfiri ; Manuel Ruiz Marín

Chaos 28, 103123 (2018)
https://doi.org/10.1063/1.5047429



CHAOS **28**, 103123 (2018)

Inference of time-varying networks through transfer entropy, the case of a Boolean network model

Maurizio Porfiri[1,a)] and Manuel Ruiz Marín[2]
[1]*Department of Mechanical and Aerospace Engineering, New York University, Tandon School of Engineering, Brooklyn, New York 11201, USA*
[2]*Department of Quantitative Methods and Informatics, Technical University of Cartagena, Calle Real 3, 30201, Cartagena, Spain*

(Received 6 July 2018; accepted 26 September 2018; published online 31 October 2018)

Inferring network topologies from the time series of individual units is of paramount importance in the study of biological and social networks. Despite considerable progress, our success in network inference is largely limited to static networks and autonomous node dynamics, which are often inadequate to describe complex systems. Here, we explore the possibility of reconstructing time-varying weighted topologies through the information-theoretic notion of transfer entropy. We focus on a Boolean network model in which the weight of the links and the spontaneous activity periodically vary in time. For slowly-varying dynamics, we establish closed-form expressions for the stationary periodic distribution and transfer entropy between each pair of nodes. Our results indicate that the instantaneous weight of each link is mapped into a corresponding transfer entropy value, thereby affording the possibility of pinpointing the dominant weights at each time. However, comparing transfer entropy readings at different times may provide erroneous estimates of the strength of the links in time, due to a counterintuitive modulation of the information flow by the non-autonomous dynamics. In fact, this time variation should be used to scale transfer entropy values toward the correct inference of the time evolution of the network weights. This study constitutes a necessary step toward a mathematically-principled use of transfer entropy to reconstruct time-varying networks. *Published by AIP Publishing.* https://doi.org/10.1063/1.5047429

From ecology to political science, complex systems find in networks a viable framework for modeling and analysis. Yet, seldom do researchers have precise knowledge of the topology of these networks. Which predator is challenging the survival of an endangered species? Which legal organization is influencing the adoption of a policy in a given community? These questions exemplify the chief objective of the field of network inference, that is, to establish efficacious techniques to reconstruct the links of a network from knowledge about the times series of the individual units. Here, we investigate an information-theoretic approach to deal with seasonal effects, eliciting the time evolution of the links and non-autonomous dynamics of the network. We focus on a Boolean network model, in which the spontaneous activity of each node is stimulated by the influence of neighboring network nodes. Through perturbation theory and computer simulations, we demonstrate the feasibility of accurately inferring each link in the network as a function of time. This effort offers a unique mathematical basis for addressing network inference in the presence of seasonal effects, laying the foundations for data-driven modeling of time-varying networked dynamical systems.

I. INTRODUCTION

The problem of inferring the topology of a network from the time series of its units is central to our understanding of biological and social networks. For example, our understanding of food webs in ecological systems demands the inference of interactions among species from observations of their abundances[1]. Similarly, elucidating the determining factors of leadership in animal collective behavior and political science requires objective methods to estimate social influence from individual responses, be it the locomotory pattern of an animal[2] or the legal activity of a state.[3,4]

Information theory offers a mathematically-principled framework for model-free inference of networks from raw time series. In an information-theoretic sense, the information encoded in a random variable can be interpreted as "a measure of how much choice is involved in the selection of the event or of how uncertain we are of the outcome"[5] such that high uncertainty is associated with more information. Through the lens of information theory, a network is seen as the medium that supports the propagation of information among its units. In this vein, one attempts at discovering potential links between nodes from the information flow between them.

Recently, we have witnessed a surge of new information-theoretic methods to enable network inference,[6–11] supported by efficacious computational suites and rigorous statistical methods.[12–16] Within this field of investigation, transfer entropy, originally introduced in Ref. 17, has emerged as a promising tool for network inference in the presence of nonlinear interactions and multiple time delays. Put simply, given two time series, transfer entropy scores the reduction in predicting the future of a time series from its past, given additional knowledge about the past of the other time series.

[a)]Author to whom correspondence should be addressed: mporfiri@nyu.edu

An excellent review on the theory, application, and implementation of transfer entropy can be found in Ref. 18. For example, through transfer entropy, researchers have reconstructed climate and financial networks, shedding light on the critical role of oceanic surface circulation on global temperature[19] and pinpointing the most vulnerable financial companies in countries suffering from credit crisis.[20] Despite significant progress, our ability to perform network inference through transfer entropy is vastly limited to units with time-invariant, autonomous, dynamics that are connected by static networks.

In many biological and social systems, neither of these assumptions should be considered valid.[21–23] For example, the abundance of species in ecological settings is often controlled by seasonal effects, which determine the availability of resources in the environment, and, in turn, the interactions among the species.[24] Similarly, policies can only be discussed or passed at certain times of the year, thereby introducing seasonal effects in the process of policy diffusion, which act together with time-varying interactions due to changes in political climate.[25]

Here, we propose a first study on the use of transfer entropy to reconstruct time-varying networks, in the presence of time-varying dynamics. Toward a mathematically-principled treatment of the problem, we focus on a variant of the Boolean network model proposed in Ref. 26. This model affords a minimalistic representation of the process of policy diffusion, while allowing for the derivation of closed-form analytical results on transfer entropy. In the model, each node is assigned to a Boolean variable, whose probability to activate at a given time step depends on (i) the activity of its neighboring nodes at the previous time step, weighted by a constant parameter quantifying social influence and (ii) its internal dynamics, encoded by a constant parameter measuring the spontaneous activity rate.

Different from Ref. 26, we hypothesize that both the network and the individual dynamics vary in time, with a known periodicity. More specifically, the weight of the links in the network evolves in time—potentially causing links to switch on and off—together with the spontaneous activity of each node. Taken *in toto*, the extended model represents a first-order Markov chain with periodic transition matrix that is amenable to a thorough mathematical treatment. For slowly-varying dynamics where nodes are sporadically active, we employ Floquet theory to calculate the unique periodic stationary distribution of the Boolean network model. At the leading order, such a distribution evolves in time with the spontaneous activity such that the probability that a node is active depends exclusively on the value of the spontaneous activity rate at the previous time step. By combining the periodic stationary distribution with the transition matrix, we demonstrate a closed-form periodic expression for transfer entropy between each pair of nodes.

In agreement with our intuition, transfer entropy depends on the instantaneous value of the weight of the link between the nodes, and, to a first degree of approximation, it is independent of any other link in the network and any weight that was attained at previous time steps. Therefore, at a given time step, we can utilize transfer entropy to guide the process of network inference, pinpointing the strongest links at that particular time step. However, comparing mere transfer entropy values across time steps does not assist in the process of estimating the evolution of a link over time. In fact, our results indicate that transfer entropy is modulated by past values of the spontaneous activity of the network, which must be taken into account to estimate the evolution of links through transfer entropy.

The rest of the paper is organized as follows. In Sec. II, we summarize preliminary results on the ergodic behavior of periodic Markov chains and briefly introduce the notion of transfer entropy for periodic stochastic processes. In Sec. III, we present the Boolean network model that is used in the analysis. In Sec. IV, we present our analytical treatment of the problem toward a closed-form expression for transfer entropy. In Sec. V, we illustrate a series of examples that demonstrate the feasibility of network inference from transfer entropy estimates based on raw time series of the Boolean network model. The examples are purposefully designed to highlight the key steps of our approach to network inference and warn against naive solutions that may lead to inaccurate or erroneous claims. Finally, Sec. VI summarized the key findings of our study and outlines potential avenues for further research.

II. MATHEMATICAL PRELIMINARIES

A. Periodic Markov chains

Throughout this paper, we are concerned with the study of a first-order Markov chain with a periodic transition matrix of period $\tau \in \mathbb{Z}_+$. The framework and notation we employ is based on Ref. 27, although we consider a discrete- rather than a continuous-time setting.

More specifically, we consider a finite-state Markov chain $Z(t)$, $t \in \mathbb{N}$, evolving in the sample space $\mathcal{Z}$ whose generic element is denoted as z_i, $i = 1, \ldots, |\mathcal{Z}|$, where $|\mathcal{Z}|$ is the cardinality of the set; here and henceforth, we use lower case letters to denote realizations of random variables. The transition matrix of the Markov chain is the τ-periodic matrix function $P(t) \in \mathbb{R}_+^{|\mathcal{Z}|\times|\mathcal{Z}|}$, whose ij-th entry at time step t, $P_{ij}(t)$, represents the probability that the Markov chain will transition from state z_i to z_j at t. In formulas, we write

$$P_{ij}(t) = \Pr[Z(t+1) = z_j | Z(t) = z_i], \tag{1}$$

where $\Pr(\cdot)$ indicates probability. Starting from an initial distribution ν_0 at $t = 0$, the time evolution of the distribution of the Markov chain is given by

$$\nu(t+1) = P^{\mathrm{T}}(t)\nu(t), \tag{2}$$

with $\nu(0) = \nu_0$, $t \in \mathbb{N}$, and T indicating matrix transposition.

By construction, the recursion in (2) guarantees that starting from a probability distribution, we always obtain a probability distribution. In other words, if the entries of ν_0 sum to one and all its entries are non-negative, then the same holds true for $\nu(t)$ at any value of t. However, depending on the choice of the initial distribution, one will obtain different evolutions for the probability distribution. Following Ref. 27, under mild conditions, we can show that there is a unique periodic stationary distribution and that, in the long run, the distribution of the Markov chain converges to that exponentially fast. In particular, we can state the following theorem.

Theorem 1. *Given a Markov chain $Z(t)$, $t \in \mathbb{N}$, with τ-periodic probability transition matrix $P(t)$ of all nonzero entries, there is a unique stationary periodic distribution $\pi(t)$. For any choice of the initial distribution ν_0, the probability distribution $\nu(t)$ converges to $\pi(t)$ at an exponential rate.*

Proof. For completeness, we briefly sketch the proof, adapted from the arguments in the continuous-time case in Theorem 1.2 of Ref. 27. To demonstrate the existence and uniqueness of the periodic stationary distribution, we consider τ consecutive iterations of (2) such that

$$\pi(\tau) = \Psi\pi(0), \tag{3}$$

where Ψ is the so-called monodromy matrix used in the study of periodic systems[28] and defined as

$$\Psi = P^{\mathrm{T}}(\tau - 1)\dots P^{\mathrm{T}}(0). \tag{4}$$

To ensure that $\pi(t)$ is τ-periodic, we must enforce $\pi(\tau) = \pi(0)$, which implies that $\pi(0)$ is an eigenvector of Ψ with unitary eigenvalue. By applying the Perron-Frobenius theorem,[29] the largest eigenvalue of Ψ is simple and equal to one; in addition, the entries of the corresponding eigenvector are all nonzero. Hence, there exists a unique distribution $\pi(0) = \pi(\tau)$ that satisfies (3), and the associated periodic stationary distribution is simply constructed by iterating (2) with $\nu_0 = \pi(0)$.

Following the line of argument in Ref. 27, we decompose $\mathbb{R}^{|\mathcal{Z}|}$ as the direct sum of Span$[\pi(0)]$ and the $(|\mathcal{Z}| - 1)$-dimensional vector space $V = \mathrm{Im}(\Psi - I_{|\mathcal{Z}|})$, where Span$(\cdot)$, Im$(\cdot)$, and $I_{|\mathcal{Z}|}$ refer to the span of a set of vectors, range of a linear mapping, and identity matrix in $\mathbb{R}^{|\mathcal{Z}|\times|\mathcal{Z}|}$. By construction, V is stable under Ψ such that $\forall v \in V$, $\Psi v \in V$. In fact, since $v \in \mathrm{Im}(\Psi - I_{|\mathcal{Z}|})$, there exists $u \in \mathbb{R}^{|\mathcal{Z}|}$ for which $v = (\Psi - I_{|\mathcal{Z}|})u$; therefore, $\Psi v = \Psi(\Psi - I_{|\mathcal{Z}|})u = (\Psi - I_{|\mathcal{Z}|})\Psi u$, which is an element of V. Given any initial distribution ν_0, we can write[30] $\nu_0 = \pi(0) + v$ for some $v \in V$ such that after $n\tau$ time steps with $n \in \mathbb{N}$, $\nu(n\tau) - \pi(0) = \Psi^n v$. In V, the largest eigenvalue of Ψ is always less than one by the Perron-Frobenius theorem,[29] which implies that $\nu(n\tau)$ converges to $\pi(0)$ exponentially fast at a rate given by the second largest eigenvalue of Ψ. Notably, the eigenvalues of such a matrix are termed the Floquet exponents associated with (2). By simply offsetting the initial time of the chain, the same reasoning can be pursued to show that $\nu(n\tau + k)$ converges to $\pi(k)$ exponentially fast for any $k = 0, \dots, \tau - 1$. The convergence is, again, at the rate given by the second largest Floquet exponent. □

B. Elements of information theory

Within the field of information theory, the amount of uncertainty of a discrete random variable X is quantified through the entropy $H(X)$ defined as[31]

$$H(X) = -\sum_{x\in\mathcal{X}} \Pr(X = x) \log \Pr(X = x), \tag{5}$$

where $\mathcal{X}$ is the sample space of X and we use natural logarithm so that entropy is measured in "nats." Equation (5) is the expectation of $-\log \Pr(X)$, from which the notion of joint and conditional entropies of two random variables X and Y follows[31]

$$H(X, Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} \Pr(X = x, Y = y) \log \Pr(X = x, Y = y), \tag{6a}$$

$$H(X|Y) = -\sum_{x\in\mathcal{X},y\in\mathcal{Y}} \Pr(X = x, Y = y) \log \Pr(X = x|Y = y), \tag{6b}$$

where $\mathcal{Y}$ is the sample space of Y.

In its original incarnation, the information-theoretic notion of transfer entropy is defined for stationary processes,[17] but its extension to stationary periodic processes, often called cyclostationary processes,[32] is straightforward.[33] Toward this aim, consider two cyclostationary processes $X(t)$ and $Y(t)$, $t \in \mathbb{N}$, with period $\tau \in \mathbb{Z}_+$, taking values in $\mathcal{X}$ and $\mathcal{Y}$, respectively. Transfer entropy, $\mathrm{TE}^{Y\to X}(t)$, is defined as

$$\begin{aligned}\mathrm{TE}^{Y\to X}(t) &= H[X(t+1)|X(t)] - H[X(t+1)|X(t), Y(t)]\\ &= \sum_{\substack{x_+\in\mathcal{X}\\ x\in\mathcal{X}\\ y\in\mathcal{Y}}} \Big\{\Pr[X(t+1) = x_+, X(t) = x, Y(t) = y]\\ &\quad\times \log \frac{\Pr[X(t+1) = x_+|X(t) = x, Y(t) = y]}{\Pr[X(t+1) = x_+|X(t) = x]}\Big\}.\end{aligned} \tag{7}$$

Transfer entropy (7) measures the reduction in the uncertainty of predicting $X(t+1)$ from $X(t)$, due to the additional knowledge about $Y(t)$, that is, the conditional mutual information[31] of $X(t+1)$ and $Y(t)$ given $X(t)$. By construction, transfer entropy is τ-periodic and non-negative.

For classical stationary processes, that is, $\tau = 1$, transfer entropy is often estimated from time series by using binning methods or kernel density estimators, see, for example, Refs. 17, 34, and 35. Similar approaches can be leveraged to estimate transfer entropy from time series in the more general case of $\tau > 1$, although care should be placed in implementing the estimation using triplets (x_+, x, y) that are downsampled from the original time series according to the underlying periodicity of the two processes. Notably, the open source MATLAB toolbox TRENTOOL features the possibility of computing transfer entropy for cyclostationary processes.[33]

III. MODEL FORMULATION

To demonstrate the use of transfer entropy (7) to assist in the inference of time-varying networks, we focus on a generalized form of the Boolean network model proposed in Ref. 26. The model was originally put forward as a minimalistic representation of the process of policy diffusion among N legal entities, constituting the nodes of a network. In this sense, the legal activity of node i, with $i = 1, \dots, N$, is modeled as a binary stochastic process $X^i(t)$, $t \in \mathbb{N}$ that takes values in $\{0, 1\}$. If a law is implemented, repealed, or substantively changed at time t, then $X^i(t) = 1$, otherwise $X^i(t) = 0$.

The activity of node i is the result of the interplay between spontaneous activity and interactions among states such that the probability that a node is active at $t+1$ is

$$\Pr[X^i(t+1) = 1 | X^1(t) = x^1, \ldots, X^N(t) = x^N] = \Theta f(t) \left[1 + \sum_{j=1}^{N} W_{ij}(t) x^j \right]. \quad (8)$$

Here, $\Theta > 0$ measures the spontaneous activity of a node, $f(t) \geq 0$ is a τ-periodic function modulating spontaneous activity, and $W_{ij}(t)$ is the generic τ-periodic, non-negative, weight of the link from the j-th to the i-th node, encapsulating political, geographical, and ideological relationships between legal entities.[3] We contemplate the possibility of having self-loops in the network such that $W_{ii}(t) > 0$. These quantities might be associated with inherent variations in the internal dynamics of each of the nodes, which differentially use their present state to activate. When self-loops are absent, a node entirely relies on the present state of its neighbors as in Ref. 26. To ensure that the right hand side of (8) is less than 1, we should enforce that for every node at all times, the sum of the weights is less than $\frac{1}{\Theta f(t)} - 1$.

Compared to Ref. 26, the model is expanded along three different directions. First, the network is assumed to be time-varying so that interactions between nodes can change over time to reflect political and ideological changes. Second, the spontaneous activity is assumed to also be time-varying to encapsulate seasonal effects in law making, where legal entities may favor particular times to enact changes. Elucidating the consequences of these two extensions on the information flow in the network constitutes the chief technical challenge of this endeavor. Finally, we expand the model to allow for self-loops and weighted interactions between the nodes. While the inclusion of self-loops and weights does not pose considerable technical challenges, it allows for strengthening the mathematical rigor of Ref. 26 and opens the door for the proposed model to transition to other technical domains, such as in theoretical neuroscience.[36]

We close the section by commenting on the value of Boolean network models versus higher-order discrete or continuous models. Working with a Boolean network model offers a vantage point toward the development of closed-form results, which may be difficult to derive for more general dynamics. Within a Boolean network model, we can establish closed-form results for the long-run probability distribution and the associated transfer entropy between any pairs of nodes. The availability of closed-form predictions allows for a transparent assessment of the key factors that shape information transfer in time-varying networks, without confounding effects associated with computational analyses, such as statistical power of simulations and incomplete exploration of model parameters. With the exception of linear autoregression processes, similar closed-form analyses would be difficult to pursue.[37] At the same time, Boolean network models are ubiquitous in the study of real-world phenomena,[38–40] and symbolic dynamics is often used to create Boolean representations of continuous processes.[41–43]

IV. ANALYSIS

To cast the proposed mathematical model within the framework of τ-periodic Markov chains in Sec. II, we assemble all the nodes in the N-dimensional vector $Z(t) = \left[X^1(t), \ldots, X^N(t)\right]^{\mathrm{T}}$. Thus, the transition probability in (8) can be written as

$$\Pr[X^i(t+1) = 1 | Z(t) = z] = \Theta f(t) \left[1 + e_i^{\mathrm{T}} W(t) z\right], \quad (9)$$

where z is a generic 2^N-dimensional Boolean vector and e_i is the i-th element of the natural basis in N dimensions (e_i is equal to zero everywhere except the i-th entry that is equal to one). For both realizations of $X^i(t+1)$, we can write

$$\Pr[X^i(t+1) = x_+^i | Z(t) = z] = (1 - x_+^i) + \Theta f(t)(2x_+^i - 1)\left[1 + e_i^{\mathrm{T}} W(t) z\right]. \quad (10)$$

At any time step, a fraction of the nodes will be active, while the others will remain inactive such that the sample space $\mathcal{Z}$ of $Z(t)$ comprises $|\mathcal{Z}| = 2^N$ Boolean vectors $z_1, \ldots, z_{|\mathcal{Z}|}$. From (10), we can use a counting argument to construct the τ-periodic transition matrix of the first order Markov chain underlying the evolution of $Z(t)$, that is,

$$P_{ij}(t) = \Pr[Z(t+1) = z_j | Z(t) = z_i] = \prod_{k=1}^{N} \left\{ \left(1 - e_k^{\mathrm{T}} z_j\right) + \Theta f(t) \left(2 e_k^{\mathrm{T}} z_j - 1\right) \left[1 + e_k^{\mathrm{T}} W(t) z_i\right] \right\}. \quad (11)$$

For convenience, we order the states with the standard convention of Boolean strings such that z_1 is the vector of all zeros.

In principle, we could use (11) to calculate the stationary periodic distribution, which according to Theorem 1 exists and is unique, and then tackle the computation of transfer entropy. With the exception of dyadic interactions ($N = 2$) and binary switching ($\tau = 2$), these calculations are opaque to interpretation or even unfeasible. To address this issue, we focus on the case of slowly-varying dynamics, $\Theta \ll 1$, for which we establish a closed-form solution for the stationary periodic distribution, along with a closed-form expression for the transfer entropy between any pair of nodes.

A. Stationary periodic distribution for slowly-varying dynamics

By factoring powers Θ in (11), the transition matrix can be written as follows:

$$P(t) = P^{(0)} + \Theta P^{(1)}(t) + \Theta^2 P^{(2)}(t) + \mathcal{O}(\Theta^3), \quad (12)$$

where $\mathcal{O}$ is the Landau symbol. Through rather lengthy and cumbersome computations, the three summands can be written in compact forms, as explained below.

The first term has a very simple expression such that

$$P_{ij}^{(0)} = \prod_{k=1}^{N}\left[\left(1-e_k^{\mathrm{T}}z_j\right)\right] = \begin{cases} 1 & \|z_j\| = 0, \\ 0 & \|z_j\| > 0, \end{cases} \tag{13}$$

where we use $\|\cdot\|$ for the infinity norm, which, for Boolean vectors simply count the number of ones. Note that $P^{(0)}$ is a constant idempotent matrix of rank one, which is all zero, except for the first column that corresponds to the state z_1 where none of the nodes is activated at the next time step.

The second term in (12) has a slightly lengthier expression, that is,

$$P_{ij}^{(1)}(t) = f(t)\sum_{r=1}^{N}\left\{\left(2e_r^{\mathrm{T}}z_j - 1\right)\left[1+e_r^{\mathrm{T}}W(t)z_i\right]\prod_{\substack{k=1\\k\neq r}}^{N}\left(1-e_k^{\mathrm{T}}z_j\right)\right\}$$

$$= \begin{cases} -f(t)\left[N + 1_N^{\mathrm{T}}W(t)z_i\right] & \|z_j\| = 0, \\ f(t)\left[1+z_j^{\mathrm{T}}W(t)z_i\right] & \|z_j\| = 1, \\ 0 & \|z_j\| > 1, \end{cases} \tag{14}$$

where 1_N is the N-dimensional vector of all ones. Different from (13), $P^{(1)}(t)$ is a function of time, and it has in general $N+1$ columns that are different from zero, those associated with the states in which at most one node is activated at the next time step. Interestingly, $P^{(1)}(t)$ is a zero row-sum matrix and its first row is independent of the weights of the network, since z_1 contains only zeros.

The third term in (12) has a richer structure such that

$$P_{ij}^{(2)}(t) = f^2(t)\sum_{\substack{r,s=1\\r\neq s}}^{N}\left\{\left(2e_r^{\mathrm{T}}z_j-1\right)\left[1+e_r^{\mathrm{T}}W(t)z_i\right]\left(2e_s^{\mathrm{T}}z_j-1\right)\left[1+e_s^{\mathrm{T}}W(t)z_i\right]\prod_{\substack{k=1\\k\neq r,s}}^{N}\left(1-e_k^{\mathrm{T}}z_j\right)\right\}$$

$$= \begin{cases} f^2(t)\sum_{\substack{r,s=1\\r>s}}^{N}\left\{\left[1+e_r^{\mathrm{T}}W(t)z_i\right]\left[1+e_s^{\mathrm{T}}W(t)z_i\right]\right\} & \|z_j\| = 0, \\ -f^2(t)\left[1+z_j^{\mathrm{T}}W(t)z_i\right]\left[N-1+\left(1_N^{\mathrm{T}}-z_j^{\mathrm{T}}\right)W(t)z_i\right] & \|z_j\| = 1, \\ f^2(t)\left\{1+z_j^{\mathrm{T}}W(t)z_i+\left[e_{\mathcal{I}_1(z_j)}^{\mathrm{T}}W(t)z_i\right]\left[e_{\mathcal{I}_2(z_j)}^{\mathrm{T}}W(t)z_i\right]\right\} & \|z_j\| = 2, \\ 0 & \|z_j\| > 2, \end{cases} \tag{15}$$

where $\mathcal{I}_1(z_j)$ and $\mathcal{I}_2(z_j)$ are used to identify the two entries of z_j that are different from zero for the case $\|z_j\| = 2$. This instance describes the possibility of activating two nodes simultaneously, which is not present in either (13) or (14). Similar to (14), $P^{(2)}(t)$ is also a zero row-sum matrix and its first row is independent of the weights of the network.

Within the second-order expansion in (12), the entries of $P(t)$ corresponding to the simultaneous activation of three or more nodes at the next time step are neglected. In the context of policy diffusion, slow dynamics ($\Theta \ll 1$) pertains to policies that address specific local problems that occur only seldom or policies that have a high start-up cost and only long-term benefits.[44,45]

Based on the second-order approximation in (12), we can compute a closed-form expression for the unique stationary periodic distribution $\pi(t)$. From (4) and (12), we establish

$$\Psi_{ij} = \Psi_{ij}^{(0)} + \Theta\Psi_{ij}^{(1)} + \Theta^2\Psi_{ij}^{(2)} + \mathcal{O}(\Theta^3), \tag{16}$$

where

$$\Psi^{(0)} = \left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau}, \tag{17a}$$

$$\Psi^{(1)} = \left[P^{(1)}(\tau-1)\right]^{\mathrm{T}}\left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-1} + \cdots + \left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-1}\left[P^{(1)}(0)\right]^{\mathrm{T}}, \tag{17b}$$

$$\begin{aligned}\Psi^{(2)} = {} & \left[P^{(2)}(\tau-1)\right]^{\mathrm{T}}\left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-1} + \cdots \\ & + \left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-1}\left[P^{(2)}(0)\right]^{\mathrm{T}} \\ & + \left[P^{(1)}(\tau-1)\right]^{\mathrm{T}}\left[P^{(1)}(\tau-2)\right]^{\mathrm{T}}\left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-2} + \cdots \\ & + \left[\left(P^{(0)}\right)^{\mathrm{T}}\right]^{\tau-2}\left[P^{(1)}(1)\right]^{\mathrm{T}}\left[P^{(1)}(0)\right]^{\mathrm{T}}.\end{aligned} \tag{17c}$$

These matrices can be considerably simplified by recalling that $P^{(0)}$ is idempotent, $P^{(0)}$ is zero everywhere except the first column that has all ones, and both $P^{(1)}$ and $P^{(2)}$ have zero row-sum. The latter claim implies that premultiplying $\left(P^{(1)}\right)^{\mathrm{T}}$ or $\left(P^{(2)}\right)^{\mathrm{T}}$ by $\left(P^{(0)}\right)^{\mathrm{T}}$ will yield the null matrix. Thus, we obtain the following compact expressions:

$$\Psi^{(0)} = \left(P^{(0)}\right)^{\mathrm{T}}, \tag{18a}$$

$$\Psi^{(1)} = \left[P^{(1)}(\tau-1)\right]^{\mathrm{T}}\left(P^{(0)}\right)^{\mathrm{T}}, \tag{18b}$$

$$\begin{aligned}\Psi^{(2)} = {} & \left[P^{(2)}(\tau-1)\right]^{\mathrm{T}}\left(P^{(0)}\right)^{\mathrm{T}} \\ & + \left[P^{(1)}(\tau-1)\right]^{\mathrm{T}}\left[P^{(1)}(\tau-2)\right]^{\mathrm{T}}\left(P^{(0)}\right)^{\mathrm{T}}.\end{aligned} \tag{18c}$$

Now, we are ready to calculate a closed-form expression for the unique stationary periodic distribution $\pi(t)$ at $t=0$. Toward this aim, we write

$$\pi(0) = \pi^{(0)}(0) + \Theta\pi^{(1)}(0) + \Theta^2\pi^{(2)}(0) + \mathcal{O}(\Theta^3) \tag{19}$$

and replace this expression into (3), with the monodromy matrix given by (16). By equating powers of the same order, we obtain

$$\pi^{(0)}(0) = \Psi^{(0)}\pi^{(0)}(0), \tag{20a}$$

$$\pi^{(1)}(0) = \Psi^{(1)}\pi^{(0)}(0) + \Psi^{(0)}\pi^{(1)}(0), \tag{20b}$$

$$\pi^{(2)}(0) = \Psi^{(2)}\pi^{(0)}(0) + \Psi^{(1)}\pi^{(1)}(0) + \Psi^{(0)}\pi^{(2)}(0). \tag{20c}$$

To describe a probability distribution, we must enforce that the sum of the entries of $\pi^{(0)}(0)$ sum to one, while the other vectors, $\pi^{(1)}(0)$ and $\pi^{(2)}(0)$, must have entries that sum to zero. Given (19), ensuring that the sum of the entries of $\pi^{(1)}(0)$ and $\pi^{(2)}(0)$ is zero corresponds to setting to zero the last summands on the right-hand sides of (20b) and (20c).

From (20a), we directly find

$$\pi_i^{(0)}(0) = \begin{cases} 1 & \|z_i\| = 0, \\ 0 & \|z_i\| > 0, \end{cases} \tag{21}$$

which identifies a vector with all zeros except the first entry that is equal to one. Should we change the time from 0 to an arbitrary time t, such a vector will not change; therefore, we drop the dependence on time and just write $\pi_i^{(0)}$. Next, we substitute (21) in (20b) to determine that $\pi^{(1)}(0)$ is simply equal to the first row of $P^{(1)}(\tau-1)$, that is,

$$\pi_i^{(1)}(0) = \begin{cases} -f(\tau-1)N & \|z_i\| = 0, \\ f(\tau-1) & \|z_i\| = 1, \\ 0 & \|z_i\| > 1. \end{cases} \tag{22}$$

Note that, by construction, the sum of the entries of $\pi_i^{(1)}$ is equal to zero. Should we change the time from 0 (corresponding to $t=\tau$ due to periodicity) to t, we would simply evaluate the function that modulates the spontaneous activity at $t-1$, that is, we would replace $f(\tau-1)$ with $f(t-1)$.

Finally, we can evaluate the last term of the expansion in (20c) from (21) and (22). On account of the structure of the terms of the monodromy matrix in (18), we determine

$$\pi^{(2)}(0) = \left[P^{(2)}(\tau-1)\right]^{\mathrm{T}}\pi^{(0)} + \left[P^{(1)}(\tau-1)\right]^{\mathrm{T}} \times \left\{\left[P^{(1)}(\tau-2)\right]^{\mathrm{T}}\pi^{(0)} + \left(P^{(0)}\right)^{\mathrm{T}}\pi^{(1)}(0)\right\}. \tag{23}$$

By recalling the periodicity of the problem, we note that the term in curly brackets is simply $\pi^{(1)}(\tau-1)$ from (19) adapted at $\tau-1$. Thus, using (21) and (20c) and carrying out the lengthy algebra associated with (14) and (15), we ultimately retrieve the following expression for the last term in the expansion of the stationary periodic distribution:

$$\pi_i^{(2)}(0) = \begin{cases} f^2(\tau-1)\frac{N(N-1)}{2} & \\ -f(\tau-1)f(\tau-2)1_N^{\mathrm{T}}W(\tau-1)1_N & \|z_i\| = 0, \\ -f^2(\tau-1)(N-1) & \\ +f(\tau-1)f(\tau-2)z_i^{\mathrm{T}}W(\tau-1)1_N & \|z_i\| = 1, \\ f^2(\tau-1) & \|z_i\| = 2, \\ 0 & \|z_i\| > 2. \end{cases} \tag{24}$$

The same argument could be replicated by offsetting the initial time of the Markov chain such that the stationary periodic distribution is equal to[46]

$$\pi_i(t) = \begin{cases} 1 - \Theta f(t-1)N + \Theta^2 f^2(t-1)\frac{N(N-1)}{2} - \Theta^2 f(t-1)f(t-2)1_N^{\mathrm{T}}W(t-1)1_N + \mathcal{O}(\Theta^3) & \|z_i\| = 0, \\ \Theta f(t-1) - \Theta^2 f^2(t-1)(N-1) + \Theta^2 f(t-1)f(t-2)z_i^{\mathrm{T}}W(t-1)1_N + \mathcal{O}(\Theta^3) & \|z_i\| = 1, \\ \Theta^2 f^2(t-1) + \mathcal{O}(\Theta^3) & \|z_i\| = 2, \\ \mathcal{O}(\Theta^3) & \|z_i\| > 2. \end{cases} \tag{25}$$

By marginalizing (25), we can compute the stationary periodic distribution of each node, which simply reads

$$\Pr[X_i(t) = x_i] = (1-x_i) + \Theta f(t-1)(2x_i-1) + \Theta^2 f(t-1)f(t-2)e_i^{\mathrm{T}}W(t-1)1_N(2x_i-1) + \mathcal{O}(\Theta^3). \tag{26}$$

A few comments are warranted on (25) and (26). First, at the first order in Θ, the stationary periodic distribution is independent of the network, in that every node has the same probability of being active, irrespective of its topological properties. Only accounting for the second order power in Θ, we register a dependence on topological properties. Second, the probability of being active at a given time step depends on the spontaneous activity and weights at previous time steps. Third, up to the third power in Θ, the nodes are mutually independent, since it is easy to show that for any state z, (25) can be written as the products of all the marginals in (26) up to $\mathcal{O}(\Theta^3)$.

B. Transfer entropy for slowly-varying dynamics

From the established closed-form expression for (25) and the underlying Markov model (8), we establish a leading order expression for the transfer entropy between any pairs of nodes in the network. Without lack of generality, we focus on nodes 1 and 2 to evaluate transfer entropy from node 2 to node 1 from (7),

$$\mathrm{TE}^{2\to 1}(t) = \sum_{x_+^1, x^1, x^2}\left\{\Pr[X^1(t+1) = x_+^1, X^1(t) = x^1, X^2(t) = x^2] \times \log\frac{\Pr[X^1(t+1) = x_+^1 | X^1(t) = x^1, X^2(t) = x^2]}{\Pr[X^1(t+1) = x_+^1 | X^1(t) = x^1]}\right\}, \tag{27}$$

where probabilities are computed with respect to the stationary periodic distribution of the Markov chain $\pi(t)$. In contrast to Ref. 26, (27) has an explicit dependence on time through the underlying model (8) and the stationary distribution (25).

From (10) and (25), along with the mutual independence of the nodes up the third power in Θ, we can derive the following expression for the conditional probability of the future state of node 1 given its present and the present of node 2:

$$\begin{aligned}&\Pr[X^1(t+1)=x^1_+|X^1(t)=x^1,X^2(t)=x^2]\\&=(1-x^1_+)+\Theta f(t)(2x^1_+-1)\left[1+W_{11}(t)x^1+W_{12}(t)x^2\right]\\&\quad+\Theta^2 f(t)f(t-1)(2x^1_+-1)\sum_{j=3}^{N}W_{1j}(t)+\mathcal{O}(\Theta^3).\end{aligned}\tag{28}$$

Through similar steps, we can derive a compact expression for the conditional probability of the future state of node 1 given its present, namely,

$$\begin{aligned}&\Pr[X^1(t+1)=x^1_+|X^1(t)=x^1]\\&=(1-x^1_+)+\Theta f(t)(2x^1_+-1)\left[1+W_{11}(t)x^1\right]\\&\quad+\Theta^2 f(t)f(t-1)(2x^1_+-1)\sum_{j=2}^{N}W_{1j}(t)+\mathcal{O}(\Theta^3).\end{aligned}\tag{29}$$

If self-loops are absent, then up to the third power in Θ, $X^1(t+1)$ and $X^1(t)$ are marginally independent and conditionally independent given $X^2(t)$. In this case, the expression for transfer entropy in (27) can be approximated by the mutual information[31] between $X^1(t+1)$ and $X^2(t)$,

$$\begin{aligned}&I[X^1(t+1);X^2(t)]=H[X^1(t+1)]-H[X^1(t+1)|X^2(t)]\\&=\sum_{x^1_+,x^2}\Big\{\Pr[X^1(t+1)=x^1_+,X^2(t)=x^2]\\&\quad\times\log\frac{\Pr[X^1(t+1)=x^1_+|X^2(t)=x^2]}{\Pr[X^1(t+1)=x^1_+]}\Big\}.\end{aligned}\tag{30}$$

More specifically, we can write[47]

$$\mathrm{TE}^{2\to1}(t)=I[X^1(t+1);X^2(t)]+\mathcal{O}(\Theta^3).\tag{31}$$

In the general case when self-loops are present, the present state of node 1 enters the transition probabilities (28) and (29) with the first power in Θ and, therefore, cannot be neglected. Thus, we should resort to the complete expression for transfer entropy in (27). Therein, we replace the joint probability with

$$\begin{aligned}&\Pr[X^1(t+1)=x^1_+,X^1(t)=x^1,X^2(t)=x^2]\\&\quad=\Pr[X^1(t+1)=x^1_+|X^1(t)=x^1,X^2(t)=x^2]\\&\quad\times\Pr[X^1(t)=x^1]\Pr[X^2(t)=x^2]+\mathcal{O}(\Theta^3),\end{aligned}\tag{32}$$

where the marginal probabilities are given in (26) and the conditional probability is in (28).

Now, we can compute a leading order expansion for transfer entropy by substituting in (27), the above expressions for conditional, (28) and (29), and joint, (32), distributions. Carrying out the expansion in Θ and using i and j rather than 1 and 2, we establish the following compact expression for transfer entropy

$$\mathrm{TE}^{j\to i}(t)=\Theta^2 f(t)f(t-1)G\left[W_{ij}(t)\right]+\mathcal{O}(\Theta^3),\tag{33}$$

where the function

$$G(x)=-x+(1+x)\log(1+x)\tag{34}$$

measures the dependence of transfer entropy on the weight of the link between i and j. For $\tau=1$ and unweighted links, this expression is equivalent to the one we derived in Ref. 26. We can formalize the entire analysis in the following proposition.

Proposition 1. *Consider the Boolean network model (8), describing the dynamics of N Boolean units with τ-periodic spontaneous activity, coupled through a weighted τ-periodic network. For any pair of distinct nodes i and j at time t, transfer entropy (27) is given by (33).*

Predictably, at the leading order in Θ, transfer entropy at time t depends on the instantaneous of the weight between nodes i and j at the same time t. In fact, (33) measures the reduction in the uncertainty in the prediction of the future of node i at time $t+1$ from its present at time t, due to additional knowledge about the present of node j. Based on the model in (8), useful information about $X^i(t+1)$ that are contained in $X^j(t)$ must be mediated by $W_{ij}(t)$, thus explaining the observed dependence. For slowly-varying dynamics, spurious connections between the nodes, passing through other nodes and involving past values of the weights of the network, will have a secondary effect.

In partial disagreement with our expectations, transfer entropy does not depend only on the instantaneous value of the function modulating the individual activity but also on the previous one. This phenomenon is related to the periodicity of the stationary distribution in (25), which depends on the past value of the spontaneous activity, rather than the present. This surprising dependence of transfer entropy on time warrants care in comparing the strengths of links across different time steps. More specifically, values of transfer entropy must be scaled by $f(t)f(t-1)$ to infer the weight of links in the network, through

$$W_{ij}(t)\simeq G^{-1}\left[\frac{\mathrm{TE}^{j\to i}(t)}{\Theta^2 f(t)f(t-1)}\right].\tag{35}$$

If the argument in (35) is less than 0.1, then $G^{-1}(y)$ can be approximated as $\sqrt{2y}$ within a 10% error, otherwise a numerical inversion is needed for the computation.

Another counterintuitive implication of Proposition 1 lies in the effect of self-loops, which at the leading order have no effect on transfer entropy between two nodes. Thus, although $X^i(t+1)$ cannot be assumed to be independent of $X^i(t)$ or conditionally independent of $X^i(t)$ given $X^j(t)$, mutual information between $X^i(t+1)$ and $X^j(t)$ is equal to transfer entropy between the two nodes at the leading order in Θ. From an information-theoretic point of view,[31] this indicates that mutual information of the triplet $X^i(t+1)$, $X^i(t)$, and $X^j(t)$ is negligible, to a first degree of approximation. This insight could not have been achieved by extending the first-order perturbation argument in Ref. 26.

Obviously, transfer entropy from a node to itself is zero since no additional information is provided to improve the prediction of future of a node from its present. However, should one be interested in recovering the weights of the self-loops, then it would be sufficient to compute mutual information between $X^i(t+1)$ and $X^i(t)$,

$$I[X^i(t+1);X^i(t)] = \sum_{x_+^i, x^i} \Big\{ \Pr[X^i(t+1) = x_+^i, X^i(t) = x^i] \times \log \frac{\Pr[X^i(t+1) = x_+^i | X^i(t) = x^i]}{\Pr[X^i(t+1) = x_+^i]} \Big\}. \quad (36)$$

Following analogous steps as those that led to (33), this quantity can be expanded as

$$I[X^i(t+1);X^i(t)] = \Theta^2 f(t) f(t-1) G[W_{ii}(t)] + \mathcal{O}(\Theta^3). \quad (37)$$

This claim can be proved by replacing for (26) and (29) and computing the leading order expansion.

A potential line of further inquiry is the extension of the approach to general time-varying dynamics, beyond periodic systems. Prior work[9,11] may help steering research along this line of research. In fact, the notions of local[9] and relative[11] transfer entropies are similar to (33), in that local transfer entropy also accounts for previous time steps in the system dynamics and relative transfer entropy is normalized by a metric of the system dynamics across multiple time steps.

V. NETWORK RECONSTRUCTION: NUMERICAL ILLUSTRATION

Here, we illustrate the possibility of accurately reconstructing the time-varying weights between each pair of nodes from the time series of the individual nodes. We concentrate on three cases of growing complexity: dyadic interactions in pairs of nodes, small networks ($N = 5$), and large networks ($N = 100$). Self-loops are not considered in Examples 1 and 2 such that nodes are equivalently driven by their neighbors, without a propensity to spontaneously activate based on their present state. Instead, in Example 3, we consider self-loops to account for variations among the nodes.

For each exemplary problem, we run the model in (8) for a total of T time steps starting from a homogenous initial condition and store the time series of all the N nodes, that is, $\{x_t^i\}_{t=0}^{T-1}$ with $i = 1, \ldots, N$. These time series, along with known values of the period, τ, and the spontaneous activity, Θ and $f(1), \ldots, f(\tau)$, are used as the input to our approach to network reconstruction, which outputs the time evolution of the weight of every link in the network.

From the time series of each pair of nodes $i, j = 1, \ldots, N$, we estimate τ values of the joint probability distribution $\Pr\left[X^i(t+1) = x_+^i, X^i(t) = x^i, X^j(t) = x^j\right]$, $t = 0, \ldots, \tau - 1$ by simply counting symbols every τ time steps. For example, to estimate $\Pr\left[X^i(1) = x_+^i, X^i(0) = x^i, X^j(0) = x^j\right]$, we will count triplets in the sequence $\{(x_1^i, x_0^i, x_0^j), (x_{1+\tau}^i, x_\tau^i, x_\tau^j), \ldots\}$. From Eq. (7), transfer entropy for each pair of distinct nodes i and j is then estimated as

$$\widetilde{\mathrm{TE}}^{j\to i}(t) = \sum_{x_+^i, x^j} \widetilde{\Pr}[X^i(t+1) = x_+^i(t), X^i(t) = x^i, X^j(t) = x^j] \log \frac{\widetilde{\Pr}[X(t+1)^i = x^i, X^i(t) = x^i, X^j(t) = x^j]\widetilde{\Pr}[X^i(t) = x^i]}{\widetilde{\Pr}[X^i(t) = x^i, X^j(t) = x^j]\widetilde{\Pr}[X^i(t+1) = x_+^i, X^i(t) = x^i]}, \quad (38)$$

where we have used superimposed tilde to identify quantities that are estimated from raw data, all of which can be derived by marginalizing the probability of triplets. From the numerical values of transfer entropy, we infer the weights of any potential link between two distinct nodes through (35). Consistent with (38), we use a superimposed tilde to identify our inferences of the network weights. When inferring the weights of self-loops, we adopt an equivalent approach, by estimating mutual information in (36) from the time series, similar to (38), and then using (37) to estimate the weight.

A. Dyadic interactions

1. Example 1

We begin with the case of two nodes coupled by an unweighted network that periodically switches with $\tau = 2$ such that only one link is active at each time, namely, $W_{12}(0) = W_{21}(1) = 1$ and $W_{21}(0) = W_{12}(1) = 0$. We further assume that the spontaneous activity is constant in time such that $\Theta = 0.1$ and $f(0) = f(1) = 1$. We consider time series of length $T = 100\,000$ samples to estimate transfer entropy. From (38), we determine $\widetilde{\mathrm{TE}}^{2\to1}(0) = 4.48 \times 10^{-3}$, $\widetilde{\mathrm{TE}}^{2\to1}(1) = 4.21 \times 10^{-5}$, $\widetilde{\mathrm{TE}}^{1\to2}(0) = 2.04 \times 10^{-5}$, and $\widetilde{\mathrm{TE}}^{1\to2}(1) = 3.93 \times 10^{-3}$, which indicate the ability of transfer entropy to pick up the periodic rewiring of the link.

From the estimates of transfer entropy, we can use (35) to infer the time evolution of the weights in the network. By performing the computation, we determine $\widetilde{W}_{12}(0) = 1.08$, $\widetilde{W}_{12}(1) = 0.09$, $\widetilde{W}_{21}(0) = 0.06$, $\widetilde{W}_{21}(1) = 1.01$, which confirm that the approach is successful in quantifying the dominant links and help isolating the secondary ones across the entire period of the dynamics.

Should one ignore the periodicity of the dynamics and opt for a crude computation of transfer entropy using the entire time series, then $\widetilde{\mathrm{TE}}^{2\to1} = 1.13 \times 10^{-3}$ and $\widetilde{\mathrm{TE}}^{1\to2} = 1.23 \times 10^{-3}$, which would mask the existence of a directional interaction between the nodes. In this case, one would likely conclude that the two nodes are bidirectionally coupled, which could be considered a valid statement in an average sense, although inaccurate at any instant in time.

2. Example 2

Here, we consider the case in which the nodes are statically coupled by a constant bidirectional unweighted link, but the spontaneous activity changes in time with a period $\tau = 4$. We maintain $\Theta = 0.1$ and $T = 100\,000$, and we list the values

TABLE I. Model parameters and transfer-entropy based inferences for Example 2.

	$t=0$	$t=1$	$t=2$	$t=3$
$f(t)$	0.5	1	0.5	2
$W_{12}(t)$	1	1	1	1
$W_{21}(t)$	1	1	1	1
$\widetilde{\mathrm{TE}}^{1\to2}(t)$	3.09×10^{-3}	4.22×10^{-3}	2.75×10^{-3}	4.79×10^{-3}
$\widetilde{\mathrm{TE}}^{2\to1}(t)$	3.18×10^{-3}	2.36×10^{-3}	1.61×10^{-3}	6.11×10^{-3}
$\widetilde{W}_{12}(t)$	0.90	1.12	0.90	1.29
$\widetilde{W}_{21}(t)$	0.88	1.56	1.21	1.13

of the function $f(t)$ in Table I. From the numerical computation of transfer entropy in Table I, we confirm that transfer entropy from node 1 to 2 is highly comparable with transfer entropy from node 2 to 1 for any time t, in agreement with the underlying bidirectional interaction. However, comparing transfer entropy readings in time, we evince a remarkable time dependence, whereby at $t=3$ we register a twofold to fourfold increase with respect to $t=2$.

Should one opt for using directly the transfer entropy values for inferring the weight of the link would probably lead to the erroneous prediction that the weight is changing in time. On the other hand, taking into consideration the time variation of $f(t)$ through (35) in Table I, we predict that the weights are approximately constant in time, with statistical variations associated with the finiteness of the time series and inherent discrepancies associated with the assumption of slowly-varying dynamics.[48]

3. Example 3

Now, we consider a more general case than Examples 1 and 2, featuring weighted, switching links and non-autonomous dynamics with a period $\tau=4$. We maintain $\Theta=0.1$ and $T=100\,000$, and we employ the model parameters listed in Table II. Transfer entropy readings demonstrate a wide variation of the information flow between the nodes as a function of time, in agreement with our expectations based on the two prior examples. By applying our approach for network reconstruction, we successfully infer the time evolution of both the weights, although we acknowledge numerical differences in Table II. Again, these differences are the combination of statistical uncertainty and the assumption of slowly-varying dynamics.[49]

TABLE II. Model parameters and transfer-entropy based inferences for Example 3.

	$t=0$	$t=1$	$t=2$	$t=3$
$f(t)$	0.5	1	0.5	2
$W_{12}(t)$	1	0	2	1
$W_{21}(t)$	1	1	0	2
$\widetilde{\mathrm{TE}}^{1\to2}(t)$	2.17×10^{-3}	2.73×10^{-3}	4.82×10^{-6}	2.09×10^{-2}
$\widetilde{\mathrm{TE}}^{2\to1}(t)$	3.31×10^{-3}	1.31×10^{-4}	6.00×10^{-3}	5.56×10^{-3}
$\widetilde{W}_{12}(t)$	0.92	0.24	1.91	1.23
$\widetilde{W}_{21}(t)$	0.73	1.21	0.04	2.66

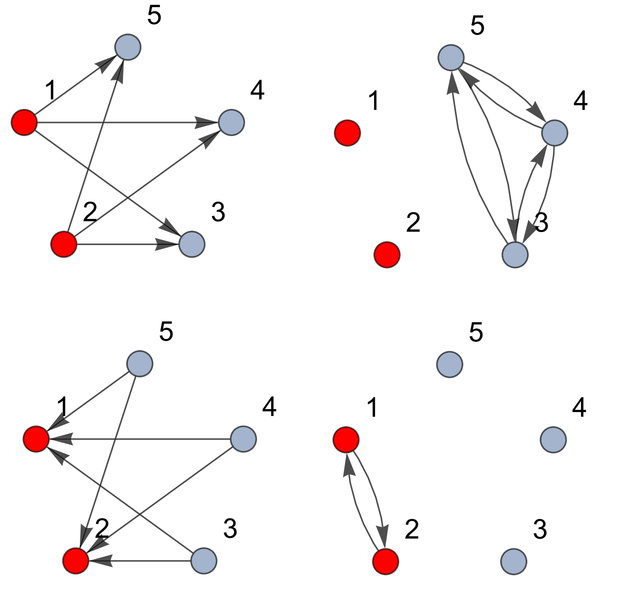


FIG. 1. Switching topology for an example with five nodes and period $\tau=4$: $t=0$ (top left); $t=1$ (top right); $t=2$ (bottom left); and $t=3$ (bottom right). In red, we show the nodes belonging to group A, and in gray those in group B.

Interestingly, if we were to perform the inference by discounting the periodicity of the dynamics, as we illustrated in Example 1, we would discover $\widetilde{\mathrm{TE}}^{2\to1}=1.09\times10^{-3}$ and $\widetilde{\mathrm{TE}}^{1\to2}=2.11\times10^{-3}$, which would lead to the erroneous prediction that on average the two weights are different, with $W_{21}(t)$ exceeding $W_{12}(t)$. In this example, the two weights' patterns only differ by a permutation, but the function $f(t)$

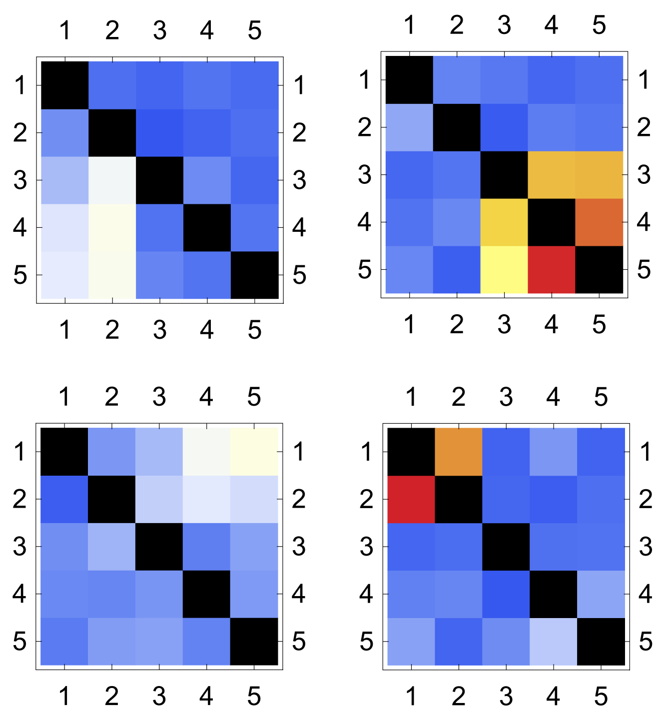


FIG. 2. Inferred weight matrices for an example with five nodes and period $\tau=4$: $t=0$ (top left); $t=1$ (top right); $t=2$ (bottom left); and $t=3$ (bottom right). Colors vary from blue to red, with blue being zero weight and red the largest weight inferred throughout the entire data set, which is 1.21. The diagonal is colored in black, as no inference on self-loops is attempted.

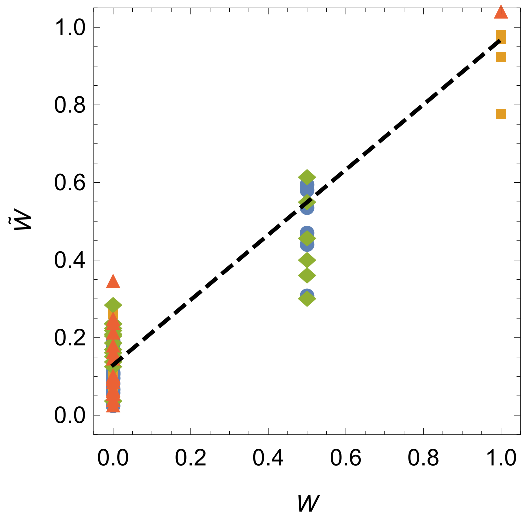


FIG. 3. Inferred weights between distinct nodes versus true weights for an example with five nodes and period $\tau = 4$: $t = 0$ (blue dots); $t = 1$ (light orange squares); $t = 2$ (green diamond); and $t = 3$ (dark orange triangles). The dashed line is the best linear fit $\tilde{W} = 0.13 + 0.84W$ computed using the $\tau N(N-1) = 80$ samples.

differentially modulates the information transfer in the two directions. While $W_{21}(t)$ attains its largest value at $t = 3$ such that $f(t)f(t-1) = 1$, $W_{12}(t)$ has its largest peak at $t = 2$ such that $f(t)f(t-1) = 0.5$.

B. Small network

Now, we consider the case of a small network of $N = 5$ nodes, in which nodes 1 and 2 are assigned to group A and nodes 3, 4, and 5 are assigned to group B. The switching topology of the network is shown in Fig. 1, where we alternate between cases in which nodes in one of the groups bidirectionally interact with each other and instances in which every node in a group unidirectionally connects with all the nodes in the other group. To differentiate the two groups, we hypothesize that group A is more active such that $f(0) = f(3) = 1$ and $f(1) = f(2) = 0.5$. Also, to distinguish between cases in which nodes communicate only within their group versus those in which interactions occur across groups, we use a unitary weight for the former and 0.5 for the latter. Similar to the previous examples, we set $\Theta = 0.1$ and $T = 100\,000$.

By applying our approach to network inference, we predict the weight matrices shown in Fig. 2, which compares very well with the real weights of the underlying model. More specifically, in the top left panel, we note the presence of the six, weak, links from group A to group B. In the top right panel, we evidence the six, strong, links within group B. In the bottom left panel, we see the six, weak, links from group B to A. Finally, the bottom right panel shows the strong bidirectional interaction within group A.

A more quantitative measure of the quality of the inference can be garnered by aggregating all the predictions on the weights between the nodes with the true values of the weights between them, as shown in Fig. 3. Therein, we use four different markers to label the time steps within the period. The coefficient of determination of the least square fit is 0.90, indicating an excellent correlation between true values and transfer-entropy based inferences. Notably, this remarkable correlation is accompanied by the absence of outliers, thereby offering support to the robustness of the proposed approach.

C. Large network

Here, we attempt at demonstrating the feasibility of network inference on a network of $N = 100$ nodes. Again, we focus on $\tau = 4$ and consider time series of length $T = 100\,000$. The parameter Θ is kept at 0.1 and the four values of $f(t)$ are drawn from a uniform distribution between

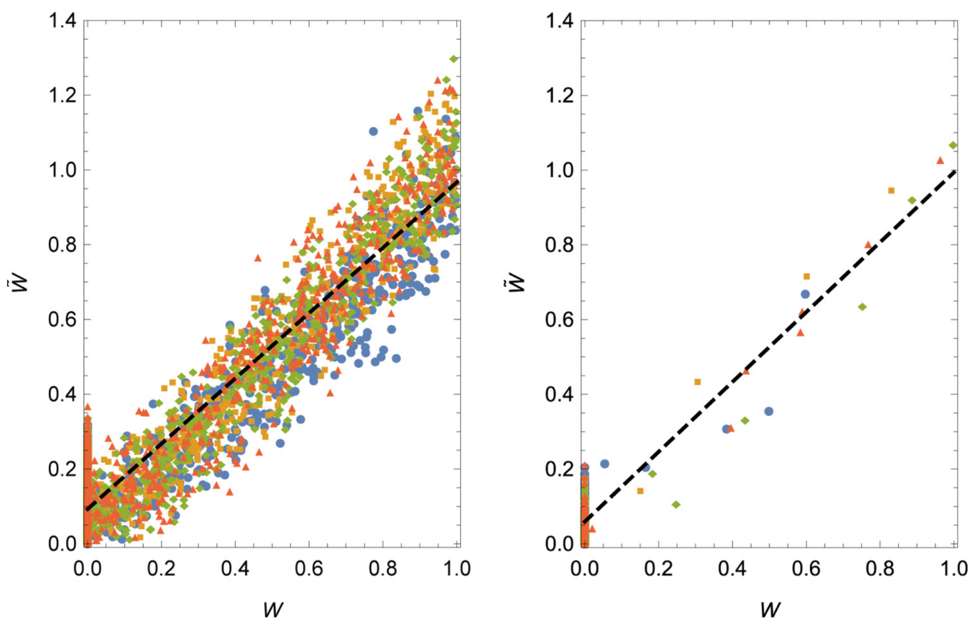


FIG. 4. Inferred weights versus true weights for an example with five nodes and period $\tau = 4$: $t = 0$ (blue dots); $t = 1$ (light orange squares); $t = 2$ (green diamond); and $t = 3$ (dark orange triangles). Panel (a): transfer entropy-based inference of weights of links between distinct nodes; the dashed line is the best linear fit $\tilde{W} = 0.09 + 0.88W$ computed using $\tau N(N-1) = 39\,600$ samples. Panel (b): mutual information-based inference of weights of self-loops; the dashed line is the best linear fit $\tilde{W} = 0.06 + 0.94W$ computed using $\tau N = 400$ samples.

0 and 1 as $f(0) = 0.810$, $f(1) = 0.953$, $f(2) = 0.845$, and $f(3) = 0.772$. Each of the four sets of weights is generated as follows. Given $t = 0, \ldots, \tau - 1$, a link between two distinct nodes or a self-loop has probability 0.05 to be present; if present, it is assigned a weight that is drawn uniformly between 0 and 1.

Mirroring the previous analysis, in Fig. 4(a), we display the inferred weights between any pair of distinct nodes as functions of their true values for the entire period. For the considered large network, the coefficient of determination slightly drops to 0.81, which is still excellent, especially considering that the network can attain an in-degree above ten. Increasing the connectivity of the nodes is, in fact, expected to challenge the accuracy of the leading order approximation for transfer entropy such that indirect connections will play a tangible role in the information flow between two nodes. This is more visible for the links with the smallest weight which may be incorrectly identified due to indirect connections through the other nodes in the network.

For completeness, we also demonstrate the possibility of inferring self-loops from mutual information, computed on individual time series. Figure 4(b) illustrates the accuracy of the prediction, by showing the inferred weights of all the self-loops as a function of their true values for the whole period. The coefficient of determination is 0.87, consistent with the accuracy of the weights of links between distinct nodes.

VI. CONCLUSIONS

Transfer entropy is opening the door to a new, data-driven, model-free approach to the inference of networks from the time series of their constituting nodes. Despite significant progress, our approach to network inference has been hindered by the lack of a methodology to systematically tackle time-varying interactions and seasonal effects. Tackling these phenomena is expected to contribute new tools that could help ongoing research in a number of technical domains where knowledge about social and biological networks is needed. For example, such tools could help refine methods to measure interactions between multiple species in real ecological settings, dominated by seasonal effects,[50] or study the diffusion of policies among legal entities,[44,45] obeying prescribed time lines and affected by the political climate.

In this paper, we attempted at bridging this gap through the mathematical analysis of a Boolean network model, where the spontaneous activity of each node is mediated by its neighbors, which stimulate the node dynamics depending on the instantaneous value of the corresponding weight. The spontaneous activity of each node is, in turn, controlled by a time-varying dynamics, common to the entire network. Both the evolution of the network and the spontaneous activity are assumed to be periodic functions, resulting into a periodic Markov chain with a unique periodic stationary distribution.

For slowly-varying dynamics, we have established closed-form expressions for the stationary periodic distribution of the Markov chain and the instantaneous value of transfer entropy between any pairs of nodes. In this sense, transfer entropy is also a periodic function, whose computation requires the use of the periodic stationary distribution and the underlying periodic Markov model. Our results indicate that transfer entropy can be used to reconstruct the time evolution of each weight in the network. Such a reconstruction should, however, take into consideration the spontaneous activity, which, in fact, modulates the instantaneous value of the transfer entropies in a counterintuitive manner. While transfer entropy only depends on the instantaneous value of the weight of the corresponding link, it depends on both the present and past values of the spontaneous activity.

Through a number of case studies, spanning dyadic interactions, small networks, and large networks, we have demonstrated the feasibility of reconstructing time-varying weighted networks for the selected Boolean model. To simulate a real-world setting in which one has limited to no knowledge of a model for the system, we numerically computed transfer entropy from the Boolean time series of their nodes. The computation uses the periodicity of the dynamics such that the time series are cogently downsampled to evaluate transfer entropy during an entire period. Predictably, a naive approach, in which one computes a single value of transfer entropy for each pair on nodes from their entire time series, leads to insufficient and, sometimes, erroneous inferences.

In its current incarnation, our approach to network inference assumes complete knowledge about the period of the dynamics and the spontaneous activity function. In many applications, the period is known in advance, such as in ecology and political science, where seasonality effects should follow yearly cycles. The time history of the spontaneous activity is more difficult to access, although one may attempt at estimating from local or global activities in the network.

Another limitation of the work is the lack of delays in the interactions between the nodes. It is tenable that these delays could further exacerbate the complexity of the time dependence of transfer entropy, thereby challenging our approach to network inference. Future work could attempt at extending the Boolean network model to account for time delays and establishing leading order solutions for transfer entropy. The model could also be extended to account for negative weights, associated with inhibitory interactions. In this case, it is likely that transfer entropy should be completed with some other information-theoretic measure, since the leading order expansion of transfer entropy might contemplate negative and positive solutions for the same value of transfer entropy.

Overall, this study constitutes a first, necessary, step toward the systematic inference of time-varying networks through transfer entropy. Focusing on a minimalistic model for the dynamics of the network has allowed for a rigorous, transparent, mathematical treatment, which lays the foundation for the analysis of real data sets from social and biological networks.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant Nos. CMMI 1433670, CMMI 1561134, and CBET 1547864, the US Army Research Office under Grant No. W911NF-15-1-0267 with Dr. Samuel C. Stanton and Dr. Alfredo Garcia as the program managers, and Ministerio de Economía y Competitividad de Espana and FEDER funds under Grant No. ECO2015-65637-P. This study is part of the collaborative activities carried out under the program Groups of Excellence of the Region of Murcia, the Fundacion Seneca, Science and Technology Agency of the Region of Murcia Project No. 19884/GERM/15. The authors are thankful to Drs. Shinnosuke Nakayama and Rifat Sipahi for useful discussions on seasonal effects and Mr. Alain Boldini for help in checking the calculations. The authors want to express their gratitude to one anonymous reviewer whose constructive feedback has greatly contributed to improve the technical rigor of the paper.

[1]A. Ives, B. Dennis, K. Cottingham, and S. Carpenter, Ecol. Monogr. **73**, 301 (2003).
[2]A. Strandburg-Peshkin, D. Papageorgiou, M. C. Crofoot, and D. R. Farine, Philos. Trans. R. Soc. B Biol. Sci. **373**, 20170006 (2018).
[3]N. Abaid, J. Macinko, D. Silver, and M. Porfiri, PLoS ONE **10**, e0123339 (2015).
[4]S. Roy and N. Abaid, R. Soc. Open Sci. **4**, 170130 (2017).
[5]C. E. Shannon, ACM SIGMOBILE Mobile Comput. Commun. Rev. **5**, 3 (2001).
[6]C. Ma, H.-S. Chen, Y.-C. Lai, and H.-F. Zhang, Phys. Rev. E **97**, 022301 (2018).
[7]O. Stetter, D. Battaglia, J. Soriano, and T. Geisel, PLoS Comput. Biol. **8**, e1002653 (2012).
[8]M. Lungarella, T. Pegors, D. Bulwinkle, and O. Sporns, Neuroinformatics **3**, 243 (2005).
[9]J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Phys. Rev. E **77**, 026110 (2008).
[10]A.-K. Seghouane and S.-i. Amari, Neural Comput. **24**, 1722 (2012).
[11]C. Wang, H. Yu, R. W. Grout, K.-L. Ma, and J. H. Chen, in *Visualization Symposium (PacificVis), 2011 IEEE Pacific* (IEEE, 2011), pp. 99–106.
[12]M. Lindner, R. Vicente, V. Priesemann, and M. Wibral, BMC Neurosci. **12**, 119 (2011).
[13]D. G. Moore, G. Valentini, S. I. Walker, and M. Levin, Front. Rob. AI **5**, 60 (2018).
[14]J. T. Lizier, Front. Robot. AI **1**, 11 (2014).
[15]A. Montalto, L. Faes, and D. Marinazzo, PLoS ONE **9**, e109462 (2014).
[16]B. L. Ruddell, and P. Kumar, Water Resour. Res. **45**, (2009).
[17]T. Schreiber, Phys. Rev. Lett. **85**, 461 (2000).
[18]T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier, *An Introduction to Transfer Entropy: Information Flow in Complex Systems* (Springer, Berlin, 2016).
[19]J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, Europhys. Lett. **87**, 48007 (2009).
[20]L. Sandoval, Entropy **16**, 4443 (2014).
[21]P. Holme and J. Saramäki, Phys. Rep. **519**, 97 (2012).
[22]I. Belykh, M. Di Bernardo, J. Kurths, and M. Porfiri, Physica D **267**, 1 (2014).
[23]N. Masuda and P. Holme, *Temporal Network Epidemiology* (Springer, 2017).
[24]G. A. Polis and K. O. Winemiller, *Food Webs: Integration of Patterns & Dynamics* (Springer Science & Business Media, 2013).
[25]W. D. Berry, E. J. Ringquist, R. C. Fording, and R. L. Hanson, Am. J. Pol. Sci. **42**, 327–348 (1998).
[26]M. Porfiri and M. R. Marín, IEEE Trans. Netw. Sci. Eng. **5**, 42 (2018).
[27]S. Herrmann and D. Landon, Stoch. Dyn. **15**, 1550022 (2015).
[28]W. J. Rugh, *Linear System Theory* (Prentice Hall, Upper Saddle River, NJ, 1996), Vol. 2
[29]R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).
[30]We have used the fact that ν_0 is a probability distribution to set the coefficient of $\psi(0)$ on the right hand side to one.
[31]T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, 2012).
[32]W. A. Gardner, A. Napolitano, and L. Paura, Signal Process. **86**, 639 (2006).
[33]P. Wollstadt, M. Martínez-Zarzuela, R. Vicente, F. J. Díaz-Pernas, and M. Wibral, PLoS ONE **9**, e102833 (2014).
[34]J. Lee, S. Nemati, I. Silva, B. A. Edwards, J. P. Butler, and A. Malhotra, Biomed. Eng. Online **11**, 19 (2012).
[35]M. Vejmelka and M. Paluš, Phys. Rev. E **77**, 026214 (2008).
[36]A. Roxin, V. Hakim, and N. Brunel, J. Neurosci. **28**, 10734 (2008).
[37]D. A. Smirnov, Phys. Rev. E **87**, 042917 (2013).
[38]E. Ackermann, T. P. Peixoto, and B. Drossel, New J. Phys. **14**, 123029 (2012).
[39]A. T. Hartnett, E. Schertzer, S. A. Levin, and I. D. Couzin, Phys. Rev. Lett. **116**, 038701 (2016).
[40]J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, "The Information Dynamics of Phase Transitions in Random Boolean Networks," in *Proc. Eleventh International Conference on the Simulation and Synthesis of Living Systems*, I (MIT Press, Cambridge, MA, 2008), pp. 374–381, available at http://lizier.me/joseph/publications/2008-Lizier-InfoDynamicsInRbns.pdf.
[41]H. Dickten and K. Lehnertz, Phys. Rev. E **90**, 062706 (2014).
[42]M. Porfiri and M. R. Marín, J. Theor. Biol. **435**, 145 (2017).
[43]M. Staniek and K. Lehnertz, Phys. Rev. Lett. **100**, 158101 (2008).
[44]C. Grabow, J. Macinko, D. Silver, and M. Porfiri, Chaos Interdiscip. J. Nonlin. Sci. **26**, 083113 (2016).
[45]R. P. Anderson, G. Jimenez, J. Y. Bae, D. Silver, J. Macinko, and M. Porfiri, SIAM J. Appl. Dyn. Syst. **15**, 1384 (2016).
[46]In agreement with our intuition, (25) indicates that at the leading order the probability of k nodes activating at the same time is proportional to $\Theta^k f^k(\tau-1)$ – this claim can be readily shown from (3) upon solving for a generic entry with k ones.
[47]For stationary processes, this expression was also used in Ref. 26, although inaccurately presented as an exact result for any network size. While for dyadic interactions, without self-loops, mutual information coincides with transfer entropy, their equivalence is only valid asymptotically in Θ for arbitrary network sizes.
[48]Should the inference be performed by utilizing the exact probability mass functions computed from the study of the Markov chain, the two weights will be equal and they will attain the following values as t goes from 0 to 3: 0.91, 1.15, 0.99, and 1.21.
[49]Should the inference be performed by utilizing the exact probability mass functions computed from the study of the Markov chain, $W_{12}(t)$ would be inferred to be 0.93, 0.00, 1.97, and 1.15, while for $W_{21}(t)$, we would obtain 0.91, 1.16, 0.00, and 2.70, as t goes from 0 to 3.
[50]L. Moniz, E. Cooch, S. Ellner, J. Nichols, and J. Nichols, Ecol. Modell. **208**, 145 (2007).