

UCLA

UCLA Previously Published Works

Title

Detecting causality in policy diffusion processes

Permalink

<https://escholarship.org/uc/item/5x8313zn>

Journal

Chaos An Interdisciplinary Journal of Nonlinear Science, 26(8)

ISSN

1054-1500

Authors

Grabow, Carsten
Macinko, James
Silver, Diana
[et al.](#)

Publication Date

2016-08-01

DOI

10.1063/1.4961067

Peer reviewed

Detecting causality in policy diffusion processes

Carsten Grabow,¹ James Macinko,² Diana Silver,³ and Maurizio Porfiri^{1,a)}

¹*Department of Mechanical and Aerospace Engineering, New York University, Tandon School of Engineering, Brooklyn, New York 11201, USA*

²*Department of Community Health Sciences and Department of Health Policy and Management, Fielding School of Public Health, University of California, 650 Charles Young Dr., Los Angeles, California 90095, USA*

³*Department of Nutrition, Food Studies, and Public Health, 411 Lafayette Street, New York University, Steinhardt School of Culture, Education, and Human Development, New York, New York 10003, USA*

(Received 11 May 2016; accepted 1 August 2016; published online 19 August 2016)

A universal question in network science entails learning about the topology of interaction from collective dynamics. Here, we address this question by examining diffusion of laws across US states. We propose two complementary techniques to unravel determinants of this diffusion process: information-theoretic union transfer entropy and event synchronization. In order to systematically investigate their performance on law activity data, we establish a new stochastic model to generate synthetic law activity data based on plausible networks of interactions. Through extensive parametric studies, we demonstrate the ability of these methods to reconstruct networks, varying in size, link density, and degree heterogeneity. Our results suggest that union transfer entropy should be preferred for slowly varying processes, which may be associated with policies attending to specific local problems that occur only rarely or with policies facing high levels of opposition. In contrast, event synchronization is effective for faster enactment rates, which may be related to policies involving Federal mandates or incentives. This study puts forward a data-driven toolbox to explain the determinants of legal activity applicable to political science, across dynamical systems, information theory, and complex networks. *Published by AIP Publishing.*

[<http://dx.doi.org/10.1063/1.4961067>]

Over the past three decades, US states have expanded the use of public health policies (laws, regulations, taxes) to facilitate the change of complex health and social behaviors. In spite of growing evidence of the effectiveness of individual policies, little is known about what motivates any given state to adopt these new policies, let alone aid in predicting state uptake of any new policy initiative. Dynamical systems, information theory, and complex networks may be particularly applicable to the analysis of policymaking among states, whereby they could help dissect dynamic interactions within a group of interconnected units (state governments) that are responsible for the production of an observable and emergent phenomenon (state health policy landscape). Here, we investigate two complementary techniques grounded on information theory and nonlinear time series analysis for the study of policy diffusion among US states. Both approaches are demonstrated on synthetic data, generated using a novel modeling framework that can be tuned to proxy important features of several real-world scenarios. This effort offers an unprecedented mathematical basis for an improved understanding of the determinants of policy diffusion across governmental units over time, laying the foundations for predictive modeling.

I. INTRODUCTION

Policy diffusion is a phenomenon that describes the spreading of public policies from one government to another.¹ Policy diffusion in American politics has been studied since the seminal paper of Walker² in 1969. Studies include a variety of different policy topics ranging from welfare reform,³ criminal justice,⁴ and public health^{5–7} to Native American gaming⁸ and school choice.⁹ Initially, most researchers focused on a single policy using event history analysis following the approach first introduced in a classic study on lottery adoptions.¹⁰ However, a number of new developments in the last decade, including key explanatory concepts,^{5,11} policy outcomes,^{12,13} and methods tailored for studying innovation,^{14,15} have allowed researchers to examine how states may learn from and influence one another. Seen in this form, policymaking can be modeled as an ensemble of individual units interacting with rules and resources, often with competing goals. With this representation in mind, we seek to understand how modern network analysis techniques could provide unique insights in terms of our understanding of the emergent properties of government systems. Accordingly, here, we further extend these approaches to unravel the topology of causal influences among states from their legal activity.

Understanding the relationship between topology and dynamics is a central research question in the fields of complex networks and dynamical systems.^{16–18} Networks are ubiquitous in nature and technology. For example, networks have been used to model interactions between friends and collaborators,^{19,20} patterns of gene expression,²¹ neuronal circuits in the brain,^{22–24} transportation systems,²⁵ power grids,²⁶ and

^{a)} Author to whom correspondence should be addressed. Electronic mail: mporfiri@nyu.edu

the internet.²⁷ Across these systems, local interactions among units influence the collective dynamics and the function of the entire network.

Due to rapid technological advancement, our knowledge about individual units in natural networks, such as the brain or the human genome, is continuously growing.^{23,28} But, apprehending the topology of the interaction networks remains an elusive problem. For example, the key constituents of cellular gene, protein, and metabolite networks are well characterized, but the pathways involved in their interaction are not fully elucidated.²¹ Similarly, in neuroscience, one can simultaneously measure the activity of at least one hundred neurons,²³ but when seeking to uncover synaptic connections of a neuronal circuit through anatomical methods, only individual synapses can be deciphered. Finally, while dyadic social interactions are relatively well understood, predicting the effect of complex networks on human behavior is an open question.^{20,29} Establishing reliable and effective methods to enable the reconstruction of interaction networks from local dynamical data may considerably advance several fields of investigation, by empowering researchers with refined data-driven descriptions of network structures underlying collective phenomena.

Here, we demonstrate the use of the recently introduced union transfer entropy (UTE)⁶ and event synchronization (ES)³⁰ to reconstruct network topologies underlying policy diffusion as understood in political science. UTE is an information-theoretic construct that we have recently proposed to elucidate causal relationships in the US legal landscape, in the form of information sharing between states toward the adoption or abrogation of a specific public health policy. Such a measure is implemented on the individual time series of legal activity of each state, by leveraging well-established tools in information theory, such as entropy and conditional mutual information.³¹ UTE differs from classical transfer entropy methods^{32–34} that have been successfully adapted to neuroscience,^{35–37} climate science,^{38–41} and animal collective behavior,^{42–45} due to its focus on slowly evolving dynamical systems. This method could be applied to the analysis of dynamical systems with rare and discrete events, in the presence of multidimensional time series, beyond the policy domain considered in this work.

ES was originally proposed to measure synchronization and time-delay patterns between signals.³⁰ This quantity should be considered a measure of similarity for event time series that takes into account varying delays and temporal

ordering. Recently, it was applied in the field of climate science to analyze the spatial structure of temporal similarity of extreme events at different locations. In particular, ES was successfully employed to investigate spatio-temporal patterns during the Indian summer monsoon⁴⁶ and to study the origin and propagation of extreme rainfall in South America.⁴⁷ The application of ES only requires the presence of extreme events that could be isolated from the individual time series, and is thereby suitable for the study of a number of other fields, including social behavior of animal groups, plasma, turbulence, cardiology, and brain research.^{45,48}

Table I synoptically compares selected methods used for network reconstruction, with emphasis on features such as implementation on directed networks to infer causality; applicability to discrete datasets similar to the public health policy data that we use in this paper; and code availability that would allow a fair comparison based on the original implementation. Other than UTE, which was specifically developed for discrete, slowly evolving dynamical systems, only ES seems amenable to such an analysis, whereby its input is a discrete binary series used to infer causality. In the category of transfer entropy, we include recent extensions that specifically address triplets, such as causation entropy³⁴ and conditional transfer entropy.⁴⁹ In Ref. 50, a new causality metric for datasets that do not have temporal information attached to them was proposed and implemented on synthetic and real datasets.

As a first, necessary step toward a systematic reconstruction of influences in the field of public policy, we posit a reduced order, minimalistic model for policy diffusion. The model is based on recent empirical observations in Ref. 56 and extends our previous work in Ref. 6, which focused on one-out regular,⁵⁷ cyclic ring-like topologies, where each state is influenced by only one neighboring state. Here, we seek to identify complex network topologies of causal influences. The model is motivated by so-called stochastic binary neurons in the context of theoretical neuroscience.⁵⁸

Synthetic data generated through the model are utilized to test the accuracy of the proposed network reconstruction techniques and offer a rigorous assessment of their potential, prior to implementation on real data. The accuracy is evaluated using a function of the receiver operating characteristic (ROC) curve, which we propose as an objective measure to capture the performance of network reconstruction. We begin the analysis with principal motifs and then extend our study to larger networks, in which individual units are

TABLE I. Select methods for network reconstruction compared on the basis of features relevant to slowly evolving discrete dynamical systems.

Method \ Feature	Non-linear	Model-free	Works with discrete data	Directed links	Code availability	Treatment of triplets	Select references
Cross-correlation	No	No	Yes	Yes	Yes	No	Refs. 45 and 51
Event synchronization	Yes	Yes	Yes	Yes	Yes	No	Refs. 30 and 45
Functional clustering algorithm	Yes	Yes	Yes	No	No	No	Ref. 52
Granger causality	No	No	Yes	Yes	Yes	No	Ref. 53
Mutual information	Yes	Yes	Yes	No	Yes	No	Ref. 54
Transfer entropy and recent developments	Yes	Yes	Yes	Yes	Yes	Yes	Refs. 32, 34, 45, 49, and 55
Union transfer entropy	Yes	Yes	Yes	Yes	Yes	No	Ref. 6

regularly or randomly interconnected. Finally, we demonstrate the application of our data-driven approaches to study potential determinants of policy diffusion across US states. Specifically, we generate two network topologies based on competing explanations of the process of policy diffusion (which is driven primarily by ideology or geography).⁷ From synthetic data, we demonstrate the possibility of precisely isolating each of these factors using our data-driven approaches.

The paper is organized as follows. We first succinctly review the data-driven measures used for network reconstruction, namely, UTE in Section II and ES in Section III. Our new model to generate legal activity data is presented in Section IV, grounded on our previous work in Ref. 6. Section V provides a background on the ROC analysis⁵⁹ used to estimate the performance of the applied methods. In Section VI, we present our results for small motifs along with larger regular and random networks. In addition, we investigate the possibility of isolating determinant factors driving policy diffusion using ES and UTE based on real-world examples. We conclude in Section VII with a summary of our main results and future work.

II. UNION TRANSFER ENTROPY FOR INFERRING CAUSALITY

To begin, we consider independent draws of a discrete random variable X described by a probability density function $p(x) = \Pr(X = x)$. The so-called Shannon entropy⁶⁰ is defined as

$$H_X = - \sum p(x) \log p(x), \quad (1)$$

where the summation is over all states x of the process and the base of the logarithm is chosen to be equal to 2, following the standard practice in information theory, see, for example, Ref. 31.

The mutual information of two random variables X and Y with joint probability $p_{XY}(x, y)$ can be calculated as

$$M_{XY} = \sum p(x, y) \log \frac{p_{XY}(x, y)}{p_X(x) p_Y(y)}, \quad (2)$$

where the notation directly follows from Eq. (1). Mutual information is a non-directional (symmetric under the exchange of X and Y) measure of the deviation from independence of two random variables.

The above information-theoretic constructs can be extended to stochastic processes. In a stationary Markov process of order k , the conditional probability of process X assuming state x_{n+1} at time step $n+1$ is independent of its state x_{n-k} at time step $n-k$, that is, $p(x_{n+1}|x_n, \dots, x_{n-k+1}) = p(x_{n+1}|x_n, \dots, x_{n-k})$. We use the shorthand notation $x_n^{(k)} = (x_n, \dots, x_{n-k+1})$ for words of length k . Given two stochastic processes X and Y , the so-called transfer entropy³² measures their deviation from the generalized Markov property, which reads $p(x_{n+1}|x_n^{(k)}) = p(x_{n+1}|x_n^{(k)}, y_n^{(k)})$. Thus, transfer entropy is given by

$$T_{X \rightarrow Y} = \sum p(x_{n+1}, x_n^{(k)}, y_n^{(k)}) \log \frac{p(x_{n+1}|x_n^{(k)}, y_n^{(k)})}{p(x_{n+1}|x_n^{(k)})}. \quad (3)$$

This quantity can be used to measure directed influences between two processes, that is, the degree of dependence of process X on process Y , but not vice versa, since it is not symmetric.

Next, we consider a networked dynamical system composed of N distinct stochastic processes X^i with $i \in \{1, \dots, N\}$, each taking the form of a discrete binary-valued time series of length T , namely, x_1^i, \dots, x_T^i . With respect to our target application in policy diffusion, each x_n^i encodes the law activity of state i on day n . Specifically, x_n^i takes value 1 if a law was implemented, repealed, or experienced material change, and it is 0 otherwise. Without loss of generality, we focus on legal activity corresponding to a specific law; the approach can be readily extended to examine laws in different domains following Ref. 6.

We introduce the so-called union transfer entropy⁶ to estimate the magnitude and direction of influences among states from either real or simulated law activity. Union transfer entropy $I_{X^j \rightarrow X^i}$ measures the directed influence that state j has on state i , and is given by

$$I_{X^j \rightarrow X^i} = \sum p(x_{n+1}^i, \hat{x}_{n,\tau}^i, \hat{x}_{n,\tau}^j) \log \frac{p(x_{n+1}^i | \hat{x}_{n,\tau}^i, \hat{x}_{n,\tau}^j)}{p(x_{n+1}^i | \hat{x}_{n,\tau}^i)}, \quad (4)$$

where the summation is taken over all possible realizations of X^i and X^j and

$$\hat{x}_{n,\tau}^i = \left| \bigcup_{r=\max\{n-\tau, 1\}}^{n-1} \{x_r^i\} \right| - 1, \quad (5a)$$

$$\hat{x}_{n,\tau}^j = \left| \bigcup_{r=\max\{n-\tau, 1\}}^{n-1} \{x_r^j\} \right| - 1, \quad (5b)$$

are scalar quantities which are obtained from legal activity of states i and j . They indicate whether or not any law activity occurred in state i and state j , respectively, within a sliding time sliding window of length τ .

These quantities are used together with the binary random variable associated with the law activity of state i at time n to offer an empirical, probabilistic description of whether or not legal activity tends to occur in one state on a particular day with respect to any recent law activity in both states. Specifically, we compute the following joint probability density function

$$p(x_{n+1}^i, \hat{x}_{n,\tau}^i, \hat{x}_{n,\tau}^j) = \frac{1}{T} \sum_{n=1}^T x_{n+1}^i \hat{x}_{n,\tau}^i \hat{x}_{n,\tau}^j. \quad (6)$$

Considering the unions of law activity $\hat{x}_{n,\tau}^i$ and $\hat{x}_{n,\tau}^j$ over the previous τ days allows for modeling processes with very low event rates, where only a limited number of events occur in the selected time window. Based on the time span required

for states' law activities to influence each other,⁶ we choose window sizes of $\tau \in \{2, 3, 4, 5\}$ years.

Averaging over these window sizes guarantees an amplification of short time delays between different states' activities: events included in smaller time windows are automatically included in larger windows as well. The data aggregation in the window τ is the main difference between UTE and transfer entropy in Eq. (3). The possibility of consolidating different law domains in a single time series is another difference between the two methods, which is however not detailed in our systematic analysis of synthetic data.

In order to incorporate these joint distributions into the calculation of union transfer entropy, we rewrite Eq. (4) as the difference of two conditional entropies,³¹ that is,

$$I_{X^j \rightarrow X^i} = H_{X^i | \hat{X}^i} - H_{X^i | \hat{X}^i, \hat{X}^j}, \quad (7)$$

whereby the two conditional entropies are computed as

$$H_{X^i | \hat{X}^i} = - \sum p(x_{n+1}^i | \hat{x}_{n,\tau}^i) \log p(x_{n+1}^i | \hat{x}_{n,\tau}^i), \quad (8a)$$

$$H_{X^i | \hat{X}^i, \hat{X}^j} = - \sum p(x_{n+1}^i | \hat{x}_{n,\tau}^i, \hat{x}_{n,\tau}^j) \log p(x_{n+1}^i | \hat{x}_{n,\tau}^i, \hat{x}_{n,\tau}^j). \quad (8b)$$

These two conditional entropies can be calculated from the empirical joint distribution in Eq. (6) by applying pertinent marginalizations and conditional probability formulae, see, for example, Ref. 61. The first conditional entropy in Eq. (8a) can be interpreted as the uncertainty in the prediction of the law activity in state i , given the union of all law activity over a historical period in that state. The second conditional entropy in Eq. (8b) refers to the union of historical law activity of another state.

Thus, $I_{X^j \rightarrow X^i}$ can be interpreted in the following way: if the uncertainty of the prediction of the law activity of state i on day n , x_n^i , given the union of all recent law activity of that state i , is reduced by conditioning on the union of all law activity of state j , then state j is said to have influence over state i . Therefore, the notion of union transfer entropy may be associated with a form of predictive causality. To summarize, $I_{X^j \rightarrow X^i}$ is directed (asymmetric under the exchange of states, such that $I_{X^j \rightarrow X^i} \neq I_{X^i \rightarrow X^j}$ in general). A larger value of $I_{X^j \rightarrow X^i}$ implies that knowledge of historical law activity of state j improves the prediction of activity of state i . A smaller value of $I_{X^j \rightarrow X^i}$ indicates the independence of state i 's law activity from state j 's. We remark that the proposed measure is tailored to pairwise interactions³³ which afford the explicit treatment of higher order interactions but may require more dense datasets.

III. DIRECTED EVENT SYNCHRONIZATION FOR QUANTIFYING SYNCHRONICITY

We consider two processes of length T , encapsulating the time series of law activities for two states i and j , given by x_n^i and x_n^j . Next, we define "useful" events in the form of implementation, repeal, or experiencing a material change of a law in state i ($x_n^i = 1$) and in state j ($x_n^j = 1$). The

corresponding event times are denoted by t_r^i and t_s^j with $r \in \{1, \dots, m_i\}$ and $s \in \{1, \dots, m_j\}$, where m_i and m_j denote the number of events that occur in the corresponding time series.

In the following, we count the fraction of event pairs that match in time, such that a pair of totally identical time series leads to a synchronicity value equal to one. By way of example, we consider two event times t_r^i and t_s^j , with $t_s^j \leq t_r^i$ and $0 \leq r, s \leq \max\{m_i, m_j\}$. Since the laws enacted in different states are rare events and their rates may change over time, we introduce a local dynamic definition of the time scale τ_{rs} to decide the association between two events. Specifically, we compute the dynamical delay for each event pair (r, s) as

$$\tau_{rs} = \min\{t_{r+1}^i - t_r^i, t_r^i - t_{r-1}^i, t_{s+1}^j - t_s^j, t_s^j - t_{s-1}^j\} / 2, \quad (9)$$

that is, the minimum of the differences of pre- and post-event times as depicted in Fig. 1. Delays are further filtered by selecting a minimum and maximum allowable delay (τ_{\min}, τ_{\max}) between t_r^i and t_s^j . In this study, we choose $\tau_{\min} = 0$ and $\tau_{\max} = 5 \times 365$ days, in line with our assumptions for UTE and that a year corresponds to 365 days.

The number of times an event occurs in process X^i shortly after an event occurs in X^j can then be calculated as

$$c(X^i | X^j) = \sum_{r,s} S_{rs}, \quad (10)$$

where

$$S_{rs} = \begin{cases} 1 & \text{if } 0 \leq \tau_{\min} < t_r^i - t_s^j \leq \tau_{rs} \leq \tau_{\max} \\ 1/2 & \text{if } t_r^i = t_s^j \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, $c(X^j | X^i)$ counts the number of times an event occurs in X^j shortly before an event in X^i . Originally, the symmetric and asymmetric quantities

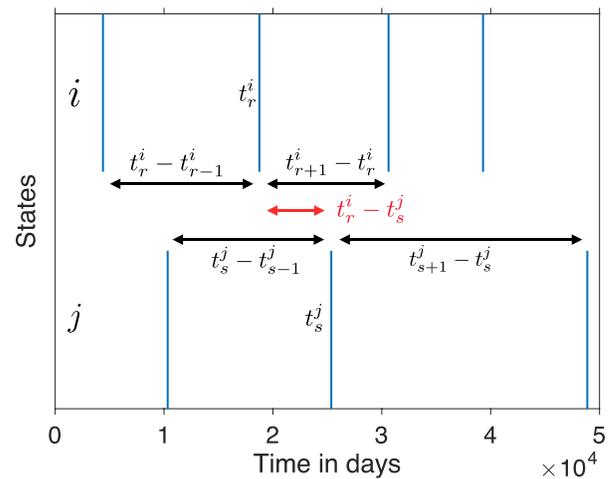


FIG. 1. Inter-event times involved in calculating the dynamical delay for ES between law activities of states i and j . Two events at t_r^i and t_s^j can be uniquely interpreted as synchronized if their time lapse is smaller than the dynamical delay, that is, the minimum of time lapses to preceding and succeeding events divided by two, as defined in Eq. (9).

$$Q = \frac{c(X^j|X^i) + c(X^i|X^j)}{\sqrt{(m_i - 2)(m_j - 2)}}, \quad (11)$$

and

$$q = \frac{c(X^j|X^i) - c(X^i|X^j)}{\sqrt{(m_i - 2)(m_j - 2)}}, \quad (12)$$

were introduced in Ref. 30 to capture the event synchronization and event delay between the pair of time series. In particular, $0 \leq Q \leq 1$, where $Q = 1$ implies that the events are completely synchronous, and $-1 \leq q \leq 1$, where $q = 1$ implies that events in X^i always occur prior to those in X^j . To facilitate the process of network reconstruction, where pairwise interactions between states should be compared across the network, we define the directed event synchronization between two time series X^i and X^j as

$$Q_{X^i \rightarrow X^j} = \frac{c(X^j|X^i)}{\sqrt{(m_i - 2)(m_j - 2)}}. \quad (13)$$

For a pair of identical time series, $Q_{X^i \rightarrow X^i} = 1/2$ and $Q_{X^i \rightarrow X^i} = 1/2$. For events in i that always anticipate events in j , our measure is identical to the originally introduced event delay in Eq. (12). By calculating directed event synchronization for all pairs of states i and j , we define a matrix $Q_{X^i \rightarrow X^j}$, from which we infer influences among states on a slow time-scale. For low event rates, inter-event times for synchronized events might be in the order of years and direction of influence associated with chronology of events is of particular importance.

IV. MODELING LAW ACTIVITY

Recalling our previously introduced notation, the output of state i is described by a binary stochastic variable

$X_n^i \in \{0, 1\}$, which defines whether a law is implemented, repealed, or substantively changed at time n , that is, $X_n^i = 1$, and we refer to it as experiencing a law event, or not, that is, $X_n^i = 0$. The law activity of state i is given by the time series x_1^i, \dots, x_T^i .

In the absence of any interactions with neighboring states, the dynamics of law adoption of a given state are fully defined by the probability that X_{n+1}^i takes value one. This probability is, in turn, proportional to the law event rate of the state and is given by

$$\Pr(X_{n+1}^i = 1) = \Theta_i, \quad (14)$$

where Θ_i is the law event rate of state i .

Interactions between states are encoded by a directed network of N nodes i with $i \in \{1, \dots, N\}$, whose topology is described by an adjacency matrix A . Its matrix elements are zeros and ones to specify which states influence others and which do not, such that

$$A_{ij} = \begin{cases} 1 & \text{if state } j \text{ influences state } i \neq j \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

An exemplary network along with the model mechanisms involved is depicted in Fig. 2. For state i in this network, A_i represents the total input and consists of contributions A_{ij} from all neighboring states that have recently (since state i had the last law activity) had a law event. The time of the last law event of state i is a crucial quantity in our model, namely,

$$\tau_n^i = \max\{\ell \leq n : X_\ell^i = 1\}. \quad (16)$$

To encode whether or not law activity occurred in the previous τ_n^i days following day n , we recall the definition of the scalar quantities from Eqs. (5a) and (5b)

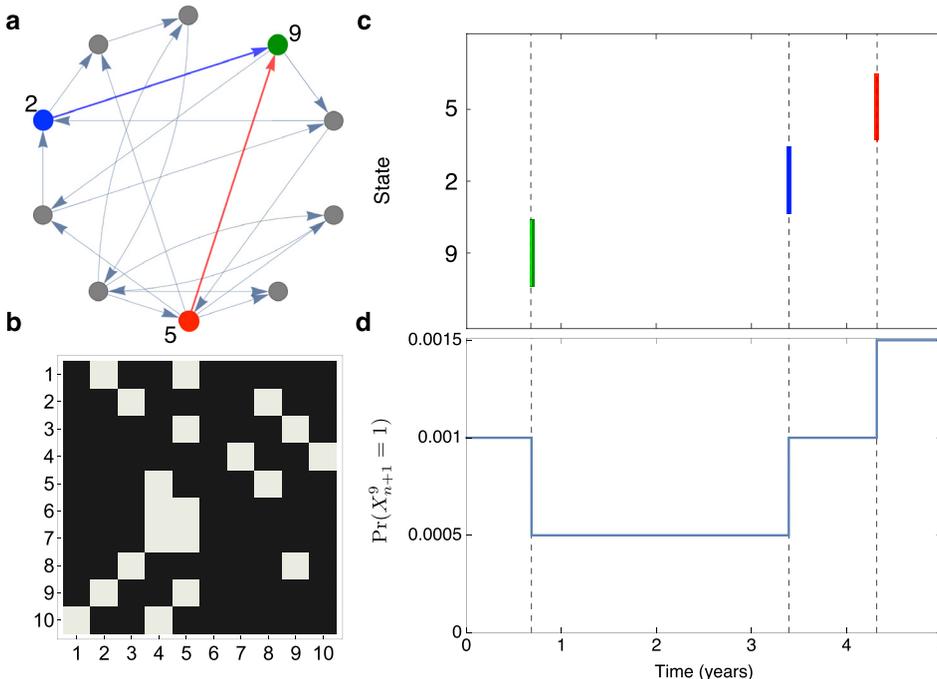


FIG. 2. Network of influences among states determines their law activity dynamics. (a) An exemplary network of $N = 10$ nodes and in-degree $k = 2$ for all nodes. State 9 (green) receives incoming links from state 5 (red) and state 2 (blue). (b) The adjacency matrix A defined in Eq. (15) encodes the network structure displayed in (a), with light gray color indicating the presence of a link. (c) Law activity of state 2 (blue bar) and state 5 (red bar) follows the last law event of state 9 (green bar). (d) Law events in the neighboring states 2 and 5 (blue and red bar in (c)) increase the law event probability of state 9. Law activity was generated for $\Theta = 0.0005$ and $\gamma = 1$.

$$\hat{x}_{n,\tau_n}^j = \left| \bigcup_{r=\max\{n-\tau_n, 1\}}^{n-1} \{x_r^j\} \right| - 1, \quad (17)$$

that now evaluates whether law activity has happened in state j since state i has had its last law event. We define the input that state i receives at time step n as

$$A_{i,n} = \sum_{j \neq i} A_{ij} \hat{x}_{n,\tau_n}^j, \quad (18)$$

and hypothesize that the probability that state i has a law event at time step $n+1$ increases linearly with its input $A_{i,n}$. Thus, the law activity of any state is mediated by both its inherent behavior and its interactions with neighboring states. The numbers of neighboring states that influence state i and the number of those that are influenced by state j are encoded by the in-degree k_i^{in} and the out-degree k_j^{out} , given by

$$k_i^{\text{in}} = \sum_j A_{ij}, \quad k_j^{\text{out}} = \sum_i A_{ij}. \quad (19)$$

The degree k_i of node i is defined as the sum of in- and out-degree.

To simplify the analysis, we assume that spontaneous enactment rates are the same for all states, that is, $\Theta_i = \Theta$ for any state i . Thus, we propose the following model for the probability of any law activity in state i at the next time step

$$\Pr(X_{n+1}^i = 1) = \gamma \Theta \sum_{j \neq i} A_{ij} \hat{x}_{n,\tau_n}^j + \Theta, \quad (20)$$

where γ is a scaling factor that defines the relative weight of interactions with respect to spontaneous activity. The first summand on the right hand side of Eq. (20) quantifies the probability of enacting a new law exclusively through network interactions, while the second summand quantifies the probability of spontaneous law enactment. For the case of in-degree k_i^{in} equal to zero, the first summand is zero and the law activity of state i is exclusively determined by the second summand. Different values of the policy adoption frequency Θ and the intensity of state-to-state diffusion γ correspond to different real-world scenarios as shown in Table II. While the model in Eq. (20) uses a linear interaction function,

alternative selections could be contemplated to model nonlinearities, including saturation phenomena.⁶²

Equation (20) for the probability of law activity defines the whole model dynamics, since the remaining probabilities for no law activity $\Pr(X_{n+1}^i = 0)$ simply follow as $\Pr(X_{n+1}^i = 0) = 1 - \Pr(X_{n+1}^i = 1)$ for each state. All parameters should be chosen such that $\Pr(X_{n+1}^i = 1)$ and $\Pr(X_{n+1}^i = 0)$ are within $[0, 1]$. This condition is readily fulfilled as we study rare events in the realm of policy diffusion for which $\Pr(X_{n+1}^i = 1) \ll 1$. Specifically, law activity in the public health law domain based on real data was found to be in the order of magnitude of $\Pr(X_{n+1}^i = 1) \approx 0.0001$, see Ref. 6.

V. DETECTING DETERMINANTS OF STATES' LEGAL ACTIVITY

To establish an understanding of both methods, ES and UTE, we quantify the robustness of the reconstruction through a ROC analysis.⁵⁹ The ROC is a parametric curve that quantifies relationship between the true and the false positive links found in the corresponding binary matrices for different thresholding, associated with the number of links in the network. The use of ROC to quantify the accuracy of network reconstruction is a contribution of this study and builds on prior work on the assessment of leader–follower relationships in coupled dynamical systems, describing the social behavior of zebrafish.⁴⁵ In order to estimate each method's ability to detect causal patterns of the underlying network structures, we compute UTE and ES for each directed pair of states i and j from Eq. (4) for the former method and Eq. (13) for the latter. This process results in the matrices $Q_{X^i \rightarrow X^j}$ for directed synchronicity and $I_{X^i \rightarrow X^j}$ for directed causal influences based on UTE.

Afterwards, we construct two binary matrices $T^{\text{UTE}}(z)$ and $T^{\text{ES}}(z)$, respectively, based on the UTE and ES scores. Here, the corresponding matrix elements take the value 1 if they are among the fraction z of links with the highest scores and 0 otherwise. Thus, a fraction z equal to 1 leads to the maximum number of $N(N-1)$ ones. As the matrix A in Eq. (20) denotes the binary connectivity matrix of the real network, then the true positive rates $R_{\text{tp}}^{\text{UTE}}$ and $R_{\text{tp}}^{\text{ES}}$ are defined as the numbers of links in $T^{\text{UTE}}(z)$, or $T^{\text{ES}}(z)$, respectively, that are present in A with respect to the total number of existing links. The false positive rates $R_{\text{fpr}}^{\text{UTE}}$ and $R_{\text{fpr}}^{\text{ES}}$ are the

TABLE II. Real-world scenarios that can be simulated by differentially selecting the system's parameters of policy adoption frequency Θ and intensity of state-to-state diffusion γ . These associations are hypothesized on the basis of previous studies on policy diffusion.^{63–65}

	High γ	Low γ
High Θ	Federal incentives to adopt a set of new laws. Policy diffusion is based on pilot experiences in leading states. Example: Provisions of the Affordable Care Act	Federal pressure and/or incentives to adopt a set of new laws. Due to time pressures and unique needs, every state moves ahead more or less independently. Example: Bioterrorism preparedness
Low Θ	Policies that have a high start-up cost and unknown, highly targeted, or only long-term benefits. States wait for early adopters to demonstrate effectiveness (and political support). Leading states provide evidence for the political/economic or overall effectiveness of the policy, eventually stimulating diffusion. Example: Medical marijuana laws	Policies that attend to specific local problems that occur only rarely, or policies that face high levels of opposition. Example: Localized disasters might trigger policy adoption around specific risk factors that would not inform policy development in other contexts

fractions of links in $T^{\text{UTE}}(z)$ or $T^{\text{ES}}(z)$, respectively, which do not match the original links. Thus, the true positive and false negative rates for UTE can be defined as

$$R_{\text{tpr}}^{\text{UTE}}(z) = \frac{\sum_{i,j} T_{ij}^{\text{UTE}}(z) A_{ij}}{\sum_{i,j} A_{ij}}, \quad (21a)$$

$$R_{\text{fpr}}^{\text{UTE}}(z) = \frac{\sum_{i,j} T_{ij}^{\text{UTE}}(z) (1 - A_{ij})}{\sum_{i,j} (1 - A_{ij})}, \quad (21b)$$

and the ones for ES as

$$R_{\text{tpr}}^{\text{ES}}(z) = \frac{\sum_{i,j} T_{ij}^{\text{ES}}(z) A_{ij}}{\sum_{i,j} A_{ij}}, \quad (22a)$$

$$R_{\text{fpr}}^{\text{ES}}(z) = \frac{\sum_{i,j} T_{ij}^{\text{ES}}(z) (1 - A_{ij})}{\sum_{i,j} (1 - A_{ij})}. \quad (22b)$$

This means that the true positive rates are estimates of the probabilities for reconstructing a link that is actually present in the real network, while the false positive rates are estimates of the probabilities for reconstructing a link that is not. Note that we have a zero denominator for either empty or complete networks. We exclude these trivial cases from our reconstruction. In the following, we apply this analysis to

different network structures encoded in the connectivity matrix A .

We illustrate our analysis in the case of the directed sample network depicted in Fig. 2 before we systematically investigate the methods' performance as functions of the main model parameters Θ and γ . Stochastic law activity is generated based on Eq. (20) with the connectivity matrix A chosen as shown in Fig. 2. As described above, we then apply ES and UTE to generate weighted matrices as depicted in Figs. 3(c) and 3(d). Through thresholding, we construct ROC curves that correspond to different realizations of the generated law activity. The thresholds are nonlinearly related to the fractions of links with the highest scores $z \in [0, 1]$.

We quantitatively compare classifiers by reducing ROC performance to a single scalar value representing the expected performance. A common method is to calculate the area under the ROC curve (AUC).^{66,67} AUC is a measure in $[0, 1]$ with 1 being the optimal performance and 0.5 chance. The perfect classifier is thus the rectangular curve yielding unit area with the top left corner denoting the perfect classification with false positive rate equal to 0 and true positive rate equal to 1. For matrices $I_{X^j \rightarrow X^i}$ and $Q_{X^j \rightarrow X^i}$ with zero entries only, that is, no dependence between time series could be detected, the performance takes value 0. We estimate the AUC-value by averaging over 10 different realizations, similar to the exemplary ROC curves in Fig. 3.

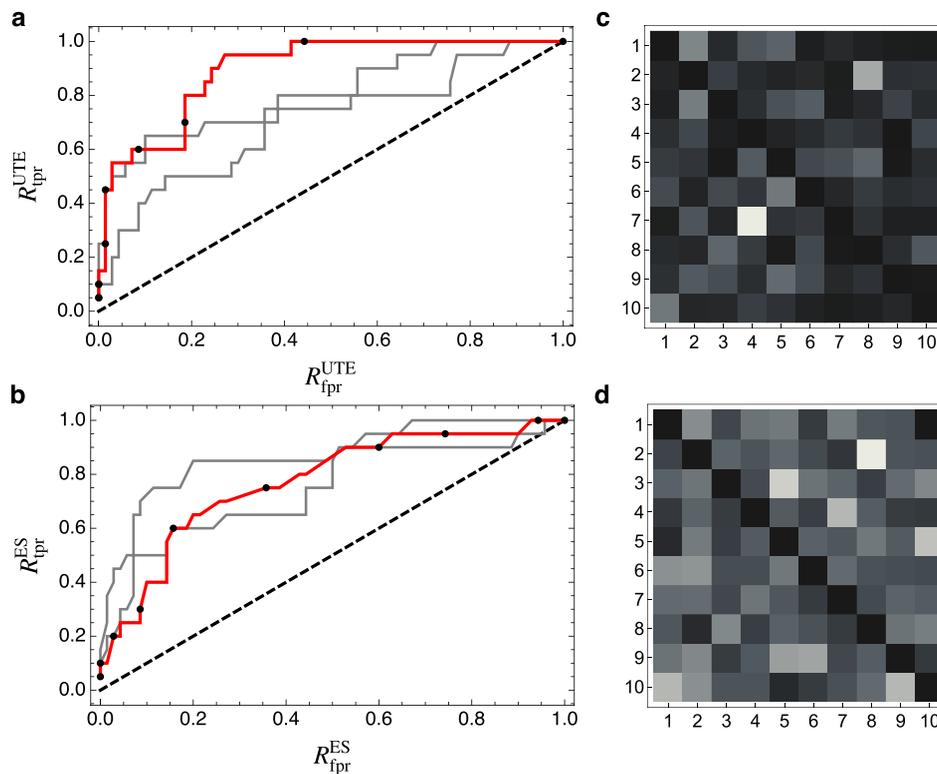


FIG. 3. ROC curves for (a) UTE and (b) ES applied to the exemplary network in Fig. 2(a). (a) Exemplary ROC curves for UTE associated with different realizations of the law activity dynamics follow from Eqs. (21a) and (21b). Each tenth data point on the red curve is marked in black; from right to left, they correspond to the values $z = 1, 43/45, 71/90, 2/3, 4/9, 23/90, 2/15, 1/15, 1/45, 1/90$. (b) Exemplary ROC curves for ES associated with different realizations of the law activity dynamics follow from Eqs. (22a) and (22b). Each tenth data point on the red curve is marked in black; from right to left, they correspond to the values $z = 1, 17/30, 3/10, 1/5, 1/9, 1/15, 1/45, 1/45, 1/90, 1/90$. (c) Exemplary matrix $I_{X^j \rightarrow X^i}$ for directed causal influences based on UTE as defined in Eq. (7) that leads to the curve highlighted in red in (a) via thresholding. Matrix values range from 0 (black) to 0.01 (white). (d) Exemplary matrix $Q_{X^j \rightarrow X^i}$ for directed synchronicity based on ES as defined in Eq. (13) that leads to the curve highlighted in red in (b) via thresholding. Matrix values range from 0 (black) to 0.5 (white). Law activity was generated over a time span of 50 000 days and with parameters $\Theta = 0.00005$ and $\gamma = 20$.

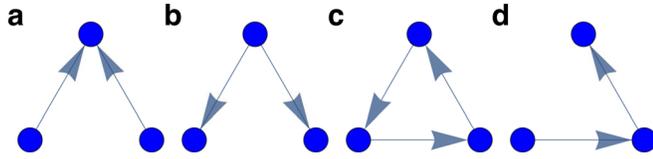


FIG. 4. Motif networks characterize four principal cases: (a) two states both influence a common neighbor; (b) two states are both influenced by a common neighbor state; (c) three states influence each other in a cycle; and (d) three states influence each other in a chain.

We first find the maxima of matrices $Q_{X_i \rightarrow X_j}$ for directed synchronicity based on ES and $I_{X_i \rightarrow X_j}$ for directed causal influences based on UTE. For each ROC curve, we then create a vector of 100 linearly spaced points between 0 and the corresponding maximum. For each of these threshold values, we calculate the true positive and false negative rates in Eqs. (21a) and (21b) for UTE and in Eqs. (22a) and (22b) for ES, respectively. Based on this procedure, ROC curves are generated. Each tenth z -value is highlighted in Figs. 3(a) and 3(b). Here,

$z = 1$ corresponds to a true positive rate equal to 1 and simultaneously to a false positive rate equal to 1. A z -value of 0 corresponds to true and false positive rates both equal to zero.

VI. RESULTS

A. Motif networks

To systematically investigate the performance of ES and UTE, we start with motifs as depicted in Fig. 4. We generate data for different Θ -values and γ -values over a time of $T = 50$ years. In the simulations, we neglect results for an initial time span of $T/10$ years, that is, for the first 5 years, in order to exclude transient effects. Furthermore, we generate 10 different stochastic law activity data for each pair of parameters, keeping the underlying network encoded in the adjacency matrix A fixed. We then average over these realizations to obtain the mean AUC-value corresponding to each (Θ, γ) pair. To generate the plots, we choose Θ -values within the interval $[10^{-5}, 10^{-2}]$ and γ -values within the

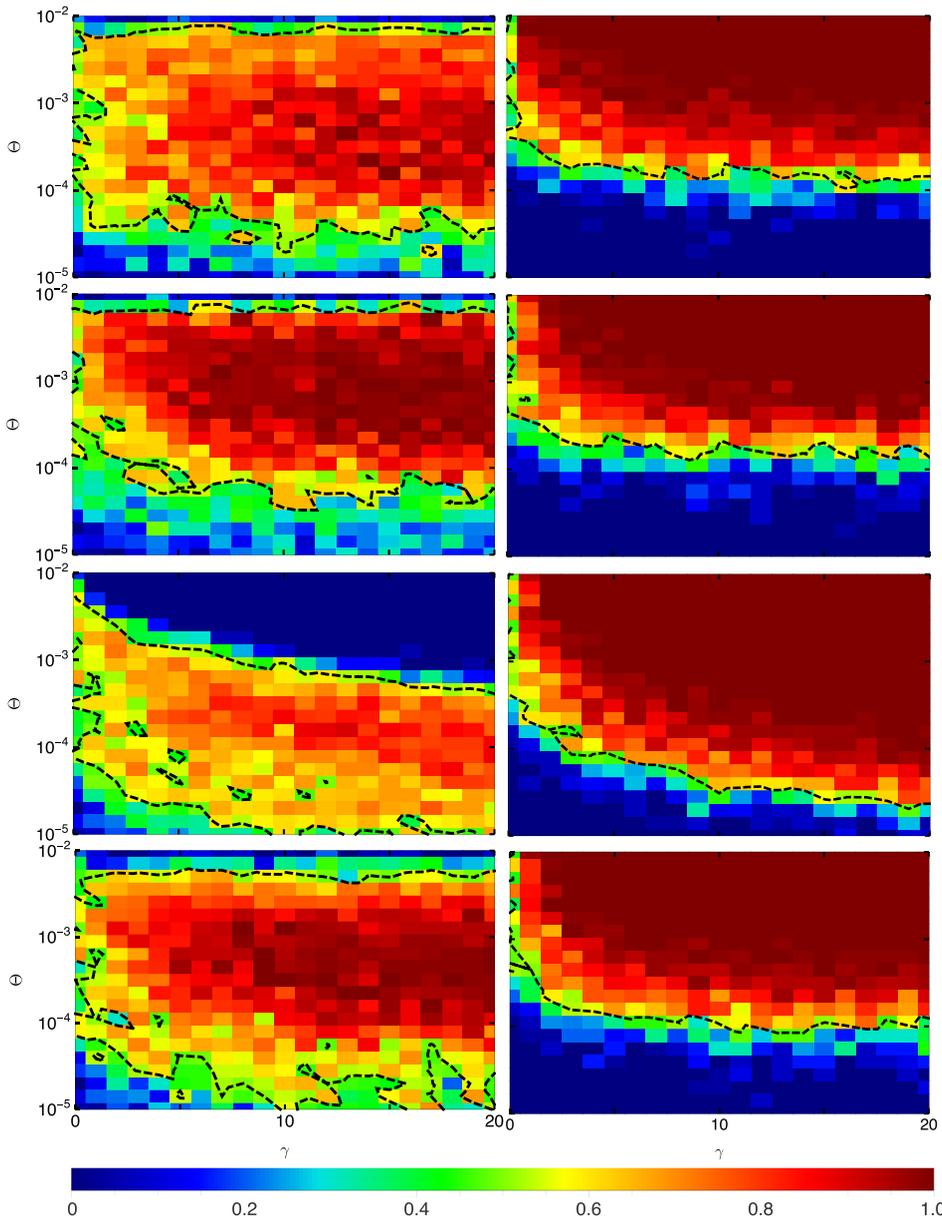


FIG. 5. Performance of UTE (left column) and ES (right column) for the four motifs in Fig. 4 (top row: motif (a); second row: motif (b); third row: motif (c); and bottom row: motif (d)). All AUC-values are averaged over 10 realizations. Contours highlighted in black identify the area where our methods perform better than chance.

interval $[0, 20]$. Based on these intervals, we create vectors of evenly spaced data values: for Θ , we generate a vector of 21 logarithmically spaced points and for γ a vector of 21 linearly spaced points.

The results for the four motifs are illustrated in Fig. 5. Therein, we show AUC-values as functions of the select control parameters, with red identifying exact reconstruction ($\text{AUC} = 1$) and green chance ($\text{AUC} = 0.5$). Thus, AUC-values below 0.5 characterize a performance worse than randomly predicting original links. As expected, the performance of both methods consistently increases with γ , whereby both methods benefit from a stronger relative weight of the states' interactions with respect to their spontaneous activity. However, increasing the spontaneous activity encoded through Θ has a differential role on the performance of the two methods. While ES seems to benefit from higher levels of legal activity, the performance of UTE is non-monotonically related to Θ . Moderate values of Θ enhance the efficacy of UTE in reconstructing the motif, but dense legal activities tend to hamper network reconstruction. This is likely related to the excessive density of the datasets, which may mask causal relationships between the time series. In this case, UTE fails to detect any causal relationship, leading to a null matrix $I_{X_i \rightarrow X_j}$. Of particular interest is the motif with directed links (Fig. 4(d)), where non-existent links may be generated due to conditional dependence. Our results in Fig. 5 show that this is not the case, since our methods predominantly detect pair-wise interactions. The reconstruction accuracy for both methods, UTE and ES, is very similar to the one achieved for motifs (a) and (b) in Fig. 4.

These data offer empirical evidence for the possibility of detecting complex interactions underlying US legal landscape using the proposed data-driven approaches. Comparing the performance of both methods, we note that UTE tends to be more effective than ES for very low event rates of the order of 10^{-5} – 10^{-4} . This is particularly evident for the ring network (motif (c)), where UTE is successful in detecting the underlying network structure for much smaller values of Θ , confirming the effectiveness of this approach in unraveling pairwise interactions in slowly evolving dynamical systems.⁶ Even for faster event rates, UTE tends to be more effective than ES across the four considered topologies, until the event rate reaches 10^{-2} and the information-theoretic approach fails to isolate causal relationships.

B. Comparing regular and random network topologies

Next, we investigate the performance of the two approaches in the reconstruction of larger networks. To shed

light on the role of network regularity on the performance of the methods, we examine a directed regular and two random networks generated using the original small-world model⁶⁸ adapted to directed networks.⁶⁹ For this purpose, we start with a graph of N nodes on a ring lattice. Each node is accessed by its k nearest neighbors on the left side. To study the impact of heterogeneity in the in-degree and out-degree distributions as defined in Eq. (19), we introduce randomness in two ways: first, we cut all the tails of the edges and then rewired them randomly, keeping in-degrees k_i^{in} constant. Second, we reverse the direction of all the rewired edges, obtaining constant out-degrees k_i^{out} .

To begin, we consider the case $k = 1$ and we focus on the exemplary networks shown in Fig. 6. For each network, we run model simulations for Θ -values in the range of $[10^{-5}, 10^{-2}]$ with 21 logarithmically spaced points and for γ -values in the range of $[0, 20]$ with 21 linearly spaced points. For each (Θ, γ) pair, data over a time of 50 years are generated using Eq. (20), whereby an initial time span of 5 years, that is, 1825 days, is discarded to exclude transient effects. The network topology given by the adjacency matrix A in our model Eq. (20) is kept fixed, and 10 different realizations of stochastic law activity data are generated.

Results for the two proposed methods are presented in Fig. 7. Therein, the mean AUC-values are averages over 10 time series generated for each network. Interestingly, as long as the in-degree remains constant for each node (upper and middle row), neither of the two methods seems affected by the network topology such that the event rate Θ and the coupling γ is nearly identical for the regular and randomized networks. This is due to the fact that the occurrence of events is dominated by the input each state receives as evident in Eq. (20), and the number of outgoing connections plays only a secondary role. Keeping the out-degree constant and varying the distribution of in-degrees instead (bottom row) improves the performance of UTE for very low event rates of the order of 10^{-5} and simultaneously for large event rates of the order of 10^{-2} , thus covering a larger region of successful causality detection. On the contrary, ES displays a consistent performance for all three cases and does not seem to be affected by heterogeneities in in- and out-degrees.

Similar to the analysis of motifs, both UTE and ES enable network reconstruction in a wide range of the parameter space. Both methods benefit from enhancing the relative weight of interactions, whereby the AUC-values increase with γ . For slow event rates, UTE tends to outperform ES, leading to exact network reconstruction for event rates as slow as 10^{-5} . But as the event rate increases above 10^{-3} , ES

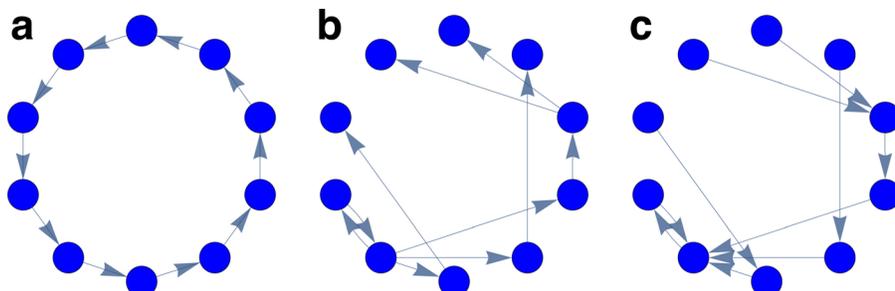


FIG. 6. Three networks with $N = 10$ nodes: (a) a regular network; (b) a random network with constant in-degree $k^{\text{in}} = 1$; and (c) a random network with constant out-degree $k^{\text{out}} = 1$.

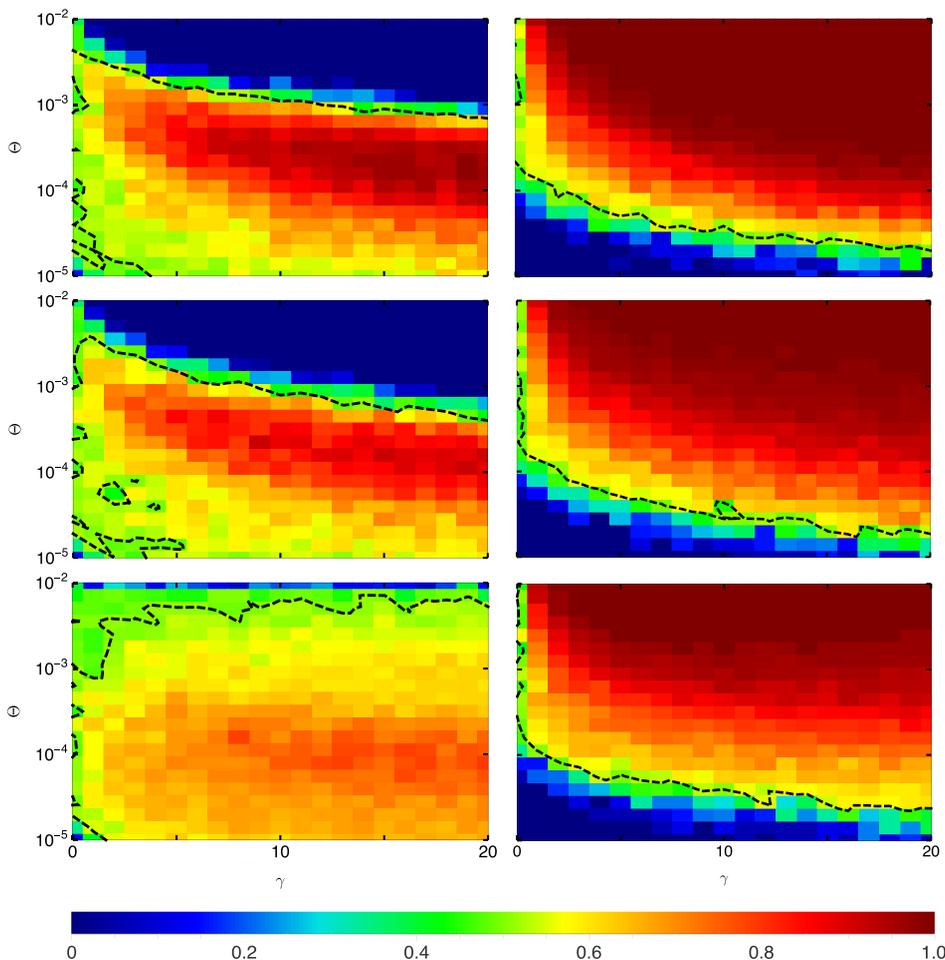


FIG. 7. AUC scores of UTE and ES for the networks depicted in Fig. 6 as functions of the parameters Θ and γ : a ring network (upper row) and randomized networks with constant in-degree (middle row) and out-degree (bottom row). All AUC-values are averaged over 10 realizations. Contours highlighted in black identify the area where our methods perform better than chance.

is capable of accurately inferring the network topology, while UTE fails to detect network links.

Next, in order to study the impact of link density on the methods’ performance, we consider the case of $k = 3$ and we focus on the exemplary networks shown in Fig. 8. For each network, we again run model simulations for a 21×21 grid of Θ -values logarithmically spaced in $[10^{-5}, 10^{-2}]$ and γ -values linearly spaced in $[0, 20]$. For each (Θ, γ) pair, we generate data based on Eq. (20) over a time of 50 years, discarding an initial time of 5 years to exclude transient effects. We generate 10 realizations of stochastic law activity data with our model Eq. (20) while keeping the networks as depicted in Fig. 8 fixed.

Results for UTE and ES are presented in Fig. 9 where the mean AUC-values are averaged over 10 realizations of stochastic time series data. The increased link density improves the performance of ES, whereby we observe a narrower

parameters’ region in which network reconstruction is not feasible (right column). Specifically, ES is successful in reconstructing networks for any choice of the model parameters, except for a small region where Θ and γ are smaller than 10^{-4} and 5, respectively. In contrast, UTE performance decreases in all three cases (left column), whereby we observe a significant reduction in the maximal AUC-values to approximately 0.7, accompanied by a wider region of parameter values in which network reconstruction is not possible. However, UTE is successful in reconstructing network topologies for small values of Θ and γ , in contrast with ES.⁷⁰

These findings suggest that UTE should be preferred for slow event rates, which do not lead to sufficiently dense datasets to enable the applicability of ES. However, UTE is computationally more challenging to implement and suffers from increased link densities, which instead seem to benefit the feasibility of ES.

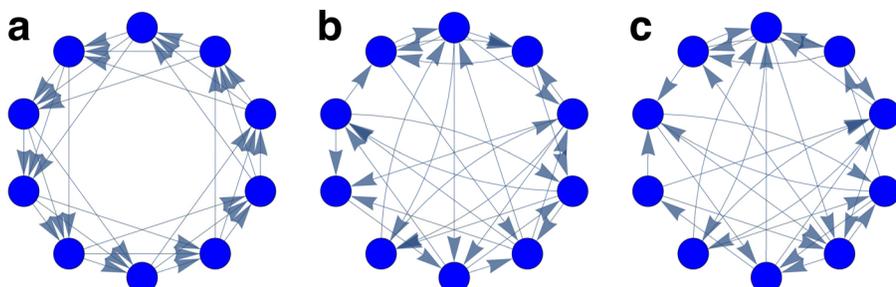


FIG. 8. Three networks with $N = 10$ nodes: (a) a regular network; (b) a random network with constant in-degree $k^{\text{in}} = 3$; and a random network with constant out-degree $k^{\text{out}} = 3$.

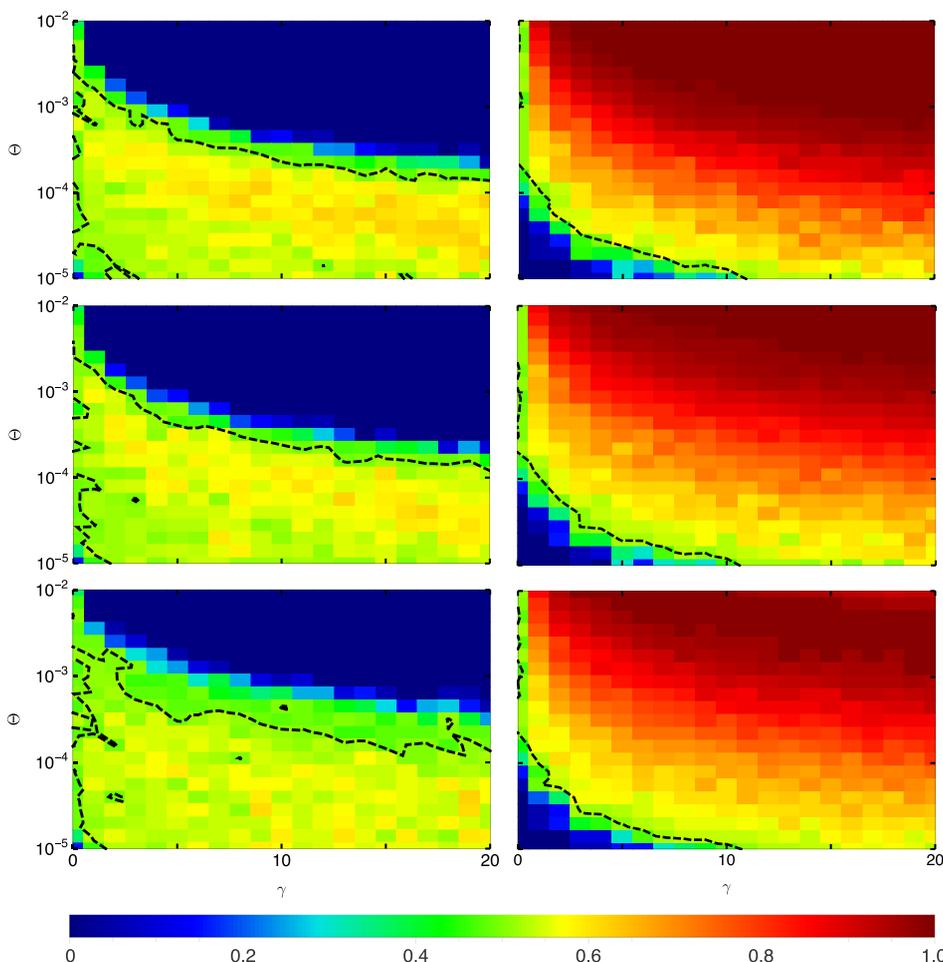


FIG. 9. AUC scores of UTE and ES for the networks depicted in Fig. 8 as functions of the parameters Φ and γ : a ring network (upper row) and randomized networks with constant in-degree (middle row) and out-degree (bottom row). All AUC-values are averaged over 10 realizations. Contours highlighted in black identify the area where our methods perform better than chance.

C. Complex networks based on real-world features

To demonstrate the potential of UTE and ES in understanding policy diffusion, we test them on synthetic data associated with the US state policy landscape. Specifically, we seek to demonstrate the potential of our data-driven approaches to help explain the political underpinning of policy diffusion, by pinpointing competing factors as the basis of states’ interactions. By applying ES and UTE to synthetic law activity data, we test the ability of detecting two different mechanisms, geography and ideology, which have been proposed as important determinants of the policy diffusion process in Refs. 6 and 7. We utilize these mechanisms to construct undirected network topologies for the 50 US states and then utilize an innovativeness index¹² to introduce directionality. Finally, we generate law activity data through our model, by using the underlying connectivity matrices encoding geography- or ideology-based interactions.

With respect to geography, we generate an undirected unweighted topology by connecting each pair of states if they share a common border. Thus, Alaska and Hawaii are decoupled from the network of continental US states. Influences among states are directed from the state with a higher legislative innovativeness index to the state with the lower one. The innovativeness index of each state is computed as an average over the years 1980–2010.¹²

With respect to ideology, our approach is motivated by the so-called Axelrod model:⁷¹ a social influence model for

the dissemination of culture; each culture (in our case each state) is characterized by certain features. Toward modeling the process of diffusion of innovations,⁷² Axelrod puts forward the notion of homophily, where cultures/states interact preferentially with similar ones. In other words, the transfer of ideas occurs most frequently between two cultures which are alike. Thus, we propose that states which are similar in political ideology, that is, more liberal or more conservative, are connected. To this aim, we calculate the total difference between the political ideologies of state pairs over the years 1980–2010, based on the average ideological position of those states’ elected officials, including the governor and the major party delegations in each house of the state’s legislative bodies.⁷³

Then, we use a simple Euclidean metric to construct a weighted connectivity matrix whose generic ij entry is

$$\frac{1}{1 + \sqrt{\Delta_{ij}^2}}, \tag{23}$$

where Δ_{ij} is the difference in ideology between states i and j . From Eq. (23), we assemble a directed binary matrix by using thresholding as we have done for ROC curves (see for example, Ref. 74) and the innovativeness index, similar to the geography-based network. For the ideology feature, the threshold for the network is chosen to be 0.985 in order to have the same number of edges (107) for the networks based on geography and ideology as shown in Fig. 10.

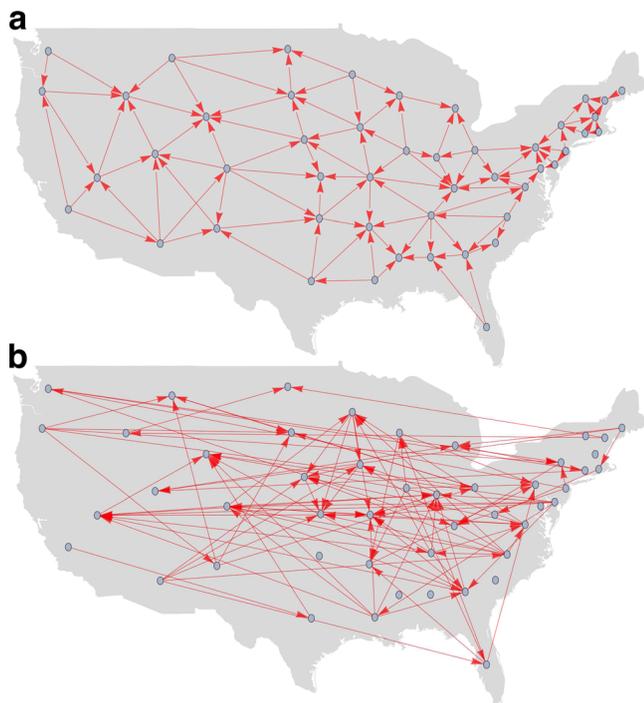


FIG. 10. Proposed networks for the 50 US states based on (a) geography and (b) ideology. The nodes represent the US states. AK and HI are omitted for clarity; in the geography-based network, AK and HI are isolated nodes, while in the ideology-based network, AK has seven incoming edges from CO, DE, GA, IN, LA, MO, and NV; and one outgoing edge to WY; and HI has one outgoing edge only to MA.

In order to exemplify the two methods’ ability to detect underlying network structures based on real-world features, we perform our simulations for Θ -values within the interval $[10^{-5}, 10^{-2}]$ with 11 evenly, logarithmically spaced data values and γ -values within the interval $[0, 30]$ with 11 evenly, linearly spaced data values. While keeping both connectivity matrices A_{Ideo} and A_{Geo} fixed, we generate 10 realizations over 1980 – 2010.

Results are depicted in Fig. 11. Although the time span of 31 years is considerably shorter and the in-degree and out-degree distributions are much more heterogeneous in comparison to the exemplary networks considered in Figs. 6 and 8, both methods are able to adequately detect causal influences in a wide range of the parameter space. While UTE outperforms ES for low values Θ and γ , smaller than 10^{-4} and 10, respectively, and succeeds in reconstructing network topologies in a wider range of parameters, ES consistently surpasses UTE for event rates larger than 10^{-4} reaching considerably high AUC-scores, as large as approximately 0.8.

These findings are in line with expectations from simulations on motifs and smaller networks. Based on Table II, our results suggest that ES does not perform well for the study of networks of influences associated with policies that occur very slowly (such as laws with extensive opposition where changes may take decades) or those that do not involve comparison with other states (such as those passed in reaction to very specific local needs). In contrast, UTE may be challenged by policies involving Federal mandates or incentives, which may mask information flow between states.

In order to demonstrate the possibility of analyzing real data, we apply the methods on a dataset consisting of law enactment data for 27 different laws in the health policy domain related to general driving regulations in all of 50 US states over a period of 31 years (1980–2010). Details on the dataset can be found in Ref. 6; each state is represented by a binary-valued discrete time-series with a resolution of a day, where 1 represents if any of the 27 laws is implemented, repealed, or substantively changed, and 0 indicates no activity.

Inferred networks using UTE and ES methods are shown in Fig. 12. The thresholds for each method were selected on the basis of achieving the same number of links as in the proposed networks for the 50 US states based on geography and ideology. The UTE based network consists of 107 links for a threshold of 0.0012, and the ES based network consists of 111 links for a threshold of 0.4330.

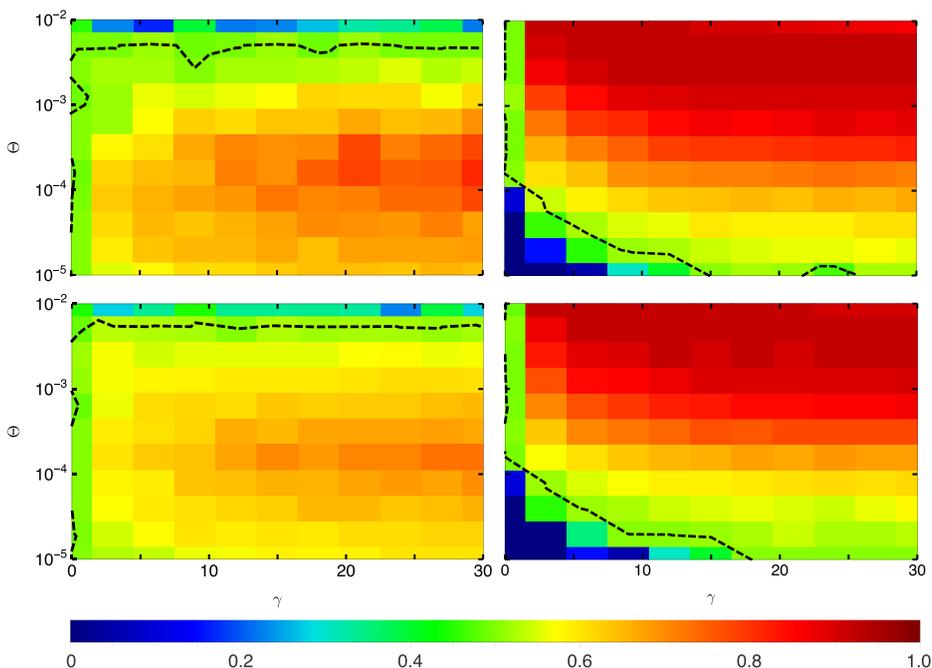


FIG. 11. AUC-scores of UTE (left column) and ES (right column) for the networks depicted in Fig. 10 as functions of the parameters Θ and γ . Networks for the 50 US states based on ideology (upper row) and based on geography (bottom row). All AUC-values are averaged over 10 realizations. Contours highlighted in black identify the area where our methods perform better than chance.

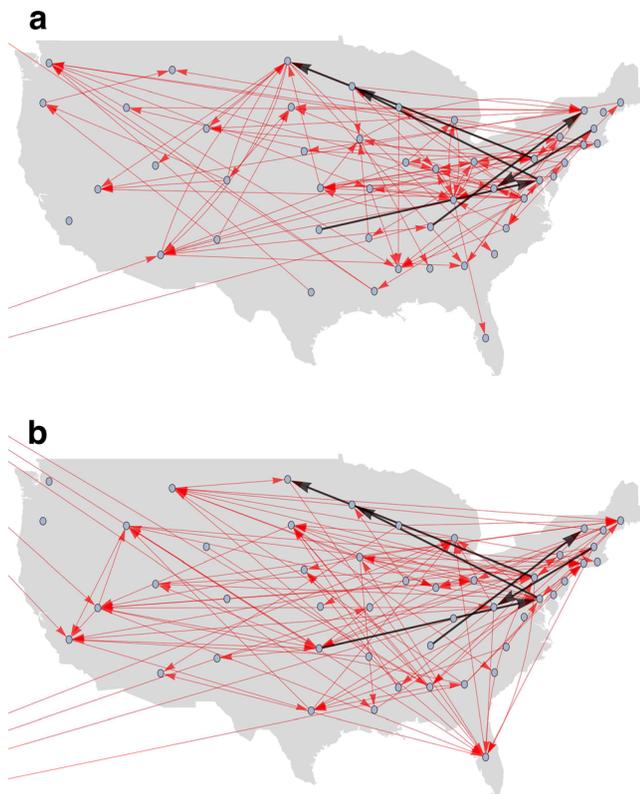


FIG. 12. Proposed networks for the 50 US states based on UTE (a) and ES (b). Links that are common between the two networks are highlighted in black. Alaska and Hawaii are not explicitly indicated on the map.

The two approaches offer a rather different perspective on what are the determining relationships underlying policy diffusion over the real dataset. Such a difference could be ascribed to the varying performance of the two methods with respect to adoption frequency and intensity of state-to-state diffusion shown in Section VI. As a result, it is difficult to rely on a specific method against the other. Future work should seek to establish objective techniques to integrate predictions from different methods and precisely quantify the possibility of false positives.

Without such techniques, we should rely on links (highlighted in Fig. 12) which are common to the two networks: one link from OK to MD, one from MD to MN, one from PA to ND, one from TN to VT, and one from MA to WV. These links reflect the directionality of policy diffusion related to regulating safer driving and provide a different perspective from analyses in other disciplines that may point to a single explanation for policy change (such as the state's political culture, or its burden of traffic accidents). They are likely to reflect the increasingly interconnected world of policy learning, which has been facilitated in recent years by national organizations that bring together health and safety officials from different states and increasingly robust community of advocates who work to disseminate information on best practices.

VII. CONCLUSIONS

The process of interstate policy diffusion in the US has been studied for decades, and US states have expanded the use of public policies (defined as laws, regulations, taxes,

and associated enforcement mechanisms) to enhance prevention and facilitate complex behaviors.¹³ Public policies are powerful public health tools, yet many aspects of their adoption and diffusion are not well understood. In spite of growing evidence for the effectiveness of numerous individual laws, a number of important questions remain unanswered about how to accelerate adoption of policies that have been found to be effective.

In this study, we introduce a novel reduced order, minimalist model for policy diffusion. The model is based on recent empirical observations in Ref. 56 and extends our previous work in Ref. 6. Our model is a first step toward systematic inferences of influences in public policy. Synthetic data generated through the model were utilized to test the accuracy of two network reconstruction techniques. We compared the performance of UTE, recently proposed by our group in Ref. 6 to study sparse and heterogeneous datasets from law enactment and ES, originally introduced to measure synchronization and time-delay patterns between signals³⁰ and later adapted to network reconstruction across a number of fields.⁴⁸ We analyzed principal motifs and then extended the scope to larger networks, in which individual units are regularly or randomly interconnected. Finally, we demonstrated our data-driven approaches to study potential determinants of policy diffusion in the US legal landscape.

We showed that both methods, UTE and ES, are successful in accurately reconstructing network topologies for a wide range of model parameters, spanning policies that rapidly spread across states with little state-to-state learning, those that emulate successful state pilot experiences, and those that fall in between these two extremes. We systematically investigated the performance of UTE and ES as functions of two parameters, Θ and γ , encapsulating the state's law enactment rate and the coupling between states, respectively. Both methods benefit from enhancing the relative weight of the interactions, which reduces the possibility of law enactment from spontaneous activity rather than coupling between states.

We found that UTE outperforms ES for slow event rates and weak interactions between states, leading to exact network reconstruction. However, ES was successful in inferring the network topology for a wider range of parameters including cases where UTE fails. We also found that the performance of ES is largely independent of the network topology, while UTE suffered from increasing link density and heterogeneous in-degree distributions. While we have focused on regular and random network topologies, future efforts should consider continuous interpolations between these topological extremes by varying the rewiring probability.^{68,75} When implemented on realistic networks constructed from ideology, geography, and innovativeness, we showed that both methods are similarly effective in isolating potential determinants that drive policy diffusion.

Overall, these findings suggest that the two methods could be synergistically integrated to unravel networks of influences for a wide range of model parameters, whereby our simulations did not hint at any selection of model parameters where both the methods fail. While policies that occur very slowly (such as laws with extensive opposition where changes may take decades) or those that do not involve comparison

with other states (such as those passed in reaction to very specific local needs) may be better tackled using UTE, ES should be preferred for policies involving Federal mandates or incentives. Through the integration of dynamical systems, information theory, and complex networks, this study contributes new tools toward an improved understanding of complex interactions within groups of interconnected units (state governments) that are responsible for an emergent phenomenon (state health policy landscape).

ACKNOWLEDGMENTS

This work was supported by the National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism under Grant No. R21-AA02523-01A1 and by New York University under a SEED grant. We thank Dr. Sachit Butail for his constructive feedback and invaluable help during the preparation of the revised version of this work based on the very constructive comments of two anonymous reviewers.

- ¹E. R. Graham, C. R. Shipan, and C. Volden, *Brit. J. Polit. Sci.* **43**, 673 (2013).
- ²J. L. Walker, *Am. Polit. Sci. Rev.* **63**, 880 (1969).
- ³M. C. Rom, P. E. Peterson, and F. S. Kenneth, Jr., *Publius J. Federalism* **28**, 17 (1998).
- ⁴T. Makse and C. Volden, *J. Polit.* **73**, 108 (2011).
- ⁵C. R. Shipan and C. Volden, *Am. J. Polit. Sci.* **50**, 825 (2006).
- ⁶R. P. Anderson, G. Jimenez, J. Bae, D. Silver, J. Macinko, and M. Porfiri, "Understanding policy diffusion in the U.S.: An information-theoretical approach to unveil connectivity structures in slowly-evolving complex systems," *SIAM J. Appl. Dyn. Syst.* **15**(3), 1384–1409.
- ⁷N. Abaid, J. Macinko, D. Silver, and M. Porfiri, *PLoS One* **10**, e0123339 (2015).
- ⁸F. J. Boehmke and R. Witmer, *Political Res. Q.* **57**, 39 (2004).
- ⁹M. Mintrom, *Am. J. Polit. Sci.* **41**, 738 (1997).
- ¹⁰F. S. Berry and W. D. Berry, *Am. Polit. Sci. Rev.* **84**, 395 (1990).
- ¹¹W. D. Berry and R. Baybeck, *Am. Polit. Sci. Rev.* **99**, 505 (2005).
- ¹²F. J. Boehmke and P. Skinner, *State Polit. Policy Q.* **12**, 303 (2012).
- ¹³G. Boushey, *Policy Diffusion Dynamics in America* (Cambridge University Press, 2010).
- ¹⁴F. J. Boehmke, *State Polit. Policy Q.* **9**, 229 (2009).
- ¹⁵C. Volden, *Am. J. Polit. Sci.* **50**, 294 (2006).
- ¹⁶S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, *Phys. Rep.* **424**, 175 (2006).
- ¹⁷G. Chen, X. Wang, and X. Li, *Introduction to Complex Networks: Models, Structures and Dynamics* (Higher Education Press, Peking, 2012).
- ¹⁸S. Tan, J. Lu, G. Chen, and D. J. Hill, *IEEE Circuits Syst. Mag.* **14**, 36 (2014).
- ¹⁹M. Newman, *Proc. Natl. Acad. Sci. U. S. A.* **98**, 404 (2001).
- ²⁰A.-L. Barabási and J. Frangos, *Linked: The New Science of Networks Science of Networks* (Basic Books, New York, 2002).
- ²¹T. S. Gardner, D. di Bernardo, D. Lorenz, and J. J. Collins, *Science* **301**, 102 (2003).
- ²²E. Bullmore and O. Sporns, *Nat. Rev. Neurosci.* **10**, 186 (2009).
- ²³E. N. Brown, R. E. Kass, and P. P. Mitra, *Nat. Neurosci.* **7**, 456 (2004).
- ²⁴D. Papo, M. Zanin, J. A. Pineda-Pardo, S. Boccaletti, and J. M. Buldú, *Philos. Trans. R. Soc. B* **369**, 20130525 (2014).
- ²⁵R. Guimera, S. Mossa, A. Turtshi, and L. Amaral, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 7794 (2005).
- ²⁶A. E. Motter, *Nat. Phys.* **9**, 191 (2013).
- ²⁷R. Albert and A. L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002).
- ²⁸B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg, *Genome Biol.* **10**, R25 (2009).
- ²⁹L. K. Gallos, D. Rybski, F. Liljeros, S. Havlin, and H. A. Makse, *Phys. Rev. X* **2**, 031014 (2012).
- ³⁰R. Quian Quiroga, T. Kreuz, and P. Grassberger, *Phys. Rev. E* **66**, 041904 (2002).
- ³¹T. M. Cover and J. A. Thomas, *Elements of Information Theory* (John Wiley & Sons, 2012).
- ³²T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).
- ³³K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, *Phys. Rep.* **441**, 1 (2007).
- ³⁴J. Sun and E. M. Bollt, *Physica D* **267**, 49 (2014).
- ³⁵T. Kreuz, J. S. Haas, A. Morelli, H. D. I. Abarbanel, and A. Politi, *J. Neurosci. Methods* **165**, 151 (2007).
- ³⁶R. Vicente, M. Wibral, M. Lindner, and G. Pipa, *J. Comput. Neurosci.* **30**, 45 (2011).
- ³⁷O. Stetter, D. Battaglia, J. Soriano, and T. Geisel, *PLoS Comput. Biol.* **8**, e1002653 (2012).
- ³⁸J. Runge, J. Heitzig, N. Marwan, and J. Kurths, *Phys. Rev. E* **86**, 061121 (2012).
- ³⁹J. Hlinka, D. Hartman, M. Vejmelka, J. Runge, N. Marwan, J. Kurths, and M. Paluš, *Entropy* **15**, 2023 (2013).
- ⁴⁰J. Runge, J. Heitzig, V. Petoukhov, and J. Kurths, *Phys. Rev. Lett.* **108**, 258701 (2012).
- ⁴¹F. H. McRobie, T. Stemler, and K. H. Wyrwoll, *Q. Sci. Rev.* **121**, 120 (2015).
- ⁴²X. R. Wang, J. M. Miller, J. T. Lizier, M. Prokopenko, and L. F. Rossi, *PLoS One* **7**, e40084 (2012).
- ⁴³Y. Sun, L. F. Rossi, C.-C. Shen, J. Miller, X. R. Wang, J. T. Lizier, M. Prokopenko, and U. Senanayake, preprint [arXiv:1407.0007](https://arxiv.org/abs/1407.0007) (2014).
- ⁴⁴N. Orange and N. Abaid, *Eur. Phys. J.-Spec. Top.* **224**, 3279 (2015).
- ⁴⁵S. Butail, V. Mwaffo, and M. Porfiri, *Phys. Rev. E* **93**, 042411 (2016).
- ⁴⁶N. Malik, B. Bookhagen, N. Marwan, and J. Kurths, *Clim. Dyn.* **39**, 971 (2012).
- ⁴⁷N. Boers, B. Bookhagen, H. M. J. Barbosa, N. Marwan, J. Kurths, and J. A. Marengo, *Nat. Commun.* **5**, 5199 (2014).
- ⁴⁸N. Marwan and J. Kurths, *Chaos* **25**, 097609 (2015).
- ⁴⁹J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, *Chaos* **20**, 037109 (2010).
- ⁵⁰M. Zanin, *PeerJ* **4**, e2111 (2016).
- ⁵¹J. VanDecar and R. Crosson, *Bull. Seismol. Soc. Am.* **80**, 150 (1990); available at <http://www.bssaonline.org/content/80/1/150.short>.
- ⁵²S. Feldt, J. Waddell, V. L. Hetrick, J. D. Berke, and M. Żochowski, *Phys. Rev. E* **79**, 056104 (2009).
- ⁵³C. Granger, *Econometrica* **37**, 424 (1969).
- ⁵⁴A. A. Margolin, I. Nemenman, K. Basso, C. Wiggins, G. Stolovitzky, R. Favera, and A. Califano, *BMC Bioinf.* **7**, S7 (2006).
- ⁵⁵G. Ver Steeg and A. Galstyan, in *Proceedings of the 21st International Conference on World Wide Web* (ACM, 2012), pp. 509–518.
- ⁵⁶J. Macinko and D. Silver, *Am. J. Public Health* **105**, 1893 (2015).
- ⁵⁷T. Fenner and A. Frieze, *Combinatorica* **2**, 347 (1982).
- ⁵⁸A. Roxin, V. Hakim, and N. Brunel, *J. Neurosci.* **28**, 10734 (2008).
- ⁵⁹P. Sonogo, A. Kocsor, and S. Pongor, *Brief Bioinf.* **9**, 198 (2008).
- ⁶⁰C. E. Shannon, *Bell Syst. Tech. J.* **27**, 379 (1948).
- ⁶¹G. Casella and R. L. Berger, *Statistical Inference* (Duxbury Resource Center, 2008).
- ⁶²K. Aihara, T. Takabe, and M. Toyoda, *Phys. Lett. A* **144**, 333 (1990).
- ⁶³F. S. Berry and W. D. Berry, *Theories of the Policy Process* (Westview Press, Boulder, CO, 1999), Vol. 169.
- ⁶⁴M. T. Heaney, *J. Health Polit. Policy Law* **31**, 887 (2006).
- ⁶⁵C. R. Shipan and C. Volden, *Public Adm. Rev.* **72**, 788 (2012).
- ⁶⁶A. P. Bradley, *Pattern Recognit.* **30**, 1145 (1997).
- ⁶⁷J. A. Hanley and B. J. McNeil, *Radiology* **143**, 29 (1982).
- ⁶⁸D. Watts and S. H. Strogatz, *Nature* **393**, 440 (1998).
- ⁶⁹G. Fagiolo, *Phys. Rev. E* **76**, 026107 (2007).
- ⁷⁰To assess the performance of ES and UTE on undirected dense networks, we analyzed a lattice with k^{in} and k^{out} equal to 6 (compared to the directed networks in Fig. 8 where either of them is 3). We find that although the performance of them is diminished, the amount of degradation is less for UTE than ES. At the same time, as ES performs better overall, the reconstruction accuracy is higher for ES than UTE over all parameter combinations.
- ⁷¹R. Axelrod, *J. Conflict Resolut.* **41**, 203 (1997).
- ⁷²P. F. C. Tilles and J. F. Fontanari, *J. Stat. Mech.: Theory Exp.* **2015**, P11026.
- ⁷³W. D. Berry, E. J. Ringquist, R. C. Fording, and R. L. Hanson, *Am. J. Polit. Sci.* **42**, 327 (1998).
- ⁷⁴J. F. Donges, Y. Zou, N. Marwan, and J. Kurths, *Europhys. Lett.* **87**, 48007 (2009).
- ⁷⁵C. Grabow, S. Grosskinsky, and M. Timme, *Phys. Rev. Lett.* **108**, 218701 (2012).