PNAS
www.pnas.org

Supplementary Information for

Inferring Dynamic Topology for Decoding Spatiotemporal Structures in Complex Heterogeneous Networks

Shuo Wang[1], Erik D. Herzog[2], István Z. Kiss[3], William J. Schwartz[4†], Guy Bloch[5], Michael Sebek[3], Daniel Granados-Fuentes[2], Liang Wang[1], and Jr-Shin Li[1]

[1]Department of Electrical and Systems Engineering, Washington University in St. Louis, USA.

[2]Department of Biology, Washington University in St. Louis, USA.

[3]Department of Chemistry, Saint Louis University, USA.

[4]Department of Neurology, University of Massachusetts Medical School, Worcester, USA.

[5]Department of Ecology, Evolution, and Behavior, Hebrew University of Jerusalem, Israel.

[†] Current address: Department of Neurology, Dell Medical School, University of Texas at Austin, USA

Jr-Shin Li
Email: jsli@wustl.edu

This PDF file includes:

Supplementary Information
Figs. S1 to S8
References Cited for SI Reference Citations

www.pnas.org/cgi/doi/10.1073/pnas.1721286115

Supplementary Information

Inferring Dynamic Topology for Decoding Spatiotemporal Structures in Complex Heterogeneous Networks

1 Methodology

Inferring topology (time-varying connections) has been an essential, but difficult, step towards understanding large, complex and diverse systems including biological, financial, and electrical networks [1–5]. In recent years, various studies based on Bayesian statistics, information theory, and spectral analysis, have been proposed to extract the dynamics of systems using their measured data [6–10]. While techniques exist for relatively small system sizes [11] (e.g., single nonlinear unit [2]), sparse networks [7, 12], or networks with predetermined topology [4], the tackling of functionality or behavior of real-world complex systems, such as the circadian clock and social synchronization, requires investigating large-scale networks (at least hundreds of nodes) and beyond logic-based (binary-valued) topology [13, 14].

The novel contribution of this work is to establish a unified, data-driven framework that provides a highly reliable and efficient approach to inferring connections of networks (ICON). We show that the ICON technique identifies the strength and direction of functional connections in diverse complex networks - from cells to societies. In this supplementary information, we provide a detailed description of the ICON technique and include numerical and statistical analysis to demonstrate the reliability and robustness of this data-driven method for unfolding the topology of large-scale networks.

1.1 Inferring Connections of Networks (ICON)

We consider the broadly-defined complex network constituted by a population of interacting dynamic units (agents). The time-evolution of such a network follows the dynamical law governed by the agent's self-dynamics and the influence from other agents, given by

$$\dot{x}_i(t) = f(x_i) + \sum_{\substack{j=1 \\ j\neq i}}^{N} K_{ij}(x_i, x_j), \quad i = 1, \ldots, N, \tag{1}$$

where the vector $x_i(t) \in \mathbb{R}^n$ denotes the state of agent i at time t, the function f represents the baseline dynamics of each agent, and K_{ij}, $i, j = 1, \ldots, N$, is the coupling function between agent i and j, which is not necessarily symmetric. Note that both f and K_{ij}'s are time-varying functions, where the time-dependency is expressed indirectly by the state-dependency, e.g., $\tilde{f}_i(t) = f(x_i(t))$ for $t \in [0, T]$, $T \in (0, \infty)$.

The central idea of ICON is to approximate the natural and coupling dynamics, f and K_{ij}, in (1) using complete orthonormal bases. Specifically, we represent f and K_{ij} as truncated series, e.g., $f \approx \sum_k a_k Q_k(x_i)$ and $K_{ij}(x_i, x_j) \approx \sum_k \sum_\ell b_{ij}^{k\ell} P_k(x_i) P_\ell(x_j)$, where $\{Q_k\}_{k=1}^{\infty}$ and $\{P_k\}_{k=1}^{\infty}$ are orthonormal bases of the respective function space containing f and K_{ij}. These approximations are always possible because any continuous function can be approximated arbitrarily well by a truncated series of orthonormal bases (e.g., Fourier series) on a compact support with a quantifiable error bound with respect to the number of truncation terms by the Stone-Weierstrass theorem [15, 16]. As a result, the dynamical law as in (1) can be approximately expressed using orthonormal bases,

e.g., for $n = 1$, by

$$\dot{x}_i(t) = \sum_{k=1}^{m} a_k Q_k(x_i) + \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k=1}^{r} \sum_{\ell=1}^{r} b_{ij}^{k\ell} P_k(x_i) P_\ell(x_j), \quad i = 1, \ldots, N, \tag{2}$$

and the coupling strength (or energy), denoted α_{ij}, between agents i and j can be defined through the coefficients of the chosen orthonormal basis by

$$\alpha_{ij} = \sqrt{\sum_{k,\ell=1}^{r} (b_{ij}^{k\ell})^2}. \tag{3}$$

In this one-dimensional case, $x_i(t) \in \mathbb{R}$ and a_k and $b_{ij}^{k\ell}$ are the scalar coefficients of the chosen orthonormal bases $\{Q_k\}$ and $\{P_k\}$, respectively. Similarly, for n-dimensional systems, i.e., $x_i(t) \in \mathbb{R}^n$ for $n = 2, 3, \ldots$, the dynamics of the u^{th}-component of agent i, denoted $[x_i]_u$, can be expressed, for $i = 1, \ldots, N$, by

$$[\dot{x}_i(t)]_u = \sum_{k_1+k_2+\cdots+k_n=1}^{m} \left(a_{k_1 \cdots k_n}^{u} \prod_{\varrho=1}^{n} Q_{k_\varrho}([x_i]_\varrho) \right) + \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k=1}^{r} \sum_{\ell=1}^{r} \sum_{v=1}^{n} b_{i_u j_v}^{k\ell} P_k([x_i]_u) P_\ell([x_j]_v), \tag{4}$$

where $a_{k_1 \cdots k_n}^{u}$ is the scalar coefficient for the basis $Q_{k_1} Q_{k_2} \ldots Q_{k_n}$ of component u of the self-dynamics f, and $b_{i_u j_v}^{k\ell}$ is the coefficient of the coupling function $K_{i_u j_v}$, namely, the coupling between the components $[x_i]_u$ and $[x_j]_v$. In particular, if the components of the self dynamics, $f(x_i) \in \mathbb{R}^n$, are independent, i.e., $[f(x_i)]_u = f([x_i]_u)$, and the interaction between two agents, say i and j, is component-wise, i.e., $K_{i_u j_v} = 0$ if $u \neq v$ for $u, v = 1, \ldots, n$, then the dynamics of each component, $[x_i]_u$, of agent i described in (4) is reduced to the same dynamical law as in (2).

Following this approach, the topology of a complex network can be effectively estimated, given the measurement time-series data of each agent i in the network. Let's now illustrate the idea using the case in which the agents $x_i(t) \in \mathbb{R}$. Let $\{\hat{x}_k^{(i)}\}$ be the available time-series data of agent i at time t_k, $i = 1, \ldots, N$ and $k = 0, 1, \ldots, M$, where N is the number of agents in the network and M is the number of data points in each time-series. Then the orthonormal basis representation in (2) enables the formulation of the complex nonlinear topology estimation as a typical linear inverse problem for each agent i, given by

$$\min_{z^{(i)}} \ \|y^{(i)} - A^{(i)} z^{(i)}\|_2, \tag{5}$$

where $y^{(i)} \in \mathbb{R}^M$ is the data vector whose elements $y_j^{(i)} = \frac{\hat{x}_j^{(i)} - \hat{x}_{j-1}^{(i)}}{\Delta t_j}$, $j = 1, \ldots, M$, denote the finite-difference approximation of the temporal dynamics of agent i at time t_j with $\Delta t_j = t_j - t_{j-1}$ being the data sampling time interval; $A^{(i)} \in \mathbb{R}^{M \times (r^2 N + m)}$ is the matrix defined by

$$A^{(i)} = \left[L^{(i)} \middle| O_{11}^{(i)} \middle| \cdots \middle| O_{1r}^{(i)} \middle| O_{21}^{(i)} \middle| \cdots \middle| O_{2r}^{(i)} \middle| \cdots \middle| O_{r1}^{(i)} \middle| \cdots \middle| O_{rr}^{(i)} \right],$$

in which $L^{(i)} \in \mathbb{R}^{M \times m}$ and $O_{k\ell}^{(i)} \in \mathbb{R}^{M \times N}$, $k, \ell = 1, \ldots, r$, are the matrices involving the orthonormal bases for the expansion of f and K_{ij}, respectively,

$$L^{(i)} = \begin{bmatrix} Q_1(\hat{x}_1^{(i)}) & \cdots & Q_m(\hat{x}_1^{(i)}) \\ \vdots & \ddots & \vdots \\ Q_1(\hat{x}_M^{(i)}) & \cdots & Q_m(\hat{x}_M^{(i)}) \end{bmatrix}, \quad O_{k\ell}^{(i)} = \begin{bmatrix} P_k(\hat{x}_1^{(i)}) P_\ell(\hat{x}_1^{(1)}) & \cdots & P_k(\hat{x}_1^{(i)}) P_\ell(\hat{x}_1^{(N)}) \\ \vdots & \ddots & \vdots \\ P_k(\hat{x}_M^{(i)}) P_\ell(\hat{x}_M^{(1)}) & \cdots & P_k(\hat{x}_M^{(i)}) P_\ell(\hat{x}_M^{(N)}) \end{bmatrix}; \tag{6}$$

and $z^{(i)} = (z_Q^{(i)}, z_{P_{11}}^{(i)}, \ldots, z_{P_{1r}}^{(i)}, z_{P_{21}}^{(i)}, \ldots, z_{P_{2r}}^{(i)}, \ldots, z_{P_{r1}}^{(i)}, \ldots, z_{P_{rr}}^{(i)})' \in \mathbb{R}^{r^2N+1}$ is the coefficient vector to be estimated for agent i, in which each $z_{P_{k\ell}}^{(i)} \in \mathbb{R}^{1\times N}$ is the coefficient vector associated with the base P_kP_ℓ, $k, \ell = 1, ..., r$. Note that technically the ICON method requires prior knowledge of the self-dynamics, e.g., its functional form, in order to uniquely determine the network topology (i.e., obtain a unique solution to (5)) with respect to the chosen sets of orthonormal basis expansions, $\{Q_k\}$, $k = 1, \ldots, m$, and $\{P_k, P_\ell\}$, $k, \ell = 1, \ldots, r$, in (2). For example, if $f(x_i)$ is a constant, e.g., representing the frequency of an oscillator in an oscillatory network, the matrix $L^{(i)}$ in (6) is a vector of ones, i.e., $L^{(i)} = [1, 1, \cdots, 1]' \in \mathbb{R}^M$, and $z_Q^{(i)} = c^{(i)}$ is a scalar for each agent i. The minimizer $z^{(i)}$ of the problem in (5) then determines the baseline dynamics, i.e., $f(x_i) = c^{(i)}$, and the connections, i.e., $K_{ij}(t) = K_{ij}(x_i(t), x_j(t)) \approx \sum_{k,\ell}[z_{P_{k\ell}}^{(i)}]_j P_k(x_i)P_\ell(x_j)$, $j = 1, \ldots, N$, of agent i, where $[z_{P_{k\ell}}^{(i)}]_j$ denotes the j^{th} element of $z_{P_{k\ell}}^{(i)}$. Most importantly, this linear formulation enables independent estimation of the time-varying couplings for each individual agent in the network, so that estimating topology of very large networks becomes possible via a parallel computation architecture.

1.2 Computing Network Topology

There exists a variety of techniques for solving the large-scale least squares, or linear inverse, problem presented in (5) [17]. Among many of these methods, a basic step is to compute the Moore-Penrose pseudoinverse using the singular value decomposition (SVD). In this work, we adopt the truncated SVD (TSVD) to solve this problem in a parallel manner for each agent i. TSVD is known as a method for dealing with ill-posed linear least squares problems [18], by which we are able to retain as much information of the reconstructed couplings as possible by controlling the cut-off of insignificant singular values. In addition, because the matrices $A^{(i)} \in \mathbb{R}^{M\times(r^2N+1)}$, $i = 1, \ldots, N$, in (5) are dense and ill-conditioned (columns may be close to be linearly dependent), TSVD provides numerically stable and accurate solutions by replacing very small singular values with zeros. For example, in our numerical implementations, the threshold for zeroing out the singular values was chosen to be 10^{-8} taking the floating error of a standard computer into account.

It is known that the theoretical solution to the problem in (5) is in terms of the pseudoinverse of $A^{(i)}$, denoted $A^\dagger$, given by

$$z^{(i)} = A^\dagger y^{(i)} = \left[(A^{(i)})'A^{(i)}\right]^{-1}(A^{(i)})'y^{(i)}, \quad i = 1, \ldots, N.$$

Although the square matrix $W^{(i)} = (A^{(i)})'A^{(i)}$ is in general positive definite ($A^{(i)}$ is composed of the chosen orthonormal bases), it may have singular values very close to zero, which results in numerical instability and computational errors. By the SVD, we can decompose $W^{(i)} = U^{(i)}\Sigma^{(i)}(V^{(i)})'$, and then obtain the solution vector $\hat{z}^{(i)}$ of (5) synthesized by the truncated singular values and singular vectors,

$$\hat{z}^{(i)} = \sum_{j=1}^{\mathcal{N}(\epsilon)} \frac{(y^{(i)})'A^{(i)}\bar{u}_j^{(i)}}{s_j^{(i)}}\bar{v}_j^{(i)}, \tag{7}$$

where $\bar{u}_j^{(i)}$ and $\bar{v}_j^{(i)}$ are the columns of $U^{(i)}$ and $V^{(i)}$, respectively, called the left and the right singular vectors, $s_j^{(i)}$ are the singular values that are the elements of the diagonal matrix $\Sigma^{(i)}$, and the positive integer $\mathcal{N}(\epsilon)$ denotes the number of singular values that is greater than $\epsilon > 0$, where ϵ is the specified truncation threshold. Clearly, one can observe $\hat{z}^{(i)} \to z^{(i)}$ as $\mathcal{N}(\epsilon) \to \text{rank}(W^{(i)})$ or as $\epsilon \to 0$.

We would like to emphasize that the efficiency and accuracy of TSVD may be degraded with the increase of the dimension of $A^{(i)}$, i.e., the size of the network. Alternative methods such as iterative

shrinkage-thresholding (IST) algorithms [19, 20] that effectively solve large-scale inverse problems in an iterative manner can be adopted to compromise between efficiency and accuracy. Moreover, when the network is sparse (i.e., $z^{(i)}$ is sparse) or the focus is placed on certain interactions so that the matrices $A^{(i)}$ become sparse, the methods of compressive sensing [21], e.g., ℓ_1 optimization, are suitable for constructing sparse coupling topology effectively.

1.3 A Special Case: Weakly-Coupled Oscillatory Networks

Many natural and engineered rhythmic systems are weakly-connected and weakly-forced by environmental fluctuations, such as neurons, circadian clock networks, and bees in a colony [1, 22–24]. The techniques of phase model reduction have been widely used to describe the dynamics of this class of nonlinear oscillatory systems [25–28], by which a network of weakly-coupled oscillators can be modeled through their averaged phase drift dynamics [29] as,

$$\dot{\varphi}_i = \omega_i + \sum_{j=1}^{N} K_{ij}(\varphi_j - \varphi_i), \quad i = 1, \dots, N. \tag{8}$$

In this equation, ω_i is the natural oscillation frequency of oscillator i, K_{ij} is the coupling function between oscillators i and j, which depends on their phase difference and is 2π-periodic, and N is the number of oscillators in the network.

We now illustrate the application of ICON to recover the topology of such a weakly-coupled oscillatory network. Since each coupling function K_{ij}, for $i, j = 1, \dots, N$, is 2π-periodic, it can be represented as a truncated Fourier series, i.e., $K_{ij}(\Delta\varphi_{ij}) = \sum_{k=1}^{r} \left[\xi_{ij}^k \sin(k\Delta\varphi_{ij}) + \eta_{ij}^k \cos(k\Delta\varphi_{ij})\right]$ with $\Delta\varphi_{ij} = \varphi_j - \varphi_i$, so that (8) can be expressed as

$$\dot{\varphi}_i = \omega_i + \sum_{\substack{j=1 \\ j\neq i}}^{N} \sum_{k=1}^{r} \Big[\xi_{ij}^k \sin(k\Delta\varphi_{ij}) + \eta_{ij}^k \cos(k\Delta\varphi_{ij})\Big], \quad i = 1, \dots, N, \tag{9}$$

where ξ_{ij}^k and η_{ij}^k are the Fourier coefficients that define the coupling strength

$$\alpha_{ij} = \sqrt{\sum_{k=1}^{r} \Big[(\xi_{ij}^k)^2 + (\eta_{ij}^k)^2\Big]}. \tag{10}$$

Given the time-series data $\{\hat{\varphi}_k^{(i)}\}$ of each oscillator i at time t_k for $i = 1, \dots, N$ and $k = 0, 1, \dots, M$, the topology of an oscillatory network as in (9) can be determined by solving the linear inverse problem as in (5) with $y^{(i)} = (y_1^{(i)}, \dots, y_M^{(i)}) \in \mathbb{R}^M$, in which $y_j^{(i)} = \frac{\hat{\varphi}_j^{(i)} - \hat{\varphi}_{j-1}^{(i)}}{\Delta t_j}$, $j = 1, \dots, M$, and the matrix $A^{(i)} = \left[L \middle| S_1^{(i)} \middle| \cdots \middle| S_r^{(i)} \middle| C_1^{(i)} \middle| \cdots \middle| C_r^{(i)}\right] \in \mathbb{R}^{M\times(2rN+1)}$, in which $L = [1, 1, \cdots, 1]' \in \mathbb{R}^M$, and

$$S_k^{(i)} = \begin{bmatrix} \sin k\Delta\hat{\varphi}_{i1}(t_1) & \sin k\Delta\hat{\varphi}_{i2}(t_1) & \cdots & \sin k\Delta\hat{\varphi}_{iN}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ \sin k\Delta\hat{\varphi}_{i1}(t_M) & \sin k\Delta\hat{\varphi}_{i2}(t_M) & \cdots & \sin k\Delta\hat{\varphi}_{iN}(t_M) \end{bmatrix} \in \mathbb{R}^{M\times N},$$

$$C_k^{(i)} = \begin{bmatrix} \cos k\Delta\hat{\varphi}_{i1}(t_1) & \cos k\Delta\hat{\varphi}_{i2}(t_1) & \cdots & \cos k\Delta\hat{\varphi}_{iN}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ \cos k\Delta\hat{\varphi}_{i1}(t_M) & \cos k\Delta\hat{\varphi}_{i2}(t_M) & \cdots & \cos k\Delta\hat{\varphi}_{iN}(t_M) \end{bmatrix} \in \mathbb{R}^{M\times N},$$

where $\Delta\hat{\varphi}_{ij} = \hat{\varphi}_j - \hat{\varphi}_i$ and $z^{(i)} = (\omega_i, z^{(i)}_{S_1}, \ldots, z^{(i)}_{S_r}, z^{(i)}_{C_1}, \ldots, z^{(i)}_{C_r})' \in \mathbb{R}^{2rN+1}$ is the Fourier coefficient vector that is being estimated, in which $z^{(i)}_{S_k} = (\xi^k_{i1}, \cdots, \xi^k_{iN}) \in \mathbb{R}^{1\times N}$ and $z^{(i)}_{C_k} = (\eta^k_{i1}, \cdots, \eta^k_{iN}) \in \mathbb{R}^{1\times N}$ for $k = 1, \ldots, r$. Hence, the estimated time-varying coupling function K_{ij} can be expressed as,

$$K_{ij}(t) \approx \sum_{k=1}^{r} \Big[[z^{(i)}_{S_k}]_j \sin(k\Delta\varphi_{ij}(t)) + [z^{(i)}_{C_k}]_j \cos(k\Delta\varphi_{ij}(t)) \Big],$$

where $[z^{(i)}_{S_k}]_j$ and $[z^{(i)}_{C_k}]_j$ denote the j^{th} element of $z^{(i)}_{S_k}$ and $z^{(i)}_{C_k}$, respectively.

Remark 1 *The choice of suitable, or optimal, basis functions would be highly dependent on the data structure. For example, Fourier series are natural and the best choice for approximating a periodic function. Other widely-used orthogonal polynomials such as Legendre or Chebyshev polynomials exhibit fast convergence rate to approximate a continuous function with high precision, i.e., they may be used to approximate a function very accurately with only a few expansion terms. In particular, it is known that Chebyshev polynomials are the optimal basis functions for approximating a continuous function with minimum interpolating error based on the Chebyshev points [30]. Theoretically, the expansion of a function in terms of a sequence of orthogonal (orthonormal) functions, called spectral approximation, inherits spectral accuracy. Namely, for the expansion $f(x_i) \approx \sum_{k=1}^{\infty} a_k Q_k(x_i)$, where $\{Q_k\}$ is a set of orthonormal bases, the k^{th} coefficient a_k decays faster than k^{-n} for all $n \geq 1$ for smooth functions [30].*

2 Implementation of ICON

In this section, we present various numerical examples to illustrate the implementation of the ICON technique using synthetic networks on different orders of size from 3 to 30000 nodes. We also provide numerical, statistical, and information-theoretic analysis to demonstrate the robustness, reliability, and efficiency of ICON, and to quantify the sufficiency of data in order to capture precise network dynamics.

2.1 Definition of Phase from Experimental Data

The phase was obtained from experimental data using peak finding and a linear interpolation method [31]. In this technique, the maximum of each cycle is located with a peak finding method, which searches for maxima of data points above a threshold (set to 50% above the mean level relative to the maximum data point). At the k^{th} oscillatory peak, the phase is set to $k \times 2\pi$; and for other times, linear interpolation is applied.

2.2 Generation of Synthetic Data

To numerically examine the performance of ICON, we generate time-series data by constructing synthetic networks with specified size, sparsity, and noise intensity. Here, the sparsity of the network is defined as the percentage of the nonzero entries in the off-diagonal elements of the adjacency matrix, and the noise intensity is defined as the ratio of the noise amplitude to the maximum coupling strength in the network. We generate synthetic data for testing ICON using two different ways, which are classified as autonomous (phase-dependent) and non-autonomous (time-dependent) data.

2.2.1 Generating Phase-dependent (Autonomous) Data

In this scenario, we synthesize data for a network of N agents by first randomly generating a coefficient matrix of size $N \times 2rN$ that contains the Fourier coefficients ξ^k_{ij} and η^k_{ij}, as defined in

(9), within a compact set $K \in \mathbb{R}^+$, for $k = 1, \ldots, r$ and $i, j = 1, \ldots, N$, given by

$$C = \left[\begin{array}{ccc|c|ccc|ccc|c|ccc} \xi^1_{11} & \cdots & \xi^1_{1N} & \cdots & \xi^r_{11} & \cdots & \xi^r_{1N} & \eta^1_{11} & \cdots & \eta^1_{1N} & \cdots & \eta^r_{11} & \cdots & \eta^r_{1N} \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ \xi^1_{N1} & \cdots & \xi^1_{NN} & & \xi^r_{N1} & \cdots & \xi^r_{NN} & \eta^1_{N1} & \cdots & \eta^1_{NN} & & \eta^r_{N1} & \cdots & \eta^r_{NN} \end{array}\right],$$

where the first rN columns represent the coefficients of the $\sin(k\Delta\varphi_{ij})$ series and the rest rN columns represent the coefficients of the $\cos(k\Delta\varphi_{ij})$ series. The sparsity $s \in [0, 1]$ can be specified when generating the coefficient matrix C. Then, the phase for each agent i can be obtained by evolving the differential equation (9) with a specified frequency ω_i. After obtaining the phase data for each agent φ_i, white noise with a specified intensity $\lambda \in [0, 1]$ can be added to φ_i.

2.2.2 Generating Time-dependent (Non-autonomous) Data

An alternative way to test ICON using synthetic networks is to generate explicit time-dependent data (i.e., $f(t)$ and $K_{ij}(t)$ instead of $f(x_i)$ and $K_{ij}(x_i, x_j)$) of a general autonomous system as in (1) or (2), and show that the autonomous dynamics can be reconstructed through the orthonormal basis representation of the dynamics. This is to avoid the possibility of obtaining a trivial solution for the network reconstruction, namely, generating autonomous data using a given (phase) model system as in (9) and then reconstructing the parameters (coefficients of the orthonormal basis functions) for this known system. Our procedure is to synthesize the couplings $K_{ij}(t)$ as functions of time by using interpolating polynomials. In this way the system dynamics are in general non-periodic. Then the interactions are recovered in the form $K_{ij}(x_i(t), x_j(t))$.

We construct a network of N agents, with sparsity $s \in [0, 1]$ and noise intensity ratio $\lambda \in [0, 1]$, in two steps. First, we generate the adjacency matrix $\mathcal{M}$ with $m_0 \doteq [sN(N-1)]$ non-zero entries, where $[sN(N-1)]$ denotes the closest integer to $sN(N-1)$. Each non-zero entry of $\mathcal{M}$, denoted α_{ij}, $i, j = 1, \ldots, N$, represents the coupling strength of K_{ij}, which is randomly assigned over a compact set $K \subset \mathbb{R}^+$ (e.g., using 'rand' in Matlab). Second, we synthesize the time-varying function $K_{ij}(t)$ using interpolating polynomials. To achieve this, we generate a sequence of random points $\{t_k\}$ that partitions the time interval $[0, T]$, i.e., $0 = t_0 < t_1 < \cdots < t_q = T$, and then use these points to synthesize an interpolating polynomial $L_{ij}(t)$ by

$$L_{ij}(t) = \sum_{p=0}^{q} \Big(\prod_{\substack{0 \le k \le q \\ k \ne p}} \frac{t - t_k}{t_p - t_k} \Big) L_{ij}(t_p), \qquad t \in [0, T].$$

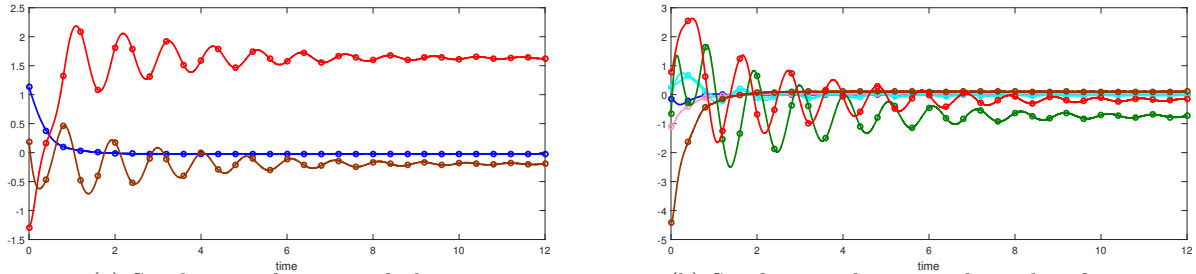
The time-varying coupling function $K_{ij}(t)$ can then be synthesized by rescaling $L_{ij}(t)$ according to the assigned coupling strength α_{ij}, that is,

$$K_{ij}(t) = \frac{\alpha_{ij} L_{ij}(t)}{\int_0^T L_{ij}(\sigma)\, d\sigma}. \tag{11}$$

Once the coupling functions are synthesized, the time-series data $\{\hat{x}^{(i)}_k\}$ for agent i at time t_k, $i = 1, \ldots, N$, can be generated by evolving (1) using the K_{ij} in (11).

2.3 Quantification of Data Sufficiency and Spectral Approximation

The sufficiency of the available data is crucial to formulate a validate estimation problem in (5). In this problem, the dimension of the matrices $A^{(i)}$ is $M \times (r^2N + 1)$, and thus it is necessary to have $M > r^2N + 1$ in order for it to be an inconsistent (overdetermined) inverse problem. Moreover,



(a) Synthetic and estimated phases (b) Synthetic and estimated coupling functions

Figure S1: (a) The time-series of the synthetic (solid line) and estimated (circled) states φ_i, $i = 1, 2, 3$, of the network in (12) for the time-horizon $t \in [0, 12]$ recovered by ICON. (b) The coupling functions of the synthetic (solid) and estimated (circled) networks.

the bound on the sufficiency of the amount of data, i.e., the length of the time measurement data, for extracting significant information to achieve a precise representation of the network dynamics may be computed and quantified [32].

On the other hand, with sufficient time measurement data $\{\hat{x}_k^{(i)}\}$ for each node x_i, $i = 1, \ldots, N$, ICON can be applied to reconstruct the continuous nonlinear coupling functions K_{ij} for $i,\ j = 1, \ldots, N$. The accuracy of the recovery relies on the number of truncation terms r of the chosen orthonormal basis, as the expansion in (2). Theoretically, the larger r leads to the better approximation to the true network dynamics, and the sufficient number of r with respect to a desired precision of the estimation can be quantified by using the tools from information theory.

Given the time-series data $\{\hat{x}_k^{(i)}\}$ of a network of N agents as in (1), and given a prescribed tolerance threshold δ, there exists an integer $r^* > 0$ such that the error of the topology recovery is bounded by δ by approximating the network dynamics using r^* terms of the chosen orthonormal basis functions as in (2). It is equivalent to saying that these r^* terms contain sufficient information of the original nonlinear dynamics. It can be shown by Shannon information theory [33] that this number r^* is proportional to the information entropy, i.e., the nonlinearity and noisiness of the time measurement data, given by,

$$r^* \propto \frac{W \log\left(\dfrac{I_{\{\hat{x}_k^{(i)}\}} + I_{\text{Noise}}}{I_{\text{Noise}}}\right)}{\delta},$$

where W is the data bandwidth, i.e., the width of the frequency range of the time-series data $\{\hat{x}_k^{(i)}\}$, $i = 1, \ldots, N$, which characterizes the nonlinearity of the data; $I_{\{\hat{x}_k^{(i)}\}}$ and I_{Noise} denote the intensities of $\{\hat{x}_k^{(i)}\}$ and the noise, respectively. The lower bound of r^* can be derived, and interested readers may refer to the works in [15, 16, 32, 33].

2.4 Numerical Examples

In this section, we illustrate the implementation procedure of the ICON technique explicitly through the topology estimation of a small network consisting of three coupled oscillators. This network is synthesized using the procedure described in Section 2.2. Additional examples of large synthetic networks with more complicated and heterogeneous structures will also be presented to address to robustness, reliability, and applicability of ICON. All of the numerical experiments in this work were implemented on a standard desktop PC (Intel Core i5 3.5 GHz, 32 GB DDR2 RAM).

2.4.1 Small Networks - an Illustrative Example

We consider a small all-to-all coupled network of three agents of the form as in (1), given by

$$\begin{aligned} \dot{x_1} &= \omega_1 + K_{12} + K_{13}, \\ \dot{x_2} &= \omega_2 + K_{21} + K_{23}, \\ \dot{x_3} &= \omega_3 + K_{31} + K_{32}, \end{aligned} \tag{12}$$

where the drift dynamics, $\omega_1 = 2.63$, $\omega_2 = 0.97$, and $\omega_3 = 2.4$, are assumed to be constant, which may represent the frequencies of the individual systems, and $K_{ij}(t)$, $i, j = 1, 2, 3$, are the coupling functions. We synthesize the time-series data, $\{\hat{x}_k^{(1)}\}$, $\{\hat{x}_k^{(2)}\}$, and $\{\hat{x}_k^{(3)}\}$, $k = 0, 1, \ldots, M = 1500$, for this networked system using the procedure described in Section 2.2 over the time period $t \in [0, 12]$ partitioned by a uniform grid of 25 points, i.e., $0 = t_0 < t_1 < \cdots < t_{24} = 12$. The synthetic data are shown in Figure S1.

Now, we demonstrate the application of ICON to recover the time-varying connections K_{ij} of this network based on the synthetic data $\{\hat{x}_k^{(1)}\}$, $\{\hat{x}_k^{(2)}\}$, and $\{\hat{x}_k^{(3)}\}$. Here, we use the Legendre polynomials of order $r = 5$ to expand the coupling functions K_{ij}, i.e., $K_{ij}(t) = K_{ij}(x_i, x_j) = \sum_{k=1}^{5}\sum_{\ell=1}^{5} b_{ij}^{k\ell} P_k(x_i) P_\ell(x_j)$ for $i, j = 1, 2, 3$ and $i \neq j$, where $\{P_k P_\ell\}$ forms a two-dimensional Legendre basis. Then, we constitute the matrix $A^{(i)} \in \mathbb{R}^{1500\times 76}$, $i = 1, 2, 3$, as

$$A^{(i)} = \left[\begin{array}{c|ccc|c|ccc} 1 & P_1(\hat{x}_1^{(i)})P_1(\hat{x}_1^{(1)}) & \cdots & P_1(\hat{x}_1^{(i)})P_1(\hat{x}_1^{(3)}) & \cdots & P_1(\hat{x}_1^{(i)})P_5(\hat{x}_1^{(1)}) & \cdots & P_1(\hat{x}_1^{(i)})P_5(\hat{x}_1^{(3)}) \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ 1 & P_1(\hat{x}_M^{(i)})P_1(\hat{x}_M^{(1)}) & \cdots & P_1(\hat{x}_M^{(i)})P_1(\hat{x}_M^{(3)}) & & P_1(\hat{x}_M^{(i)})P_5(\hat{x}_M^{(1)}) & \cdots & P_1(\hat{x}_M^{(i)})P_5(\hat{x}_M^{(3)}) \\ \hline \cdots & P_5(\hat{x}_1^{(i)})P_1(\hat{x}_1^{(1)}) & \cdots & P_5(\hat{x}_1^{(i)})P_1(\hat{x}_1^{(3)}) & \cdots & P_5(\hat{x}_1^{(i)})P_5(\hat{x}_1^{(1)}) & \cdots & P_5(\hat{x}_1^{(i)})P_5(\hat{x}_1^{(3)}) \\ & \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ & P_5(\hat{x}_M^{(i)})P_1(\hat{x}_M^{(1)}) & \cdots & P_5(\hat{x}_M^{(i)})P_1(\hat{x}_M^{(3)}) & & P_5(\hat{x}_M^{(i)})P_5(\hat{x}_M^{(1)}) & \cdots & P_5(\hat{x}_M^{(i)})P_5(\hat{x}_M^{(3)}) \end{array}\right].$$

We also have the data vector $y^{(i)} \in \mathbb{R}^M$ with the elements $y_j^{(i)} = \frac{\hat{x}_j^{(i)} - \hat{x}_{j-1}^{(i)}}{\Delta t_j}$ for $i = 1, 2, 3$ and $j = 1, \ldots, 1500$. Now we solve the linear inverse problem as in (5) using TSVD, which yields $\hat{z}^{(i)} = (\hat{\omega}_i, z_{P_{11}}^{(i)}, \ldots, z_{P_{15}}^{(i)}, z_{P_{21}}^{(i)}, \ldots, z_{P_{25}}^{(i)}, \ldots, z_{P_{51}}^{(i)}, \ldots, z_{P_{55}}^{(i)})' \in \mathbb{R}^{76}$, $i = 1, 2, 3$, where the constant coefficients, $\hat{\omega}_1 = 2.6299$, $\hat{\omega}_2 = 0.9685$, $\hat{\omega}_3 = 2.4019$, are the estimated drift dynamics (natural frequencies) and the cut-off threshold for the singular values is set to be $\epsilon = 10^{-8}$. These estimates $\hat{z}^{(i)}$ define the network dynamics, and the reconstructed network is shown in Figure S1. The true (synthetic) and the estimated state trajectories are plotted in Figure S1(a), and the coupling functions are displayed in Figure S1(b). The estimated topology and coupling strength show excellent agreement with that of the true network. This result is also confirmed by the comparison of the adjacency matrices of the true ($\mathcal{M}$) and estimated ($\hat{\mathcal{M}}$) networks,

$$\mathcal{M} = \begin{bmatrix} 0.0000 & 0.6167 & 1.2229 \\ 0.8710 & 0.0000 & 3.8409 \\ 4.9523 & 2.8912 & 0.0000 \end{bmatrix}, \qquad \hat{\mathcal{M}} = \begin{bmatrix} 0.0000 & 0.6175 & 1.2229 \\ 0.8713 & 0.0000 & 3.8409 \\ 4.9523 & 2.8912 & 0.0000 \end{bmatrix},$$

in which the ij^{th} elements of $\mathcal{M}$ and $\hat{\mathcal{M}}$, α_{ij} and $\hat{\alpha}_{ij}$, represent the synthetic and estimated coupling strength between oscillators i and j, respectively, as defined in (3). This shows a perfect network reconstruction by ICON with 100% recovery in connections and almost perfect (99.8%) recovery in coupling strength, where the recovery rate in coupling strength is calculated by

$$\Big(\sum_{i,j=1}^{3}\sum_{k,\ell=1}^{5}(|\hat{b}_{ij}^{k\ell}|)\Big)\Big/\Big(\sum_{i,j=1}^{3}\sum_{k,\ell=1}^{5}(|b_{ij}^{k\ell}|)\Big) \approx 99.8\%.$$

Remark 2 (Network Inference based on autonomous and non-autonomous data) *The estimation based on non-autonomous data generated by the procedure described in Section 2.2 is valid and causes minor discrepancy compared to the use of autonomous (state-dependent) data. The difference results from the step of approximating $K_{ij}(t)$ by $K_{ij}(x_i(t), x_j(t))$. To illustrate the subtle distinction between the two scenarios, we also synthesize the phase data of the example above through an autonomous approach. That is, we first generate the coupling functions as Fourier series, i.e., $K_{ij}(x_i, x_j) = \sum_{j=1,j\neq i}^{N} \sum_{k=1}^{r} \sum_{\ell=1}^{r} b_{ij}^{k\ell} P_k(x_i) P_\ell(x_j)$, where P_k and P_ℓ are Fourier bases, $N = 3$, $r = 5$, and the coefficients $b_{ij}^{k\ell}$ are randomly selected over $[0, 2]$ using the Matlab built-in function 'rand'. Then, we obtain the phase data $\hat{x}_i$, $i = 1, 2, 3$, by evolving (12). The true and the recovered networks are represented by their adjacency matrices, $\mathcal{M}$ and $\hat{\mathcal{M}}$, respectively, i.e.,*

$$\mathcal{M} = \begin{bmatrix} 0.0000 & 0.6150 & 1.2500 \\ 0.8750 & 0.0000 & 3.8500 \\ 5.0000 & 3.0000 & 0.0000 \end{bmatrix}, \qquad \hat{\mathcal{M}} = \begin{bmatrix} 0.0000 & 0.6148 & 1.2498 \\ 0.8749 & 0.0000 & 3.8500 \\ 4.9999 & 3.0000 & 0.0000 \end{bmatrix}.$$

This result of 99.9% recovery rate shows excellent estimation performance of ICON based on autonomous data.

2.4.2 Non-oscillatory Networks of Multi-dimensional Agents

In this section, we illustrate the generality of the ICON technique for non-oscillatory networks consisting of multi-dimensional agents, i.e., $n > 1$. To fix ideas, we consider an example involving a network of 10 non-periodic two-dimensional oscillators, where the self-dynamics of each agent is of the form

$$f(x_i) = \begin{bmatrix} \xi_i & -\omega_i \\ \omega_i & \eta_i \end{bmatrix} x_i, \qquad i = 1, 2, \ldots, 10.$$

We generate the synthetic data of this networked system, $\left\{ \begin{bmatrix} \hat{x}_{k1}^{(i)} \\ \hat{x}_{k2}^{(i)} \end{bmatrix} \right\}$, $k = 0, \ldots, M$, according to the procedure described in Section 2.2 with $M = 201$ (number of time sampling points) over $T = 2$ (time duration), 80% sparsity rate, and the coupling strength $\alpha_{ij} \in [-5, 5]$, where $k = 1, \ldots, 201$ and $i, j = 1, \ldots, 10$. The coupling between dimension u of agent i and dimension v of agent j can be represented by a tensor product, given by

$$K_{i_u j_v} = \sum_{k=1}^{r} \sum_{\ell=1}^{r} \sum_{\substack{j=1 \\ j\neq i}}^{N} b_{i_u j_v}^{k\ell} P_k\Big(\begin{bmatrix} \hat{x}_{1u}^{(i)} \\ \vdots \\ \hat{x}_{Mu}^{(i)} \end{bmatrix} \Big) \otimes P_\ell\Big(\begin{bmatrix} \hat{x}_{1v}^{(j)} \\ \vdots \\ \hat{x}_{Mv}^{(j)} \end{bmatrix} \Big). \tag{13}$$

The corresponding linear inverse problem (5) is formulated as $\min_{z_s^{(i)}} \|y_s^{(i)} - A_s^{(i)} z_s^{(i)}\|_2$, $s = 1, 2$, where the elements of $y_s^{(i)} \in \mathbb{R}^M$, i.e., $y_{ks}^{(i)} = \frac{\hat{x}_{ks}^{(i)} - \hat{x}_{(k-1)s}^{(i)}}{\Delta t_k}$ for $k = 1, \ldots, M$, denote the finite-difference approximation of the temporal dynamics of dimension s, $s = 1, 2$, of agent i at time t_k with $\Delta t_k = t_k - t_{k-1}$; $A_s^{(i)} \in \mathbb{R}^{M \times (nr^2 N + 2)}$ is the matrix defined by

$$A_s^{(i)} = \Big[L \big| O_{11_s}^{(i)} \big| \cdots \big| O_{1r_s}^{(i)} \big| O_{21_s}^{(i)} \big| \cdots \big| O_{2r_s}^{(i)} \big| \cdots \big| O_{r1_s}^{(i)} \big| \cdots \big| O_{rr_s}^{(i)} \Big],$$

in which $L \in \mathbb{R}^{M\times 2}$ denotes the baseline dynamics defined by $L = \begin{bmatrix} \hat{x}_{11}^{(1)} & \hat{x}_{12}^{(2)} \\ \vdots & \vdots \\ \hat{x}_{M1}^{(1)} & \hat{x}_{M2}^{(2)} \end{bmatrix}$, and $O_{k\ell_s}^{(i)} \in \mathbb{R}^{M\times nN}$ $(n = 2)$, $k, \ell = 1, \ldots, r$, are the matrices involving the orthonormal bases,

$$O_{k\ell_s}^{(i)} = \begin{bmatrix} P_k(\hat{x}_{1s}^{(i)})P_\ell(\hat{x}_{11}^{(1)}) & \cdots & P_k(\hat{x}_{1s}^{(i)})P_\ell(\hat{x}_{11}^{(N)}) & P_k(\hat{x}_{1s}^{(i)})P_\ell(\hat{x}_{12}^{(1)}) & \cdots & P_k(\hat{x}_{1s}^{(i)})P_\ell(\hat{x}_{12}^{(N)}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ P_k(\hat{x}_{Ms}^{(i)})P_\ell(\hat{x}_{M1}^{(1)}) & \cdots & P_k(\hat{x}_{Ms}^{(i)})P_\ell(\hat{x}_{M1}^{(N)}) & P_k(\hat{x}_{Ms}^{(i)})P_\ell(\hat{x}_{M2}^{(1)}) & \cdots & P_k(\hat{x}_{Ms}^{(i)})P_\ell(\hat{x}_{M2}^{(N)}) \end{bmatrix};$$

and $z_s^{(i)} = (c_s^{(i)}, z_{P_{11_s}}^{(i)}, \ldots, z_{P_{1r_s}}^{(i)}, z_{P_{21_s}}^{(i)}, \ldots, z_{P_{2r_s}}^{(i)}, \ldots, z_{P_{r1_s}}^{(i)}, \ldots, z_{P_{rr_s}}^{(i)})' \in \mathbb{R}^{r^2nN+2}$ is the coefficient vector to be estimated for agent i, in which $c_s^{(i)} \in \mathbb{R}^{1\times 2}$ and $z_{P_{k\ell_s}}^{(i)} \in \mathbb{R}^{1\times nN}$ are the coefficient vectors associated with L and the base P_kP_ℓ for dimension s, respectively.

Similarly, these least-squares problems can be solved by using the TSVD (in general can be solved by using the tensor (multilinear) SVD [34]), and the estimated parameters $c_s^{(i)}$ containing $\hat{\xi}_i$, $\hat{\omega}_i$, and $\hat{\eta}_i$, are shown in the following table, which agree with the true values ξ, ω, and η, respectively.

ξ_i	-2.0920	-0.2996	-0.0851	-1.3978	-0.9021	-1.4808	-0.1442	-1.9975	-2.8041	-0.0253
$\hat{\xi}_i$	-2.0921	-0.2967	-0.0050	-1.3978	-0.9031	-1.4808	-0.1447	-1.9975	-2.8041	-0.0268
ω_i	-0.5229	1.2028	0.4439	-0.0352	-0.2805	-1.4783	0.3712	-0.5778	0.2201	0.4783
$\hat{\omega}_i$	-0.5234	1.2025	0.4437	-0.0329	-0.0800	-1.4783	0.3722	-0.5778	0.2200	0.4785
η_i	-0.0871	-0.0070	-0.2458	-0.5732	-1.1532	-0.8467	-1.0793	-0.7473	-1.3874	-0.0170
$\hat{\eta}_i$	-0.0875	-0.0100	-0.2460	-0.5732	-1.1532	-0.8467	-1.0793	-0.7474	-1.3874	-0.0182

In addition, the adjacency matrices of the true ($\mathcal{M}$) and estimated ($\hat{\mathcal{M}}$) networks shown below also demonstrate the accuracy of the ICON estimation,

$$\mathcal{M} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.1268 & 0.1065 & 0.0083 & 3.8192 & 0.6517 & 0.4258 & 0.0000 \\ 0.6995 & 0.0000 & 4.5908 & 0.0000 & 0.4477 & 0.0021 & 0.0000 & 1.7608 & 0.7478 & 4.4420 \\ 0.6095 & 1.4748 & 0.0000 & 1.9085 & 0.7482 & 0.0305 & 1.0076 & 1.5516 & 0.0000 & 0.4166 \\ 0.0000 & 0.2893 & 0.4904 & 0.0000 & 0.8956 & 0.5898 & 2.5807 & 1.5577 & 0.8336 & 2.3556 \\ 2.2007 & 2.6630 & 0.0000 & 0.0564 & 0.0000 & 3.4327 & 0.0162 & 0.0000 & 0.0618 & 0.8156 \\ 1.8485 & 0.0000 & 1.4446 & 1.0147 & 0.3760 & 0.0000 & 0.0549 & 0.0000 & 3.1003 & 5.3161 \\ 0.2140 & 0.0000 & 0.0000 & 1.6154 & 0.4362 & 0.1139 & 0.0000 & 0.0377 & 2.2530 & 0.7163 \\ 0.0000 & 0.0000 & 0.5203 & 5.3331 & 0.0000 & 0.9635 & 1.1648 & 0.0000 & 0.0000 & 0.0000 \\ 0.1439 & 0.1821 & 0.4999 & 2.4388 & 0.8186 & 1.0682 & 0.0698 & 4.1223 & 0.0000 & 3.6328 \\ 1.2805 & 1.3370 & 0.0000 & 0.0000 & 0.9856 & 3.3314 & 2.0202 & 0.1025 & 0.0115 & 0.0000 \end{bmatrix},$$

$$\hat{\mathcal{M}} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0013 & 0.1260 & 0.1117 & 0.0143 & 3.8190 & 0.6517 & 0.4268 & 0.0001 \\ 0.7009 & 0.0000 & 4.5908 & 0.0000 & 0.4478 & 0.0000 & 0.0000 & 1.7612 & 0.7478 & 4.4420 \\ 0.6096 & 1.4748 & 0.0000 & 1.9090 & 0.7485 & 0.0296 & 1.0050 & 1.5516 & 0.0000 & 0.4167 \\ 0.0000 & 0.2901 & 0.4902 & 0.0000 & 0.8956 & 0.5910 & 2.5807 & 1.5577 & 0.8336 & 2.3563 \\ 2.2007 & 2.6633 & 0.0005 & 0.0550 & 0.0000 & 3.4300 & 0.0000 & 0.0101 & 0.0618 & 0.8150 \\ 1.8485 & 0.0002 & 1.4444 & 1.0177 & 0.3760 & 0.0000 & 0.0549 & 0.0000 & 3.1003 & 5.3161 \\ 0.2151 & 0.0001 & 0.0001 & 1.6154 & 0.4364 & 0.1201 & 0.0000 & 0.0357 & 2.2530 & 0.7163 \\ 0.0000 & 0.0000 & 0.5201 & 5.3331 & 0.0000 & 0.9635 & 1.1649 & 0.0000 & 0.0000 & 0.0005 \\ 0.1436 & 0.1827 & 0.4995 & 2.4388 & 0.8178 & 1.0682 & 0.0711 & 4.1223 & 0.0000 & 3.6328 \\ 1.2805 & 1.3370 & 0.0000 & 0.0000 & 0.9856 & 3.3314 & 2.0202 & 0.0993 & 0.0115 & 0.0000 \end{bmatrix},$$

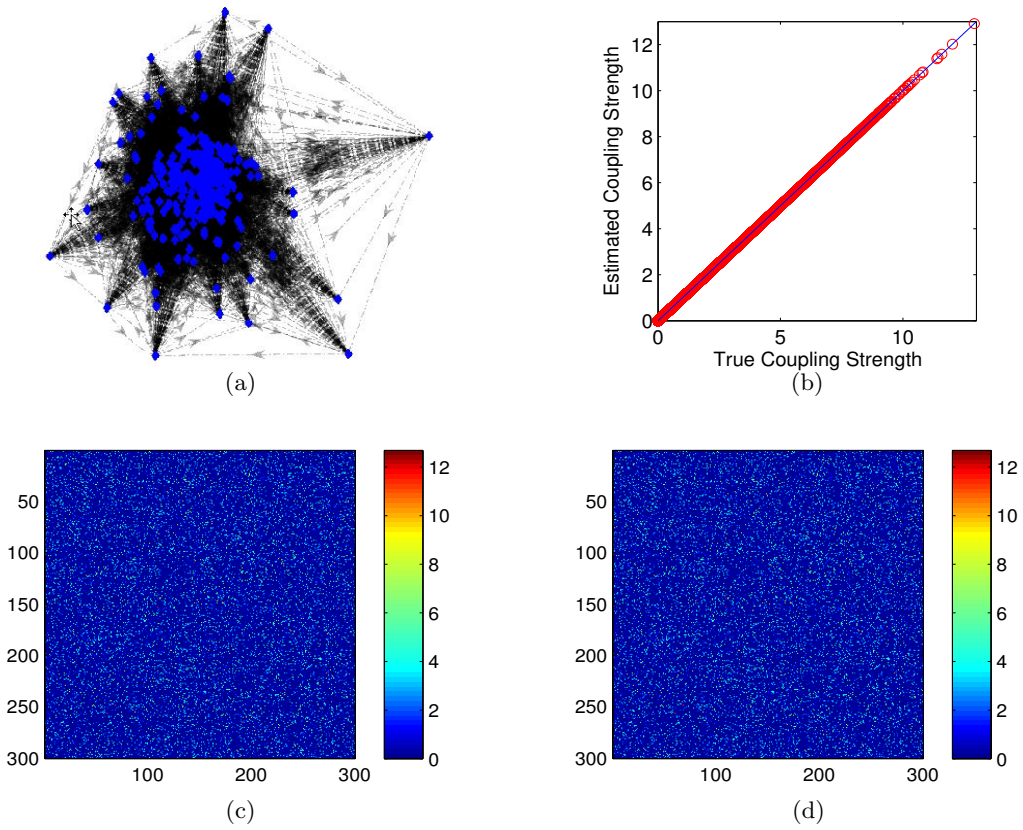


Figure S2: (a) The graph representation of the topology of a time-varying dense noisy network of 300 oscillators (20.1% sparsity) synthesized by using the procedure described in Section 2.2.2 over the time period $t \in [0, 20]$. The coupling functions are generated by interpolating polynomials with weights α_{ij} based on a uniform grid of 81 points, and the polynomials are sampled by 15000 points with additive white noise of 1% intensity. The baseline dynamics $f(x_i) = \omega_i$ of each agent and the coupling strength α_{ij} for each K_{ij} were randomly selected within $(0, 5)$ and $[0, 12]$, respectively. (c) The heat map showing the coupling topology of the network. (d) The estimated coupling topology of the network with a full topology recovery (17986/17986 connections identified) and 96.1% accuracy on recovering the coupling strength as illustrated in (b).

where the ij^{th} elements, α_{ij} and $\hat{\alpha}_{ij}$, of $\mathcal{M}$ and $\hat{\mathcal{M}}$, respectively, represent the synthetic and the estimated coupling strength between oscillators i and j, defined by

$$\alpha_{ij} \doteq \sqrt{\sum_{u,v=1}^{n} \alpha_{i_u j_v}^2}, \qquad \hat{\alpha}_{ij} \doteq \sqrt{\sum_{u,v=1}^{n} \hat{\alpha}_{i_u j_v}^2},$$

where $\alpha_{i_u j_v}$, $\hat{\alpha}_{i_u j_v}$ are defined in (3).

2.4.3 Large-scale Networks

A dense network of 300 agents: In Figure S2, we show the topology estimation for a dense synthetic network (20.1% sparsity rate) of 300 oscillators. The network shown in Figure S2(a) was synthesized using the procedure described in Section 2.2.2 over the time period $t \in [0, 20]$. In particular, the coupling functions are generated by interpolating polynomials with weights α_{ij} based on a uniform grid of 81 points, i.e., $0 = t_0 < t_1 < \cdots < t_{80} = 20$. These polynomials are then sampled by 15000 points with additive white noise of 1% intensity. The baseline dynamics $f(x_i) = \omega_i$ of each agent and the coupling strength α_{ij} (the ij^{th} element of $\mathcal{M}$) for each K_{ij} were randomly selected within $(0, 5)$ and $[0, 12]$, respectively. The heat map in Figure S2(a) illustrates

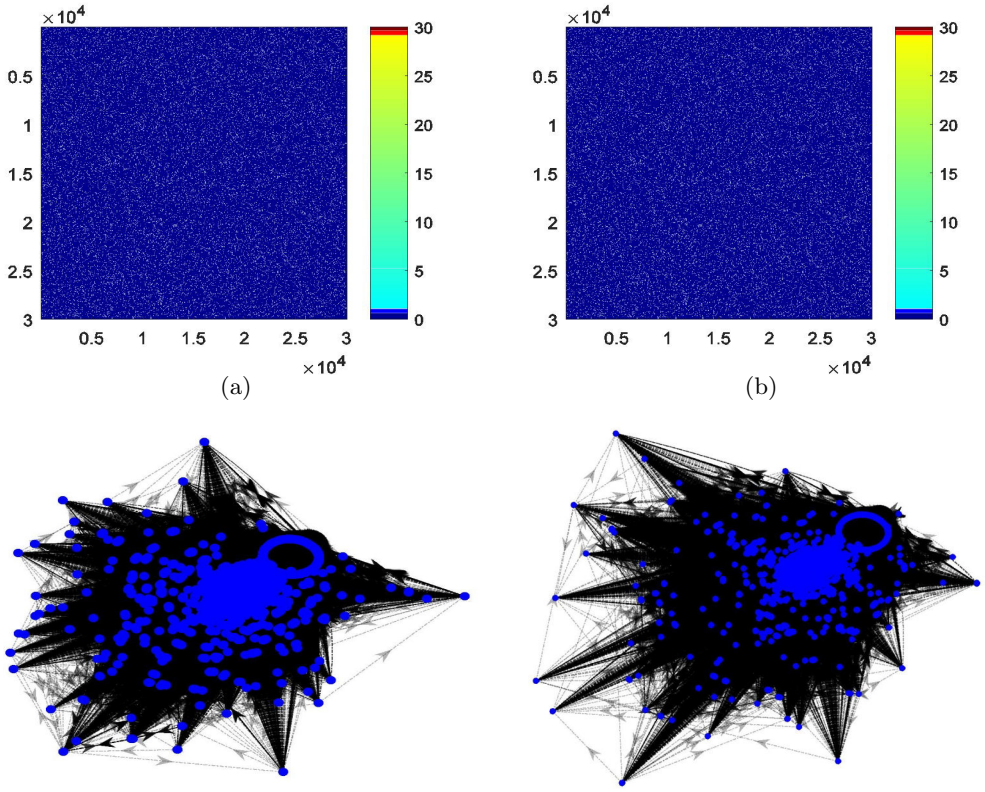


Figure S3: (a) The coupling topology of a time-varying noisy network of 30000 oscillators (1.2% sparsity and 1% noise intensity) synthesized using the procedure described in Section 2.2.1 over the time period $t \in [0, 20]$ (60000 sampling points) in the presence of additive white noise of 1% intensity. The baseline dynamics $f(x_i) = \omega_i$ of each agent and the coupling strength α_{ij} for each K_{ij} were randomly selected within $(0, 10)$ and $[0, 32]$, respectively. In the estimation, the time-varying coupling functions K_{ij} were approximated by the two-dimensional Legendre polynomials of order $r = 1$. (b) The estimated coupling topology of the network with 87.9% (9465526/10768516 identified). (c) The graph representation of the network in which nodes with stronger connectivity are plotted in the center of the graph and nodes with weaker connectivity are plotted off the center. (d) The estimated connectivity topology showing good agreement with that of the true (synthetic) network shown in Figure 3(c).

the coupling topology of the network. In the estimation, the time-varying coupling functions K_{ij} were approximated by two-dimensional Fourier series of order $r = 3$. Figure S2(d) shows a perfect recovery in topology (17986/17986 connections identified) of this dense network using ICON with 96.1% accuracy rate for recovering the coupling strength α_{ij}. The comparison between the true and estimated coupling strength is shown in Figure S2(b). The computation time for this topology estimation was 101.0 minutes and 51.64 seconds using parallel computing.

A myriad network of 30000 agents: In Figure S3, we show the capability and robustness of ICON for inferring a very large network of 30000 agents (1.2% sparsity rate). The synthetic network shown in Figure S3(a) (coupling topology) and Figure S3(c) (graph representation) was synthesized using the procedure described in Section 2.2.1 over the time period $t \in [0, 20]$ (60000 sampling points) in the presence of additive white noise of 1% intensity, and the coupling functions were expanded using Fourier series of order $r = 1$. The baseline dynamics $f(x_i) = \omega_i$ of each agent and the coupling strength α_{ij} (the ij^{th} element of $\mathcal{M}$) for each K_{ij} were randomly selected within $(0, 10)$ and $[0, 32]$, respectively. In the estimation, the time-varying coupling functions K_{ij} were approximated by the Fourier series of order $r = 1$. The recovery rate in connections was 87.9% (9465526/10768516 identified) (see Figure S3(b)). Figure S3(d) illustrates the estimated

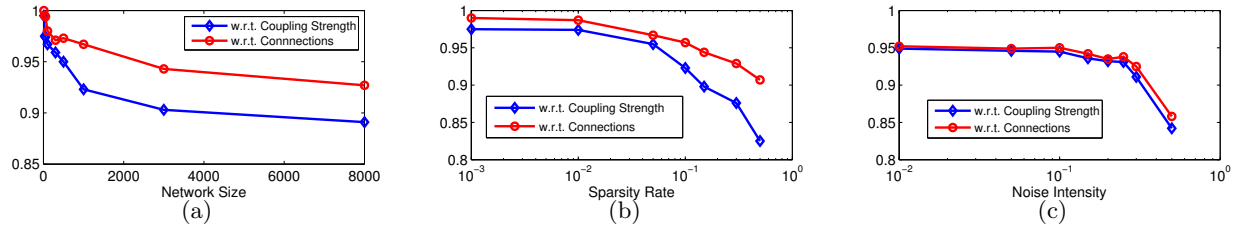


Figure S4: (a) The averaged accuracy rates of ICON applied to estimate networks of different sizes (200 trials for each size of networks) with the same sparsity (1.5%) and noise intensity (1%) for the time period $t \in [0, 20]$. The performance is plotted in terms of the recovery of coupling strength as well as connections. (b) The averaged accuracy rates of ICON applied to estimate 500-node networks of various sparsity levels (200 trials for each case) in the presence of 1% noise intensity for $t \in [0, 20]$. (c) The averaged accuracy rates of ICON applied to estimate 500-node networks (1.5% sparsity) in the presence of different noise intensity ratios (200 trials for each size of networks) for $t \in [0, 20]$. The accuracy of recovering the connections (blue diamonds) and the coupling strength (red pentagrams) are both displayed.

connectivity topology (nodes with stronger connectivity are plotted in the center of the graph, and nodes with weaker connectivity are plotted off the center), which agrees with that of the true (synthetic) network shown in Figure 3(c).

3 Robustness, Reliability, and Efficiency of ICON

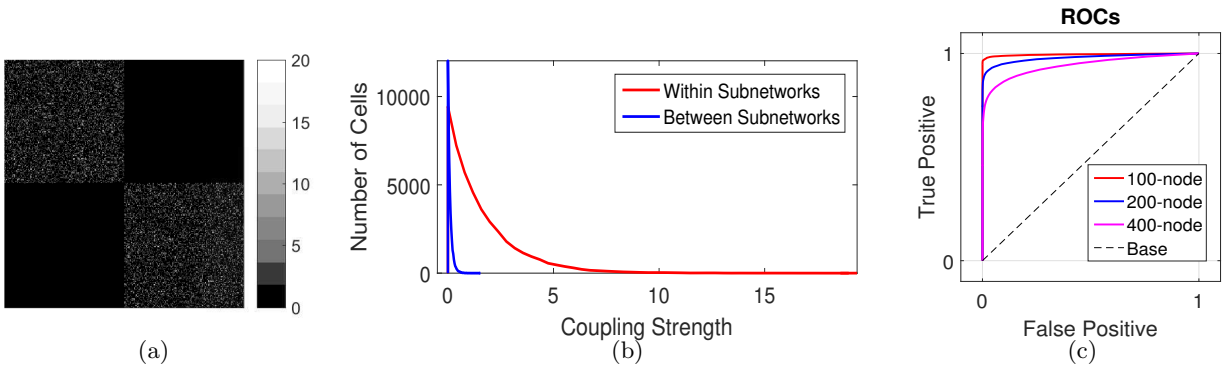
In this section, we apply ICON to infer topology of synthetic networks in different sizes with different noise intensities or sparsity levels. We use these estimation results to illustrate the reliability and robustness of ICON through a statistical hypothesis test.

3.1 Robustness with respect to Network Properties

Here, we demonstrate the robustness of ICON to some essential network properties, including the size, noise intensity, and sparsity. Figure S4(a) displays the averaged accuracy rates of ICON (in terms of recovering the coupling strength or the number of connections) applied to estimate networks of different sizes (200 trials for each size of networks) with the same sparsity (1.5%) and noise intensity (1%). These networks were synthesized in the same way as those generated in Figure S2. It is shown that the performance of the estimation remains excellent with the increase of the network size (88% averaged accuracy over 200 trials for estimating networks of 8000 oscillators). Similar performance analysis of ICON was conducted and shown in Figure S4(b) and S4(c) with respect to different sparsity levels (network size = 500; noise intensity = 1%) and noise intensity ratios (network size = 500; sparsity = 1.5%), respectively. While the accuracy of topology estimation was slightly degraded with the increase of the sparsity and noise intensity of the network, the ICON technique was demonstrated to be sufficiently reliable and robust.

3.2 Reliability Analysis by Hypothesis Test

In addition to the validation through various numerical experiments presented in Section 3.1, the reliability of ICON can be further illustrated statistically through a hypothesis test. To achieve this, we created two independent networks of 400 nodes, each of which is densely connected, and then considered the network of 800 nodes composed of these two subnetworks of 400 nodes (see Figure S5(a)). The distributions of the coupling strength of estimated connections within each of and between the two 400-node subnetworks, shown in Figure S5(a), were obtained based on the estimation result by ICON, which are plotted in Figure S5(b). Based on these distributions, we obtained the Receiver Operating Characteristic (ROC) curves for testing whether a connection is within one of the two subnetworks (true positive) or not based on its strength. The ROC



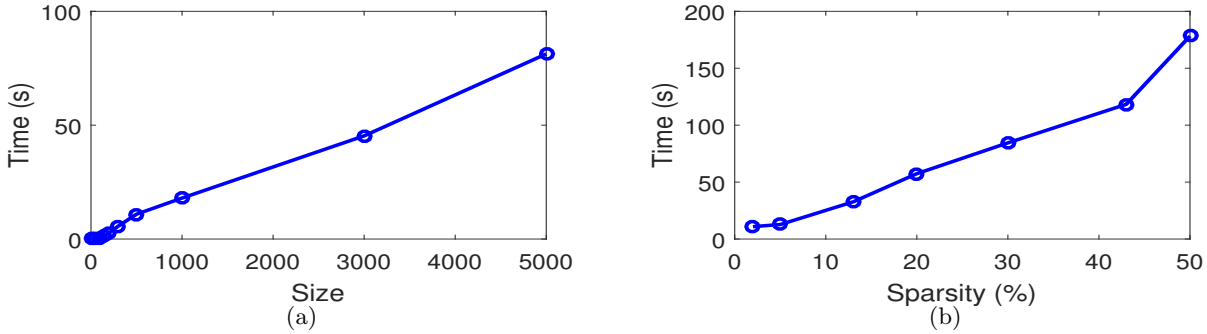
(a) (b) (c)

Figure S5: (a) The topology of the synthetic network constituted by two 400-node densely-connected networks. (b) The distributions of the estimated coupling strength of connections within each of and between the two 400-node subnetworks shown in (a). (c) The receiver operating characteristic (ROC) curves for testing whether a connection is within one of the two subnetworks (true positive) of size 100, 200, and 400, based on its coupling strength.

curves plotted in Figure S5(c) are for networks of different sizes, where we observe high sensitivity (probability of detection) of the estimation results. This hypothesis test demonstrates the reliability of the ICON technique. Note that it is expected and reasonable to observe slight declination in the sensitivity with the increase of the network size.

3.3 Computational Efficiency and Time Complexity of ICON

The ICON technique is computationally efficient owing to the linear formulation and its parallel structure as presented in (5). Here, we illustrate the computational efficiency of ICON by using the numerical experiments for estimating topology of the synthetic networks in different sizes and sparsity levels created in Section 3.1. Figure S6(a) and S6(b) show the averaged running time per agent over 200 trials for each experiment (circled) with respect to the change in the network size and sparsity, respectively. In these figures, we observe sublinear time complexity of ICON in both cases.



(a) (b)

Figure S6: (a) The averaged one-agent running time over 200 trials for each experiment (circled) on estimating networks (created in Section 3.1, Figure 4(a)) of different sizes, where the sparsity of the networks is set to be 1.5%, the noise intensity is 1%, and the time horizon is $t \in [0, 20]$. (b) The averaged running time over 200 trials for each experiment (circled) on estimating 100-node networks (created in Section 3.1, Figure 4(b)) of different sparsity levels in the presence of additive white noise of 1% intensity using parallel computing, where the time-horizon is $t \in [0, 20]$.

4 Social Synchronization in Groups of Mice

Figure S7 illustrates the estimated network topology for 7 different quintets of mice using ICON. The synchronization behavior of each group of mice can then be observed by evolving these estimated

social networks with time. The distinct colored nodes in the figure represent five different mice in each quintet and the edges represent the coupling (solid lines for strong connections and dashed lines for weak connections with thickness denoting the coupling strength). The networks illustrated in D, E, F, and G, form connected graphs over the period of cohabitation, which imply the inclination of synchrony, ρ, as also reflected by the calculated sync indices (Kuramoto order [25, 35]), and the dominant eigenvalues λ_2 of the Laplacian matrix of each of these connected networks. The increasing and stabilizing trend of ρ and the values of λ_2 (0.0010, 0.0030, 0.0304, and 0.0734 for the networks D to G, respectively) indicate that they will be synchronized (in the order of synchronizability) after 68 days of cohabitation; while the oscillatory behavior of ρ and the negligible values of λ_2 conclude asynchrony for the other three cases A, B, and C. Such predictions of social synchronization from data using ICON match perfectly with the experimental observations [36,37].

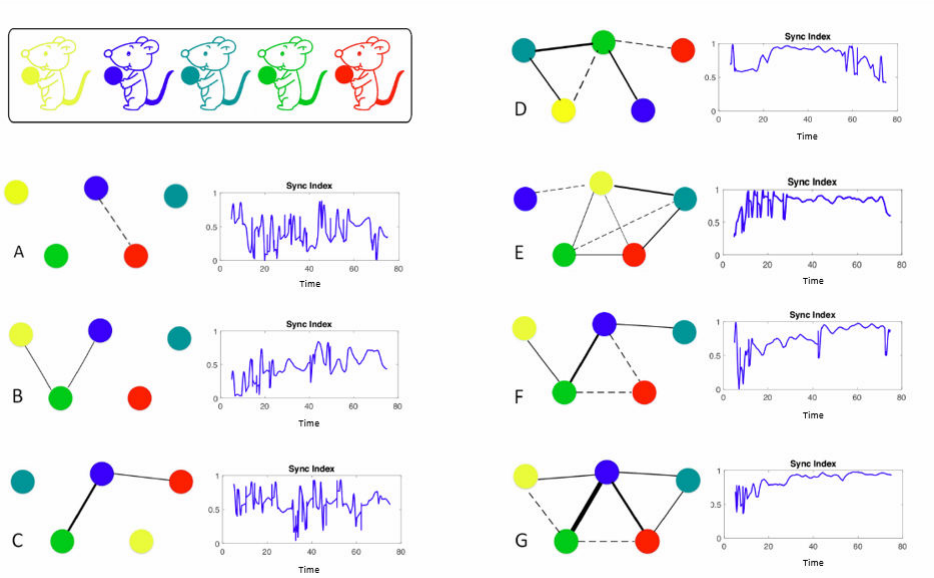


Figure S7: Social Networks of Laboratory Mice. Subfigures A to G show the estimated network topology of 7 quintets of laboratory mice using ICON based on the measurement data of their body temperature. These results are aligned in the ascending order with respect to the inclination of synchrony based on the trend of the sync indices and the values of the dominant eigenvalues of the Laplacian matrix of the networks. Networks D to G are predicted to be synchronous, while A to C are not after the period of cohabitation, which match perfectly with the experimental observations.

The estimation was based on 68 days of temperature recording, which was set by the Ibutton's maximum capacity at the 15-min sampling rate (a trade-off we made between frequency of sampling and length of recording). The table below shows the λ_2 and σ values plotted in Figure 4E in the manuscript:

λ_2	0.0000	0.0000	0.0001	0.0002	0.0003	0.0306	0.0734
σ	0.3560	0.5036	0.5036	0.5609	0.6737	0.9171	0.9436

The Pearson correlation coefficient between these two variables is given by

$$r_{\text{Pearson}} = \frac{\sum(\sigma - M_\sigma)(\lambda_2 - M_{\lambda_2})}{\sqrt{\Big(\sum(\sigma - M_\sigma)^2\Big)\Big(\sum(\lambda_2 - M_{\lambda_2})^2\Big)}} = 0.8301.$$

Also, from the table above, we obtain the ranks of the values of λ_2 and σ as follows:

r_{λ_2}	1.5	1.5	3	4	5	6	7
r_σ	1	2.5	2.5	4	5	6	7

with which the Spearman correlation coefficient can be computed, given by

$$r_{\text{Spearman}} = \frac{\sum(r_\sigma - M_{r_\sigma})(r_{\lambda_2} - M_{r_{\lambda_2}})}{\sqrt{\Big(\sum(r_\sigma - M_{r_\sigma})^2\Big)\Big(\sum(r_{\lambda_2} - M_{r_{\lambda_2}})^2\Big)}} = 0.9727,$$

and the two-tailed value is 0.00023. This simple statistical analysis suggests that λ_2 and σ are monotonically correlated based on the given small sample size (7 data points from the experiments).

5 Circadian Neurons

Figure S8 is the graph representation of the estimated SCN network of 541 cells, in which there are 173,627 connections. In this experiment, the raw data for the Per2::Luc SCN recordings were the complete measurement following TTX treatment.

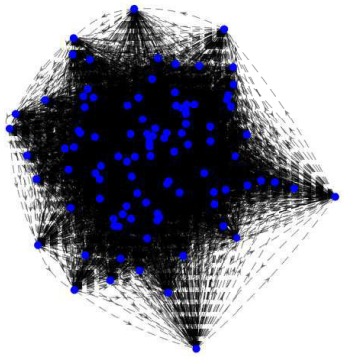


Figure S8: Graph representation of the estimated SCN network.

References Cited

[1] J. H. Abel, K. Meeker, D. Granados-Fuentes, P. C. St. John, T. J. Wang, B. B. Bales, F. J. Doyle, E. D. Herzog, and L. R. Petzold, "Functional network inference of the suprachiasmatic nucleus," *Proceedings of the National Academy of Sciences*, vol. 113, no. 16, pp. 4512–4517, 2016.

[2] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *PNAS*, vol. 113, no. 15, pp. 3932–3937, 2016.

[3] A. J. Majda, C. Franzke, and D. Crommelin, "Normal forms for reduced stochastic climate models," *Proceedings of the National Academy of Sciences*, vol. 106, no. 10, pp. 3649–3653, 2009.

[4] Y. H. Chang, J. W. Gray, and C. J. Tomlin, "Exact reconstruction of gene regulatory networks using compressive sensing," *BMC Bioinformatics*, vol. 15, no. 1, p. 400, 2014.

[5] M. C. Antle and R. Silver, "Orchestrating time: arrangements of the brain circadian clock," *Trends Neurosci.*, vol. 28, no. 3, pp. 145–151, 2005. Review 909MV English Times Cited:20 Cited References Count:69.

[6] D. Glannakis and A. J. Majda, "Data-Driven Methods for Dynamical Systems: Quantifying Predictability and Extracting Spatiotemporal Patterns," *Mathematical and Computational Modeling: With Applications in Natural and Social Sciences, Engineering, and the Arts*, pp. 137–191, 2015.

[7] K. Ota and T. Aoyagi, "Direct extraction of phase dynamics from fluctuating rhythmic data based on a bayesian approach," *arXiv:1405.4126*, 2014.

[8] S. G. Shandilya and M. Timme, "Inferring network topology from complex dynamics," *New Journal of Physics*, vol. 13, p. 013004, 2011.

[9] W.-X. Wang, R. Yang, Y.-C. Lai, V.Kovanis, and M. A. F. Harrison, "Time-seriesbased prediction of complex oscillator networks via compressive sensing," *EPL*, vol. 94, no. 4, p. 48006, 2011.

[10] M. Timme and J. Casadiego, "Revealing networks from dynamics: an introduction," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 343001, 2014.

[11] T. Stankovski, A. Duggento, P. V. E. McClintock, and A. Stefanovska, "Inference of time-evolving coupled dynamical systems in the presence of noise," *PRL*, vol. 109, no. 024101, 2012.

[12] B. Kralemann, A. Pikovsky, and M. Rosenblum, "Reconstructing phase dynamics of oscillator networks," *Chaos*, vol. 21, p. 025104, 2011.

[13] R. D. Smet and K. Marchal, "Advantages and limitations of current network inference methods," *Nature Reviews: Microbiology*, vol. 8, p. 717, 2010.

[14] D. Henriques, A. F. Villaverde, M. Rocha, J. Saez-Rodriguez, and J. R. Bangal, "Data-driven reverse engineering of signaling pathways using ensembles of dynamic models," *PLoS Computational Biology*, vol. 13, p. e1005379, 2017.

[15] W. Rudin, *Principles of Mathematical Analysis (3rd Edition)*. St. Louis: McGraw-Hill Inc., 1976.

[16] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*. Berlin: Springer-Verlag, 2006.

[17] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems*. Amsterdam: Elsevier Academic Press, 2013.

[18] P. C. Hansen, "The truncatedsvd as a method for regularization," *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534–553, 1987.

[19] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[20] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. LVII, p. 14131457, 2004.

[21] Y. C. Elder and G. Kutyniok, *Compressed Sensing: Theory and Applications*. New York: Cambridge University Press, 2012.

[22] D. Hansel, G. Mato, and C. Meunier, "Phase dynamics for weakly coupled hodgkin-huxley neurons," *EPL (Europhysics Letters)*, vol. 23, no. 5, p. 367, 1993.

[23] B. Pfeuty, Q. Thommen, and M. Lefranc, "Robust entrainment of circadian oscillators requires specific phase response curves," *Biophysical Journal*, vol. 100, pp. 2557–2565, 06 2011.

[24] G. Bloch, "The social clock of the honeybee," *Journal of Biological Rhythms*, vol. 25, pp. 307–317, OCT 2010.

[25] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*. New York: Springer, 1984.

[26] I. Z. Kiss, C. G. Rusin, H. Kori, and J. L. Hudson, "Engineering complex dynamical structures: Sequential patterns and desynchronization," *Science*, vol. 316, no. 5833, pp. 1886–1889, 2007.

[27] A. Nabi and J. Moehlis, "Charge-balanced optimal input for phase models of spiking neurons," *Discrete and Continuous Dynamical Systems*, vol. 28, no. 4, 2010.

[28] A. Zlotnik and J.-S. Li, "Optimal entrainment of neural oscillator ensembles," *Journal of Neural Engineering*, vol. 9, no. 4, p. 046015(12), 2012.

[29] F. Hoppensteadt and E. Izhikevich, *Weakly connected neural networks*. New Jersey: Springer-Verlag, 1997.

[30] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods*. 2006.

[31] A. S. Pikovsky, M. Rosenblum, and J.Kurths, *Synchronization: A Universal Concept in Nonlinear Science*. Cambridge: Cambridge University Press, 2001.

[32] R. M. Sammer and J.-J. E. Slotine, "Stable adaptive control and recursive identification using radial gaussian networks," *Proceedings of the 30th CDC*, 1991.

[33] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. USA: University of Illinois Press, 1963.

[34] J. Chen and Y. Saad, "On the tensor svd and the optimal low rank orthogonal approximation of tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 4, pp. 1709–1734, 2009.

[35] I. T. Tokuda, S. Jain, I. Z. Kiss, and J. L. Hudson, "Inferring phase equations from multivariate time series," *PRL*, vol. 99, no. 064101, 2007.

[36] T. L. Leise, P. Indic, M. J. Paul, and W. J. Schwartz, "Wavelet meets actogram," *Journal of Biological Rhythms*, vol. 28, no. 1, pp. 62–68, 2013.

[37] M. J. Paul, P. Indic, and W. J. Schwartz, "Social synchronization of circadian rhythmicity in female mice depends on the number of cohabiting animals," *Biology Letters*, vol. 11, no. 6, 2015.