

the dynamical law governed by the agent's self-dynamics and the influence from other agents,

$$\dot{x}_i(t) = f(x_i) + \sum_{j \neq i} K_{ij}(x_i, x_j), \quad i = 1, \dots, N, \quad [1]$$

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, e.g., frequency, of each agent, and K_{ij} , where $i, j = 1, \dots, N$, is the coupling function between agent i and j (K_{ij} can be different from K_{ji}). The dynamics, i.e., f and K_{ij} , and thus the topology of such a directed network are problematic to infer from data, because the inherent nonlinearity within and between individual agents as well as the large-scale nature of these systems deteriorate the computational efficiency and accuracy for estimation (10); moreover, state observations may be incomplete or only partially available in practice and contaminated by intrinsic and external noise (14).

The central idea of our approach is to approximate the agent's self-dynamics and coupling dynamics, f and K_{ij} , respectively, by using complete orthogonal bases, e.g., the Legendre or Chebyshev polynomials, or the Fourier basis. Specifically, we represent each of these functions as a truncated series, i.e., $f(x_i) \approx \sum_k a_k Q_k(x_i)$ 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 $\{P_k\}_{k=1}^\infty$ constitute orthonormal bases of the respective function space containing f and K_{ij} . Then, the dynamical law in ref. 1 can be expressed as (for $n = 1$)

$$\dot{x}_i(t) = \sum_k a_k Q_k(x_i) + \sum_{j \neq i} \sum_k \sum_\ell b_{ij}^{k\ell} P_k(x_i) P_\ell(x_j), \quad [2]$$

where a_k and $b_{ij}^{k\ell}$ are the scalar coefficients, and the coupling strength α_{ij} is defined by $\alpha_{ij} = \sqrt{\sum_k \sum_\ell (b_{ij}^{k\ell})^2}$. Following this strategy based on the orthonormal basis representation, given the measurement data (time series) of each agent i in the network, the topology estimation can be formulated as a simple linear inverse problem for each agent i , given by

$$\min_{z^{(i)}} \|y^{(i)} - A^{(i)} z^{(i)}\|_2, \quad [3]$$

where $y^{(i)}$ is the data vector; $A^{(i)} \in \mathbb{R}^{M \times (n^2 N + m)}$ is a matrix composed of the orthonormal bases, in which M is the number of data points in the time series for agent i and r and m are the numbers of expansion terms in the truncated series $\{P_k\}$ and $\{Q_k\}$, respectively; and $z^{(i)}$ is the coefficient vector that is being determined (see *SI Appendix, sections 1 and 2* for detailed formulation and examples for the case of $n > 1$).

Most importantly, this formulation enables independent estimation of the coupling dynamics for each individual agent in the network, so that estimating topology of very large networks becomes possible via a parallel computation architecture. Note that there exists a variety of techniques for solving the large-scale least squares (linear inverse) problem as in Eq. 3 (15), for example, the truncated singular value decomposition (TSVD) for regularization of ill-conditioned matrix $A^{(i)}$ (15, 16), compressive sensing for very large-scale networks with sparse topology (17), and iterative methods, such as the iterative recursive thresholding and iterative shrinkage-thresholding algorithms for large-scale inverse problems with dense matrix data (18, 19). One may employ a specific method depending on the size and structure of the network. The optimized coefficient vectors $z^{(i)}$ for $i = 1, \dots, N$ define the time-varying functions f and K_{ij} in Eq. 1, and thus identify the dynamic network topology from data.

Results

Complex systems constituted by a network of rhythmic components appear in many engineered and living systems at different scales, such as circadian neurons, chemical oscillators, power grids, and animal societies (20–24), in which the network emerging properties are functionally significant, such as sleep–wake cycles and jet lag, battery pack charge–discharge cycles, and short-term memory and communication through social behaviors (25). Revealing topology and connectivity is essential to characterize the dynamical structures and functions of such networks, which, in turn, leads to the fundamental understanding of many real-world complex systems, such as functional connections in the brain, synergy of circadian clocks, and social synchronization in groups of animals. We apply the ICON technique to determine topology of networks of oscillators with different periodicities, degree nodes, coupling functions, and time scales, arising *in silico*, and in electrochemistry, neuronal networks, and groups of mice. The excellent agreement of the estimation with the experiments over such an extremely diverse set of systems validates our strategy and suggests that our data-driven methodology provides a general framework that can be applied to numerous additional systems of interacting dynamic units.

Synthetic Networks. We first tested the accuracy of ICON to estimate the topology of a synthetic network with oscillatory system dynamics. The techniques of phase model reduction have been widely used to describe the dynamics of coupled nonlinear oscillatory systems (22, 26–28), e.g., circadian clock networks. Here, we consider topology estimation of a network of slightly heterogeneous, weakly coupled, generalized Kuramoto-type nonlinear oscillators (26), whose dynamics are described by $\dot{\phi}_i(t) = \omega_i + \sum_{j \neq i} K_{ij}(\phi_i - \phi_j)$, where $\phi_i \in [0, 2\pi)$ and $\omega_i > 0$ denote the

phase and the natural oscillation frequency of oscillator i , respectively, and the coupling function K_{ij} , $i, j = 1, \dots, N$, is 2π -periodic. Fig. 1 shows the estimation of a representative network consisting of 600 oscillatory units in the presence of white noise. This synthetic network (Fig. 1A) was created using the above Kuramoto model, with ω_i randomly sampled over $(0, \pi)$ and the coupling functions K_{ij} , $i, j = 1, \dots, 600$, randomly synthesized, and following the procedure presented in *SI Appendix, section 2*. The reconstructed network (represented using α_{ij} over the sampling time interval) is shown in Fig. 1A and B. The estimation, using the TSVD with $A^{(i)} \in \mathbb{R}^{12,000 \times 6,001}$, showed excellent agreement with the true network (Fig. 1A and B) with precise recovery in the natural oscillation frequency (Fig. 1C) and the coupling strength (Fig. 1D), where K_{ij} was approximated using a Fourier series of 10 terms. The computational time for estimating the entire network simultaneously was 381.3 min (65.3 s for estimating an agent) on a standard desktop computer (time was averaged over estimating 10 such randomly synthesized networks of 600 nodes).

Electrochemical Oscillators. We chose a network of oscillatory chemical reactions to validate the ICON method in an experimental system where a complex coupling topology can be experimentally designed with a relatively large number of nodes. In this experiment, 15 corroding nickel wires immersed in sulfuric acid represent the nodes of the network; the reactions are coupled by resistors to produce the network structure (Fig. 2A) (29). Because of inherent heterogeneities in the uncoupled network, the oscillatory reactions have slightly different frequencies on each node. When electrical coupling occurs through the resistors with sufficiently strong coupling, the network can exhibit synchronized oscillations. We applied a synchronization engineering technique with a global linear feedback (22) to desynchronize the oscillations, and recorded the relaxation of the system to the synchronized

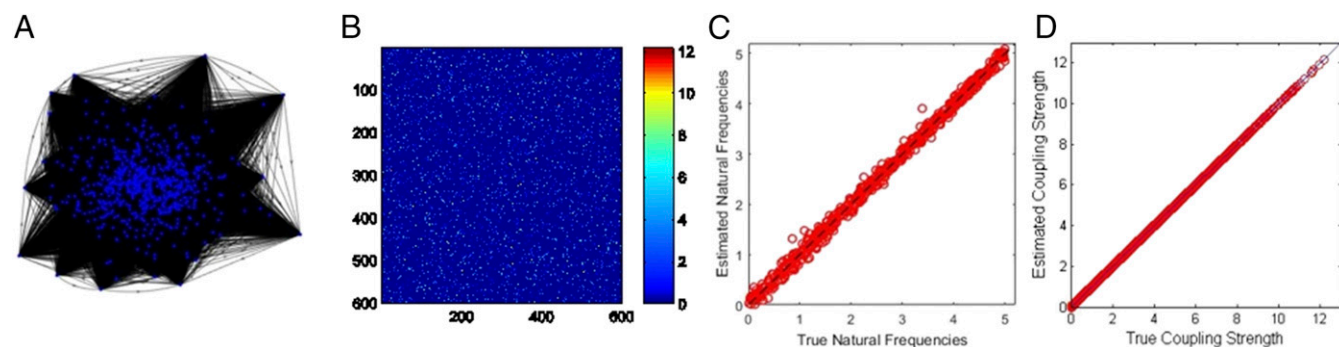


Fig. 1. Demonstration of inferring the connections, and their strengths, of a synthetic network. (A) The graph representation of a synthetic noisy random network of 600 oscillators created using Eq. 2 with additive white noise (1% intensity with respect to the maximum coupling strength) over the time horizon $t \in [0, 20]$, where the frequencies of the oscillators ω_i were randomly sampled over $(0, 5)$, the coupling strength of the functions K_{ij} were randomly selected over $[0, 12]$ (each colored dot represents the coupling strength α_{ij} of the time-varying coupling K_{ij}), and the matrix $A^{(t)}$ in Eq. 3 is of the dimension $12,000 \times 6,001$. (B) The heat map of the coupling topology. (C) The accuracy of the estimated natural oscillation frequencies. (D) The accuracy of the estimated coupling strength. In this estimation, ICON accurately recovered 95.1% of connections (26,561/27,930) with 93.1% accuracy rate on recovering the coupling strength in this representative network.

behavior (about 33 cycles of data collection, 37,600 sample points) and applied the ICON technique to reconstruct the dynamics of the system.

The estimated natural frequencies (Fig. 2C) precisely matched (with 0.05% relative SD) the values measured in independent experiments. Fig. 2A shows that the ICON technique identified all of the 14 existing network connections properly out of the 105 possible links. The technique also found two connections between nodes that were not coupled (6–7 and 10–11). Because these two predicted links were weak and between nodes that were coupled indirectly through another node, the impact of such mismatches on the overall network behavior was small. We validated the reliability of the frequency and topology estimation by looking at the synchronization pattern at a 44% weakened coupling strength. At this coupling strength, the network splits into three frequency clusters with two, three, and six elements, respectively; the rest of the nodes were phase-drifting (Fig. 2B). As shown in Fig. 2D, the phase model extracted from the ICON technique (adjusted for the weakened coupling strength) properly predicted all of the pairwise synchronization indices σ_{ij} between the elements, and thus properly predicts the synchronization pattern of the system. Such high-fidelity prediction of the spatiotemporal pattern was not previously possible, e.g., with methods based on kinetic models (29), because of the inability to extract theoretical and experimental

information on the origin of heterogeneities and the timing of the interactions between the regulated chemical reactions.

Circadian Neurons. The suprachiasmatic nucleus (SCN) of the mammalian hypothalamus has been referred to as the master circadian pacemaker that drives daily rhythms in behavior and physiology (4). When isolated from their network, SCN neurons can express sustained or damped circadian oscillations, or even arrhythmic patterns. The SCN network can be pharmacologically perturbed while monitoring the component cells for desynchronization and resynchronization of their circadian gene expression (4, 30–32). We tested the ability of ICON to estimate the connections and coupling functions that underlie circadian resynchronization by analyzing PERIOD2 (PER2) protein levels during and after SCN explants were treated with the voltage-gated Na^+ channel blocker, tetrodotoxin (TTX). Using recordings of PER2-driven bioluminescence (2,850 sample points) from 541 SCN cells (Fig. 3A), ICON recovered a time-dependent network. We found that the influence of individual cells on the network resynchronization (i.e., the sum of the coupling strengths of all of its outgoing connections) varied ~ 10 -fold, depending on the cell and its location within the SCN (Fig. 3B–D). The 14 and 16 hub cells (the cells with strong outgoing connections) identified in the left and right SCN, respectively, could indicate the presence of supercells with greater impact on circadian synchrony in the SCN. The coupling

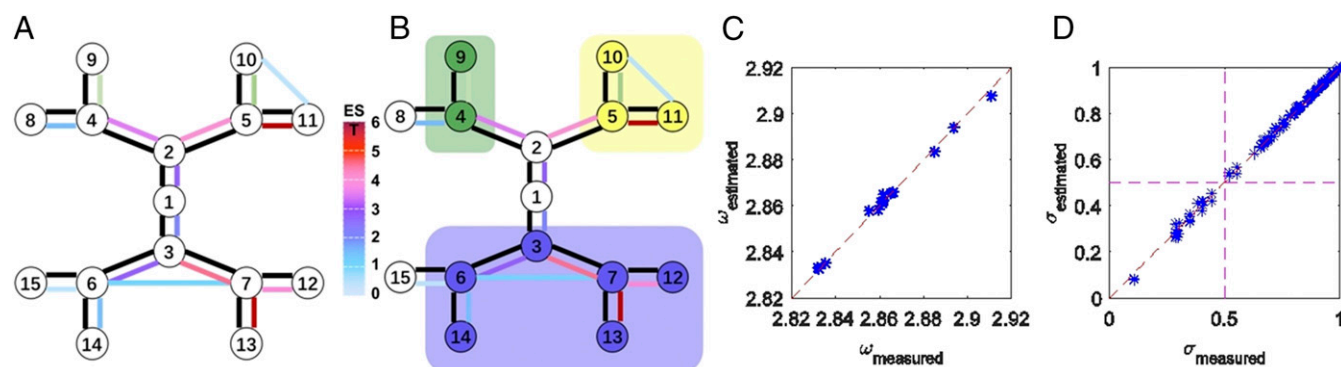


Fig. 2. ICON-derived estimation of interactions between chemical oscillators. (A) Experimentally designed (black connections) and estimated (colored connections) oscillator networks. (B) Experimentally observed (colored nodes) and predicted (colored shades) synchronization subgroups of the respective networks in A with reduced coupling strength. (C) Experimentally measured and estimated oscillation frequencies of the oscillators. (D) Pairwise synchronization parameters of experimental designed and estimated networks.

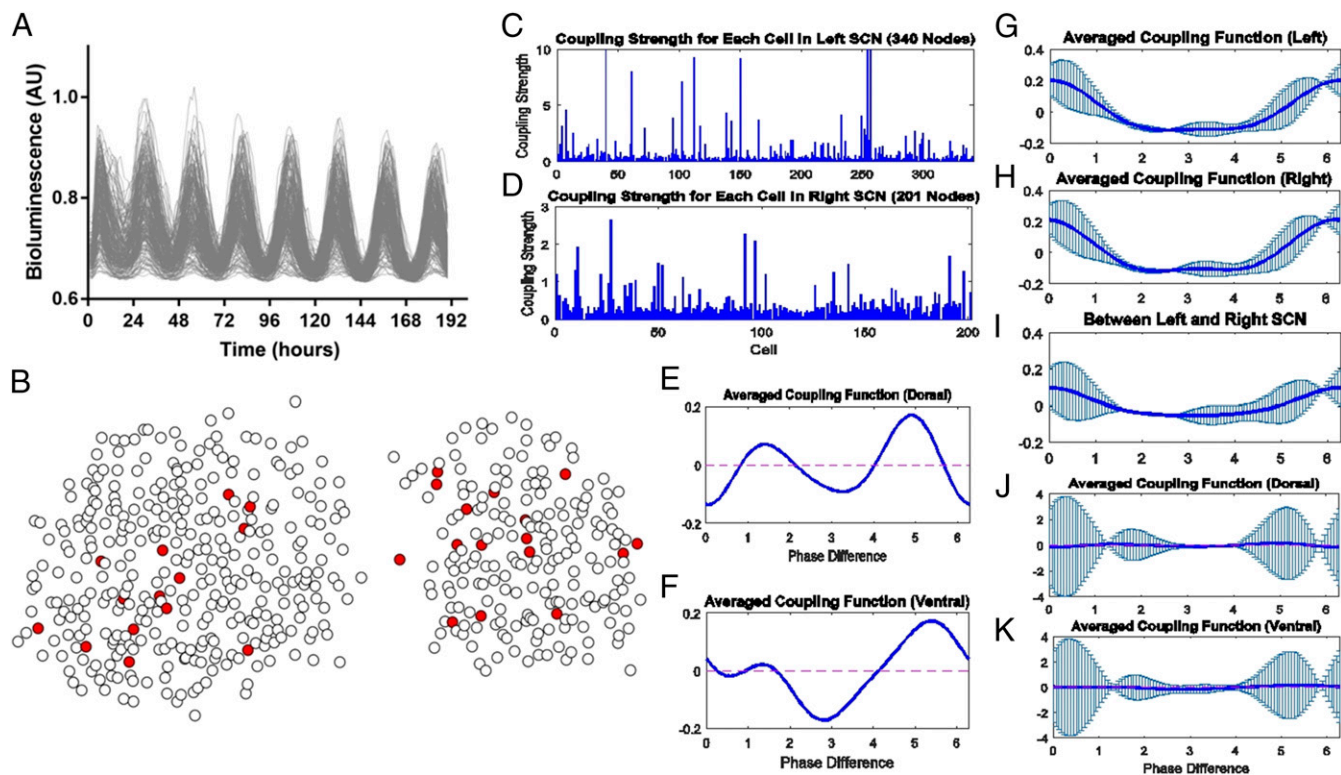


Fig. 3. ICON-derived intercellular communication in a circadian tissue. (A) Bioluminescence shows PER2 expression in over 500 SCN cells used for connectivity estimation. TTX was washed out at day 0 to allow the cells to resynchronize their daily rhythms. (B) ICON identified hub cells (red dots) as 5% of the recorded cells and localized throughout the SCN. The estimated total outgoing coupling strength for each cell revealed similar distributions of strongly coupled cells (14 and 16 hub cells, respectively) in the (C) left and (D) right SCN. The corresponding averaged coupling functions between cells were similar within (G) left and (H) right SCN, but different between (I) left and right SCN, and (E and J) dorsal and (F and K) ventral SCN (error bars show SD).

dynamics can be further analyzed by comparing the shape of the averaged coupling functions between different SCN regions. Specifically, we found similar numbers and kinds of connections within left and right SCN (Fig. 3 G and H), but weaker connections between than within the two SCN (Fig. 3 G–I). Coupling functions within the bilateral dorsal SCN (Fig. 3 E and J) were more similar to each other than to those in the bilateral ventral SCN (Fig. 3 F and K), consistent with prior predictions based on anatomy and physiology (33–36). We note that the averaged interaction functions often have large cosine harmonics, which results in nearly zero slope at zero phase difference. Such large nonisochronicity is an important oscillatory property, which was computationally predicted by a biomolecular model of the SCN (33) and supported in Fig. 3 E and F.

Social Synchronization in Groups of Mice. Social synchronization of animal and human activities is a central phenomenon that characterizes the temporal order and structures in both social and biological systems (14, 30, 37). We utilized ICON to reconstruct the topology of social networks of groups of cohoused laboratory mice, and to analyze and predict synchronization of their circadian rhythms within each group. Using previously reported recordings of body temperature of inbred female mice housed in groups of five (38) (15-min resolution over 10 d before they were grouped, 68 d of cohabitation, and 7 d after separation), we convolved the data (8,192 sample points) with the complex-valued Morlet continuous wavelet functions to measure period and amplitude over time (39). Daily phases of the circadian peak were determined for each mouse and then interpolated with a peak-finding approach.

The network estimations using ICON differed depending on the synchronization behavior of the seven recorded quintets of mice (Fig. 4 and *SI Appendix*, section 4). The recorded synchronization

behavior of each group of mice and predicted coordination of the inferred network, i.e., the sync indices calculated based on the experimental data (measured sync indices, red triangles in Fig. 4) and the estimated phase data (estimated sync indices, blue curves in Fig. 4), reliably and quantifiably identified mutual entrainment in the same cohabitating groups as demonstrated by the positive correlation between the sync index, σ , and the dominant eigenvalue, λ_2 , of the Laplacian matrix of each of the seven estimated networks. [A relationship between sync index and λ_2 could be expected based on the Wu–Chua conjecture (40), which was experimentally confirmed by electrochemical oscillators (29).] Moreover, our results suggest that the synchrony of such social networks was not induced by the distribution of the natural frequencies of the mice (Fig. 4F), but by the interactions among mice within each quintet (Fig. 4E).

Nonoscillatory Networks. The ICON technique is not restricted to the inference of the topology of networks with oscillatory coupling dynamics. Fig. 5 shows the estimation of a network with nonperiodic coupling dynamics consisting of 60 nonlinear dynamical units. This synthetic network was created using Eq. 1, with constant self-dynamics, $f(x_i) = \omega_i$, randomly sampled over $(-\pi/2, \pi/2)$, and the coupling functions $K_{ij}(x_i, x_j)$, $i, j = 1, \dots, 60$, randomly synthesized using the procedure presented in *SI Appendix*, section 2, where $x_i(t) \in \mathbb{R}$. The reconstructed network (Fig. 5 A and B), using TSVD, showed excellent agreement with the true network with precise recovery in both the natural oscillation frequency (Fig. 5C) and the coupling strength (Fig. 5D), where K_{ij} was approximated using a 2D Fourier series of 10 terms. More analysis of the ICON estimation on multidimensional ($n = 2$) nonoscillatory and larger networks (up to 30,000 agents) with respect to different network properties (e.g., sparsity and noise

the SCN. The notion of the orthonormal basis representation introduced in this paper enables the linear and parallel formulation of the nonlinear topology estimation problem, which lays the foundation for analyzing real-world complex networks of tremendous size, such as brain, transportation, internet, power grids, and social networks (14, 24). Importantly, the ICON technique provides guidelines for the refinement of experimental designs toward a comprehensive understanding of complex heterogeneous networks. We note that, in this work, we assumed that the network dynamics are additive between the self-dynamics and the coupling dynamics and within the coupling dynamics, and considered the networks in the presence of white noise. Although these are valid assumptions in many practical scenarios, future studies should test ICON on networks that reorganize (e.g., add or lose nodes or change coupling) or have multiplicative and colored noise in their self-dynamics or coupling dynamics. Here, ICON was used on networks where we have some prior knowledge of the agent's self-dynamics [i.e., the functional form of $f(x_i)$] and with data recorded from nodes while the network relaxed from desynchronized to synchronized states. These expectations also need further evaluation. Ultimately, constructing reliable network topology is the essential next step for the control of collective behavior of networked systems, such as the design of entrainment waveforms for asymptotic synchronization of a nonsynchronous

network, pinning controls for stabilization of an oscillatory network, or desynchronization stimulus for suppression of unwanted synchrony in brain networks.

Materials and Methods

See *SI Appendix* for detailed materials and methods. Detailed formulation and implementation of the ICON technique can be found in *SI Appendix, sections 1 and 2*. Especially, the detailed method for weakly coupled oscillatory networks is provided in *SI Appendix, section 1*. The robustness, reliability and efficiency of the ICON technique can be numerically validated by numerous synthetic networks, and the results are included in *SI Appendix, section 3*. The phase definition of the experimental data and the detailed analysis of the mouse and SCN data are also provided in *SI Appendix, sections 2, 4, and 5*, respectively. All SCN and mouse data came from previous publications (4, 38).

ACKNOWLEDGMENTS. This collaboration was supported a National Academy of Science Keck Future Initiative Seed Grant for research on Collective Behavior from Cells to Societies (to G.B., E.D.H., and J.-S.L.), and by the National Science Foundation (NSF) Award ECCS-1509342 (to J.-S.L.), the Air Force Office of Scientific Research Grant FA9550-17-1-0166 (to J.-S.L.), and National Institutes of Health Grants 1R21-EY027590-01 (to J.-S.L.), U01EB021956 and NS09536702 (to E.D.H.), and GM094109 (to W.J.S.). G.B. and W.J.S. thank Clark Way Harrison Visiting Professorships for supporting their stays at Washington University in St. Louis. I.Z.K. acknowledges support from NSF Grant CHE-1465013.

- Strogatz SH (2001) Exploring complex networks. *Nature* 410:268–276.
- Brockmann D, Helbing D (2013) The hidden geometry of complex, network-driven contagion phenomena. *Science* 342:1337–1342.
- Gao J, Barzel B, Barabási A-L (2016) Universal resilience patterns in complex networks. *Nature* 530:307–312.
- Abel JH, et al. (2016) Functional network inference of the suprachiasmatic nucleus. *Proc Natl Acad Sci USA* 113:4512–4517.
- Sorrentino F, Pecora LM, Hagerstrom AM, Murphy TE, Roy R (2016) Complete characterization of the stability of cluster synchronization in complex dynamical networks. *Sci Adv* 2:e1501737.
- Aton SJ, Herzog ED (2005) Come together, right...now: Synchronization of rhythms in a mammalian circadian clock. *Neuron* 48:531–534.
- Glannakis D, Majda AJ (2015) Data-driven methods for dynamical systems. *Mathematical Computational Modeling* (John Wiley, Hoboken, NJ), pp 135–191.
- Ota K, Aoyagi T (2014) Direct extraction of phase dynamics from fluctuating rhythmic data based on a Bayesian approach. arXiv:1405.4126.
- Brunton SL, Proctor JL, Kutz JN (2016) Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proc Natl Acad Sci USA* 113:3932–3937.
- Chang YH, Gray JW, Tomlin CJ (2014) Exact reconstruction of gene regulatory networks using compressive sensing. *BMC Bioinformatics* 15:400.
- De Smet R, Marchal K (2010) Advantages and limitations of current network inference methods. *Nat Rev Microbiol* 8:717–729.
- Henriques D, Villaverde AF, Rocha M, Saez-Rodriguez J, Banga JR (2017) Data-driven reverse engineering of signaling pathways using ensembles of dynamic models. *PLOS Comput Biol* 13:e1005379.
- Kralemann B, Pikovsky A, Rosenblum M (2011) Reconstructing phase dynamics of oscillator networks. *Chaos* 21:025104.
- Fuchikawa T, Eban-Rothschild A, Nagari M, Shemesh Y, Bloch G (2016) Potent social synchronization can override photic entrainment of circadian rhythms. *Nat Commun* 7:11662.
- Aster RC, Borchers B, Thurber CH (2013) Linear regression. *Parameter Estimation and Inverse Problems* (Academic, Boston), 2nd Ed, pp 25–53.
- Hansen PC (1987) The truncated SVD as a method for regularization. *BIT* 27:534–553.
- Eldar YC, Kutyniok G (2012) *Compressed Sensing: Theory and Applications* (Cambridge Univ Press, Cambridge, UK).
- Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J Imaging Sci* 2:183–202.
- Daubechies I, Defrise M, De Mol C (2004) An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun Pure Appl Math* 57:1413–1457.
- Hansel D, Mato G, Meunier C (1993) Phase dynamics for weakly coupled Hodgkin-Huxley neurons. *EPL* 23:367–372.
- Igarashi H, et al. (2011) Vasoactive intestinal peptide (VIP) and VIP receptors: elucidation of structure and function for therapeutic applications. *Int J Clin Med* 2:500–508.
- Kiss IZ, Rusin CG, Kori H, Hudson JL (2007) Engineering complex dynamical structures: Sequential patterns and desynchronization. *Science* 316:1886–1889.
- Dörfler F, Chertkov M, Bullo F (2013) Synchronization in complex oscillator networks and smart grids. *Proc Natl Acad Sci USA* 110:2005–2010.
- Bloch G (2010) The social clock of the honeybee. *J Biol Rhythms* 25:307–317.
- Buzsáki G, Mizuseki K (2014) The log-dynamic brain: How skewed distributions affect network operations. *Nat Rev Neurosci* 15:264–278.
- Kuramoto Y (1984) *Chemical Oscillations, Waves, and Turbulence* (Springer, New York).
- Nabi A, Moehlis J (2009) Charge-balanced optimal inputs for phase models of spiking neurons. *ASME 2009 Dynamic Systems and Control Conference* (Am Soc Mech Eng, New York), pp 685–687.
- Zlotnik A, Li J-S (2012) Optimal entrainment of neural oscillator ensembles. *J Neural Eng* 9:046015.
- Wickramasinghe M, Kiss IZ (2013) Spatially organized dynamical states in chemical oscillator networks: Synchronization, dynamical differentiation, and chimera patterns. *PLoS One* 8:e80586.
- Yamaguchi S, et al. (2003) Synchronization of cellular clocks in the suprachiasmatic nucleus. *Science* 302:1408–1412.
- Buhr ED, Yoo S-H, Takahashi JS (2010) Temperature as a universal resetting cue for mammalian circadian oscillators. *Science* 330:379–385.
- Abraham U, et al. (2010) Coupling governs entrainment range of circadian clocks. *Mol Syst Biol* 6:438.
- Myung J, et al. (2015) GABA-mediated repulsive coupling between circadian clock neurons in the SCN encodes seasonal time. *Proc Natl Acad Sci USA* 112:E3920–E3929.
- Mieda M, Okamoto H, Sakurai T (2016) Manipulating the cellular circadian period of arginine vasopressin neurons alters the behavioral circadian period. *Curr Biol* 26:2535–2542.
- Taylor SR, Wang TJ, Granados-Fuentes D, Herzog ED (2017) Resynchronization dynamics reveal that the ventral entrains the dorsal suprachiasmatic nucleus. *J Biol Rhythms* 32:35–47.
- Schroder EA, Esser KA (2013) Circadian rhythms, skeletal muscle molecular clocks, and exercise. *Exerc Sport Sci Rev* 41:224–229.
- Castillo-Ruiz A, Paul MJ, Schwartz WJ (2012) In search of a temporal niche: social interactions. *Progress in Brain Research*, eds Kalsbeek A, Mellow M, Roenneberg T, Russell GF (Elsevier, New York), Vol 199, pp 267–280.
- Paul MJ, Indic P, Schwartz WJ (2015) Social synchronization of circadian rhythmicity in female mice depends on the number of cohabiting animals. *Biol Lett* 11:20150204.
- Leise TL, Indic P, Paul MJ, Schwartz WJ (2013) Wavelet meets actogram. *J Biol Rhythms* 28:62–68.
- Wu CW, Chua LO (1996) On a conjecture regarding the synchronization in an array of linearly coupled dynamical systems. *IEEE Trans Circuits Syst* 43:161–165.
- Freeman GM, Jr, Krock RM, Aton SJ, Thaben P, Herzog ED (2013) GABA networks destabilize genetic oscillations in the circadian pacemaker. *Neuron* 78:799–806.
- Schmal C, Herzog ED, Herzog H (2018) Measuring relative coupling strength in circadian systems. *J Biol Rhythms* 33:84–98.