Cross-linked structure of network evolution

Danielle S. Bassett,[1,2,3,a)] Nicholas F. Wymbs,[4] Mason A. Porter,[5,6] Peter J. Mucha,[7,8] and Scott T. Grafton[4]

[1]*Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[2]*Department of Physics, University of California, Santa Barbara, California 93106, USA*
[3]*Sage Center for the Study of the Mind, University of California, Santa Barbara, California 93106, USA*
[4]*Department of Psychology and UCSB Brain Imaging Center, University of California, Santa Barbara, California 93106, USA*
[5]*Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, Oxford OX2 6GG, United Kingdom*
[6]*CABDyN Complexity Centre, University of Oxford, Oxford, OX1 1HP, United Kingdom*
[7]*Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, Chapel Hill, North Carolina 27599, USA*
[8]*Department of Applied Physical Sciences, University of North Carolina, Chapel Hill, North Carolina 27599, USA*

We study the temporal co-variation of network co-evolution via the *cross-link structure* of networks, for which we take advantage of the formalism of hypergraphs to map cross-link structures back to network nodes. We investigate two sets of temporal network data in detail. In a network of coupled nonlinear oscillators, hyperedges that consist of network edges with temporally co-varying weights uncover the driving co-evolution patterns of edge weight dynamics both within and between oscillator communities. In the human brain, networks that represent temporal changes in brain activity during learning exhibit early co-evolution that then settles down with practice. Subsequent decreases in hyperedge size are consistent with emergence of an autonomous subgraph whose dynamics no longer depends on other parts of the network. Our results on real and synthetic networks give a poignant demonstration of the ability of cross-link structure to uncover unexpected co-evolution attributes in both real and synthetic dynamical systems. This, in turn, illustrates the utility of analyzing cross-links for investigating the structure of temporal networks. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4858457]

Networks provide a useful framework for gaining insights into a wide variety of social, physical, technological, and biological phenomena.[1] As time-resolved data become more widely available, it is increasingly important to investigate not only static networks but also temporal networks.[2,3] It is thus critical to develop methods to quantify and characterize dynamic properties of nodes (which represent entities) and/or edges (which represent ties between entities) that vary in time. In the present paper, we describe methods for the identification of cross-link structures in temporal networks by isolating sets of edges with similar temporal dynamics. We use the formalism of hypergraphs to map these edge sets to network nodes, thereby describing the complexity of interaction dynamics in system components. We illustrate our methodology using temporal networks that we extracted from synthetic data generated from coupled nonlinear oscillators and empirical data generated from human brain activity.

INTRODUCTION

Many complex systems can be represented as temporal networks, which consist of components (i.e., nodes) that are connected by time-dependent edges.[2,3] The edges can appear, disappear, and change in strength over time. To obtain a deep understanding of real and model networked systems, it is critical to try to determine the underlying drivers of such edge dynamics. The formalism of temporal networks provides a means to study dynamic phenomena in biological,[4–6] financial,[7,8] political,[9–11] social,[12–18] and other systems.

Capturing salient properties of temporal edge dynamics is critical for characterizing, imitating, predicting, and manipulating system function. Let us consider a system that consists of the same N components for all time. One can parsimoniously represent such a temporal network as a collection of edge-weight time series. For undirected networks, we thus have a total of $N(N-1)/2$ time series, which are of length T. The time series can either be inherently discrete or they can be obtained from a discretization of continuous dynamics (e.g., from the output of a continuous dynamical system). In some cases, the edge weights that represent the connections are binary, but this is not true in general.

Several types of qualitative behavior can occur in time series that represent *edge dynamics*.[19,20] For example, unvarying edge weights are indicative of a static system, and independently varying edge weights indicate that a system does not exhibit meaningfully correlated temporal dynamics. A much more interesting case, however, occurs when there are meaningful transient or long-memory dynamics. As we illustrate in this article, one can obtain interesting insights in

a)Author to whom correspondence should be addressed. Electronic mail: dsb@seas.upenn.edu

such situations by examining network *cross-links*, which are defined via the temporal co-variation in edge weights. Illuminating the structure of cross-links has the potential to enable predictability.

To gain intuition about the importance of analyzing cross-links, it is useful to draw an analogy from biology. The cellular cytoskeleton[21] is composed of actin filaments that form bridges (edges) between different parts (nodes) of a cell. Importantly, the bridges are themselves linked to one another via actin-binding proteins. Because the network edges in this system are not independent of each other, the structure of cross-links has important implications for the mechanical and transport properties of the cytoskeleton. Similarly, one can think of time-dependent relationships between edge weights as cross-links that might change the temporal landscape for dynamic phenomena like information processing, social adhesion, and systemic risk. Analyzing cross-links allows one to directly investigate time-dependent correlations in a system, and it thereby has the potential to yield important insights on the (time-dependent) structural integrity of a diverse variety of systems.

In this article, we develop a formalism for uncovering the structure in time-dependent networks by extracting groups of edges that share similar temporal dynamics. We map these cross-linked groups of edges back to the nodes of the original network using hypergraphs.[22] We define a co-evolution hypergraph via a set of hyperedges that captures cross-links between network edges. (In this paper, we use the term "co-evolution" to indicate temporal co-variation of edge weights in time. The term co-evolution has also been used in other contexts in network science.[39,40]) Each hyperedge is given by the set of edges that exhibit statistically significant similarities to one another in the edge-weight time series (see Fig. 1). A single temporal network can contain multiple hyperedges, and each of these can capture a different temporal pattern of edge-weight variation.

We illustrate our approach using ensembles of time-dependent networks extracted from a nonlinear oscillator model and empirical neuroscience data.

CROSS-LINK STRUCTURE

To quantify network co-evolution, we extract sets of edges whose weights co-vary in time. For a temporal network $\mathbf{A}_t$, where each t indexes a discrete sequence of $N \times N$ adjacency matrices, we calculate the $E \times E$ adjacency matrix

Λ, where the matrix element Λ_{ab} is given by the Pearson correlation coefficient between the time series of weights for edge a and that for edge b. Note that $E = N(N-1)/2$ is the total number of possible (undirected) edges per layer in a temporal network. The layers can come from several possible sources: data can be inherently discrete, so that each layer represents connections at a single point in time; the output of a continuous system can be discretized (e.g., via constructing time windows), etc. We identify the statistically significant elements of the edge-edge correlation matrix Λ (see the supplementary material[23]), and we retain these edges (with their original weights) in a new matrix Λ'. We set all other elements of Λ' to 0.

We examine the structure of the edge-edge co-variation represented by the $E \times E$ matrix Λ' by identifying sets of edges that are connected to one another by significant temporal correlations (i.e., by identifying *cross-links*; see Fig. 1). If Λ' contains multiple connected components, then we study each component as a separate edge set. If Λ' contains a single connected component, then we extract edge sets using community detection. (See the supplementary material[23] for a description of the community-detection techniques that we applied to the edge-edge association matrix.) We represent each edge set as a hyperedge, and we thereby construct a co-evolution hypergraph $\mathbf{H}$. The nodes are the original N nodes in the temporal network, and they are connected via a total of η hyperedges that we identified from Λ'. The benefit of treating edge communities as hyperedges is that one can then map edge communities back to the original network nodes. This, in turn, makes it possible to capture properties of edge-weight dynamics by calculating network diagnostics on these nodes.

Diagnostics

To evaluate the structure of co-evolution hypergraphs, we compute several diagnostics. To quantify the extent of co-evolution, we define the *strength* of co-variation as the sum of all elements in the edge-edge correlation matrix: $\nu_{A_t} = \sum_{a,b} \Lambda'_{ab}$. To quantify the breadth of a single co-variation profile, we define the *size* of a hyperedge as the number of cross-links that comprise the hyperedge: $s(h) = \frac{1}{2} \sum_{a,b \in \lambda} [\Lambda'_{ab} > 0]_{\lambda}$, where the square brackets denote a binary indicator function (i.e., 1 if is true and 0 if it is false) and λ indicates the set of edges that are present in the hyperedge h of the matrix Λ'. To quantify the prevalence of hyperedges in a single node in the network, we define the *hypergraph degree* of a node i to be equal to the number of hyperedges η_i associated with node i.

NETWORKS OF NONLINEAR OSCILLATORS

Synchronization provides an example of network co-evolution, as the coherence (represented using edges) between many pairs of system components (nodes) can increase in magnitude over time.[24,25] Pairs of edge-weight time series exhibit temporal co-variation (i.e., they have nontrivial cross-links) because they experience such a trend. Perhaps less intuitively, nontrivial network co-evolution can also occur even without synchronization. To illustrate this
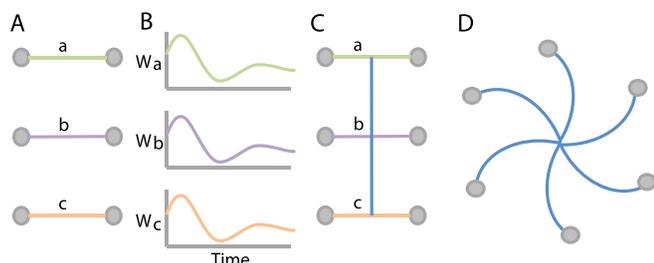

FIG. 1. Co-evolution cross-links and hyperedges. A set of (a) node-node edges with (b) similar edge-weight time series are (c) cross-linked to one another, which yields (d) a hyperedge that connects them.

phenomenon, we construct temporal networks from the time-series output generated by interacting Kuramoto oscillators,[26] which are well-known dynamical systems that have been studied for their synchronization properties (both with and without a nontrivial underlying network structure).[24,25,27–32] By coupling Kuramoto oscillators on a network with community structure,[31] we can probe the co-evolution of edge weight time series both within and between synchronizing communities.

In Fig. 2(a), we depict the block-matrix community structure in a network of 128 Kuramoto oscillators with 8 equally sized communities. The phase $\theta_i(t)$ of the ith oscillator evolves in time according to

$$\frac{d\theta_i}{dt} = \omega_i + \sum_j \kappa C_{ij} \sin(\theta_j - \theta_i), \quad i \in \{1, ..., N\}, \quad (1)$$

where ω_i is the natural frequency of oscillator i, the matrix $\mathbf{C}$ gives the binary-valued (0 or 1) coupling between each pair of oscillators, and κ (which we set to 0.2) is a positive real constant that indicates the strength of the coupling. We draw the frequencies ω_i from a Gaussian distribution with mean 0 and standard deviation 1. Each node is connected to 13 other nodes (chosen uniformly at random) in its own community and to one node outside of its community. This external node is chosen uniformly at random from the set of all nodes from other communities.

To quantify the temporal evolution of synchronization patterns, we define a set of temporal networks from the time-dependent correlations (which, following Ref. 31, we use to measure synchrony) between pairs of oscillators: $A_{ij}(t) = \langle |\cos[\theta_i(t) - \theta_j(t)]| \rangle$, where the angular brackets indicate an average over 20 simulations. We perform simulations, each of which uses a different realization of the coupling matrix $\mathbf{C}$ (see the supplementary material[23] for details of the numerics). Importantly, edge weights not only vary (see Fig. 2(b)) but they also *co-vary* with one another (see Fig. 2(c)) in time: the strength of network co-evolution, which we denote by ν_{A_t}, is greater than that expected in a null-model network in which each edge-weight time series is independently shuffled so that the time series are drawn uniformly at random.

In this example, the cross-links given by the non-zero elements of $\mathbf{\Lambda}'$ form a single connected component due to the extensive co-variation. One can distinguish cross-links according to their roles relative to the community structure in Fig. 2(a):[33] (i) pairs of within-community edges, (ii) pairs of between-community edges, and (iii) pairs composed of one within-community edge and one between-community edge. Assortative pairings [i.e., cases (i) and (ii)] are significantly more represented than disassortative pairings [i.e., case (iii)] (see Fig. 2(d)). The assortative nature of cross-links might be driven by the underlying block structure in Fig. 2(a): within-community edges are directly connected to one another via shared nodes, whereas between-community edges are more distantly connected to one another via a common input (e.g., sparse but frequently-updating representations of states of other oscillators).
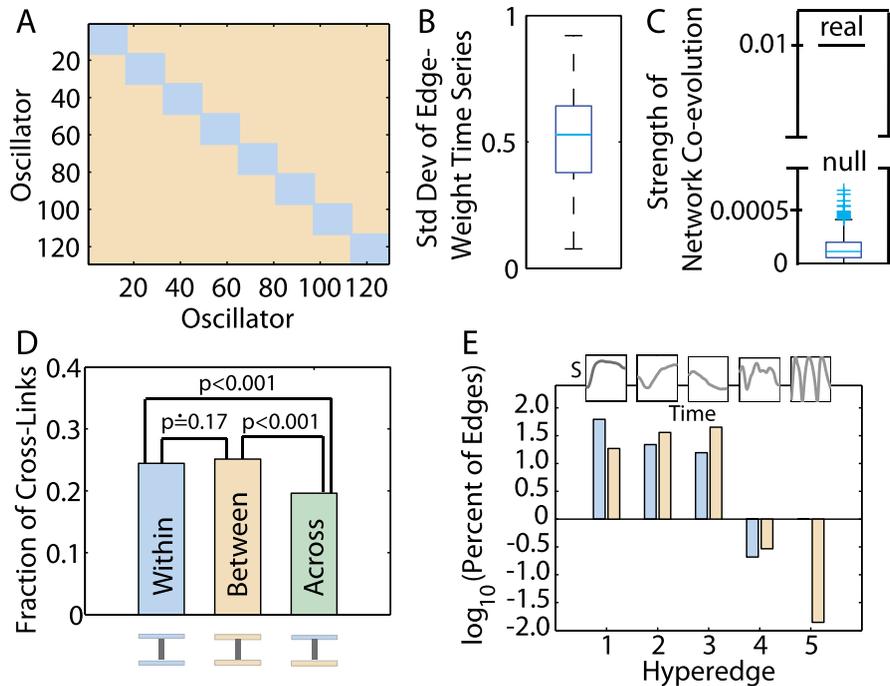

FIG. 2. Co-evolution properties of Kuramoto oscillator network dynamics. (a) Community structure in a network of Kuramoto oscillators. (b) A box plot of the standard deviation in edge weights over time for a temporal network of Kuramoto oscillators. (c) Strength of network co-evolution ν_{A_t} of the real temporal network and a box plot indicating the distribution of ν_{A_t} obtained from 1000 instantiations of a null-model network. (d) Fraction of significant edge-edge correlations (i.e., *cross-links*) that connect a pair of within-community edges ("*Within*"), that connect a pair of between-community edges ("*Between*"), and that connect a within-community edge to a between-community edge ("*Across*"). We calculated the statistical significance of differences in these fraction values across the 3 cross-link types by permuting labels uniformly at random between each type of pair. (e) Fraction of (blue) within-community and (peach) between-community edges in each of the 5 edge sets extracted from $\mathbf{\Lambda}'$ using community detection. We give values on a logarithmic scale. *Insets*: Mean synchronization [$S(t) = \sum_{(i,j) \in h} A_{ij}(t)$] of these edges as a function of time for each hyperedge h.

Using community detection, we identified 5 distinct edge sets (i.e., *hyperedges*) in Λ' with distinct temporal profiles (see Fig. 2(e)). The first hyperedge tends to connect within-community edges to each other. On average, they tend to synchronize early in our simulations. The second and third hyperedges tend to connect between-community edges to each other. The second hyperedge connects edges that tend to exhibit a late synchronization, and the third one connects edges that tend to exhibit an initial synchronization followed by a desynchronization. The fourth and fifth hyperedges are smaller in size (i.e., contain fewer edges) than the first three, and their constituent edges oscillate between regimes with high and low synchrony. The edges that constitute the fifth hyperedge oscillate at approximately one frequency, whereas those in the fourth hyperedge have multiple frequency components. See the supplementary material[23] for a characterization of the temporal profiles and final synchronization patterns of hyperedges in the network of Kuramoto oscillators.

Together, our results demonstrate the presence of multiple co-evolution profiles: early synchronization, late synchronization, desynchronization, and oscillatory behavior.[28] Moreover, the assortative pairing of cross-links indicates that temporal information in this system is segregated not only within separate synchronizing communities but also in between-community edges.

NETWORKS OF HUMAN BRAIN AREAS

Our empirical data capture the changes in regional brain activity over time as experimental subjects learn a complex motor-sequencing task that is analogous to playing complicated keyboard arpeggios. Twenty individuals practiced on a daily basis for 6 weeks, and we acquired MRI brain scans of blood oxygenated-level-dependent (BOLD) signal at four times during this period. We extracted time series of MRI signals from $N = 112$ parts of each individual's brain.[34] Co-variation in BOLD measurements between brain areas can indicate shared information processing, communication, or input; and changes in levels of coherence over time can reflect the network structure of skill learning. We summarize such functional connectivity[35] patterns using an $N \times N$ coherence matrix,[4,5] which we calculate for each experimental block. We extract temporal networks, which each consist of 30 time points, for naive (experimental blocks corresponding to 0–50 trials practiced), early (60–230), middle (150–500), and late (690–2120) learning.[34] We hypothesize that learning should be reflected in changes of hypergraph properties over the very long time scales (6 weeks) associated with this experiment.

Temporal brain networks exhibit interesting dynamics: all four temporal networks exhibit a non-zero variation in edge weights over time (see Fig. 3(a)). Importantly, edge weights not only vary but also *co-vary* in time: the strength of network co-evolution ν_{A_t} is greater in the 4 real temporal networks than expected in a random null-model network in which each edge-weight time series is independently permuted uniformly at random (see Fig. 3(b)). The magnitude of temporal co-variation between functional connections is modulated by learning: it is smallest prior to learning and largest during early learning (i.e., amidst most performance gains). These results are consistent with the hypothesis that
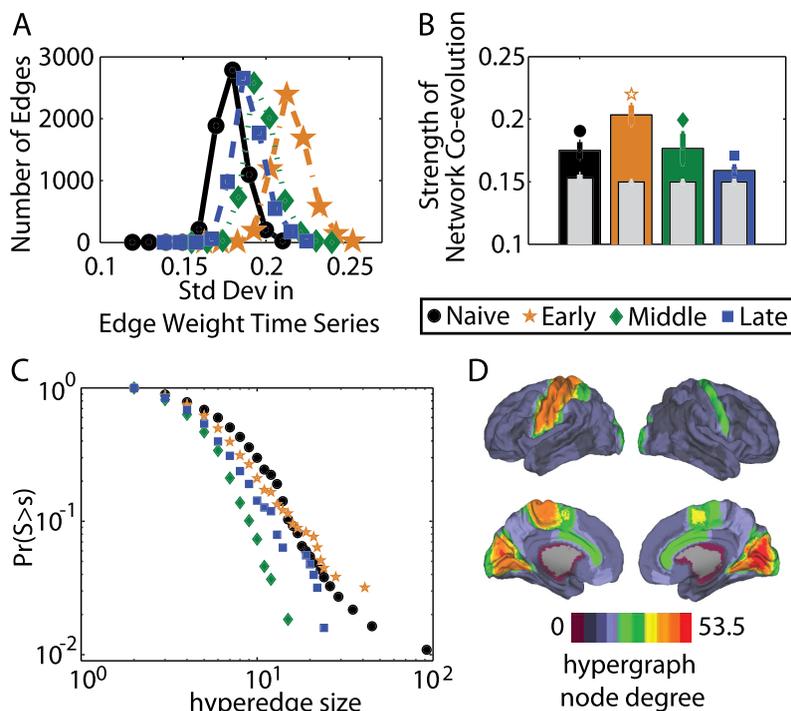

FIG. 3. Co-evolution properties of brain network dynamics. (a) A histogram of the number of edges as a function of the standard deviation in edge weights over time for the 4 temporal networks. (b) Strength of network co-evolution ν_{A_t} of 4 temporal networks and the respective null-model networks (gray). Error bars indicate standard deviation of the mean over study participants. (c) Cumulative probability distribution *Pr* of the size *s* of hyperedges in the 4 learning hypergraphs. (d) Anatomical distribution of early-learning hypergraph node degree (averaged over the 20 participants). We obtain qualitatively similar results from the early, middle, late, and extended learning temporal networks. In panels (a)-(c), color and shape indicate the temporal network corresponding to (black circles) naive, (orange stars) early, (green diamonds) middle, and (blue squares) late learning.

the adjustment of synaptic weights during learning alters the synchronization properties of neurophysiological signals,[4] which could manifest as a steep gain in the co-evolution of synchronized activity of large-scale brain areas.

To uncover groups of co-evolving edges, we study the edge-edge correlation matrix $\mathbf{\Lambda}'$, whose density across the 4 temporal networks and the 20 study participants ranged from approximately 1% to approximately 95%. We found that the significant edges were already associated with multiple connected components, so we did not further partition the edge sets into communities. The distribution of component sizes s is heavy-tailed (see Fig. 3(c)), which perhaps reflects inherent variation in the communication patterns that are necessary to perform multiple functions that are required during learning.[4] With long-term training, hyperedges decrease in size (see Fig. 3(c)), which might reflect an emerging autonomy of sensorimotor regions that can support sequential motor behavior without relying on association cortex.

Hyperedges indicate temporal co-variation of putative communication routes in the brain and can be distributed across different anatomical locations. The hypergraph node degree quantifies the number of hyperedges that are connected to each brain region. We observe that nodes with high hypergraph degree are located predominantly in brain regions known to be recruited in motor sequence learning:[36] the primary sensorimotor strip in superior cortex and the early visual areas located in occipital cortex (see Fig. 3(d)).

METHODOLOGICAL CONSIDERATIONS AND FUTURE DIRECTIONS

The approach that we have proposed in this paper raises several interesting methodological questions that are worth additional study.

First, there are several ways (e.g., using the edge-edge correlation matrix $\mathbf{\Lambda}$) to define the statistical significance of a single element in a large matrix that is constructed from correlations or other types of statistical similarities between time series (see the supplementary material[23]). Naturally, one should not expect that there is a single "best-choice" correction for false-positive (i.e., Type I) errors in these matrices that is applicable to all systems, scales, and types of association. In the future, rather than using a single threshold for statistical significance to convert $\mathbf{\Lambda}$ to $\mathbf{\Lambda}'$, it might be advantageous to use a range of thresholds—perhaps to differentially probe strong and weak elements of a correlation matrix, as has been done in the neuroimaging literature[37]— to characterize the organization of the hypergraphs on different geometrical scales (i.e., for different distributions of edge-weight values).

Second, the dependence of the hypergraph structure on the amount of time T that we consider is also a very interesting and worthwhile question. Intuitively, the hypergraph structure seems to capture transient dependencies between edges for small T but to capture persistent dependencies between edges for large T. A detailed probing of the T-dependence of the hypergraph structure could be particularly useful for studying systems that exhibit (i) temporally independent state transitions based on their

cross-linked structures and (ii) co-evolution dynamics that occur over multiple temporal scales.

Finally, the approach that we have proposed in this paper uses hypergraphs to connect dependencies between interactions to the components that interact. Alternatively, one can construe the interactions themselves as one network and the components that interact as a second network. This yields a so-called *interconnected network* (which is a type of *multilayer network*[38]), and the development of techniques to study such networks is a burgeoning area of research. Using this lens makes it clear that our approach can also be applied "in the other direction" to connect sets of components that exhibit similar dynamics (one network) to interactions between those components (another network). This yields a simple multilayer structure in which a single set of components is connected by two sets of associations (similarities in dynamics and via a second type of interaction). However, we believe that the "forward" direction that we have pursued is the more difficult of the two directions, as one needs to connect a pair of networks whose edges are defined differently and whose nodes are also defined differently. Hypergraphs provide one solution to this difficulty because they make it possible to bridge these two networks. Moreover, many dynamical systems include both types of networks: a network that codifies dependencies between nodes and a network that codifies dependencies between node-node interactions.

CONCLUSION

Networked systems are ubiquitous in technology, biology, physics, and culture. The development of conceptual frameworks and mathematical tools to uncover meaningful structure in network dynamics is critical for the determination and control of system function. We have demonstrated that the cross-link structure of network co-evolution, which can be represented parsimoniously using hypergraphs, can be used to identify unexpected temporal attributes in both real and simulated dynamical systems. This, in turn, illustrates the utility of analyzing cross-links for investigating the structure of temporal networks.

ACKNOWLEDGEMENTS

[1]M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, 2010).
[2]P. Holme and J. Saramäki, Phys. Rep. **519**, 97 (2012).
[3]*Temporal Networks*, edited by P. Holme and J. Saramäki (Springer, 2013).

[4]D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, Proc. Natl. Acad. Sci. U.S.A. **108**, 7641 (2011).

[5]D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, and P. J. Mucha, Chaos **23**, 013142 (2013).

[6]N. F. Wymbs, D. S. Bassett, P. J. Mucha, M. A. Porter, and S. T. Grafton, Neuron **74**, 936 (2012).

[7]D. J. Fenn, M. A. Porter, M. McDonald, S. Williams, N. F. Johnson, and N. S. Jones, Chaos **19**, 033119 (2009).

[8]D. J. Fenn, M. A. Porter, S. Williams, M. McDonald, N. F. Johnson, and N. S. Jones, Phys. Rev. E **84**, 026109 (2011).

[9]A. S. Waugh, L. Pei, J. H. Fowler, P. J. Mucha, and M. A. Porter, "Party polarization in Congress: A network science approach," preprint arXiv:0907.3509 (2012).

[10]P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, Science **328**, 876 (2010).

[11]K. T. Macon, P. J. Mucha, and M. A. Porter, Physica A **391**, 343 (2012).

[12]T. J. Fararo and J. Skvoretz, *Status, Network, and Structure: Theory Development in Group Processes* (Stanford University Press, 1997) pp. 362–386.

[13]A. Stomakhin, M. B. Short, and A. L. Bertozzi, Inverse Prob. **27**, 115013 (2011).

[14]J.-P. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A. L. Barabási, Proc. Natl. Acad. Sci. U.S.A. **104**, 7332 (2007).

[15]Y. Wu, C. Zhou, J. Xiao, J. Kurths, and H. J. Schellnhuber, Proc. Natl. Acad. Sci. U.S.A. **107**, 18803 (2010).

[16]S. González-Bailón, J. Borge-Holthoefer, A. Rivero, and Y. Moreno, Sci. Rep. **1**, 197 (2011).

[17]N. A. Christakis and J. H. Fowler, New Eng. J. Med. **357**, 370 (2007).

[18]T. A. B. Snijders, C. E. G. Steglich, and M. Schweinberger, in *Longitudinal Models in the Behavioral and Related Sciences*, edited by K. Van Montfort, H. Oud, and A. Satorra (Lawrence Erlbaum, 2007), pp. 41–71.

[19]J.-J. Slotine and Y.-Y. Liu, Nat. Phys. **8**, 512 (2012).

[20]T. Nepusz and T. Vicsek, Nat. Phys. **8**, 568 (2012).

[21]R. A. Peters, *Biochemical Lesions and Lethal Synthesis* (Pergamon Press, Oxford, 1963).

[22]B. Bollobás, *Modern Graph Theory* (Springer Verlag, 1998).

[23]See supplementary material at http://dx.doi.org/10.1063/1.4858457 for additional methodological details and supporting results.

[24]A. Pikovsky and M. Rosenblum, Scholarpedia **2**, 1459 (2007).

[25]A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, 2003).

[26]Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Springer-Verlag, 1984).

[27]S. H. Strogatz, Physica D **143**, 1 (2000).

[28]A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Phys. Rep. **469**, 93 (2008).

[29]S. I. Shima and Y. Kuramoto, Phys. Rev. E **69**, 036213 (2004).

[30]D. M. Abrams and S. H. Strogatz, Phys. Rev. Lett. **93**, 174102 (2004).

[31]A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).

[32]J. Stout, M. Whiteway, E. Ott, M. Girvan, and T. M. Antonsen, Chaos **21**, 025109 (2011).

[33]R. Guimerà and L. A. N. Amaral, Nature **433**, 895 (2005).

[34]D. S. Bassett, N. F. Wymbs, M. P. Rombach, M. A. Porter, P. J. Mucha, and S. T. Grafton, PLOS Comp. Biol. **9**, e1003171 (2013).

[35]K. J. Friston, Hum Brain Mapp **2**, 56 (1994).

[36]E. Dayan and L. G. Cohen, Neuron **72**, 443 (2011).

[37]D. S. Bassett, B. G. Nelson, B. A. Mueller, J. Camchong, and K. O. Lim, Neuroimage **59**, 2196 (2012).

[38]M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," preprint arXiv:1309.7233 (2013).

[39]Y. B. Xie, W. X. Wang, and B. H. Wang, Phys. Rev. E **75**, 026111 (2007).

[40]J. Y. Kim and K.-I. Goh, Phys. Rev. Lett. **111**, 058702 (2013).

Supplemental Material for

"Cross-Linked Structure of Network Evolution"

Danielle S. Bassett[1,2,3,*], Nicholas F. Wymbs[4], Mason
A. Porter[5,6], Peter J. Mucha[7,8], Scott T. Grafton[4]

[1]*Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA;*

[2]*Department of Physics, University of California, Santa Barbara, CA 93106, USA;*

[3] *Sage Center for the Study of the Mind,*

University of California, Santa Barbara, CA 93106;

[4] *Department of Psychology and UCSB Brain Imaging Center,*

University of California, Santa Barbara, CA 93106, USA;

[5] *Oxford Centre for Industrial and Applied Mathematics,*

Mathematical Institute, University of Oxford, Oxford OX1 3LB, UK;

[6] *CABDyN Complexity Centre, University of Oxford, Oxford, OX1 1HP, UK;*

[7]*Carolina Center for Interdisciplinary Applied Mathematics,*

Department of Mathematics, University of North Carolina, Chapel Hill, NC 27599, USA;

[8]*Department of Applied Physical Sciences,*

University of North Carolina, Chapel Hill, NC 27599, USA;

[*]Corresponding author. Email address: dsb@seas.upenn.edu

(Dated: November 7, 2013)

In this supplementary document, we include the following material to support the work described in the main text.

1. A detailed description of statistical corrections for edge-edge association matrices.

2. A description of the community-detection techniques that we applied to the edge-edge association matrix.

3. A characterization of the temporal profiles and final synchronization patterns of hyperedges in the network of Kuramoto oscillators.

4. A comparison to null models based on surrogate data.

5. A note on numerical implementation.

6. Figure S1: Hyperedge Identification in a Network of Kuramoto Oscillators.

7. Figure S2: Final Synchronization Patterns and Temporal Profiles of Hyperedges.

STATISTICAL CORRECTIONS FOR EDGE-EDGE ASSOCIATION MATRICES

In the main text, we describe a method for extracting cross-links from temporal networks. For a temporal network $\mathbf{A_t}$, we calculate the $E \times E$ adjacency matrix $\mathbf{\Lambda}$, where the matrix element Λ_{ab} gives the Pearson correlation coefficient between the time series of weights for edge a and the time series of weights for edge b. Note that $E = N(N-1)/2$ is the total number of possible (undirected) edges per layer in a temporal network. (Each layer can come from a single point in time, aggregation over a given time window, etc.) For simplicity, we employ a correlation coefficient as a measure of statistical association to examine linear relationships in ensembles of edge-weight time series [1]. Because we seek to determine sets of edges that might have a common driver, we do not employ sparse network methods such as the graphical lasso [2] or Bayesian network [3] methods that attempt to estimate pairwise relationships between time series in a manner that is independent of other variables.

Given the very large number of statistical tests that the above procedure entails, we threshold the edge-edge correlation matrix $\mathbf{\Lambda}$ to retain only statistically significant connections, which we determine by estimating the p-value associated with the Pearson coefficient r for each edge-edge correlation. Using a false-positive correction for multiple comparisons, we threshold $\mathbf{\Lambda}$ by identifying significant matrix elements as those whose associated p-value satisfies

$$p < \frac{1}{M} = \frac{2}{E(E-1)}, \tag{1}$$

where M is the number of tests that were performed. We retain the original weights of significant matrix elements in a new matrix $\mathbf{\Lambda}'$ and set nonsignificant matrix elements to 0.

The type of multiple comparisons correction that one uses to control for Type I errors (i.e., false positives) in correlation matrices derived from (both real and simulated) dynamical systems is itself interesting [4, 5]. The false-positive correction of $p < 1/M$ that we applied is an increasingly common choice in the study of correlation matrices in the neuroscience literature [6–12]. It has been argued that alternative choices, such as the false-discovery rate [13, 14] and Bonferroni-correction methods [15–17], are too stringent for situations like correlation matrices in which variables are highly inter-dependent [5], and they can lead to an overly large number of Type II errors (i.e., false negatives) [18].

After performing the statistical correction to obtain the weighted thresholded matrix $\mathbf{\Lambda}'$, we wish to extract cohesive sets of co-evolving edges. Two potential cases are apparent. The

3

simpler case occurs when $\mathbf{\Lambda}'$ is composed of disconnected components that each contain a set of co-evolving edges. We illustrate this scenario in the main manuscript using networks of brain regions. In a second case, $\mathbf{\Lambda}'$ contains a single large connected component — which can but need not include all of a network's edges — from which one must further extract sets of co-evolving edges. We illustrate this scenario, which arises from extensive and broadly distributed temporal covariance, in the main manuscript using networks of Kuramoto oscillators.

To study the second scenario, we need to use a method for extracting sets of strongly cross-linked edges in $\mathbf{\Lambda}'$. One possible approach is to choose a more stringent statistical threshold for creating $\mathbf{\Lambda}'$ in the first place. For example, one could tune the threshold so that it fragments $\mathbf{\Lambda}$ into several disconnected components. However, such an approach requires the choice of an arbitrarily stringent threshold on the p-value p and entails the risk of Type II errors (i.e., false negatives) [5]. In this paper, we employ an alternative approach: we extract sets of strongly cross-linked edges using community detection [19, 20]. An advantage of this approach is that we can exploit the complete information housed in $\mathbf{\Lambda}'$ by using community-detection methods that account for cross-link weights and their signs [21].

COMMUNITY DETECTION ON EDGE-EDGE ASSOCIATION MATRICES

Methods for detecting communities in networks make it possible to algorithmically extract groups of nodes that are highly and mutually interconnected [19, 20, 22]. In this paper, we seek sets of edges that are strongly and densely cross-linked to one another[23]. We identify such "communities" (or "modules") of edges by optimizing a *modularity quality function* that is suitable for signed matrices [21]:

$$Q = \sum_{ab} \left[\Lambda'_{ab} - \gamma^+ P^+_{ab} + \gamma^- P^-_{ab} \right] \delta(g_a, g_b), \qquad (2)$$

where $\mathbf{\Lambda}' = (\Lambda'_{ab})$ is the $E \times E$ thresholded and weighted correlation matrix, edge a is assigned to community g_a, edge b is assigned to community g_b, the Kronecker delta $\delta(g_a, g_b) = 1$ if $g_a = g_b$ and it equals 0 otherwise, γ^+ and γ^- are resolution parameters, and P^+_{ab} and P^-_{ab} are the respective expected weights of the positive and negative cross-links that connect edge a and edge b via a specified null model. We employ a signed null model [21] with $\gamma^+ = \gamma^-$

(also see [24], who study the case $\gamma^+ = \gamma^- = 1$), so that

$$P_{ab}^+ = \frac{k_a^+ k_b^+}{\sum_{ab} k^+} \,, \qquad P_{ab}^- = \frac{k_a^- k_b^-}{\sum_{ab} k^-} \,, \tag{3}$$

where $k_a^\pm = \sum_b \Lambda_{ab}'^\pm$ is the strength of cross-link a in the matrix $\mathbf{\Lambda}'^\pm$. The matrix $\mathbf{\Lambda}'^+$ retains all positively weighted elements of Λ_{ab}' and sets all negatively weighted elements of Λ_{ab}' to 0. The matrix $\mathbf{\Lambda}'^-$ retains all negatively weighted elements of Λ_{ab}' and sets all positively weighted elements of Λ_{ab}' to 0.

Maximization of Q yields a hard partition of the edge-edge network into communities such that the total cross-link weight inside of communities is as large as possible (relative to the null model and subject to the limitations of the employed computational heuristics, as optimizing Q is NP-hard [19, 20, 25]). Given the near-degeneracy of the landscape of the modularity function Q [26], we perform 100 optimizations of Eq. 2 and obtain consensus partitions over these optimizations via a comparison to an appropriate null model. (See Ref. [27] for a detailed description of the method.)

The structural resolution parameter $\gamma = \gamma^+ = \gamma^-$ is a tunable scalar that sets the size of the communities in the (near) optimal partition [27, 28]. Small values of γ produce a few large communities, whereas large values of γ produce many small communities. By tuning γ, one can therefore examine the community structure at different scales [29–34] of both real [35–39] and simulated [40, 41] dynamical systems.

For simplicity, we focus on a single resolution-parameter value for detailed investigation. We choose a value that provides insight into the relationship between the community structure of edges and the community structure of nodes, and later we discuss at length the procedure that we used to select this value. In Fig. 1A, we show a template block structure that summarizes the community structure of nodes in a network of Kuramoto oscillators. Each block contains edges that are located either (i) within communities (template blocks 1–8) or (ii) between communities (template blocks 9–36). We characterize the similarity between this template (which yields a network partition that we label by α) and the community structure of edges at a given value of the structural resolution parameter γ (which yields a partition that we label by β) using the z-score of the Rand coefficient [42]. We use w_{11} to denote the count of edge pairs that are classified together in both partitions (e.g., α and β). We use w_{10} to denote the count of edge pairs that are classified together in the first partition but classified separately in the second partition, and we define w_{01} analogously

5

as the count of edge pairs that are classified separately in the first partition but classified together in the second partition. We use w_{00} to denote the count of edge pairs that are classified separately in both partitions. The total number R of node pairs is then given by the sum of these quantities: $R = w_{11} + w_{10} + w_{01} + w_{00}$. We calculate the Rand z-score in terms of the network's total number of node pairs R, the number of pairs R_α classified the same way in partition α, the number of pairs R_β classified the same way in partition β, and the number of node pairs $w_{\alpha\beta}$ that are assigned to the same community both in partition α and in partition β. The z-score of the Rand coefficient comparing these two partitions is

$$z_{\alpha\beta} = \frac{1}{\sigma_{w_{\alpha\beta}}} \left(w_{\alpha\beta} - \frac{R_\alpha R_\beta}{R} \right) , \tag{4}$$

where $\sigma_{w_{\alpha\beta}}$ is the standard deviation of $w_{\alpha\beta}$ (as in [42]).

In the resolution parameter range $\gamma \in [0.2, 4]$, the z-score of the Rand coefficient between the template and partitions into communities of edges appears to have two regimes (see Fig. 1B). For $\gamma \lesssim 1.8$, the z-score exhibits are large variability over multiple optimizations of the modularity quality function in Eq. 2, which suggests that the optimization landscape of Q is replete with local maxima [26]. However, for $\gamma \gtrsim 1.8$, the z-score has a much smaller variability over the multiple optimizations, which suggests that the partitions in this regime are relatively robust [27]. In this second regime, ($\gamma \gtrsim 1.8$), the mean z-score also decreases with increasing γ, which indicates that partitions with a large number of small communities (i.e., for γ values closer to 4) exhibit less similarity to the template than partitions with a small number of large communities (i.e., γ values closer to 1.8).

We choose to examine the community structure in the edge-edge correlation matrix at the resolution parameter $\gamma = 1.8$ for two reasons: (i) at this resolution-parameter value, partitions are more robust (i.e., less variable) over multiple optimizations than they are at lower values of γ; and (ii) this approximately maximizes the similarity, as measured by the Rand z-score, between the community structure of edges and the community structure of nodes. To visualize the cross-linked edge communities that are present at $\gamma = 1.8$, we construct a consensus partition [43] over the 100 optimizations using a method that corrects for statistical noise in sets of partitions defined in comparison to a null model [27]. The consensus partition assigns each edge to one of 5 communities of varying sizes (see Fig. 1C). Each community yields a hyperedge, and we note that the pattern of hyperedges in the network has an inherently different structure than the final synchronization pattern of the

network of Kuramoto oscillators (compare Figs. 1C and D) [44]. In the next section, we characterize the differences between these two structures in greater detail.

TEMPORAL PROFILES AND FINAL SYNCHRONIZATION PATTERNS OF HYPEREDGES IN THE NETWORK OF KURAMOTO OSCILLATORS

Each hyperedge that we identified in the network of Kuramoto oscillators consists of a set of edges with a different temporal weight profile (see Fig. 2A). Edges are cross-linked based on the similarity in their temporal weight profile, and community detection makes it possible to extract cohesive groups of edges with similar profiles. The first two hyperedges, whose dynamics we show in the left two panels of Fig. 2, tend to consist of between-community edges (see Fig. 2B) and exhibit either late increases in weight (which yields late synchronizaton) or decreases in weight (desynchronization) over time (see Fig. 2A). The hyperedge whose dynamics we show in the center panel of Fig. 2 includes the majority of the within-community edges and exhibits a strong increase in weight (and hence oscillator synchronization) early in the simulation. The final two hyperedges, whose dynamics we show in the right two panels of Fig. 2, consist of edges that exhibit high-frequency oscillatory behavior in their weights.

Our investigation of cross-links and subsequent hyperedge extraction identifies similarities between edges that are based on their temporal profiles and can therefore be different from their final synchronization values. For example, hyperedges 1–3 in Fig. 2 each include edges with a wide range of final synchronization values that range from very strong ($A_{ij} \doteq 0.9$) to very weak ($A_{ij} \doteq -0.2$). Each hyperedge instead captures a property of edge dynamics: the trajectory that that edge followed to attain a given final synchronization value.

COMPARISON TO NULL MODELS BASED ON SURROGATE DATA

When examining networks that are extracted from real data, it is important to determine when observed structures different significantly from those in a relevant null-model system [27]. Specifically, for networks constructed from statistical similarities between time series (such as the brain networks that we examine in this paper), one can construct null models based on surrogate time series. By comparing cross-link structure in the real and null-model systems, one can probe potentially meaningful features of co-evolution in the real network.

7

We employ a surrogate-data generation method that has been used previously to construct covariance matrices [45] and to characterize static [46] and temporal [27] networks. The *Fourier transform (FT) surrogate* scrambles the phase of time series in Fourier space [47] and thereby preserves the mean, variance, and autocorrelation function of the original time series. We assume that the linear properties of the time series are specified by the squared amplitudes of the discrete Fourier transform

$$|S(u)|^2 = \left| \frac{1}{\sqrt{V}} \sum_{v=0}^{V-1} s_v e^{i2\pi uv/V} \right|^2 , \tag{5}$$

where s_v denotes an element in a time series of length V. (That is, V is the number of elements in the time-series vector.) We construct surrogate data by multiplying the Fourier transform by phases chosen uniformly at random and then transforming back to the time domain:

$$\bar{s}_v = \frac{1}{\sqrt{V}} \sum_{v=0}^{V-1} e^{ia_u} |S_u| e^{i2\pi kv/V} , \tag{6}$$

where $a_u \in [0, 2\pi)$ are chosen independently and uniformly at random [48]

We construct FT surrogate time series from the original time series that we extracted from each brain region of each subject during each scanning session. Using identical procedures to those that we employed to study the real time series, we cut each surrogate time series into time windows that correspond to trial blocks, compute the coherence between pairs of surrogate time series, calculate the thresholded edge-by-edge correlation matrix $\mathbf{\Lambda}'$, and extracted hyperedges defined as the connected components $\mathbf{\Lambda}'$. In contrast to the heavy-tailed hyperedge-size distributions that we observe in the real data (see Fig. 3C of the main manuscript), we find that the size distributions extracted from the surrogate data are narrow and peaked: $s \approx 2.09 \pm 0.29$ (mean $\pm$ standard deviation) for naive learning, $s \approx 2.09 \pm 0.29$ for early learning, $s \approx 2.13 \pm 0.34$ for middle learning, and $s \approx 2.20 \pm 0.40$ for late learning. The maximum hyperedge size is 3 and the minimum is 2. These results demonstrate that the learning-related human brain co-evolution structure that we report in the main manuscript cannot be attributed to the mean, variance, or autocorrelation function of the original time series.

8

A NOTE ON NUMERICAL SIMULATION

To simulate the dynamics of the network of Kuramoto oscillators, we solve the discrete-time equation

$$\theta_t = \theta_{t-1} + \tau\omega_i + \sum_j \kappa C_{ij}\sin(\theta_j - \theta_i)\,, \tag{7}$$

where ω_i is the natural frequency of oscillator i, the matrix $\mathbf{C}$ gives the binary-valued (0 or 1) coupling between each pair of oscillators, τ (which we set to 0.1) is a positive real constant that indicates the time step, and κ (which we set to 0.2) is a positive real constant that indicates the strength of the coupling. We solve equation (7) for $t \in \{1, \ldots, T\}$ for a maximum of $T = 101$ time points. We base our simulation method on the implementation in Ref [49]. Each matrix in the temporal network $\mathbf{A}_t$ gives the time-dependent correlations, measured at time point t, between pairs of oscillators.
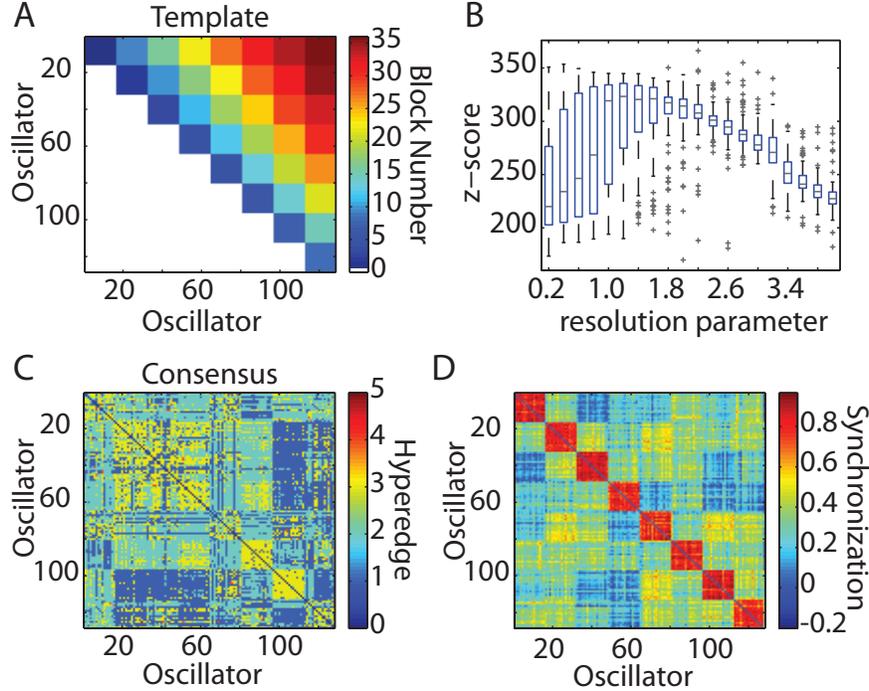
FIG. 1. **Hyperedge Identification in a Network of Kuramoto Oscillators**. *(A)* Template indicating the block structure of the community structure of nodes in a network of 128 Kuramoto oscillators. Blocks 1–8 contain within-community edges, and blocks 9–36 contain between-community edges. Color indicates the block number. *(B)* The z-score of the Rand coefficient between the upper triangle of the template in *(A)* and the partition of the thresholded and weighted edge-edge correlation matrix $\mathbf{\Lambda'}$ into communities of edges. Box plots indicate quartiles and 95% confidence intervals over the 100 optimizations of the signed modularity quality function in Eq. 2. *(C)* Consensus over partitions obtained from 100 optimizations at $\gamma = 1.8$. Each community of edges constitutes a *hyperedge*, and color indicates hyperedge number. *(D)* The final synchronization pattern of the network of Kuramoto oscillators at the final time ($T = 100$), which is reminiscent of the community structure of the network (which we show in Fig.2A in the main manuscript). Color indicates time-dependent correlation between pairs of oscillators (which we use to indicate their level of synchrony, following [40]): $A_{ij}(t) = \langle |\cos[\theta_i(t) - \theta_j(t)]| \rangle$, where the angular brackets indicate an average over 20 simulations.
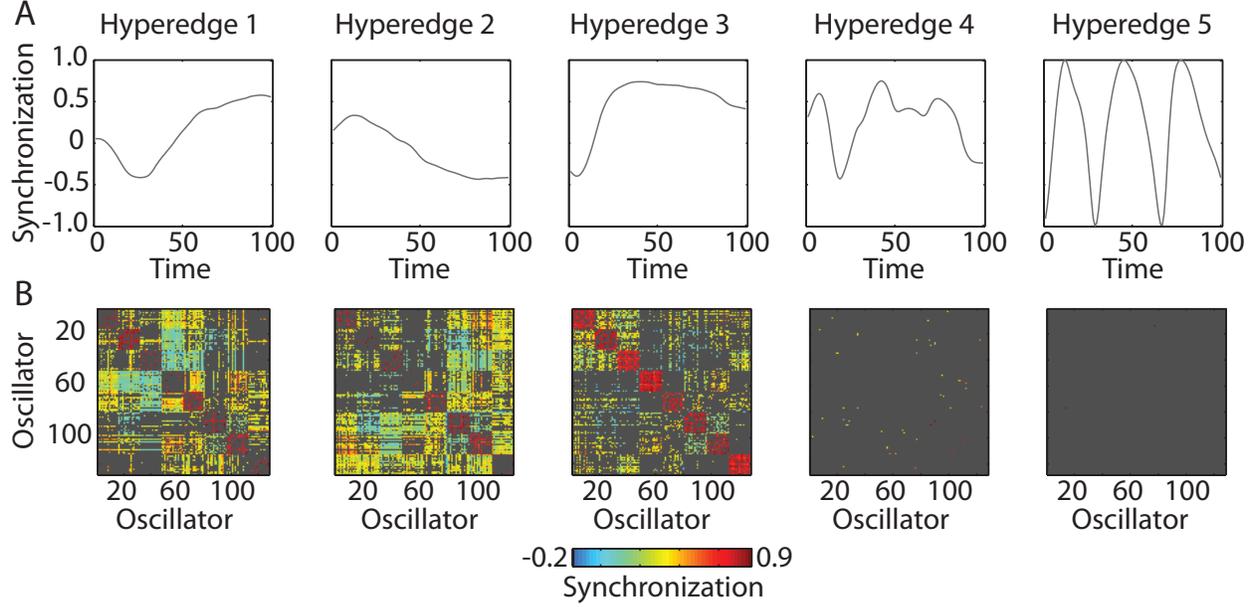
FIG. 2. **Temporal Profiles and Final Synchronization Patterns of Hyperedges**. *(A)* The mean synchronization of edges as a function of time $[S(t) = \sum_{(i,j) \in h} A_{ij}(t)]$ and *(B)* the final synchronization weights of each edge. From left to right, we plot these for hyperedge 1 (in the left panel) to hyperedge 5 (right panel). Color indicates time-dependent correlation between pairs of oscillators: $A_{ij}(t) = \langle |\cos[\theta_i(t) - \theta_j(t)]| \rangle$, where the angular brackets indicate an average over 20 simulations. Matrix elements highlighted in gray indicate edges that are members of a hyperedge other than their own.

[1] However, the methods developed in this paper also work for other choices of statistical association between edge-weight time series.

[2] J. Friedman, T. Hastie, and R. Tibshirani, Biostatistics **9**, 432 (2008).

[3] G. E. Hinton and Z. Ghahramani, Philos Trans R Soc Lond B Biol Sci **352**, 1177 (1997).

[4] L. Prignano and A. Díz-Guilera, Phys Rev E **85** (2012).

[5] A. Fornito, A. Zalesky, and M. Breakspear, NeuroImage **Epub ahead of print** (2013).

[6] A. F. Alexander-Bloch, N. Gogtay, D. Meunier, R. Birn, L. Clasen, F. Lalonde, R. Lenroot, J. Giedd, and E. T. Bullmore, Front Syst Neurosci **4** (2010).

[7] A. Messé, S. Caplain, M. Pélégrini-Issac, S. Blancho, R. Lévy, N. Aghakhani, M. Montreuil, H. Benali, and S. Lehéricy, PLoS ONE **8**, e65470 (2013).

[8] D. S. Bassett, A. Meyer-Lindenberg, D. R. Weinberger, R. Coppola, and E. Bullmore, Proc Natl Acad Sci USA **106**, 11747 (2009).

[9] M. E. Lynall, D. S. Bassett, R. Kerwin, P. McKenna, U. Muller, and E. T. Bullmore, J Neurosci **30**, 9477 (2010).

[10] S. A. Weiss, D. S. Bassett, D. Rubinstein, T. Holroyd, J. Apud, D. Dickinson, and R. Coppola, Front Hum Neurosci **5** (2011).

[11] S. H. Jin, P. Lin, and M. Hallett, PLoS ONE **6**, e28682 (2011).

[12] W. Liao, Z. Zhang, D. Mantini, Q. Xu, Z. Wang, G. Chen, Q. Jiao, Y. F. Zang, and G. Lu, Brain Connect **Epub ahead of print** (2013).

[13] Y. Benjamini and Y. Hochberg, J R Stat Soc Ser B **57**, 289 (1995).

[14] Y. Benjamini and Y. Yekutieli, Ann Stat **29**, 1165 (2001).

[15] Y. Hochberg, Biometrika **75**, 800 (1988).

[16] O. J. Dunn, J. Amer. Stat. Assoc. **56**, 52 (1961).

[17] C. W. Dunnett, Journal of the American Statistical Association **50**, 1096 (1955).

[18] P. Roxy and J. L. Devore, , 464 (2011).

[19] M. A. Porter, J.-P. Onnela, and P. J. Mucha, Not Amer Math Soc **56**, 1082 (2009).

[20] S. Fortunato, Phys Rep **486**, 75 (2010).

[21] V. A. Traag and J. Bruggeman, Phys Rev E **80**, 036115 (2009).

[22] M. E. J. Newman, Nat Phys **8**, 25 (2012).

[23] It is important to note that the identification of communities of edges based on the temporal covariance of their edge-weight time series is different from the identification of "edge communities", which consist of set of edges in static undirected networks that share common nodes [50].

[24] S. Gomez, P. Jensen, and A. Arenas, Phys Rev E , 016114 (2009).

[25] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hoefer, Z. Nikoloski, and D. Wagner, IEEE Trans on Knowl Data Eng **20**, 172 (2008).

[26] B. H. Good, Y. A. de Montjoye, and A. Clauset, Phys Rev E **81**, 046106 (2010).

[27] D. S. Bassett, M. A. Porter, N. F. Wymbs, S. T. Grafton, J. M. Carlson, and P. J. Mucha, Chaos **23**, 013142 (2013).

[28] A. Lancichinetti and S. Fortunato, Phys Rev E **84**, 066122 (2011).

[29] J.-P. Onnela, D. J. Fenn, S. Reid, M. A. Porter, P. J. Mucha, M. D. Fricker, and N. S. Jones, Phys Rev E **86** (2012).

[30] J. M. Kumpula, J. Saramäki, K. Kaski, and J. Kertész, in *SPIE Fourth International Symposium on Fluctuations and Noise* (International Society for Optics and Photonics, 2007) pp. 660116–660116.

[31] R. Lambiotte, J.-C. Delvenne, and M. Barahona, arXiv preprint arXiv:0812.1770 (2009).

[32] J. C. Delvenne, S. N. Yaliraki, and M. Barahona, Proc Natl Acad Sci U S A **107**, 12755 (2010).

[33] M. T. Schaub, J. C. Delvenne, S. N. Yaliraki, and M. Barahona, PLoS One **7**, e32210 (2012).

[34] M. T. Schaub, R. Lambiotte, and M. Barahona, Phys Rev E **86**, 026112 (2012).

[35] D. Meunier, R. Lambiotte, A. Fornito, K. D. Ersche, and E. T. Bullmore, Front Neuroinf. **3**, 37 (2009).

[36] D. Meunier, R. Lambiotte, and E. T. Bullmore, Front Neurosci **4** (2010).

[37] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. Moore, and E. Bullmore, PLoS Comput Biol **6**, e1000748 (2010).

[38] D. S. Bassett, J. A. Brown, V. Deshpande, J. M. Carlson, and S. T. Grafton, NeuroImage **54**, 1262 (2011).

[39] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, Proc Natl Acad Sci USA **108**, 7641 (2011).

[40] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Phys Rev Lett **96**, 114102 (2006).

[41] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Physica D **224**, 27 (2006).

[42] A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, SIAM Rev **53**, 526 (2011).

[43] A. Lancichinetti and S. Fortunato, Sci Rep **2** (2012).

[44] Note that we adopt the notion of synchronization from Ref. [40] and measure synchrony between a pair of oscillators in terms of the time-dependent correlation between them.

[45] H. Nakatani, I. Khalilov, P. Gong, and C. van Leeuwen, Phys Lett A **319**, 167 (2003).

[46] A. Zalesky, A. Fornito, and E. Bullmore, NeuroImage **60**, 2096 (2012).

[47] D. Prichard and J. Theiler, Phys Rev Lett **73**, 951 (1994).

[48] The code that we used for this computation actually operates on $[0, 2\pi]$, which is the same as $[0, 2\pi)$ except for a set of measure 0.

[49] D. Sumpter, "Collective animal behaviour," (2013).

[50] Y. Y. Ahn, J. P. Bagrow, and S. Lehmann, Nature **466**, 761 (2010).