Revealing dynamics, communities and criticality from data

Deniz Eroglu[1,2,3],* Matteo Tanzi[2,4], Sebastian van Strien[2], and Tiago Pereira[1,2]

[1]*Instituto de Ciências Matemáticas e Computação, Universidade de São Paulo, São Carlos, Brazil*
[2]*Department of Mathematics, Imperial College London, London, UK*
[3]*Department of Bioinformatics and Genetics, Kadir Has University, 34083 Istanbul, Turkey and*
[4]*Courant Institute of Mathematical Sciences, New York University, New York NY, USA*

Complex systems such as ecological communities and neuron networks are essential parts of our everyday lives. These systems are composed of units which interact through intricate networks. The ability to predict sudden changes in the dynamics of these networks, known as critical transitions, from data is important to avert disastrous consequences of major disruptions. Predicting such changes is a major challenge as it requires forecasting the behaviour for parameter ranges for which no data on the system is available. We address this issue for networks with weak individual interactions and chaotic local dynamics. We do this by building a model network, termed an *effective network*, consisting of the underlying local dynamics and a statistical description of their interactions. We show that behaviour of such networks can be decomposed in terms of an emergent deterministic component and a *fluctuation* term. Traditionally, such fluctuations are filtered out. However, as we show, they are key to accessing the interaction structure. We illustrate this approach on synthetic time-series of realistic neuronal interaction networks of the cat cerebral cortex and on experimental multivariate data of optoelectronic oscillators. We reconstruct the community structure by analysing the stochastic fluctuations generated by the network and predict critical transitions for coupling parameters outside the observed range.

I. INTRODUCTION

We are surrounded by a range of complex networks composed of many units forming an intricate network of interactions. Neuron networks form an important class of examples where the interaction structure is heterogeneous [1]. Because changes in the interaction can have massive ramifications on the system as a whole, it is desirable to predict such disturbances and thus enact precautionary measures to avert potential disasters. For instance, neurological disorders such as Parkinson's disease, schizophrenia, and epilepsy, are thought to be associated with an anomalous interaction structure among neurons [2]. As in the case of neuron networks, it is impossible to directly determine the interaction structure. Therefore, a major scientific challenge is to develop techniques using measurements of the time evolution of the nodes to indirectly recover the network structure and predict the network behaviour when the interactions change.

The literature on data-based network reconstruction is vast. Reconstruction methods can be classified into *model-free* methods and *model-based* methods. The former identify the presence and strength of a connection between two nodes by measuring the *dependence* between their time-series in terms of: correlations [3, 4], mutual information [5], maximum entropy distributions [6, 7], Granger causality, and causation entropy [8, 9]. Such methods alone do not provide information on the dynamics, which is necessary to predict critical transitions. Model-based methods provide estimates (or assume a priori knowledge) of the dynamics and interactions, and use this knowledge to reconstruct the network structure. When the interactions are strong, the network structure can be recovered [10–12]. For a more extensive account of reconstruction (model-free and -based) methods see the reviews [10, 13, 14].

In many applications, the behaviour of isolated nodes is chaotic and the interaction is weak [1, 15, 16]. The network structure typically has communities and hierarchical organisations such as the rich-clubs [17]. As the interaction strength per connection is weak and the statistical behaviour of the nodes is persistent, the influence of each node on the network corresponds essentially to a random signal. Existing techniques fail to reconstruct a model from the data, as they require the interaction to be of the same magnitude as the isolated dynamics. In our setting, only the cumulative contribution of many links matter and the network signals decompose into a deterministic and a fluctuation term. The latter, which is usually filtered out, turns out to give crucial information on the network structure and is fundamental to our approach.

In this paper, we introduce the notion of an *effective network* which aims to model a complex system from observations of the nodes evolution when the network has a heterogeneous structure, the strength of interaction is small and local dynamics are highly erratic. This approach starts by reconstructing the local dynamics from observations of nodes with relatively few connections, and then recover the interaction function from observations of the highly connected nodes whose dynamics are the most affected by the interactions as a result of the multitude of connections they receive from the rest of the network [18, 19]. A key achievement is that this reconstruction enables us to identify community structures also when the coupling is only weak. Moreover, it recovers enough information to forecast and anticipate the network behaviour, even in situations where the parameters of the system change into ranges that have not been previously encountered.

A. Complex networks of nonlinear systems

We consider networks with N nodes with chaotic isolated dynamics and pairwise interactions. The network is described by its adjacency matrix $\boldsymbol{A}$, whose entry A_{ij} equals 1 if node i

receives a connection from j and equals 0 otherwise. The time evolution of the state $\boldsymbol{x}_i(t)$ of node i at time t is expressed as

$$\boldsymbol{x}_i(t+1) = \boldsymbol{F}_i(\boldsymbol{x}_i(t)) + \alpha \sum_{j=1}^{N} A_{ij} \boldsymbol{H}(\boldsymbol{x}_i(t), \boldsymbol{x}_j(t)). \quad (1)$$

When performing reconstruction, the isolated local dynamics $\boldsymbol{F}_i \colon M \to M$, the coupling function $\boldsymbol{H}$, the coupling parameter α (that is small), the adjacency matrix $\boldsymbol{A}$, and even the dimension k is the degree of the space M, are assumed to be *unknown*. These equations model important complex systems such as neuron networks [20], smart grids[21, 22], superconductors [23], and cardiac pacemaker cells [24].

B. Main assumptions

Our three assumptions are: (a) the local dynamics are close to some unknown *ergodic* and *chaotic* map $\boldsymbol{F}$ (that is, $\|\boldsymbol{F} - \boldsymbol{F}_i\| \leq \delta$, which is often the case in applications [25, 26]). (b) The network connectivity is *heterogeneous*, which means that the number of incoming connections at a node i (given by its degree $k_i = \sum_j A_{ij}$) varies widely across the network. k_i is large for a few nodes called *hubs*. (c) α is such that, denoting by $\Delta = \max_i k_i$ the maximum number of connections $\alpha\Delta$ is of the same magnitude of $\boldsymbol{DF}$. Assumptions (a) and (c) imply that only the cumulative effect of the coupling is important. A prime example is the cat cerebral cortex which possesses inter-connected regions split into communities with a hierarchical organization as well as modular and disassortative rich-clubs. This network has heterogeneous connectivity, chaotic motion and weak coupling [27–29]. Other examples include the drosophila optic lobe network [30, 31]. For a given dataset, our effective network first tests whether the underlying system satisfies assumptions $(a) - (c)$, and, if so, reconstructs the model.

We assume the availability of a time series of observations

$$y_i(t) = \phi(\boldsymbol{x}_i(t))$$

where ϕ is a projection to a variable on which unit interactions depend. This situation occurs frequently in applications; as with measurements of membrane potentials in neurons.

II. EFFECTIVE NETWORKS RECOVER STRUCTURE AND DYNAMICS

To obtain an *effective (reconstruction of the) network* from observations, we combine statistical analysis, machine-learning techniques, and dynamical systems theory for networks. An effective network provides: local evolution laws and averaged interactions for each unit that, in combination, closely approximate the unit dynamics; a network with the same degree distribution and community structures as the original system. We use the term "effective" because it gathers sufficient data to reproduce the behaviour of the original network and predict its critical transitions.

Using our assumptions for the network and local dynamics, we can show that the evolution at each node will have low-dimensional excursions over finite time scales. More precisely, the evolution rule at node i is given by

$$\boldsymbol{G}_i(\boldsymbol{x}_i) = \boldsymbol{F}_i(\boldsymbol{x}_i) + \beta_i \boldsymbol{V}(\boldsymbol{x}_i(t))$$

where $\boldsymbol{F}_i \approx \boldsymbol{F}$ is the isolated dynamics,

$$\beta_i = \alpha k_i$$

is the rescaled degree, and

$$\boldsymbol{V}(\boldsymbol{x}) = \int \boldsymbol{H}(\boldsymbol{x}, \boldsymbol{y}) d\mu(\boldsymbol{y})$$

where μ is physical measure of the isolated dynamics. $\boldsymbol{V}$ takes into account the cumulative effect of interactions on node i. The true dynamics

$$\boldsymbol{x}_i(t+1) = \boldsymbol{G}_i(\boldsymbol{x}_i(t)) + \boldsymbol{\xi}_i(t)$$

is influenced by a fluctuation term $\boldsymbol{\xi}_i(t)$ that is small for an interval of time which is exponentially large and depends on the state of neighbours of the ith node. This low-dimensional reduction has been rigorously established in test cases (see [19]). See Appendix for further information.

The approximation described above applies to the measured state variable $y_i(t)$. First of all we pre-process the data according to the system under study (see Appendix E). The processed variable is still referred to as $y_i(t)$. Takens reconstruction tells us that $y_i(t+1)$ is a nonlinear function of $k+1$ past points $y_i(t), \ldots y_i(t-k)$, for a given number k provided by the approach. Here, we focus on the case when $k = 1$, which occurs in many real-world examples, and discuss cases with $k \geq 2$ in Appendix E. This means that

$$y_i(t+1) = g_i(y(t)) + \xi_i(t) \quad (2)$$

where $g_i = f_i(y_i(t)) + \beta_i v(y_i(t))$, and v is the corresponding projection of effective coupling $\boldsymbol{V}$.

A. Reconstruction procedure

An effective network is obtained in three main steps:
Step 1: Reduced dynamics. We employ Takens reconstruction. If the time series is high dimensional, we discard it. Otherwise, once we are in the appropriate dimension, we estimate and learn the rule g_i. We decompose g_i as a linear combination of basis functions, tailored to the application. The parameters of the basis functions are obtained by performing a 10-fold cross-validation with 90% training and 10% test [32, 33] (see Appendix A). As the dynamics is low-dimensional other techniques such as compressive sensing [34, 35] or embedding [36] can be also employed.
Step 2: Isolated dynamics and effective coupling. We run a model-free estimation that coarsely classify nodes according to their degree by assigning to every pair of y_i and y_j a *Pearson distance* $s_{ij} \geq 0$ such that $s_{ij} \approx 0$ if the attractors of i and j are similar and $s_{ij} \approx 1$ if they are distinguishable. The

higher the number of nodes with behaviour different from i, the larger the intensity $S_i = \sum_j s_{ij}$. Low degree nodes have typically small S_i while for hubs this quantity is large. Notice that for the low-degree nodes, $\alpha k_i v$ is negligible and the dynamics at the low-degree nodes are close to f. Therefore we use g_i at the identified low degree nodes to obtain an approximation for $f \approx g_i$, while g_i at hub nodes allows to estimate $\beta_i v \approx g_i - f$. We estimate β_i by Bayesian inference.

Step 3: Network structure and communities. Since $\beta_i = \alpha k_i$, we can recover the network's degree distribution from β_i. Then, having the local rules g_i, we can decompose the time series in terms of a low-dimensional deterministic part and the fluctuation term ξ_i, and use this last term to recover community structures. If nodes i and j interact with the same nodes, they are subject to the same inputs and the correlation $\mathrm{Cor}(\xi_j, \xi_i)$ is high. If not, $\mathrm{Cor}(\xi_i, \xi_j)$ is nearly zero due to the decay of correlations in the deterministic part. Thus, $\mathrm{Cor}(\xi_j, \xi_i)$ is high when nodes i and j have high *matching index* (high fraction of common connections), and are likely to belong to the same cluster. Given the matrix $\rho_{ij} = \mathrm{Cor}(\xi_i, \xi_j)$, we estimate the adjacency matrix $\boldsymbol{A}$ by thresholding the correlation matrix as $A_{ij} = \Theta(\rho_{ij} > \tau)$, where Θ is a Heaviside step function and the value of the threshold τ between 0.3 and 0.6. We then apply the modularity-based Louvain method [37] on $\boldsymbol{A}$ to detect communities.

That $\mathrm{Cor}(\xi_j, \xi_i)$ is high when nodes i and j have high matching index, is true for generic coupling as shown by the following argument. In general, the coupling function is a sum of terms $h(x, y) = u(x)v(y)$. This leads to noise terms

$$\xi_i(t) = u(x_i) \left(\frac{1}{\Delta} \sum_j A_{ij} v(y_j) - k_i \int v(y) d\mu(y) \right)$$

where μ is the physical measure of the local dynamics. Given i and j the sum can be split into common connections to i and j and to the independent connections:

$$\xi_i = u(x_i)[\zeta_i(t) + w(t)] \text{ and } \xi_j = u(x_j)[\zeta_j(t) + w(t)]$$

where w is the noise due to the common connections (notice that w has zero mean), and ζ_i, ζ_j depend on different coordinates and can be assumed to be uncorrelated. Omitting the time index t, the covariance of ξ_i and ξ_j is

$$\mathrm{Cov}(\xi_i, \xi_j) \approx \mathbb{E}[(u(x_i)w)(u(x_j)w)].$$

After some manipulations, we obtain

$$\mathrm{Cov}(\xi_i, \xi_j) \approx \langle u \rangle^2 \mathrm{Var}(w) \qquad (3)$$

so, if $\int u(x) d\mu(x) = 0$, the correlation between the noise will vanish even though they have a common term. Thus, the above scheme is able to recover communities if $\langle v \rangle \neq 0$. If this condition is not met, the network reconstruction via the g_i's is not possible. We remark that $\langle v \rangle = 0$ is a special condition on the coupling that is destroyed by small perturbations.

It is crucial that the correlation analysis is restricted to fluctuations ξ_i. Since the variance of the deterministic part of y_i is larger than that of the small fluctuations ξ_i, performing a direct correlation analysis between y_i and y_j hides all the

contributions coming from the covariance between ξ_i and ξ_j. Consequently, the correlation of the deterministic part is close to zero due to the chaotic dynamics, as shown in Appendix A.

B. Benchmark model for the isolated dynamics

We present the effective network methodology applied to networks of neurons. We use synthetic time-series where each neuron is simulated using the Rulkov model, which has two variables, u and w, evolving at different time scales as described by $\boldsymbol{F}(\boldsymbol{x}) = (F_1(u, w), F_2(u, w))$ with

$$F_1(u, w) = \frac{\beta}{1 + u^2} + w \quad \text{and} \quad F_2(u, w) = w - \nu u - \sigma.$$

The fast variable u describes the membrane potential and is the state variable measured by the observed time series $y_i(t)$, while w describes the slow currents. Different combinations of parameters σ and β give rise to different dynamical states of the neuron, such as resting, tonic spiking, and chaotic bursts. To test our procedure we considered two cases: $\sigma = \nu = 0.001$ and $\beta = 5.9$, which correspond to tonic spiking, and $\beta = 4.4$ which correspond to bursting. As for the coupling, we consider *chemical synaptic coupling*, that is, $\boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (h(u_i, u_j), 0)$ with $h(u_i, u_j) = (u_i - V_s)\Gamma(u_j)$, where

$$\Gamma(u_j) = 1/(1 + \exp\{\lambda(u_j - \Theta_s)\}),$$

and *electrical synaptic coupling*, $\boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (h(u_i, u_j), 0)$ with $h(u_i, u_j) = u_j - u_i$. In the chemical coupling, V_s is a parameter called reverse potential. Choosing $V_s > u_i(t)$, the synaptic connection is excitatory. We take $V_s = 20$, $\Theta_s = -0.25$, and $\lambda = 10$. In addition to Rulkov maps, we show in the Appendix E that the approach performs well on a wide range of nonlinear local dynamics such as: doubling maps, logistic maps, Spiking Neurons, Henon maps. We also provide performance analysis for Rössler oscillators in Section III of the Supplementary Material.

III. REVEALING COMMUNITY STRUCTURE: THE RICH-CLUB MOTIF

We focus on the network structure of the cat cerebral cortex [29]. The network contains 53 meso-regions arranged in four communities that follow functional subdivisions; visual (16 nodes), auditory (7 nodes), somatomotor (16 nodes) and frontolimbic (14 nodes), as shown in Fig 1 (a). Some cortical areas (hubs) form a hidden layer called a *rich-club* and are densely connected to each other and the communities. A set of nodes form a rich-club if their level of connectivity exceeds what would be expected by chance alone. The maximum number of connections in this network is $\Delta = 37$.

The regions and their connections were discovered by using datasets from tract-tracing experiments [27, 28]. The network obtained is weighted. For simplicity and to improve the performance in detecting communities, we turn the network into
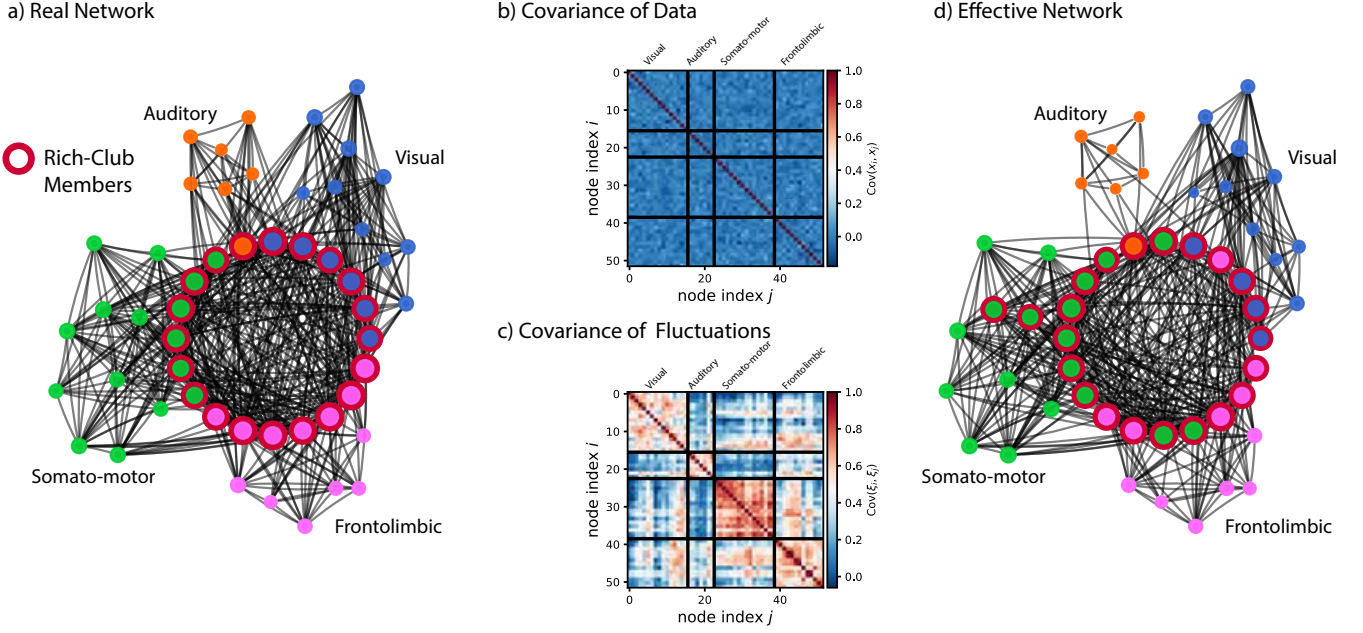
FIG. 1. **Effective network of the cat cerebral cortex.** We use the local dynamics as a spiking neuron coupled via electric synapses. (a) The cat cerebral cortex network with nodes colour coded according to the four functional modules. Rich-club members are indicated by red encircled nodes. (b) *The covariance matrix of the data* cannot detect communities. (c) *The covariance matrix of the fluctuations* can distinguish clusters. This matrix has entries color coded according to the key on the right with red entries corresponding to couple of nodes sharing a large numbers of nearest neighbours in the network, while blue nodes correspond to couple of nodes that share a small number of common neighbours. (d) A model in the cat cortex constructed via the effective network approach. From the matrix in (c) we can recover a representative effective network. The reconstructed network represents the actual network in (a) with good accuracy.

an undirected simple graph [29]. We simulate each mesoregion as a neuron interacting via electrical synapses and obtain a multivariate data $\{y_1(t), y_2(t), \ldots, y_N(t)\}$ for a time $T = 5000$. For simplicity, we will denote $y_i = \{y_i(t)\}_{t=0}^T$.

A. Comparison with previous approaches

For comparison, we recover the network using two widely employed approaches: functional networks [38–40], and sparse recovery techniques [10, 34]. The intuition behind the *functional network approach* is that nodes with similar time series have similar characteristics. The functional network can be constructed by the matrix of similarities between nodes via statistical analysis [41, 42]. As a measure of similarity, we employ a covariance analysis between the time series. The functional network cannot detect communities in this case since the time-series at different nodes are essentially uncorrelated (Fig. 1 (b)). Other similarity measures give no significant improvement. See Appendix B for the details.

The key idea in *sparse recovery techniques* is to write the dynamics as a linear combination of basis functions with unknown coefficients, and the presence of a link is determined when any coefficient of the corresponding interaction is nonzero. Thus a link is present if the estimated coefficient corresponding to the link is above a given threshold σ.

We implemented the sparse recovery method to our bench-

mark model when the strength of each connection is of order $\alpha \approx 0.015$. Hence we have chosen values of σ close to this value. The reconstructed network does not identity the clusters correctly as can be seen by comparing the blue and red markers in Figure 2. In the cases that we are studying
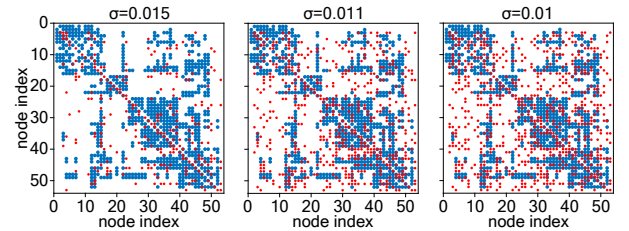

FIG. 2. **Sparse recovery method on on cat cerebral cortex.** Sparse recovery is applied to the data generated by bursting neurons electrically coupled on the cat cerebral cortex. Selecting the threshold parameter σ in the method changes the reconstructed network. Here we show the results of sparse recovery method for different enforced sparsity σ. The nonzero entries of the original network's adjacency matrix are in blue. The red filled circles represent the nonzero entries in the adjacency matrix of the network reconstructed with the sparse recovery method. As each connection is small in comparison with the isolated dynamics, the sparse recovery tends to neglect them.

here each individual link provides a negligible contribution and only the cumulative effect of many links is relevant. The coefficients to be recovered are close to zero, and cannot be

distinguished from zero terms. A discussion on sparse recovery can be found in Section I of Supplementary Materials.

B. Community structure via effective networks

Remarkably, the effective network is able to recover the community structures (Fig. 1 (c)). Using Step 1, 2, and 3 we obtain a model for the isolated dynamics, coupling function, distribution of degrees, and correlations $Cor(\xi_i, \xi_j)$. To apply the method of community detection in [37], we threshold the matrix of correlations, Fig. 1 (c), considering nodes i and j linked only when the correlations were greater than 0.5. We test threshold values ranging from 0.3 to 0.6 and obtained the similar results as the distribution of the entries of the matrix of correlations is unimodal and has a peak near 0.5. We use the algorithm in [43] to compute the rich-club coefficients for each node. The coefficient depends on the degree and is a number between 0 and 1. We assigned to the rich-club the nodes with coefficient at least 0.8. As shown by Figure 1 (d), the effective network methodology is able to classify the nodes in the network according to their function.

Notice that our model predicts the presence of a link between two nodes i and j when $Cor(\xi_i, \xi_j)$ is high. Since every node makes most of its interactions within a cluster, two nodes with highly correlated fluctuations $\xi(t)$ are likely to belong to the same community, and this can be enforced in the effective network by adding a connection between them.

C. Performance of the communities reconstruction

To quantify the effectiveness of community reconstruction, we compute $PE = \frac{m}{N}$, where N is the total number of nodes and m is the number of nodes assigned to the wrong community. We compute PE for $\Delta\alpha$ between 0.05 and 0.4. For each value of α, we considered 50 different simulations by choosing different initial conditions. The figure shows the plot of the mean of PE and a shaded region corresponding to the standard deviation. For $\Delta\alpha$ values larger than 0.4, the reconstruction procedure cannot identify the communities correctly as synchronization rich club which appears.

In the Appendix E, we analyzed synthetic networks with 100 nodes which are undirected and have a rich-club structure. We used them as benchmark to evaluate the success of the reconstruction. The ability of the reconstruction procedure to recover the community structure was tested for various coupling functions and isolated dynamics.

IV. PREDICTING CRITICAL TRANSITIONS IN RICH-CLUBS

The ability to reconstruct the network and dynamics from data can be exploited to predict critical transitions that may occur when the coupling strength varies. This is crucial for applications. For example in the cat brain, a transition to col-
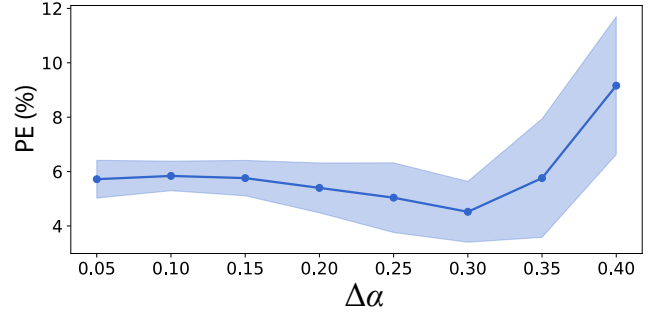

FIG. 3. **Prediction error for misidentification of communities in the reconstructed cat cerebral cortex from synthetic data.** For each realisation, the chosen parameters are the same as in Figure 1 and only the overall coupling is changed. Mean and standard deviation of prediction error (PE) computed for the network over 50 realizations for each value of α. If $\Delta\alpha > 0.42$, the system synchronizes and the procedure cannot reconstruct the community structures.

lective dynamics in the rich-club has drastic repercussions for the functionality of the network [29, 44].

The goal is to obtain and predict the onset of collective motion in the rich-club from data recorded when the network is far from a collective dynamics. The effective network can predict the onset of such collective dynamics based on a single multivariate time series for fixed coupling strength in a regime far from the synchronized state. We analyze time-series obtained simulating the dynamics for $\Delta\alpha = 0.3$, and reconstruct the network structure and the isolated dynamics.

Transitions to synchronization between the scale variable is possible while the fast spikes remain out of synchrony [45]. Notice that the slow variable w changes on a scale $1/\mu$. In the present setting we have $1/\mu = 10^3$ which is about number of points we need to apply the approach. Thus, for such short time series we can neglect the slow scale. This is also an advantage of this present approach. To estimate the transition to burst synchronization, we obtain the slow variable as a filter over the membrane potential (fast variable). Since we measure the membrane potential $y_i(t) = u_i(t)$, the slow variable is given as $z_i(t) = \mu \sum_{k=1}^{t} (y_i(k) - \sigma)$ and for a choice μ and σ this can be identified with the slow variable of the model w. In Appendix C, we derive the following equation for the slow variable of a node in the rich club:

$$z(t+1) = (\lambda - \Delta\alpha)z(t) + \mu \sum_{n=0}^{t} z(n)$$

where $\lambda = 1.42$ is estimated from the data. The equation can be used to analyze the effect of the network connectivity on the dynamics. We can use the data on the network and the dynamics recovered from the time-series recorded at $\Delta\alpha = 0.3$ to predict that at the value $\Delta\alpha \approx 0.42$ the rich-club will develop a burst synchronization (details in Appendix C).

To capture a transition to a synchronized state, we introduce a phase $\theta_j(t)$ for the slow variable. To define $\theta_j(t)$, we first smooth the time series [46]. Then, we find the time t_n of local maxima as the nth maximum point of the slow variable. We
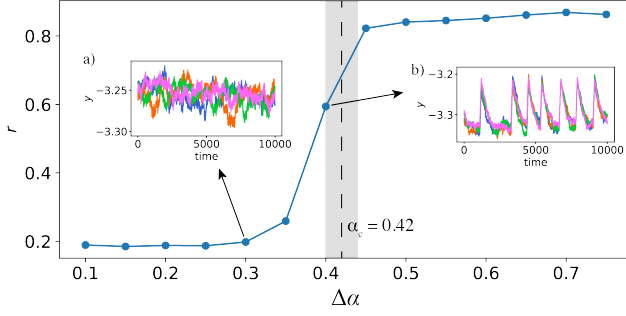
FIG. 4. **Prediction of critical transitions in the rich-club of the cat cerebral cortex.** The level of synchronization r of the rich-club is shown for different values of the coupling strength. Insets show time series of neuronal dynamics of four rich-club members and color of time series matches with the color of nodes in Fig. 1. For values in the grey shaded region, r is increasing towards close to one and the rich-club exhibits collective behavior. We can predict the critical coupling α_c (standard deviation in shaded region) by studying the effective network obtained from a time series measured at $\Delta\alpha = 0.3$.

introduce the phase variable θ as

$$\theta_j(t) = 2\pi \left(\frac{t - t_n}{t_{n+1} - t_n} + t_n \right), \quad t_n < t < t_{n+1}$$

as shown in Ref. [47]. We then compute the order parameter

$$r(t)e^{i\psi(t)} = \frac{1}{N_c} \sum_{j=1}^{N_c} e^{i\theta_j(t)}.$$

A small value of the order parameter, $r \approx 0$, means that no collective state is present, whereas $r(t) \approx 1$ means that the bursts are synchronized. Figure 4 shows that behaviour of r as a function of the coupling. The rich-club undergoes a transition to burst synchronization at $\Delta\alpha \approx 0.4$ that corresponds to an increase of roughly 40% of the coupling strength and is close to the predicted value $\Delta\alpha \approx 0.42$. In Appendix E, we show other examples where the local dynamics is chaotic.

V. OBTAINING A STATISTICAL DESCRIPTION OF THE NETWORK

The effective network can provide statistical description of the network structure. To illustrate this, we reconstruct the statistical properties of scale-free networks.

A. Scale-free networks of coupled bursting neurons

We consider coupled bursting neurons with excitatory synapses [45] in scale-free networks. A scale-free network has degree distribution $P(k) = Ck^{-\gamma}$, where $\gamma > 0$ is the characteristic exponent and C is a normalising constant. We generate a scale-free network with $N = 10^4$ nodes such that the probability of having a node of degree k is proportional to $k^{-\gamma}$, where $\gamma = 2.53$. We use a random network model which

is an extension of the Erdös-Rényi model for random graphs with a general degree distribution. More details are provided in Appendix D.

For this reconstruction we only need 2000 data points for each node. Again, to every pair of time series y_i and y_j we assign a *Pearson distance* $s_{ij} \geq 0$ and the node intensity $S_i = \sum_j s_{ij}$. The empirical distribution of the intensities S_i approximates the degree distribution of the network, see the second inset of Fig 5(a). In the example here, the estimated structural exponent from the distribution of S_i is $\gamma_{\mathrm{est}} = 3.1$, which yields a relative error of nearly 25% with respect to the true value of γ (see the plots in Figure 5 a)). The functional network therefore overestimates γ, which has drastic consequences for the predicted character of the network. For example, the number of connections of a hub for a scale-free network is concentrated at $k_{\max} \sim N^{1/(\gamma-1)}$, so the relative inaccuracy for the estimate k_{est} of the maximal degree is $k_{\max}/k_{\mathrm{est}} = N^{1/\gamma - 1/\gamma_{\mathrm{est}}}$, which is about 500%. Such inaccuracy has important repercussions for the ability to predict the emergence of collective behaviour [19, 48].

The statistical measures used for the construction of a functional network typically depend in a nonlinear way on the degrees, thus causing a distortion in the statistics. We will discuss the case of Pearson distance. Suppose that the signals $\{(y_i(t), y_i(t+1))\}$ are purely deterministic, $y_i(t+1) = g_i(y_i(t))$. The Pearson distance s_{ij} between the signal at i and j is a number between 0 and 1, depending on how close these graphs are. This distance depends nonlinearly on the degrees k_i and k_j. Devising another distance s'_{ij} without knowledge of the interaction, in general, still carries the nonlinear dependence on the degrees. Once fluctuations from the network are included the differences between time-series can be due to fluctuations rather than differences in the degrees. The decomposition of the rules in terms of interactions and fluctuations is essential to recover degree distribution accurately.

The effective network provides a better statistical description of the network structure. To compare with the functional network approach, we constructed an effective network of the same system tested for the functional network. The estimate for γ from the effective network is $\gamma_{\mathrm{est}} = 2.55$, which has an error of only 1% (inset one of Fig. 5 (a)). We repeat the analysis on a different network with different parameters γ in the degree distribution. The estimated γ_{est} values are shown in Fig. 5 (c) as a function of the true parameter γ. The relative error on the estimated exponent is within 2%.

B. Performance of the degree distribution reconstruction

In Appendix E, we present additional simulations showing how accurate the degree distribution is reconstructed for various isolated dynamics. In particular in Figure 3 we show the results for a) doubling maps with diffusive coupling, b) logistic maps with Kuramoto interactions, c) spiking neurons with electrical coupling, and d) Hénon maps with the y-component diffusive coupled with the x-component. Moreover, in Section III F we show the performance of the reconstruction for a system of differential equations coupled on scale-free networks.
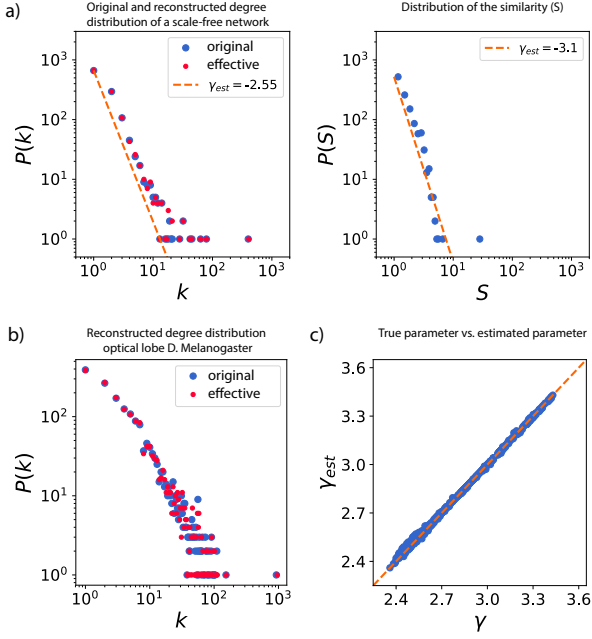
FIG. 5. **Reconstruction of structural power-law exponents γ of scale-free networks from data.** We estimated γ from the multivariate time series obtained from the dynamics random scale-free networks with degree distribution $P(k) \propto k^{-\gamma}$. The plots in panel (a) compare the functional and effective network approach. We obtain better estimates using the effective network. Panel (b) shows the degree distribution of the original system (in blue) and that estimated from an effective model (in red) for the neural network in the optical lobe of Drosophila melanogaster. We obtained an accuracy of 3% in the structural exponent γ. Panel (c) shows the true exponent γ versus γ_{est} obtained with an effective network from data for spiking neuron coupled with chemical synapses. We generated 1000 networks with distinct γ, from which the γ_{est} estimate is within 2% accuracy.

We provide a study on the effects of noise in the reconstruction. We established that for stochastically stable [49] systems such that the doubling map if the noise amplitude η_0 satisfies $\eta_0 < \alpha k_{\min}$, where $k_{\min}$ is the minimal degree the reconstruction procedure works. When the noise amplitude is of order αk_i nodes with degree less than k_i cannot be estimated.

C. The optic lobe of D. Melanogaster.

We applied our method to data simulated from the neuronal network in the Drosophila Melanogaster optic lobe, which constitutes >50% of the total brain volume and contains 1781 nodes [30]. The degree distribution has a power-law tail [31]. We used spiking neurons with chemical coupling to simulate the multivariate time series, from which we constructed an effective model and estimate the degree distribution (Fig. 5 (b)).

D. Experimental data of optoelectronic oscillators

We now apply our effective network to experimental data of networks of optoelectronic oscillators whose nonlinear component is a Mach-Zehnder intensity modulator. This data was generated in Ref. [50] where the authors studied enhancement of synchronization by structural changes in the network. The the experimental setup can also be found in Refs. [50, 51]. Each element consists of a clocked optoelectronic feedback loop. Light from a 780 nm continuous-wave laser is nonlinearly transformed as it passes through the Mach-Zehnder intensity modulator. Light intensity is converted into an electrical signal by a photoreceiver and measured by a field-programmable gate array (FPGA) via an analog-to-digital converter. The FPGA is clocked at 10 kHz, resulting in the discrete-time map dynamics of the oscillators. The FPGA controls a digital-to-analog converter that drives the modulator with a voltage $x_i(t+1) = \beta I(xi(t))$, closing the feedback loop. The elements are coupled electronically on the FPGA according to the desired coupling matrix as described in detail in Ref. [51]. The system can be modeled as

$$x_i(t+1) = \beta I(x_i(t)) + \sigma \sum_{j=1}^{n} A_{ij}[I(x_j(t)) - I(x_i(t))] \bmod 2\pi$$

where t is discrete time, β is the feedback strength, $I(x) = \sin^2(x + \delta)$ is the normalized intensity output of the Mach-Zehnder modulator, x represents the normalized voltage applied to the modulator, and δ is the operating point set to $\pi/4$. The data is acquired for $\beta = 4.5$ and 17 elements coupled through the network presented in Figure 6 left panel. The coupling strength σ varies from 0 to 1 in steps of 0.0325 starting from 0.015625. For each fixed value of σ, we obtain the experimental multivariate time series $\{x_1(t), \cdots, x_{17}(t)\}_{t=1}^{15385}$.

We discard the first 5000 data points for each $i = 1, \cdots, 17$ as a transient. We will provide an analysis for the coupling $\sigma = 0.03125$. First, we perform a functional network analysis by considering a correlation matrix Σ_x of the multivariate time series. To obtain a model of the adjacency matrix we threshold Σ_x. The value of the threshold 0.02 is chosen such that the functional network has a mean degree close to the actual network. The result is shown in Figure 6 in the middle panel and as observed the functional network does not capture the actual network structure.

Next we employ the effective network. We start by applying Step (1) to learn the function g_i and Step (2) from where we obtain the degrees and coupling strength. Once we obtain g_i, we filter the determinist part from x_i to obtain the fluctuations ξ_i. Next, we compute correlation matrix Σ_ξ for the fluctuations ξ_i. To turn this matrix into a network, we threshold it. Again the value of the threshold is fixed such that the mean degree is closed to the actual network. Here, any threshold value from 0.07 to 0.1 works. The result is shown in Figure 6, in the right panel, and shows excellent agreement with the actual network. In fact, only two links are misidentified.

We also performed the analysis for further coupling strengths σ. For large coupling strengths both functional network and effective network will capture the network misiden-

tifying on average 4 links. In these cases, the effective network has the advantage that it provides in addition to a model of the adjacency matrix also a model for the local dynamics.

VI. CONCLUSIONS

We have introduced an *effective network* obtained from time-series of a complex network observing the dynamics at each node. Our method complements the existing ones in two ways: First of all, it encompasses the case of chaotic local dynamics at each node. Secondly it deals with weak coupling among the nodes. Both cases are commonly found in applications [1, 15, 16]. Key to the success of the reconstruction is the heterogeneity of the network which allows us to perform a multi-level reduction. To recover the community structures, we use that certain noise terms associated with the time series at two nodes in the same community are correlated. By collecting data when the network is far from critical transitions, an effective network enables us to predict a critical transition.

We have compared our procedure with methodologies most relevant for the systems considered. We have excluded results tailored to specific setups or dynamics (binary dynamics [52], and see [10] for a review). We did not consider methods that rely on measurements obtained by intervening on the system with controlled inputs [13] and restrict our attention to time-series recorded under constant conditions. When the coupling is strong, sparse recovery can be applied [34]. When the coupling is weak sparse recovery cannot distinguish small parameters from those that are identically zero thus misidentifying connections between nodes. Also model-free methods are ill-suited as the influence of a single pairwise interaction on the time-series is weak and can hardly be detected.

The effective network methodology performs well when the network is heterogenous and has a few nodes making a large number of connections while most of the nodes are less connected, and the local dynamics are chaotic and their typical orbits visit most of the phase space. The effective network approach did not perform well in two cases. The first is when most of the observed time-series take values on a very restricted part of the phase space, for example if the local dynamics has a singular attractor, as an attracting fixed point, or if it spends long periods of time in a small region, like around the fixed points of the classical Lorenz attractor. This means that we don't have access to a big portion of phase space, and no prediction is possible in those regimes of coupling strength that make these portions accessible. The passage near a fixed point also suppresses the fluctuations hindering the reconstruction of communities. This is what seems to happen for example in the bursting dynamics of Rulkov maps, when the quiescent state is too long. These situations are excluded if the local dynamics is sufficiently chaotic. The second case is when the coupling is strong enough to synchronize big parts of the network. For example, a synchronous rich-club can send similar forcing to nodes in different communities resulting in high correlations between the fluctuations. Therefore our method would identify these nodes as belonging to the same community even if they are not.

Appendix A: Effective network representation from data

A summary of the effective network approach is given in Figure 7. Here we include some details that were omitted for the sake of presentation in the main text.

In Step 2 of the reconstructing procedure, we identify low degree nodes by analysing the distribution of S_i. More precisely, we use the top N_{top} nodes of the highest intensity to obtain a proxy for the isolated dynamics. We then average these rules to get $\langle g \rangle \approx f$. The choice of N_{top} is not fixed and depends on the number of nodes and the fluctuation $\sigma_g^2 = \langle (g_i - \langle g \rangle)^2 \rangle$. For scale-free (Barabasi-Albert) networks the degree of the hubs scales as $N^{1/2}$, a good heuristic is to choose N_{top} satisfying $\sigma_g^2 / N_{\text{top}}^{1/2} \ll 1$. The effective coupling function $\alpha k_i v$ can be obtained analysing the family $\{g_i - \langle g \rangle\}_{i=1}^N$ which can yield the shape of v up to a multiplicative constant via a nonlinear regression by imposing that $g_i - \langle g \rangle$ and $g_j - \langle g \rangle$ are linearly dependent. The choice of the base function for the fitting is supervised (see Appendix E).

In Step 3, after selecting a v that satisfactorily approximates $g_i - \langle g \rangle$ up to a multiplicative constant over all indices i, the parameter β_i is estimated using a dynamic Bayesian inference. Because the fluctuations $\xi_i(t)$ are close to Gaussian, we use a Gaussian likelihood function and a Gaussian prior for the distribution of the values of β_i, and hence obtain equations for the mean and variance. We split the data into epochs of 200 points and update the mean and variance iteratively.

1. Community structures

Once we obtain the rules g_i, we filter the deterministic part of the time series y_i and access the fluctuations ξ_i (recall Eq. (2)) and decompose it as $\xi_i = \xi_i^c + \xi_i^o$ where ξ_i^c is the fluctuation of the local mean field from nodes in the cluster containing i, and ξ_i^o is the contribution from outside the cluster. Since a node makes most of its connections within its cluster, $\xi_i^c \gg \xi_i^o$ with high probability, and thus if i and j belong to the same cluster $\text{Corr}(\xi_i, \xi_j) = \text{Corr}(\xi_i^c, \xi_j^c)$. The common noise is generated by the common connections between nodes i and j. For fixed isolated dynamics and coupling function

$$\text{Corr}(\xi_i^c, \xi_j^c) \propto \widehat{\mu}_{ij}.$$

$\text{Corr}(\xi_i, \xi_j)$ is related to the *matching index* [29] of the nodes i and j. This is a parameter used to quantify the number of common neighbours of two nodes. Recall that the degree of node i is $k_i = \sum_j^N A_{ij}$ and counts the number of neighbours it has. Consider the neighbourhood of node i, $\Gamma(i) = \{j \in \{1, \ldots, N\} \,|\, A_{ij} = 1\}$. This is the set of nodes that shares an edge with the node i. The matching index of nodes i and ℓ is the cardinality of the overlap of their neighbourhoods $\mu_{i\ell} = |\Gamma(i) \cap \Gamma(\ell)|$. We consider the normalised matching index:

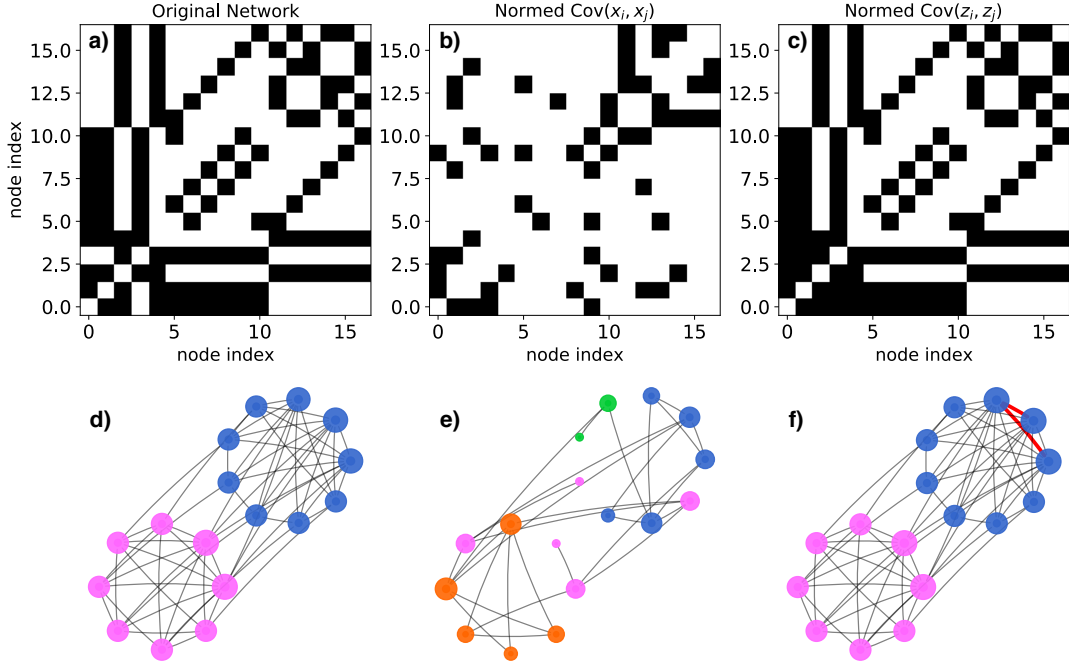$$\widehat{\mu}_{i\ell} = \frac{|\Gamma(i) \cap \Gamma(\ell)|}{|\Gamma(i) \cup \Gamma(\ell)|}$$

FIG. 6. **Effective network from experimental data of networks of optoelectronic oscillators.** We consider multivariate time series of voltages of a network of 17 weakly coupled optoelectronic oscillators (interaction corresponding to 1% of the oscillator amplitude). In the left panel, we show the actual network used to coupled the optoelectronic oscillators in Ref. [50] as adjacency matrix in a) and graph representation in d). In the mid panel, we show the reconstruction of the network by a functional network analysis in terms of its adjacency matrix in b) and graph representation in e). In the right panel, we show the reconstruction of the network from an analysis of the dynamical fluctuations by applying the effective network approach as adjacency matrix in c) and graph representation in f). The effective network provide a striking reconstruction and only two links are misidentified and are indicated in f as red links. In the graph representation, the nodes of the network are coloured according to the community obtained by a community detection algorithm [37].
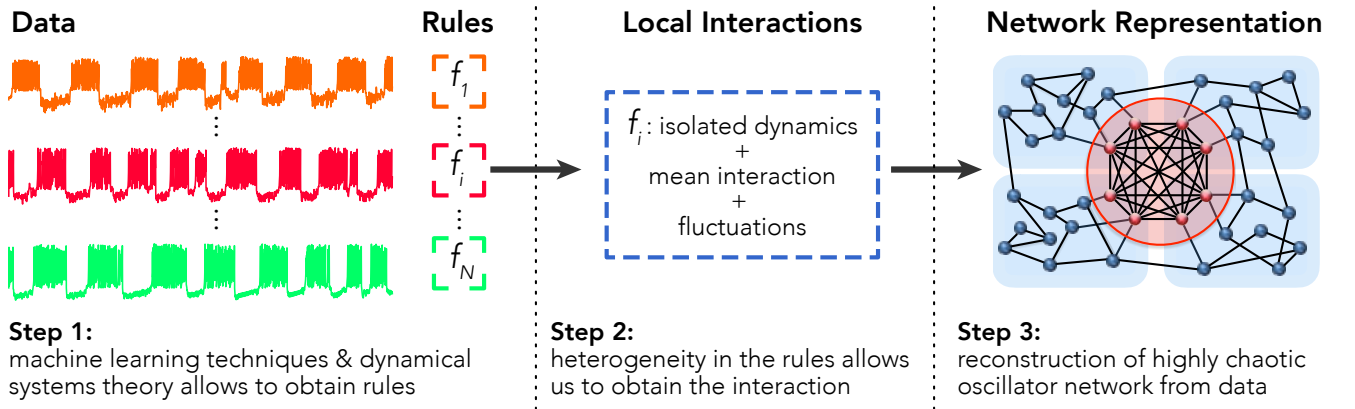

FIG. 7. Reconstruction scheme with the effective network. From the time series, we build a model for the local evolution f_i at each node. Under the assumption that such rules change from node to node depending on their connectivity, we estimate the coupling function. Using the fluctuations of the time series with respect to the low-dimensional rules, we recover the community structures. Gathering all this information, we obtain an effective network that can be used to predict critical transitions

.

or equivalently in terms of the adjacency matrix

$$\widehat{\mu}_{i\ell} = \frac{(A + A^2)_{i\ell}}{k_i + k_\ell - (A + A^2)_{i\ell}}.$$

Clearly $\widehat{\mu}_{i\ell} = 1$ if and only if i and l are connected to exactly the same nodes, and $\widehat{\mu}_{i\ell} = 0$ if they have no common neighbours. It is well known that in the cat cerebral cortex nodes in the same community have a high matching index while nodes

are distinct communities has a low matching index. This tends to be typically in modular networks [29]. For nodes in distinct clusters the component $\xi_i^c \approx 0$, so $\text{Corr}(\xi_i, \xi_j) \approx 0$. We recover the network structure from a *noise covariance analysis*.

Filtering out the deterministic part plays a major role in recovering community structures. Suppose we have two signals of the form $y_i(t) = Y_i(t) + \zeta(t)$, $i = 1, 2$, where Y_i is independent of i and $\zeta(t)$ is a common noise term. Y_i represents the superposition of the deterministic chaos and the independent fluctuations. For the correlation, we have

$$\text{Corr}(y_i, y_j) \approx \frac{\text{Cov}(\zeta, \zeta)}{\sigma_{Y_1}^2 \sigma_{Y_2}^2}$$

Hence, the large values of the variance of the time series ($\sigma_{y_i} \approx \sigma_{Y_i} \gg \sigma_\zeta$) suppress the contribution of the common noise, and an analysis solely based on the the original time series y_i will overlook the common noise contribution.

Appendix B: Functional networks

For networks of chaotic oscillators, building the functional network from the standard Pearson correlation between time series gives no meaningful results because of the decay of correlation intrinsic to dynamics. Functional networks are built using a Pearson distance $s_{ij} \geq 0$ describing the proximity of the dynamics at two nodes i and j. To do this, we consider the time series $z_i(t) := (y_i(t), y_i(t+1))$, $t = 0, \ldots, T-1$ reordered in $z_i^{\text{lex}}(t)$ according to the lexicon order; that is, according to the magnitude of the first component of $z_i(t)$. Then, let r_{ij} be the Pearson correlation, $r_{ij} = \text{Cor}(z_i^{\text{lex}}, z_j^{\text{lex}})$, so that $r_{ij} = 1$ indicates that the attractors at nodes i and j agree. Define the Pearson distance $s_{ij} = 1 - |r_{ij}|$ so that $s_{ij} = 0$ indicates agreement of the dynamics and $s_{ij} > 0$ measures the difference between the attractors.

The intensity $S_i = \sum_j s_{ij}$ approximates how many nodes have a dynamical rule different from i and helps to distinguish between poorly connected nodes and hubs. Since most of the network is composed of poorly connected nodes, they exhibit a smaller S_i than high-degree nodes, which are scarcer and have different dynamics from the low-degree nodes.

Appendix C: Predicting critical transitions

Here we explain how to gather the information for a theoretical prediction of the critical transition.
Reduction in the rich-club. Nodes in the rich-club have degrees of approximately Δ and make $\kappa\Delta$ connections inside the rich-club and $(1 - \kappa)\Delta$ connections to the rest of the network. Following our reduction scheme, the interactions within and outside the rich-club can be described by the expected value of the interactions with respect to the invariant measure associated with each of them. Let C denote the set of nodes in the rich-club, then the coupling term is

$$\sum_j A_{ij} \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{j \in C} A_{ij} \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) + \sum_{j \notin C} A_{ij} \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j)$$

However,

$$\sum_{j \notin C} A_{ij} \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) = (1 - \kappa)\Delta \int \boldsymbol{h}(\boldsymbol{x}_i, \boldsymbol{y}) d\mu(\boldsymbol{y}) + \boldsymbol{\xi}_i^o(t)$$

where μ is the invariant measure for the nodes outside the rich-club. Hence, for the rich-club we obtain

$$\boldsymbol{x}_i(t + 1) = \boldsymbol{q}_i(x_i(t)) + \sum_{j \in C} A_{ij} \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{x}_j) + \boldsymbol{\xi}_i^o(t),$$

where

$$\boldsymbol{q}_i(\boldsymbol{x}_i(t)) = \boldsymbol{F}_i(\boldsymbol{x}_i(t)) + (1 - \kappa)\Delta\alpha \int \boldsymbol{H}(\boldsymbol{x}_i, \boldsymbol{y}) d\mu(\boldsymbol{y}).$$

Predicting the transition to collective behaviour. Let us recall that when isolated $u_i(t+1) = F_{1,i}(u_i(t)) + w_i(t)$ where $F_{1,i} \approx F_1$, $w_i(t+1) = w_i(t) + \mu(w_i(t) - 1)$, and

$$w(t+1) = w_0 + \mu \sum_{n=0}^{t} (u(n) - 1)) \tag{C1}$$

Using the reduction Eq. (C1), in the network we obtain

$$u_i(t+1) = F_{1,i}(u_i(t)) + u_i(t) + \Delta\alpha[\langle u \rangle - u_i(t)] + \xi_i(t)$$

where i denotes the ith nodes in the rich-club, $\langle u \rangle$ is the mean in the rich-club and ξ_i are fluctuations. We fix two nodes $y_i = u_i$ and $y_j = u_j$ in the rich-club and consider

$$\zeta(t) = u_i(t) - u_j(t)$$

Using that $F_{1,i} \approx F_1$ by the mean value theorem we obtain

$$\zeta(t+1) = DF_1(x_i(t))\zeta(t) + \mu \sum_{n=0}^{t} \zeta(n) - \Delta\alpha\zeta(t)$$

and introducing a proxy for the dynamics of the slow variables

$$\eta(t) = \sum_{n=0}^{t} z(n)$$

and considering $\sum_{n=0}^{t} DF_1(x_i(n))\zeta(n) \approx \lambda \sum_{n=0}^{t} \zeta(n)$ where we used that $\sum_{n=0}^{t} \zeta(n)$ is a slow variable. We obtain

$$\eta(t+1) = (\lambda - \Delta\alpha)\eta(t) + \mu \sum_{n=0}^{t} \eta(n)$$

For the cat cerebral cortex $\Delta = 37$. Given the time series $\{y_i\}$ for $\Delta\alpha = 0.3$, we estimate F_1 using our method i as the slow variables are constants over short time scales, and the obtain slow variables as a filter over the fast variables. From the data, we estimate $\lambda = 1.42$ and thus we obtain $\Delta\alpha = 0.42$. At this critical value the slow variables tend the stay together due to the contraction in the dynamics. This is related to the onset of synchronization in the bursts, which is captured via a phase variable through the order parameter.

For estimation of the power-law distribution parameters, we use the maximum likelihood estimator [53, 54]. After that, we test the reliability between the data and the power law by using the goodness-of-fit method. If the resulting p-value is larger than 0.1, the power-law estimation is an appropriate hypothesis for the data. A complete procedure for the analysis of power-law data can be found in Ref. [55].

Appendix D: Dimensional reduction in heterogeneous networks

We present an informal statement of the theoretical results used in the reconstruction procedure. For a precise statement see [19]. The theorem has three main assumptions:

1) **The local dynamics must increase the distance between points** by a constant factor.

2) **The networks are heterogeneous**. Most of the nodes have small degree $\delta \sim N^{\frac{\epsilon}{2}}$, and some nodes are hubs with degree $\Delta \sim N^{\frac{1}{2}+\epsilon}$.

3) **The reduced dynamics must be hyperbolic.** The maps $\boldsymbol{G}_j$ are either expanding or to have a finite number of attracting periodic orbits. In dimension one, every map can be perturbed by an arbitrarily small amount to obtain such an hyperbolic map [56].

Under these assumptions, we have the following result

Theorem 1 ([19]) *For every hub node j, the dynamics at the hub is given by*

$$\boldsymbol{x}_j(t+1) = \boldsymbol{G}_j(\boldsymbol{x}_j(t)) + \boldsymbol{\xi}_j(t)$$

where $|\boldsymbol{\xi}_j(t)| < \xi$ for time T with $1 \leq T \leq \exp[C\xi^2\Delta]$, and a set of initial condition of measure $1 - T/\exp[C\xi^2\Delta]$, where C is constant in Δ and ξ.

Notice that one can pick the time scale T exponentially large, but such that $T/\exp[C\xi^2\Delta]$ is very small so that, for large Δ, the approximation result holds for very long time and for a large set of initial conditions.

Appendix E: The effective network for a variety of chaotic dynamics and coupling

We tested the performance of the effective network in recovering community structure and degree distribution for the systems listed below. Recovery of community structures was tested on a network of 100 nodes having five clusters of 20 nodes each. Four of these clusters are modeled as Erdös-Renyi random graph with connection probability $p = 0.3$, and the fifth, the integrating cluster, with $p = 0.8$. The coupling strength is α is of the order of 10^{-4}. Recovery of degree distribution was tested on scale-free networks with 6000 nodes and characteristic exponent γ varying between 2.4 and 3.6,

and coupling strength at $\alpha\Delta = 0.5$. Details and results of the simulations can be found in Supplementary Materials.

Doubling maps. Since the dynamics is one dimensional, we denote $\boldsymbol{x} = x$ and $\boldsymbol{F}_i(\boldsymbol{x}) = f_i(x)$ with $f_i(x) = 2x + \varepsilon_i \sin 2\pi x \mod 1$ and where we take ε_i to be i.i.d. random variables uniformly distributed on $[0, 10^{-3}]$. Likewise we write $\boldsymbol{H} = h$ with $h(x_j, x_i) = \sin 2\pi x_j - \sin 2\pi x_i$. We were able to recover all community structures, and the characteristic exponent γ within 0.5% accuracy.

Logistic map. Again, $\boldsymbol{x} = x$ and $\boldsymbol{F}_i(\boldsymbol{x}) = f_i(x)$ where $f(x) := 4x(1-x)$, and we consider $h(x_j, x_i) = \sin(2\pi x_j - 2\pi x_i)$. We were able to recover all community structures, and the characteristic exponent within 0.5% accuracy.

Spiking neurons with electrical synapses. We use the same spiking neurons as in the main body of the manuscript and denoting $\boldsymbol{x} = (u, w)$ the coupling function reads as $\boldsymbol{H}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{E}(\boldsymbol{x_j} - \boldsymbol{x_i}) = (u_j - u_i, 0)$. We were able to recover all community structures and the characteristic exponent within 2% accuracy.

Bursting neurons with electrical synapses. Our numerical investigation reveals that when the resting time is not much larger then the total bursting time the reduced dynamics is capable of extracting the relevant information of the time series. Thus, we fixed the neuron parameter $\beta = 4.4$ to obtain a bursting dynamics. The coupling is electrical as for the systems above. We were able to recover all community structures.

Henon Maps. Using the notation $\boldsymbol{x} = (u, w)$, the coupled Hénon maps we study are given by $\boldsymbol{F}(u, w) = (1 - 1.4u^2 + w, 0.3w)$ and $\boldsymbol{H}(\boldsymbol{x_i}, \boldsymbol{x_j}) = (w_j - w_i, 0)$. We assume to observe only the dynamics of the first component $y = \phi(\boldsymbol{x}) = u$. In this multidimensional case, the reconstruction will start by determining the dimension of the reduced system. Takens embedding reveals that the dimension is two for large time excursions, hence, we will aim at learning a function

$$y_i(t+1) = g_i(y_i(t), y_i(t-1)) + \xi_i(t). \tag{E1}$$

We use polynomial functions for the fitting via a 10-fold cross-validation. Our theory implies that $g_i(y_i(t), y_i(t-1)) = f(y_i(t), y_i(t-1)) + \alpha k_i v(y_i(t), y_i(t-1))$ where f models the isolated dynamics and v the coupling. We obtain f from the low-degree nodes via a similarity analysis. We learn h by $\alpha k_i v(y_i(t), y_i(t-1)) = g_i(y_i(t), y_i(t-1)) - f(y_i(t), y_i(t-1))$. We were able to recover all community structures and the characteristic exponent within 2% accuracy.

[1] Kandel, E.R., Schwartz, J.H. & Jessell, T.M. eds., *Principles of neural science*, Vol. 4. (New York: McGraw-hill, 2000).

[2] Bohland, J. W. *et al.* A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS Computational Biology* **5**, e1000334 (2009).

[3] De La Fuente, A., Bing, N., Hoeschele, I., & Mendes, P. Dis-

covery of meaningful associations in genomic data using partial correlation coefficients. *Bioinformatics*, **20**, 3565-3574 (2004).

[4] Reverter, A., & Chan, E. K. Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks. *Bioinformatics*, **24**, 2491-2497 (2008).

[5] Butte, A. J., & Kohane, I. S. Mutual information relevance net-

works: functional genomic clustering using pairwise entropy measurements. *In Biocomputing 2000* (pp. 418-429), (1999).

[6] Braunstein, A., Pagnani, A., Weigt, M., & Zecchina, R. Inference algorithms for gene networks: a statistical mechanics analysis. *Journal of Statistical Mechanics: Theory and Experiment*, **2008**, P12001 (2008).

[7] Cocco, S., Leibler, S., & Monasson, R. Neuronal couplings between retinal ganglion cells inferred by efficient inverse statistical physics methods. *Proc. Natl. Acad. Sci. USA*, **106**(33), 14058-14062 (2009).

[8] Bressler, S. L., & Seth, A. K. Wiener Granger causality: a well established methodology. *Neuroimage*, **58**(2), 323-329 (2011).

[9] Ladroue, C., Guo, S., Kendrick, K., & Feng, J. Beyond element-wise interactions: identifying complex interactions in biological processes. *PloS one*, **4**(9), e6899 (2009)

[10] Wang, W., Lai, Y., & Grebogi, C. Data based identification and prediction of nonlinear and complex dynamical systems. *Physics Reports*, **644**, 1–76 (2016).

[11] Casadiego, J., Nitzan, M., Hallerberg, S., & Timme, M. Model-free inference of direct network interactions from nonlinear collective dynamics. *Nat. Commun.* **8**, 2192 (2017).

[12] Han, X., Shen, Z., Wang, W. X., & Di, Z.. Robust reconstruction of complex networks from sparse data. *Phys. Rev. Lett.*, **114**, 028701 (2015).

[13] Nitzan, M., Casadiego, J., & Timme, M. Revealing physical interaction networks from statistics of collective dynamics. *Science advances*, **3**, e1600396 (2017).

[14] Stankovski, T., Pereira, T., McClintock, P. V., & Stefanovska, A. Coupling functions: universal insights into dynamical interaction mechanisms. *Reviews of Modern Physics*, **89**, 045001 (2017).

[15] Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* **440**, 1007–1012 (2006).

[16] Haas, J. S. A new measure for the strength of electrical synapses. *Front. Cell. Neurosci.*, **9**, 378 (2015).

[17] Van Den Heuvel, M. P., & Sporns, O. Rich-club organization of the human connectome. *J. Neurosci.*, **31**, 15775–15786. (2011).

[18] Park, H. J., & Friston, K. Structural and functional brain networks: From connections to cognition. *Science* **342**, 6158 (2013).

[19] Pereira, T., van Strien, S., & Tanzi, M. Heterogeneously coupled maps: hub dynamics and emergence across connectivity layers. *To appear in J. Eur. Math. Soc., preprint: arXiv* **1704.06163**, 1–63 (2017)

[20] Izhikevich, E. M. *Dynamical systems in neuroscience*, (MIT press, 2007).

[21] Yadav, P., McCann, J. A., & Pereira, T. Self-synchronization in duty-cycled internet of things (IoT) applications. *IEEE Internet of Things Journal* **4**, 2058–2069 (2017).

[22] Dörfler, F., Chertkov, M., & Bullo, F. Synchronization in complex oscillator networks and smart grids. *Proc. Natl. Acad. Sci. USA* **110**, 2005–2010 (2013).

[23] Watanabe, S., & Strogatz, S. H. Constants of motion for superconducting Josephson arrays. *Physica D: Nonlinear Phenomena* **74**, 197–253 (1994).

[24] Winfree, A.T. *The Geometry of Biological Time*, Interdisciplinary Applied Mathematics: Vol 12 (Springer-Verlag New York, 2001).

[25] Pinto, R. D. *et al.* Synchronous behavior of two coupled electronic neurons. *Phys. Rev. E* **62**, 2644–2656 (2000).

[26] Eroglu, D., Lamb, J. S. W., & Pereira, T. Synchronisation of chaos and its applications. *Contemporary Physics* **58**, 207–243 (2017).

[27] Scannell, J. W. & Young, M. P. The connectional organization of neural systems in the cat cerebral cortex. *Curr. Biol.* **3**, 191–200 (1993).

[28] Scannell, J. W., Blakemore, C. & Young, M. P. Analysis of connectivity in the cat cerebral cortex. *J. Neurosci.* **15**, 1463–1483 (1995).

[29] Zamora-López, G., Zhou, C., & Kurths, J. Cortical hubs form a module for multisensory integration on top of the hierarchy of cortical networks. *Frontiers in Neuroinformatics* **4**, 1–13 (2010).

[30] Takemura, S., *et al.* A visual motion detection circuit suggested by Drosophila connectomics. *Nature* **500**, 175–181 (2013).

[31] García-Pérez, G., Boguñá, M., & Serrano, M. Á. Multiscale unfolding of real networks by geometric renormalization. *Nat. Phys.* (2018).

[32] Shandilya, S. G., & Timme, M. Inferring network topology from complex dynamics. *New Journal of Physics*, **13**, 013004 (2011).

[33] James, G., Witten, D., Hastie, T., & Tibshirani, R. *An introduction to statistical learning*, Vol. 112. (New York: Springer, 2013.)

[34] Brunton, S. L., Proctor, J. L., & Kutz, J. N. Discovering governing equations from data: Sparse identification of nonlinear dynamical systems. *Proc. Natl. Acad. Sci. USA* **113**, 3932–3937 (2015).

[35] Wang, W. X., *et at.* Predicting catastrophes in nonlinear dynamical systems by compressive sensing. *Phys. Rev. Lett.* **106**, 154101 (2011).

[36] Judd K., & Mees A. Embedding as a modeling problem. *Physica D: Nonlinear Phenomena* **120**, 273 – 286 (1998).

[37] Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Étienne, L. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).

[38] Eguiluz, V. M., Chialvo, D. R., Cecchi, G. A., Baliki, M., & Apkarian, A. V. Scale-free brain functional networks. *Phys. Rev. Lett.* **94**, 018102 (2005)

[39] Bullmore, E., & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neuroscience* **10**, 186–198 (2009).

[40] Bettinardi, R. G., *et al.* How structure sculpts function: Unveiling the contribution of anatomical connectivity to the brain's spontaneous correlation structure. *Chaos* **27**, 047409 (2017).

[41] Zhang, J., & Small, M., Complex network from pseudoperiodic time series: Topology versus dynamics. *Phys. Rev. Lett.* **96**, 238701 (2006).

[42] Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. USA* **100**, 253–258 (2003).

[43] Colizza, V., Flammini, A., Serrano, M. A., & Vespignani, A. Detecting rich-club ordering in complex networks. *Nat. Phys.* **2**, 110 (2006).

[44] Lopes, M. A., *et at.* An optimal strategy for epilepsy surgery: Disruption of the rich-club? *PLoS computational biology* **13**, e1005637 (2017).

[45] Rulkov, N. F. Regularization of synchronized chaotic bursts. *Phys. Rev. Lett.* **86**, 183–186 (2001).

[46] Cleveland, W.S. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* **74** 368:829-836 (1979).

[47] Pereira, T., M. S. Baptista, & J. Kurths. Phase and average period of chaotic oscillators. *Physics Letters A* **362**, 159-165 (2007).

[48] Pereira, T. Hub synchronization in scale-free networks.

Phys. Rev. E, **82**(3), 036201(2010).

[49] Tanzi, M., Pereira T., and van Strien, S. Robustness of ergodic properties of non-autonomous piecewise expanding maps. *Ergodic Theory and Dynamical Systems* 39, 1121-1152 (2019).

[50] Hart, J. D., Zhang, Y., Roy, R., and Motter, A. E., Topological control of synchronization patterns: Trading symmetry for stability. *Phys. Rev. Lett.* **122**, 058301 (2019).

[51] Hart, J. D., Schmadel, D. C., Murphy, T. E., and Roy R., Experiments with arbitrary networks in time-multiplexed delay systems, *Chaos* **27**, 121103 (2017).

[52] Li, Y., et al. Universal style transfer via feature transforms. *Advances in neural information processing systems*, pp. 386-396, (2017).

[53] Muniruzzaman, A. N. M. On Measures of location and dispersion and tests of hypotheses in a pare to population. *Calcutta Statistical Association Bulletin* **7**, 115–123 (1957).

[54] Hill, B. M. A Simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**, 1163–1174 (1975).

[55] Clauset, A., Shalizi, C. R. & Newman M. E. J. Power-Law distributions in empirical data. *SIAM Review* **51**, 661–703 (2009).

[56] de Melo, W., and Van Strien, S., *One-dimensional dynamics*. Springer (2012).

[57] Rössler, O. E. *An equation for continuous chaos*. Physics Letters A 57, no. 5 (1976): 397-398.

Data Archival

The connection matrices of cat cortex is found at `https://sites.google.com/site/bctnet/datasets`. Connectivity of Drosophila Melanogaster is found at `https://neurodata.io/project/connectomes/`. The experimental data on the optoelectronic oscillators from Ref. [50] can be obtained by contacting Joseph Hart and R. Roy upon reasonable request.

ACKNOWLEDGMENTS