Broken detailed balance and entropy production in directed networks

Ramón Nartallo-Kaluarachchi[1,2,*], Malbor Asllani[3,§], Gustavo Deco[4,5,6], Morten L. Kringelbach[2,7,8], Alain Goriely[1,†] and Renaud Lambiotte[1,9,‡]

1. Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, United Kingdom

2. Centre for Eudaimonia and Human Flourishing, University of Oxford, 7 Stoke Pl, Oxford, OX3 9BX, United Kingdom

3. Department of Mathematics, Florida State University, Tallahassee, FL 32306, United States of America

4. Centre for Brain and Cognition, Computational Neuroscience Group, Universitat Pompeu Fabra, 08018 Barcelona, Spain

5. Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08018 Barcelona, Spain

6. Institucio Catalana de la Recerca i Estudis Avancats (ICREA), 08010 Barcelona, Spain

7. Center for Music in the Brain, Aarhus University, & The Royal Academy of Music, Aarhus/Aalborg, Denmark

8. Department of Psychiatry, University of Oxford, Oxford, OX3 7JX United Kingdom

9. The Turing Institute, British Library, 96 Euston Rd, London, NW1 2DB, United Kingdom

*{nartallokalu}†{goriely}‡{lambiotte}@maths.ox.ac.uk, §{masllani}@fsu.edu

(Dated: March 5, 2024)

The structure of a complex network plays a crucial role in determining its dynamical properties. In this work, we show that the directed, hierarchical organisation of a network causes the system to break detailed balance and dictates the production of entropy through non-equilibrium dynamics. We consider a wide range of dynamical processes and show how different directed network features govern their thermodynamics. Next, we analyse a collection of 97 empirical networks and show that strong directedness and non-equilibrium dynamics are both ubiquitous in real-world systems. Finally, we present a simple method for inferring broken detailed balance and directed network structure from multivariate time-series and apply our method to identify non-equilibrium and hierarchical organisation in both human neuroimaging and financial time-series. Overall, our results shed light on the thermodynamic consequences of directed network structure and indicate the importance and ubiquity of hierarchical organisation and non-equilibrium dynamics in real-world systems.

I. INTRODUCTION

The abstraction of large complex systems as networks of interconnected elements has been instrumental in the modelling of systems in ecology [1], economics [2], sociology [3], bio-medicine [4], neuroscience [5, 6] and beyond [7]. In particular, dynamical processes evolving on networks have become prototypical models of real-world systems in a range of diverse fields [8–10]. Reconciling the relationship between the structure of interactions and the emergent dynamical phenomena of such systems remains a outstanding challenge. Many physical, chemical and biological systems operate far from thermodynamic equilibrium [11]. These non-equilibrium systems consume energy and dissipate heat to their surroundings, producing entropy [12]. In particular, energy consumption and entropy production represent a key mechanism by which living systems are able to stave off thermodynamic equilibrium and so-called 'heat death' [13]. Equilibrium systems are often characterised by symmetric interactions between identical elements that in-turn yield time-reversible dynamics. Similarly, non-reciprocal interactions between elements cause violations of the so-called 'detailed balance condition' given, equivalently, by the equalities $P_{\boldsymbol{yx}} = P_{\boldsymbol{xy}}$ and $\pi_{\boldsymbol{y}} \cdot P_{\boldsymbol{x}|\boldsymbol{y}} = \pi_{\boldsymbol{x}} \cdot P_{\boldsymbol{y}|\boldsymbol{x}}$, and illustrated in Figure 1. Here $P_{\boldsymbol{xy}}$ is the joint transition probability from state $\boldsymbol{x}$ to $\boldsymbol{y}$, $P_{\boldsymbol{x}|\boldsymbol{y}}$ is the transition probability from $\boldsymbol{y}$ to $\boldsymbol{x}$ conditional on the system being in state $\boldsymbol{y}$, and $\pi_{\boldsymbol{y}}$ is the steady-state probability of being in state $\boldsymbol{y}$. Violation of detailed balance leads to non-equilibrium steady states and irreversibility in the system's dynamics. The degree to which a system diverges from thermodynamic equilibrium can be quantified through the rate at which it produces entropy [14, 15]. The entropy production rate (EPR) quantifies the distance of the system from equilibrium and the irreversibility of its dynamics by measuring the divergence between the probability of observing system trajectories and their time-reversals [16]. In this work, we make progress on this front by demonstrating a novel and important link between the structure of directed networks and the thermodynamics of dynamical processes evolving on them.

Broken detailed balance and non-equilibrium dynamics have been observed in a range of microscopic [17–22] and mesoscopic [23] processes at the molecular and cellular level in living systems. At the macroscopic scale, temporal irreversibility has been observed in evolutionary dynamics [24] and large-scale neural dynamics [25–27]. However, despite many advancements in modern non-equilibrium statistical physics [28],

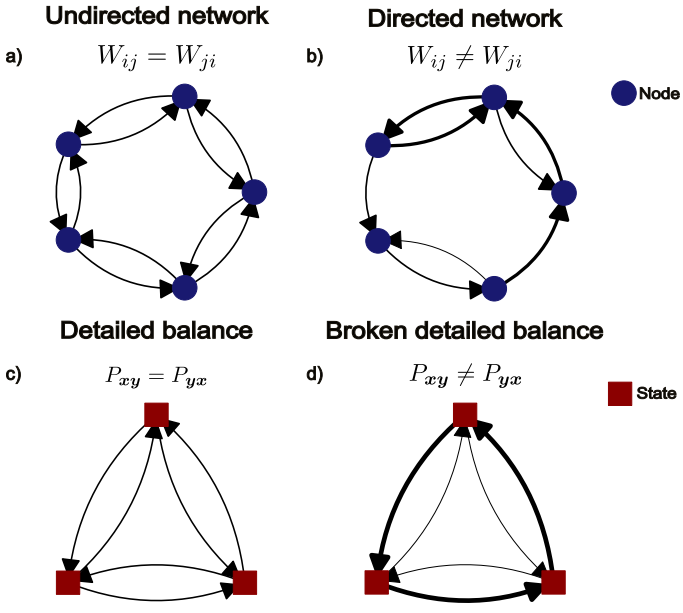

FIG. 1. **Asymmetry in complex systems** Symmetric nodal interactions leads to symmetric transition rates between system states whilst asymmetric interactions lead to broken detailed balance. **Top:** *a*) An undirected network with $W_{ij} = W_{ji}$. *b*) A directed network with $W_{ij} \neq W_{ji}$. Blue circles represents nodes or elements of the system. **Bottom:** *c*) A system in detailed balance with $P_{xy} = P_{yx}$. *d*) A system violating detailed balance with $P_{xy} \neq P_{yx}$. Red squares represent distinct, discrete system states. The thickness of connections represents the weight/probability, respectively.

results in stochastic thermodynamics have been limited to the study of small systems, with large, complex systems only attracting attention very recently [29–32]. When we abstract systems as networks of nodes and edges, symmetric interactions correspond to undirected networks, i.e., when a pair of edges exist between two nodes and they have the same weight in both directions. Conversely, asymmetric interactions correspond to directed networks, where the existence and strength of edges can vary in each direction. Previous attempts to reconcile network science and non-equilibrium thermodynamics have focused on network representations of state-space, where nodes represent mesoscopic states [33] or thermodynamic quantities [34] whilst edges represent transition rates. However, little is known about the role of network structure in real-space, where nodes and edges represent elements of the system and their interaction strengths, with the exception of chemical reaction networks [35–37]. Despite this, in systems such as the human brain, pairwise interactions have been shown to be the dominant contribution to the EPR, highlighting the importance of network structure in a system's thermodynamics [38]. Understanding the role of network topology remains an important unsolved problem in the thermodynamics of complex systems [39].

As anticipated, symmetry-breaking in the organisation of complex systems is the structural feature that drives network dynamical processes out of equilibrium. Dynamical processes on directed networks drastically differ from their undirected counterparts [40, 41] including in their phase-transitions [42], synchronisation properties [43, 44], topological resilience [45, 46] and pattern formation [47, 48]. Recent studies on extensive datasets of real-world networks from an array of disparate fields, have shown a ubiquity of strong directedness and clear signs of hierarchical organisation [41, 49–53]. As a result, the dynamics of directed networks are more indicative of the dynamics of real-world complex systems. In particular, this strong directedness results in a marked non-normality of the operators defined on such networks [50, 54]. Consequently, non-normality pulls the underlying networked systems away from equilibrium [50]. Motivated by the ubiquity of directed structures in the real world and their significance in nonlinear dynamics, this paper aims to further strengthen this link from the perspective of non-equilibrium statistical physics.

In this work, we bring new insight into the role of network structure in the emergence of broken detailed balance and irreversibility by demonstrating how the directedness of the interaction network causes a dynamical process to diverge from thermodynamic equilibrium. We first define a range of measures of directedness in networks and a scheme to smoothly parameterise the directedness of a network. Secondly, we introduce four network dynamical processes and calculate their EPR, namely the discrete- and continuous-time random walks (RW) [55, 56], Ornstein-Uhlenbeck (OU) [57] and Ising dynamics [58]. We show that increased directedness drives an increase in the EPR for all processes across network size. Next, using the Erdös-Rényi (ER) [59] as a null model, we decouple different measures of directedness and show that locally evolving processes, like the RW, produce more entropy when the system becomes more 'locally directed', whereas the EPR of globally coupled processes, like the Ising and OU dynamics, is dictated by the 'global directedness' of the underlying network. Subsequently, we consider structural data in the form of 97 real-world directed networks from a range of fields including biology, social interactions, ecology, transport, and language [52]. By considering the processes evolving on real-world topologies, we are able to further confirm the link between the directedness measures and the EPR in real systems. Finally, we describe a simple, but powerful, method for extracting directed networks and the EPR from multivariate time-series (MVTS) using a linear auto-regression technique [60]. We apply this method to MVTS from human neuroimaging and the stock-market to reveal the hierarchical organisation of brain-regions and stock price interactions as well as the non-equilibrium nature of their dynamics. In particular we confirm the increased EPR in task-based brain states,

as previously found [25, 26, 61, 62], but additionally show that a reorganisation of the hierarchy of brain regions drives this phenomena. Moreover, we show that consumer goods/services emerge as leader nodes in the hierarchy of stocks that governs market movements. Overall, this paper draws an important and fundamental link between the directed structure of a complex system and its stochastic thermodynamics. This work puts to the forefront the importance of considering asymmetries in interactions when studying the dynamics of complex networks, a consideration often overlooked in areas such as neuroscience [63, 64], and presents a new perspective for analysing both non-equilibrium systems and directed networks in both theory and empirical data.

II. MEASURES OF DIRECTEDNESS IN NETWORKS

As anticipated, non-reciprocal interactions in complex systems disrupt detailed balance, leading to an increase in the EPR. This section delineates four established metrics to quantify the overall directedness of networks and to examine their influence on EPR within both random graphs and empirical datasets. These systems are represented as directed networks with non-negative weighted adjacency matrices, $\boldsymbol{W} = (W_{ij})$, where $W_{ij} \geq 0$ signifies the strength of the directed link $i \to j$, including self-edges $W_{ii} \geq 0$ and characterized by inherent asymmetry $\boldsymbol{W} \neq \boldsymbol{W}^\top$. By exploring four distinct metrics for network directedness, we seek to elucidate its effect on EPR, highlighting the emergence of hierarchical structuring within networks – an outcome deeply influenced by non-reciprocal interactions [52, 65].

A. Irreciprocity

As a measure of the directedness of a network, we first consider the (ir)reciprocity [66][67, 68]. The reciprocity quantifies how reciprocated pairwise relationships are on average across the network by comparing each connection $i \to j$ with $j \to i$. We first define the reciprocated portion of the pairwise relationship between two nodes,

$$\overleftrightarrow{W}_{ij} = \min(W_{ij}, W_{ji}) = \overleftrightarrow{W}_{ji}, \tag{1}$$

and the unreciprocated portion in each direction,

$$\overleftarrow{W}_{ij} = W_{ji} - \overleftrightarrow{W}_{ij} = \overrightarrow{W}_{ji}. \tag{2}$$

The reciprocity is then quantified by,

$$r(\boldsymbol{W}) = \frac{\sum_{i,j\neq i} \overleftrightarrow{W}_{ij}}{\sum_{i,j\neq i} W_{ij}}, \tag{3}$$

which is in the range $[0, 1]$ with $r = 0$ corresponding to a perfectly unreciprocated network, where edges can only run in a single direction, and $r = 1$ corresponding to an undirected network. We quantify the overall directedness of the network as $1 - r$, which we define to be the *irreciprocity*. Whilst the irreciprocity gives a measure of directedness for the network as a whole, we describe it as a 'local' measure, meaning it averages over each pairwise relationship in turn without analyzing the structure of the network 'globally'. As a result, this measure fails to distinguish between 'loop-like' structures and motifs that cause the network to globally follow a single direction. Directedness and non-zero irreciprocity are equivalent.

B. Trophic directedness

The question of whether a network globally follows a direction in its structure is another important notion of directedness. This idea is intimately linked to the idea of a hierarchy within the system where nodes can be organized into levels indicating their position in the top-down organization of the network. First being put forward in the field of ecology [49], in the context of food webs, *trophic (in)coherence* is a measure quantifying how neatly a network can be organized into so-called trophic levels [51].

Inspired by the Helmholtz-Hodge decomposition [69], the trophic incoherence of a network is given by,

$$F_0(\boldsymbol{W}, \boldsymbol{h}) = \min_{\boldsymbol{h}} \frac{\sum_{i,j} W_{ij}(h_j - h_i - 1)^2}{\sum_{i,j} W_{ij}}, \tag{4}$$

where $\boldsymbol{h} = (h_1, ..., h_N)$ is the vector with entries corresponding to the trophic levels for each of the N nodes, that minimizes the cost function [70]. The trophic levels are found as solution of the linear system

$$\boldsymbol{\Lambda h} = \boldsymbol{v}, \tag{5}$$

where $v_i = \sum_j W_{ji} - W_{ij}$ and $\boldsymbol{\Lambda} = \mathrm{diag}(\boldsymbol{u}) - \boldsymbol{W} - \boldsymbol{W}^\top$ defines the (symmetric) weighted graph Laplacian with $u_i = \sum_j W_{ji} + W_{ij}$ [51].

The nodes of any network can be partitioned into weakly connected components which are disjoint sets of nodes where node i belong to a component if there is a node j in the component with $\max(W_{ij}, W_{ji}) > 0$. Furthermore, the number of weakly connected components corresponds to the nullity of the Laplacian. Consequently, the system (5) has a non-unique solution corresponding to the nullity of the Laplacian. By enforcing that the lowest trophic level in each (weakly) connected component is equal to 0, one can obtain a unique solution to the equation (5) and calculate F_0. The trophic incoherence, F_0, is restricted to the range $[0, 1]$ with $F_0 = 1$ corresponding to completely non-hierarchical (including undirected) networks and $F_0 = 0$ corresponding to networks that can be perfectly

organized into trophic levels. We will consider *trophic directedness*, $\sqrt{1-F_0}$, as a measure of directedness in a network. Trophic directedness assumes the direction of the flow to be bottom-up meaning a low trophic level corresponds to the 'top' of the hierarchy if we were to consider a more intuitive top-down visualisation. By convention, throughout this paper, we display the trophic levels of a network by plotting nodes with the lower trophic levels at the top and inverting the y-axis.

C. Non-normality

A third notion of directedness is *network non-normality* [50]. A matrix $\boldsymbol{W}$ is said to be normal if it satisfies,

$$\boldsymbol{W}\boldsymbol{W}^\top = \boldsymbol{W}^\top\boldsymbol{W}. \tag{6}$$

It is, therefore, non-normal if $\boldsymbol{W}\boldsymbol{W}^\top \neq \boldsymbol{W}^\top\boldsymbol{W}$ [54]. Symmetric matrices are normal whilst non-normal matrices are necessarily asymmetric. Correspondingly, undirected networks are normal whilst non-normal networks are directed. The significance of developing a metric to quantify non-normality stems from the prevalence of non-normal matrices across a broad spectrum of applications, encompassing both linear systems and networks [54].

The eigenvectors of a non-normal matrix do not form an orthonormal basis or in other words such matrices are not diagonalisable by unitary matrices. As a result, the autonomous system of linear differential equations, $\dot{\boldsymbol{x}}(t) = \boldsymbol{W}\boldsymbol{x}(t)$ where $\boldsymbol{W}$ is non-normal, can undergo transient growth such that small perturbations can excite the system away, temporarily, from asymptotically stable equilibria. This feature has lead to the investigation of the role and uses of non-normal interactions in linear systems [50], neuronal [71] and communication networks [72] as well as in pattern formation [48], synchronization [43], resilience to perturbation [45] and network control of instabilities [73, 74].

A range of measures can be used to quantify the non-normality of a matrix [50]. Of particular interest are those derived from its spectrum, $\sigma(\boldsymbol{W})$, which governs its behavior as a linear system, and its pseudo-spectrum, $\sigma_\epsilon(\boldsymbol{W}) = \{\sigma(\boldsymbol{W}+\boldsymbol{E}) : ||\boldsymbol{E}|| < \epsilon\}$, which governs its response to perturbations [54]. The transient behavior of non-normal linear systems cannot be explained by the traditional spectral abscissa, $\alpha(\boldsymbol{W}) = \max \Re(\sigma(\boldsymbol{W}))$, which determines the asymptotic dynamics. Instead, measures such as the numerical abscissa, $\omega(\boldsymbol{W}) = \max \sigma(\frac{1}{2}(\boldsymbol{W}+\boldsymbol{W}^\top))$, capture the transient short-term growth of non-normal systems whilst pseudo-spectral measures such as the ϵ-pseudo-spectral abscissa, $\alpha_\epsilon(\boldsymbol{W}) = \max \Re(\sigma_\epsilon(\boldsymbol{W}))$, and the Kreiss constant, $\mathcal{K}(\boldsymbol{W}) = \max_{\epsilon>0} \frac{\alpha_\epsilon(\boldsymbol{W})}{\epsilon}$, capture their sensitivity to perturbation [54]. We quantify the degree to which a network breaks the normality condition using a common spectral measure, the *Henrici index*,

$$d_H(\boldsymbol{W}) = \sqrt{||\boldsymbol{W}||_F^2 - \sum_{i=1}^{N} |\lambda_i|^2}, \tag{7}$$

where $||\boldsymbol{W}||_F = \sqrt{\sum_i \sum_j W_{ij}^2}$ is the Frobenius norm and $\{\lambda_i\}$ is the set of eigenvalues [54]. The Henrici index is 0 when a matrix is normal and positive otherwise. To facilitate the comparison of various networks in terms of non-normality, irrespective of their size, the *normalized Henrici index* has been introduced [50],

$$\hat{d}_H(\boldsymbol{W}) = \frac{d_H(\boldsymbol{W})}{||\boldsymbol{W}||_F}, \tag{8}$$

which has values between 0 and 1. We opt for the Henrici index as it captures the spectral properties of $\boldsymbol{W}$, without directly assuming linear dynamics. In order to understand the nature of non-normality as a measure of hierarchical asymmetry, we consider an unweighted adjacency matrix $\boldsymbol{A}$. The entry $(\boldsymbol{A}^\top\boldsymbol{A})_{ij}$ represents the number of common sources connecting into nodes i and j whilst $(\boldsymbol{A}\boldsymbol{A}^\top)_{ij}$ represents the number of common targets from nodes i and j [51]. Therefore, in a normal unweighted network, for every given pair of nodes, the number of common sources and targets will coincide whereas in a non-normal unweighted network they will not coincide. In the case of weighted edges, non-normality captures the hierarchical asymmetry and the quantity $|\boldsymbol{W}\boldsymbol{W}^\top - \boldsymbol{W}^\top\boldsymbol{W}|$ is maximised is the case of a directed acyclic graph (DAG) where, after node relabelling, $\boldsymbol{W}$ is upper triangular yielding a clear net direction and nodal hierarchy [50].

Nevertheless, trophic directedness and non-normality are not equivalent and there are a range of networks which are trophically undirected yet non-normal and vice-versa (see Appendix A). Previous work has shown that trophic directedness and non-normality are closely, but non-linearly, correlated [41, 51]. Furthermore, non-normality does not capture all directed networks, as the set of asymmetric matrices contains, but is much larger than, the set of non-normal matrices. There is no complete characterisation of normal, asymmetric matrices but examples include circulant, block-circulant with circulant blocks, and skew-symmetric matrices. In Appendix A, we revisit the notion of trophic flatness ($F_0 = 1$) extending the concept to show its inequivalence to normality in weighted networks. We further contribute a new example of a trophically flat, non-normal network, uniquely without using a self-loop, thus broadening the discourse beyond previous studies.

D. Parameterising directedness in networks

To investigate the dynamical affects of directedness on network dynamics, it would be preferable to continuously vary the level of asymmetry in a network. To this aim, we first generate highly non-normal, hierarchical networks using a preferential attachment (PA) scheme with weak reciprocal links (see Appendix B for details) [50]. We then linearly interpolate between this strongly directed network and its undirected Hermitian,

$$\hat{\boldsymbol{W}}(\epsilon) = (1-\epsilon)\tilde{\boldsymbol{W}} + \epsilon \boldsymbol{W}, \tag{9}$$

for $\epsilon \in [0,1]$. Here, the Hermitian network is given by $\tilde{\boldsymbol{W}} = \frac{1}{2}\left(\boldsymbol{W} + \boldsymbol{W}^\top\right)$. By increasing the parameter ϵ, we are able to continuously increase the directedness of the network up to some maximal value. For a highly non-normal, hierarchical $\boldsymbol{W}$, Figure 2 shows the measures of directedness for the interpolated networks as a function of ϵ.

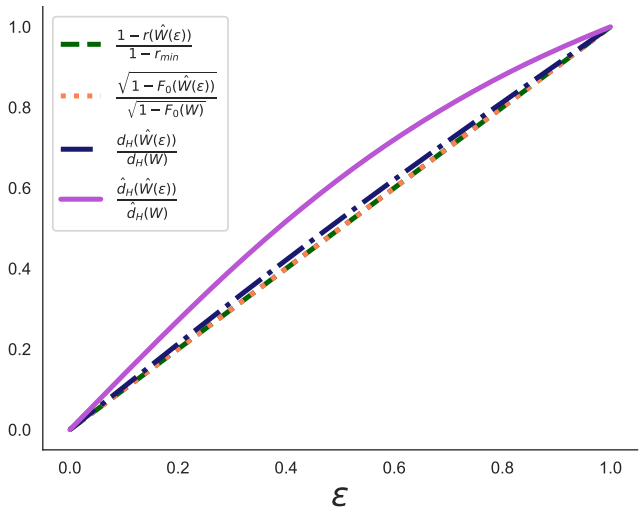

FIG. 2. **Parameterising directedness**: Beginning with a hierarchical, non-normal network $\boldsymbol{W}$, we can measure the directedness of the parameterised network as a function of ϵ, the interpolation parameter. Increasing ϵ increases the directedness almost linearly for each of the four measures. Here each measure is normalised by its maximum value which occurs at $\epsilon = 1$.

For the irreciprocity and trophic directedness, we can show that the interpolation parameterises the measure exactly linearly (for a proof see Appendix C) i.e.,

$$1 - r(\hat{\boldsymbol{W}}(\epsilon)) = \epsilon(1 - r(\boldsymbol{W})), \tag{10}$$

$$\sqrt{1 - F_0(\hat{\boldsymbol{W}}(\epsilon)} = \epsilon\sqrt{1 - F_0(\boldsymbol{W})}. \tag{11}$$

The eigenvalues of the matrix $\hat{\boldsymbol{W}}(\epsilon)$ are not calculable from $\tilde{\boldsymbol{W}}$ and $\boldsymbol{W}$ so the same cannot be done for the Henrici indices. However, the numerical experiments, shown for just one network in Figure 2, consistently show an increase with ϵ. Furthermore, we note that any weakly connected component in $\boldsymbol{W}$ becomes strongly connected in the interpolated network $\hat{\boldsymbol{W}}(\epsilon)$ for $\epsilon < 1$. Using the PA scheme detailed in Appendix B, the network is strongly connected, by construction.

III. ENTROPY PRODUCTION RATE OF DYNAMICS ON DIRECTED NETWORKS

The investigation of the effect of directedness in breaking detailed balance and increasing the EPR in networked systems, will be focused on four prototypical network dynamical processes: the random walk in discrete and continuous time [55, 75], the Orstein-Uhlenbeck [76] and the Ising model [58]. Each of the dynamics shares three important traits. First, they represent a broad spectrum of stochastic dynamics, with both continuous and symbolic variables as well as continuous and discrete time, that are well-studied and have found applications in many disparate fields. Second, under certain conditions, all the processes converge to equilibrium steady states on undirected networks and non-equilibrium, entropy-producing steady states on directed networks. Third, for each system, we can either explicitly calculate or numerically estimate the EPR with minimal sampling from the dynamics, alleviating the need for computational approaches.

A. Random walks on directed networks

We first consider the dynamics of non-interacting walkers randomly exploring the graph [7, 55, 75].

1. *Discrete-time random walk*

Initially, we consider a discrete-time random walk (DTRW) taking place on a weighted network with weight matrix $\boldsymbol{W}$. The DTRW defines an homogeneous discrete-time Markov chain over the discrete state space $\{1, ..., N\}$ of nodes. The state of the random walker, $x(t)$, at any time t is the node on which the walker is standing. At each time, the walker moves node according to the transition probability moving from node i to node j:

$$T_{ij} = \frac{W_{ij}}{\sum_j W_{ij}}. \tag{12}$$

Clearly, a walker can only move from one node to its neighbours and the probability that the walker moves to a particular node is proportional to the weight of the edge connecting their current position and the destination node. On undirected networks, the DTRW converges to an equilibrium steady state. On directed networks, instead,the graph must be strongly connected in

order for the existence of a single steady state. Assuming a population of random walkers that is originally distributed on the network with density $\boldsymbol{\pi}(0)$, then after t steps, the distribution on the network is given by,

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\boldsymbol{T}^t, \tag{13}$$

where $\boldsymbol{T} = (T_{ij})$. Therefore the steady-state satisfies

$$\boldsymbol{\pi} = \boldsymbol{\pi}\boldsymbol{T}, \tag{14}$$

namely, it is the Perron left eigenvector of the transition matrix $\boldsymbol{T}$ [7]. Consequently, the joint transition probability is given by,

$$P_{ij} = T_{ij}\pi_i, \tag{15}$$

where π_i is the steady-state probability of the walker being at node i.

The EPR for discrete Markovian dynamics is the Kullback-Leibler (KL) divergence between the forward and backward joint transition rates between all pairs of states [77]. For the RW, where nodes are states, it is given by,

$$\Phi = \sum_{i,j} P_{ij} \log \frac{P_{ij}}{P_{ji}}, \tag{16}$$

which we can now calculate explicitly given the exact steady-state probabilities. However, we note that not all states are achievable from other states. Given the existence of an edge $i \to j$ with no reciprocal link, $P_{ij} > P_{ji} = 0$, then (16) will diverge to infinity. Therefore on a graph containing at least one entirely unreciprocated edge, the EPR of the DTRW in a single time-step is, in fact, infinite. To solve this problem and to quantify the degree to which the detailed balance condition is violated under these constraints, we use the Jensen-Shannon (JS) divergence [78] instead of the KL-divergence to quantify the difference between forward and backward transition probabilities. For the DTRW, we define our JS-divergence EPR to be:

$$\Phi = \sum_{i \to j} P_{ij} \log \frac{P_{ij}}{\tilde{P}_{ij}} + P_{ji} \log \frac{P_{ji}}{\tilde{P}_{ij}}, \tag{17}$$

where the sum is over all directed edges and $\tilde{\boldsymbol{P}} = \frac{1}{2}(\boldsymbol{P} + \boldsymbol{P}^\top)$ is the Hermitian, or averaged, distribution. We can directly verify that this quantity does not diverge as $\tilde{P}_{ij}$ does not vanish on connected nodes. Importantly, unlike in the other models, directedness is not a sufficient condition to guarantee that the steady state is non-equilibrium and entropy producing, as will be shown later on.

2. *Continuous-time random walk*

In addition to the DTRW, we also consider the continuous time random walk (CTRW) in which walkers transition between nodes at event times specified by a Poisson renewal process [75]. In continuous-time, the transition probabilities between nodes become time-dependent and are given by,

$$P_t(j,i) = (e^{t\boldsymbol{L}})_{ij}, \tag{18}$$

where $P_t(j,i)$ is the probability of a walker being at node j at time t given they were at node i at time zero and $\boldsymbol{L} = \boldsymbol{T} - \boldsymbol{I}$ is the *random walk graph Laplacian*. The steady-state probabilities $\boldsymbol{\pi}$ coincide with the DTRW and are given by the Perron left-eigenvector of $\boldsymbol{T}$.

The time-dependent EPR of a continuous time Markov process is given by,

$$\Phi(t) = \frac{1}{2}\sum_{i,j}(P_t(i,j)\pi_{t,j} - P_t(j,i)\pi_{t,i}) \log\left(\frac{P_t(i,j)\pi_{t,j}}{P_t(j,i)\pi_{t,i}}\right), \tag{19}$$

where $\pi_{t,i}$ is the time-dependent probability of being in state i [79]. Assuming that the system is in the steady state, the EPR of the CTRW is given by,

$$\Phi(t) = \frac{1}{2}\sum_{i,j}(e^{tL_{ji}}\pi_j - e^{tL_{ij}}\pi_i) \log\left(\frac{e^{tL_{ji}}\pi_j}{e^{tL_{ij}}\pi_i}\right). \tag{20}$$

The EPR, $\Phi(t)$, varies in time, and so we compare the EPR, $\Phi(T)$, of different networks at some chosen final time $T > 0$.

B. The network-based multivariate Orstein-Uhlenbeck process

The Ornstein-Uhlenbeck (OU) process in one dimension is a linear stochastic dynamical system modelling a particle in Brownian motion [80]. It can be extended to the multivariate case which models a number of interacting particles [57]. In addition, it can be re-cast as a network dynamical system [76], with interactions being constrained to the weighted edges of a network, as will be considered here.

Consider a system of N interacting particles, the multivariate OU process is given by the Langevin system,

$$\frac{d\boldsymbol{x}}{dt} = -\boldsymbol{B}\boldsymbol{x}(t) + \boldsymbol{\xi}(t), \tag{21}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^N$ is the time-dependent state vector, $\boldsymbol{B} \in \mathbb{R}^{N\times N}$ is the **friction matrix**, and $\boldsymbol{\xi}(t) \in \mathbb{R}^N$ is additive white noise with covariance given by,

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^\top(t')\rangle = 2\boldsymbol{D}\delta(t-t'), \tag{22}$$

where $\boldsymbol{D} \in \mathbb{R}^{N\times N}$ is the *noise covariance matrix* which is, by definition, symmetric.

Given a weighted network, $\boldsymbol{W}$, we can constrain the interactions such that they occur along the edges of the network. In addition, we assume the additive noise to be applied independently to each node in the network. Under these assumptions, the OU is given by the Langevin system,

$$\frac{d\boldsymbol{x}}{dt} = \Theta(\gamma \boldsymbol{W} - \boldsymbol{I})\boldsymbol{x}(t) + \boldsymbol{\nu}(t), \tag{23}$$

where $\boldsymbol{I}$ is the identity matrix, Θ is the reversion rate and γ is the coupling parameter [76]. As the thermal fluctuations are assumed to be uncorrelated in time and between nodes, the additive noise satisfies,

$$\langle \boldsymbol{\nu}(t)\boldsymbol{\nu}^\top(t')\rangle = 2\sigma \boldsymbol{I}\delta(t-t'), \tag{24}$$

where σ is the noise intensity. We relate the networked system to the generalised OU process with the following relations,

$$\boldsymbol{B} := \Theta(\boldsymbol{I} - \gamma \boldsymbol{W}), \tag{25}$$

$$\boldsymbol{D} := \sigma \boldsymbol{I}. \tag{26}$$

Returning to the generalised multivariate case, we will derive the EPR rate of the OU in a steady state (see Appendix D and Ref. [57] for further details). If each eigenvalue of the friction matrix, $\boldsymbol{B}$, has positive real part, then, in the absense of noise, the system relaxes exponentially fast to $\boldsymbol{x} = 0$. Therefore, in the presence of noise, the process will relax to a steady state with Gaussian fluctuations. Generally, this is a non-equilibrium steady state with the EPR being dependent on the matrices $\boldsymbol{B}$ and $\boldsymbol{D}$. The steady state is Gaussian with mean $\boldsymbol{x} = \boldsymbol{0}$ and steady state covariance given by,

$$\boldsymbol{S} = \lim_{t\to\infty} \langle \boldsymbol{x}(t)\boldsymbol{x}^\top(t)\rangle. \tag{27}$$

It can be shown that $\boldsymbol{S}$ satisfies the following Sylvester equation [81] (see Appendix D),

$$\boldsymbol{B}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{B}^\top = 2\boldsymbol{D}. \tag{28}$$

The condition for the steady state to be reversible and in equilibrium is known to be [82],

$$\boldsymbol{B}\boldsymbol{D} = \boldsymbol{D}\boldsymbol{B}^\top, \tag{29}$$

where the covariance is given by,

$$\boldsymbol{S} = \boldsymbol{B}^{-1}\boldsymbol{D}. \tag{30}$$

In the case of the networked system, the reversibility condition becomes,

$$\boldsymbol{W} = \boldsymbol{W}^\top, \tag{31}$$

which corresponds to the interaction network being undirected. When condition (29) is not satisfied, the EPR, Φ, can be written in the form,

$$\Phi = -\operatorname{Tr}(\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{Q}). \tag{32}$$

where $\boldsymbol{Q} = \boldsymbol{B}\boldsymbol{S} - \boldsymbol{D}$. Therefore, given a matrix pair $\boldsymbol{B}, \boldsymbol{D}$ - or correspondingly, a network, $\boldsymbol{W}$, and noise intensity, σ - one has only to numerically solve the Sylvester equation (28) in order to calculate the EPR Φ. We note that a range of numerical techniques and linear algebra packages exist for the accurate and efficient solution of equations of this type [83].

C. The Ising model

The final stochastic dynamical system we consider is the ubiquitous *Ising spin-glass model* [58]. The Ising model with N nodes, is made up of N discrete spins that can take values in $\{+1, -1\}$, 'up' and 'down' spins respectively. We consider the system, in the absence of external fields, evolving in discrete time with either sequential or parallel spin updates. Given the state of the system, a spin configuration, $\boldsymbol{x}(t) = (x_1(t), ..., x_N(t))$, the spins at time $t+1$ are updated as a discrete time Markov chain,

$$P(\boldsymbol{x}(t+1)|\boldsymbol{x}(t)) = \prod_i \frac{\exp\left(x_i(t+1)\sum_j W_{ij}x_j(t)/T\right)}{2\cosh\left(\sum_j W_{ij}x_j(t)/T\right)}, \tag{33}$$

where T is the thermodynamic temperature and $\boldsymbol{W} = (W_{ij})$ are the pairwise coupling strengths defined by a weighted network. In the absence of external fields, the Ising model has reversible dynamics and is in equilibrium when the coupling strengths are symmetric i.e. $W_{ij} = W_{ji}$, corresponding to an undirected network, and irreversible dynamics when the coupling strengths are asymmetric, corresponding to a directed network [84]. The joint transition probability between two spin configurations, or states, $\boldsymbol{y}, \boldsymbol{z}$, is given by,

$$P_{\boldsymbol{y}\boldsymbol{z}} = P(\boldsymbol{x}(t+1) = \boldsymbol{z}, \boldsymbol{x}(t) = \boldsymbol{y}). \tag{34}$$

We recall the EPR for Markovian dynamics, given by,

$$\Phi = \sum_{\boldsymbol{y},\boldsymbol{z}} P_{\boldsymbol{y}\boldsymbol{z}} \log \frac{P_{\boldsymbol{y}\boldsymbol{z}}}{P_{\boldsymbol{z}\boldsymbol{y}}}. \tag{35}$$

Note that $\boldsymbol{z}, \boldsymbol{y}$ are spin configurations, not nodes as in the RW. We can factorise the joint transition probabilities as follows,

$$P_{\boldsymbol{y}\boldsymbol{z}} = P_{\boldsymbol{z}|\boldsymbol{y}} \cdot \pi_{\boldsymbol{y}}, \tag{36}$$

where $P_{\boldsymbol{z}|\boldsymbol{y}} = P(\boldsymbol{x}(t+1) = \boldsymbol{z}|\boldsymbol{x}(t) = \boldsymbol{y})$ is the conditional transition probability and $\pi_{\boldsymbol{y}}$ is the steady state probability of being in state $\boldsymbol{y}$. In order to estimate the EPR, we can estimate the steady state probabilities $\pi_{\boldsymbol{y}}$ using numerical sampling and use the conditional transition probability given by equation (33) to calculate Φ using (35). However, we note that there are 2^N possible configurations for a system with N spins. Therefore, as the system gets large, estimating the steady state probability becomes computationally challenging. For this

reason, our analysis of the Ising model bisects into the case of small networks ($N \leq 10$), where we can employ this approach, and the case of large networks ($N > 10$), where we require a mean-field approximation [85].

1. Small networks of Ising spins

As mentioned above, for small networks of spins ($N \leq 10$) we estimate the steady state probabilities by sampling from the model using Glauber dynamics [86], a sequential spin update rule given by,

$$P[x_i(t+1) = 1|\boldsymbol{x}(t)] = [1 + \exp(-\frac{2}{T}\sum_j W_{ij}x_j(t))]^{-1}. \tag{37}$$

We note that Glauber dynamics simulate the Ising model with sequential, not parallel updates as described in equation (33), but that these converge to the same steady state dynamics over time. After a sufficient burn-in period, we count the number of occurrences of each of the 2^N states and divide that by the total number of time-steps to given an estimate of the steady state probabilities, which we then use to estimate the EPR.

2. Large networks of Ising spins

The combinatorial explosion of state space with the number of spins, means that estimating the steady state probabilities is computationally challenging for large systems. Alternative approaches include coarse-graining the state space by clustering together states and estimating transition probabilities [25, 87, 88], or using mean-field approximations [85]. Whilst using a coarse-grained state-space has been shown to capture the affect on the EPR of changing the temperature of the Ising model [25], it proved to be an inaccurate when varying the directedness of the network (see Appendix G) and so we hereby opt for a mean-field approach.

In order to approximate the EPR of the system, we focus on two key statistical properties of the system,

$$m_i(t) = \sum_{\boldsymbol{x}(t)} x_i(t)P(\boldsymbol{x}(t)), \tag{38}$$

$$D_{il}(t) = \sum_{\boldsymbol{x}(t),\boldsymbol{x}(t-1)} \left[x_i(t)x_l(t-1)P(\boldsymbol{x}(t),\boldsymbol{x}(t-1)) - m_i(t)m_l(t-1)\right]. \tag{39}$$

where $\boldsymbol{m}(t) = (m_1(t), ..., m_N(t))$ is the average activation rate of the system and $\boldsymbol{D}(t)$ is the delayed correlation matrix. These are sufficient statistics to define a particular Ising model [85].

Given a particular network, $\boldsymbol{W}$, and the time-delayed correlations, $\boldsymbol{D}(t)$, we can calculate the time-dependent EPR, given by,

$$\Phi(t) = \sum_{i,j}(W_{ij} - W_{ji})D_{ij,t}. \tag{40}$$

We note that $\Phi(t) = 0$ if the network is undirected i.e. $W_{ij} = W_{ji}$. We will estimate the average activation rate and delayed correlations using the so-called **classical naive mean field** (NMF) given by,

$$m_i(t) \approx \tanh \sum_j W_{ij}m_j(t-1), \tag{41}$$

$$D_{il}(t) \approx W_{il}(1 - m_i^2(t))(1 - m_l^2(t-1)). \tag{42}$$

For a derivation see Appendix H. Beginning with all spins set to 1, for a given network, we can use the NMF to approximate the time-delayed correlations and use equation (40) to estimate the time-dependent EPR, which we iteratively perform until $\Phi(t)$ converges to a time-independent value Φ.

IV. ENTROPY PRODUCTION IN SYNTHETIC HIERARCHICAL NETWORKS

A. An exactly treatable case: 2-node networks

Before considering large directed networks, we first consider a simple 2-node network with asymmetric coupling, as shown in Figure 3.

The 2-node directed network is defined by the weight matrix,

$$\boldsymbol{W} = \begin{bmatrix} 0 & W_{12} \\ W_{21} & 0 \end{bmatrix}. \tag{43}$$

This network has only two free parameters and so we can explore the space of directed networks and see how the directedness and EPR change. Firstly, the irreciprocity

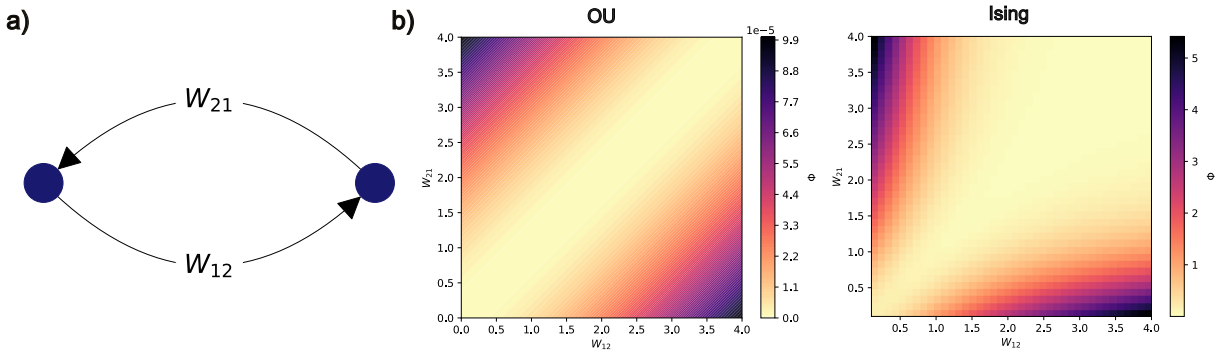

FIG. 3. **Entropy production in the 2-node network**: *a*) We first consider a minimal directed network with only 2 nodes, where we can explore the entire space of networks by varying the interaction strengths. *b*) **Left:** The exact EPR of the OU as given in equation (54) for different 2-node networks. The positive diagonal represents undirected networks which do not produce entropy. Further from this line, the EPR is higher. **Right:** The approximated EPR of the Ising model also increases as the network becomes more directed, but with a non-linear relationship.

of this network is given by,

$$1 - r = \frac{|W_{12} - W_{21}|}{W_{12} + W_{21}}. \tag{44}$$

The trophic levels, $\boldsymbol{h}$, are given by the solution to the equation,

$$\begin{bmatrix} W_{12} + W_{21} & 0 \\ 0 & W_{12} + W_{21} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} W_{21} - W_{12} \\ W_{12} - W_{21} \end{bmatrix}, \tag{45}$$

which has general solution,

$$h_1 - h_2 = \frac{W_{21} - W_{12}}{W_{21} + W_{12}}, \tag{46}$$

and trophic directedness,

$$\sqrt{1 - F_0} = \frac{|W_{12} - W_{21}|}{W_{12} + W_{21}}, \tag{47}$$

which coincides with the irreciprocity. The Henrici and normalised Henrici indices are given by,

$$d_H = |W_{12} - W_{21}|, \tag{48}$$

$$\hat{d}_H = \frac{|W_{12} - W_{21}|}{\sqrt{W_{12}^2 + W_{21}^2}}. \tag{49}$$

Each of the measures somehow captures the asymmetry between pair of weights, thereby quantifying the directedness, but with unique nuances. The irreciprocity and trophic directedness are normalised by the l_1 norm of the matrix whilst the normalised Henrici index uses the l_2 norm and the Henrici index is unnormalised. Next we consider the dynamics on this network, beginning with the RW. The RW does not converge to a steady state on the 2-node network unless the walkers are originally distributed with the steady state distribution. This is due to the fact that the network is bipartite meaning at each time-step a random walker can only move to the other node. The transition matrix is given by,

$$\boldsymbol{T} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{50}$$

Therefore, $\boldsymbol{\pi} = \left(\frac{1}{2}, \frac{1}{2}\right)$ is the Perron left eigenvector and the steady-state. Yet, despite the directedness of the network, this steady state is clearly preserves detailed balance as $P_{12} = P_{21} = \frac{1}{2}$. Next we consider the OU process. We note that the eigenvalues of the associated friction matrix,

$$\boldsymbol{B} = \Theta(\boldsymbol{I} - \gamma \boldsymbol{W}), \tag{51}$$

are given by the pair,

$$\lambda_{\pm} = \Theta \pm \Theta\gamma\sqrt{W_{12}W_{21}}. \tag{52}$$

Therefore, to guarantee decay to a steady state, we enforce the condition,

$$\gamma < \frac{1}{\sqrt{W_{21}W_{12}}}. \tag{53}$$

Handily, the Sylvester equation (D8) is explicitly solvable for systems of size $N = 2$ and so the EPR has a closed form solution [57]. For the 2-node network, the EPR becomes,

$$\Phi = \frac{\Theta\gamma^2}{2}(W_{12} - W_{21})^2, \tag{54}$$

a function of the difference, $|W_{12} - W_{21}|$, between the weights of the pair of edges. Importantly, this means the EPR increases quadratically with the (unnormalised) Henrici index, a relationship that we will show to hold consistently across experiments for the OU. Furthermore, the EPR is not normalised with respect to the total strength of the network. Trivially, when the connection is symmetrical, $W_{12} = W_{21}$, the system preserves detailed balance.

Finally, we can consider the Ising dynamics on this network. Whilst we do not have the exact solution, there are only four possible spin configurations in this system, therefore we can accurately estimate the steady

state probabilities and consequently the EPR. The EPR, Φ, for both the OU and Ising systems is plotted as a function of the network weights (W_{12}, W_{21}) as shown in Figure 3. The positive diagonal $W_{12} = W_{21}$ represents the subset of undirected networks, which preserve detailed balance. In both cases, as the asymmetry of the network increases the EPR increases. The EPR in the OU is a function purely of the distance between the two weights, whereas for the Ising model, the relationship appears to be nonlinear.

B. Entropy production in parameterised synthetic networks

Equipped with a range of stochastic network dynamics where we can estimate the EPR, as well as a mechanism to systematically vary the directedness of the network, we can investigate the affect of directedness on the degree of non-equilibrium. Considering networks of sizes $N = 10$, 100, 500 and 1000, we generate hierarchical networks using preferential attachment (see Appendix B) and then interpolate between these networks and their Hermitian as described in Section II. Recall the interpolation is given by,

$$\hat{\boldsymbol{W}}(\epsilon) = (1-\epsilon)\tilde{\boldsymbol{W}} + \epsilon \boldsymbol{W}, \tag{55}$$

for $\epsilon \in [0, 1]$, with $\epsilon = 0$ corresponding to an undirected network and $\epsilon = 1$ being maximally directed. Recall further, that Figure 2 showed that all four measures of asymmetry scaled linearly or almost linearly with the parameter ϵ. Therefore we can consider ϵ to be a normalised (relative) measure of directedness. Figure 4 shows the results of varying ϵ for all four dynamics and for different network sizes. Clearly, in each case, the directedness ϵ causes a non-linear increase in the relative EPR that remains consistent as the network size increases. Furthermore, there are clear similarities between dynamics. Firstly, we consider the two leftmost panels, corresponding to the DT/CTRW dynamics. We see that the normalised EPR $\Phi(\epsilon)/\Phi_{\max}$ increases quadratically in ϵ. In the third panel, we see an identical quadratic increase for the OU process. In fact, this quadratic increase in ϵ can be proven to be exact in the special cases of 2-node and circulant networks (see Appendices E and F). In the final panel, we show the results for the NMF Ising model. The EPR also increases with ϵ, but this increase is faster than quadratic (note the $y-$axis is logarithmic) and approaches an exponential increase as $N \to \infty$. Across both dynamics and system size, directedness causes a violation of detailed balance and drives the EPR of the system.

V. ENTROPY PRODUCTION AND DIRECTEDNESS IN ERDŐS–RÉNYI GRAPHS

In the previous section, we used the interpolation to vary the directedness of a given network. We were able to show that the EPR scales, non-linearly, with the relative directedness of the network for all four considered dynamics. However, as all the directedness measures increased (almost) linearly in ϵ, it is unclear which 'type' of directedness is actually driving the violation of detailed balance and the EPR. In an effort to decouple the different measures of directedness, we consider a different approach to generating networks. Instead of specifically generating hierarchical, directed networks and interpolating, as before, we now utilise directed Erdős–Rényi graphs as a null network model [59]. A directed Erdős–Rényi (ER) graph, $G(N, p)$, is a randomly sampled, unweighted network with N nodes. Each directed pair of nodes $i \to j$ is connected with independent probability $p \in [0, 1]$. The ER generating process does not assume or enforce a hierarchical structure, however hierarchical directedness can emerge spontaneously. For each value of p, we randomly sample ER graphs and measure their directedness with each measure. In addition, we measure the EPR of each of the four dynamics on these networks.

Figure 5 shows the behaviour of the directedness measures as a function of the ER parameter p, normalised by the maximum value across all trials, (left pane) as well as the behaviour of the EPR for each of the four systems as a function of the ER parameter p (right pane), also normalised by the maximum value. For these calculations, we consider $N = 100$ and sample 100 graphs at each value of p. The figure shows the mean, with standard deviation shading, over the 100 samples. Firstly, we note the different profiles of the directedness measures. The unique behaviour of each measure shows we have been able to decouple the different notions of directedness, allowing us to identify which one is closely correlated with the EPR. The irreciprocity, trophic directedness and normalised Henrici index all decay as a function of p but with different decay rates. On the other hand, the Henrici index has a parabolic shape, peaking at $p = 0.5$. We can compare this to the different profiles of each of the dynamics in the right pane. The DTRW curve decays linearly in p mirroring the curve of the irreciprocity, whilst CTRW curve decays exponentially mirroring the trophic directedness. We note that $\Phi_{\text{DTRW}}, \Phi_{\text{CTRW}}$ are only defined when the network is strongly connected and so are only calculated for samples generated with $p > 0.1$ where we checked that the sampled graph was strongly connected. Both the OU and Ising curves follow the parabolic shape of the Henrici index, peaking at $p = 0.5$. The Ising curve show far greater variance at each value of p suggesting that the EPR is very susceptible to small structural changes in the network or that the system is operating

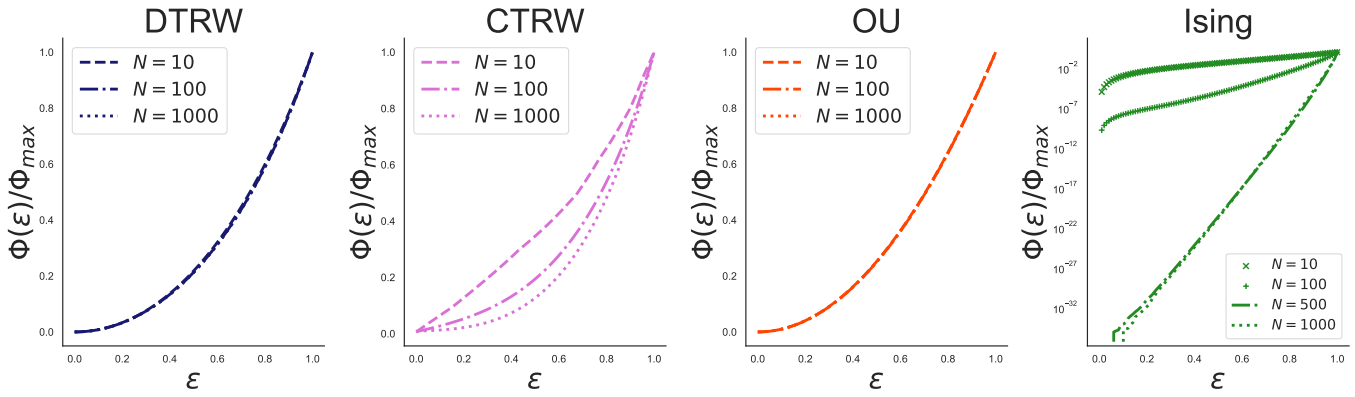

FIG. 4. **Entropy production rate in interpolated synthetic networks**: Normalised EPR, $\Phi(\epsilon)/\Phi_{\max}$, in the DT/CTRW, OU and Ising systems as a function of the directedness parameter ϵ. **Columns 1 & 2:** The first two panes show the results for the DT/CTRW. Across network sizes, $\Phi(\epsilon)/\Phi_{\max}$ scales quadratically in ϵ. **Column 3:** The third column shows the results for the OU process. The normalised EPR, $\Phi(\epsilon)/\Phi_{\max}$, increases quadratically in ϵ, as found analytically in the 2-node network. **Column 4:** The final column shows the result for the NMF Ising approximation. The normalised EPR, $\Phi(\epsilon)/\Phi_{\max}$, increases faster than quadratically and approaches an exponential increase as $N \to \infty$. Note that the $y-$axis is logarithmic from 0 to 1 for this pane.

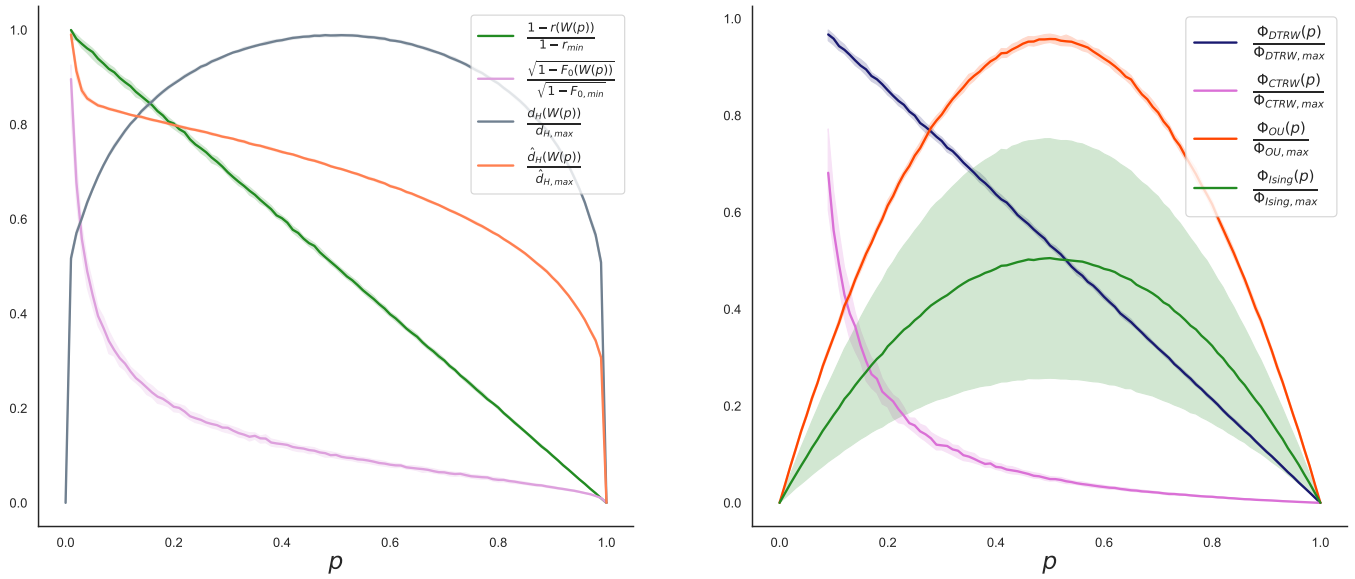

FIG. 5. **Directedness and entropy production rate in Erdős–Rényi graphs**: **Left:** This plot shows the mean, with standard deviation shading, of normalised directedness measures as a function of the ER parameter p for $N = 100$ over 100 samples. We can see that the irreciprocity, trophic directedness and normalised Henrici index decrease as a function of p, with different decay rates. On the other hand, the Henrici index has a parabolic shape and peaks at $p = 0.5$. **Right:** This plot shows the mean, with standard deviation shading, of the normalised EPR for the four systems as a function of the ER parameter p for $N = 100$ over 100 samples. The DTRW curve follows the irreciprocity. The CTRW curve follows the trophic directedness. Both the OU and Ising curves peak at $p = 0.5$, closely following the Henrici index of the network. The Ising curve shows very large variance at each value of p suggesting that small structural differences lead to large differences in Φ potentially being close to the critical temperature.

close to the critical temperature, where the EPR is maximised [29, 89].

The DTRW is a process that only sees the network locally. Therefore, in a single time-step, the dynamics, and by extension the EPR, depend only on the directedness of individual edges in or out of the walker's momentary position. As a result, the EPR is strongly correlated with, or driven by, the irreciprocity, which strictly measures differences between the forward and backward connection strength. Furthermore, DTRW dynamics are invariant to scaling of the network weights

i.e. the dynamics of the process evolving on $\boldsymbol{W}$ is the same as that of $\alpha\boldsymbol{W}$ for $\alpha > 0$. Therefore, the fact that the irreciprocity is a normalised measure also explains the strong correlation. Whilst the CTRW still sees the network locally, the extension to continuous time means that the probability of transitioning (albeit not directly) between any pair of nodes, in a finite time t, is non-zero. As a result, the global directed structure of the network plays a role in the dynamics and the EPR. The CTRW is also invariant to scaling of the network. This explains why the trophic directedness, a normalised but global measure of directedness, is strongly correlated with the EPR of the CTRW.

On the other hand, the OU and Ising dynamics, are complex interacting systems where, at any given time-point, the entire network structure is influencing the dynamics. Furthermore, they are not invariant to scaling of the network. As a result, the unnormalised Henrici's index, measuring non-normality, appears to be most strongly correlated with the EPR of the systems, indicating that the global hierarchy, as well as the magnitude of the weights, of the network are responsible for driving the interacting dynamics from equilibrium.

VI. ENTROPY PRODUCTION AND DIRECTEDNESS IN REAL WORLD NETWORKS

In this section, we consider a dataset of 97 real-world directed networks from ecology, sociology, biology, language, transport and economics [52]. For a full description of the dataset see Appendix J. We measure the directedness of these networks with the four measures and then consider the four dynamical systems evolving on these networks and measure the EPR. Furthermore, we differentiate between networks from different domains, and plot the correlations between the directedness measures and the EPR. For real-world networks, the irreciprocity, trophic directedness and non-normality are all strongly positively correlated (see Appendix A and Refs. [41, 50, 51]). In addition, most directedness is trophically or hierarchically organised, indicating a lack of 'loop-like' structures, that are asymmetric but not hierarchical, in real-world networks [90, 91]. Next, by considering the four dynamical systems of interest evolving on the real world networks, we measure the EPR and plot the correlations with the different measures of directedness.

For both the DT/CTRW dynamics, we restrict to the 9 strongly connected networks to guarantee the existence of a steady state. Figure 6 shows the EPR of each system correlated with each of the four directedness measures. The first column shows the results for the DTRW. There is a strong correlation between the normalised directedness measures and the EPR and poor correlation with the Henrici index. This supports the conclusions from Section V that indicated that normalised directedness was a better determinant of the EPR than unnormalised measures when the system is invariant to scaling the network. Due to the strong correlations between the three normalised measures (see Appendix A, Fig. 10), the hierarchical measures of trophic directedness and normalised Henrici also show strong correlations. However, the analysis in Section V suggests that it is, in fact, the irreciprocity that is driving this relationship. In the second column of Figure 6, we can see that the EPR of the CTRW is only weakly correlated with the directedness measures indicating that directedness alone is not sufficient to determine the distance of this system from equilibrium. This suggests that the thermodynamics of the CTRW are more dependent on the network substructure than the other systems. In previous experiments, using the interpolation or the ER null model, network substructure and size was not varied with the complex heterogeneity seen in real-world network data from diverse fields. The third column clearly shows that the Henrici index almost perfectly correlates with the EPR of the OU process indicating that the symmetry breaking in a non-normal network directly drives broken detailed balance in the OU model. This further supports our findings in synthetic networks in Sections IV & V. Finally the rightmost column shows the results for the Ising model. The Ising model is correlated strongly with the normalised measures of directedness. Whilst the previous section suggested that the Henrici index is also strongly correlated to the EPR in the Ising model, the high variance in the EPR also indicated that the Ising model could be susceptible to small changes in network structure, or proximity to criticality, which could explain noisier relationships in the real-world networks.

The analysis on real-world networks indicates that directedness again drives broken detailed balance in the dynamics. However, it distinguishes between systems like the OU process where the absolute non-normality, regardless of network sub-structure, dictates the distance from equilibrium, and systems like the CTRW, where the low correlations indicate that directedness alone is not enough to predict the EPR.

VII. ENTROPY PRODUCTION AND DIRECTED NETWORKS FROM MULTIVARIATE TIME-SERIES

In the previous section, we considered a plethora of real-world structural network data from a diverse range of fields and assumed the dynamics evolving on the networks. However, in many complex systems, such (directed) structural network data is either unavailable or of less interest than empirical spatiotemporal data of node activities in the form of a multivariate time-series (MVTS). Extracting information about the structural

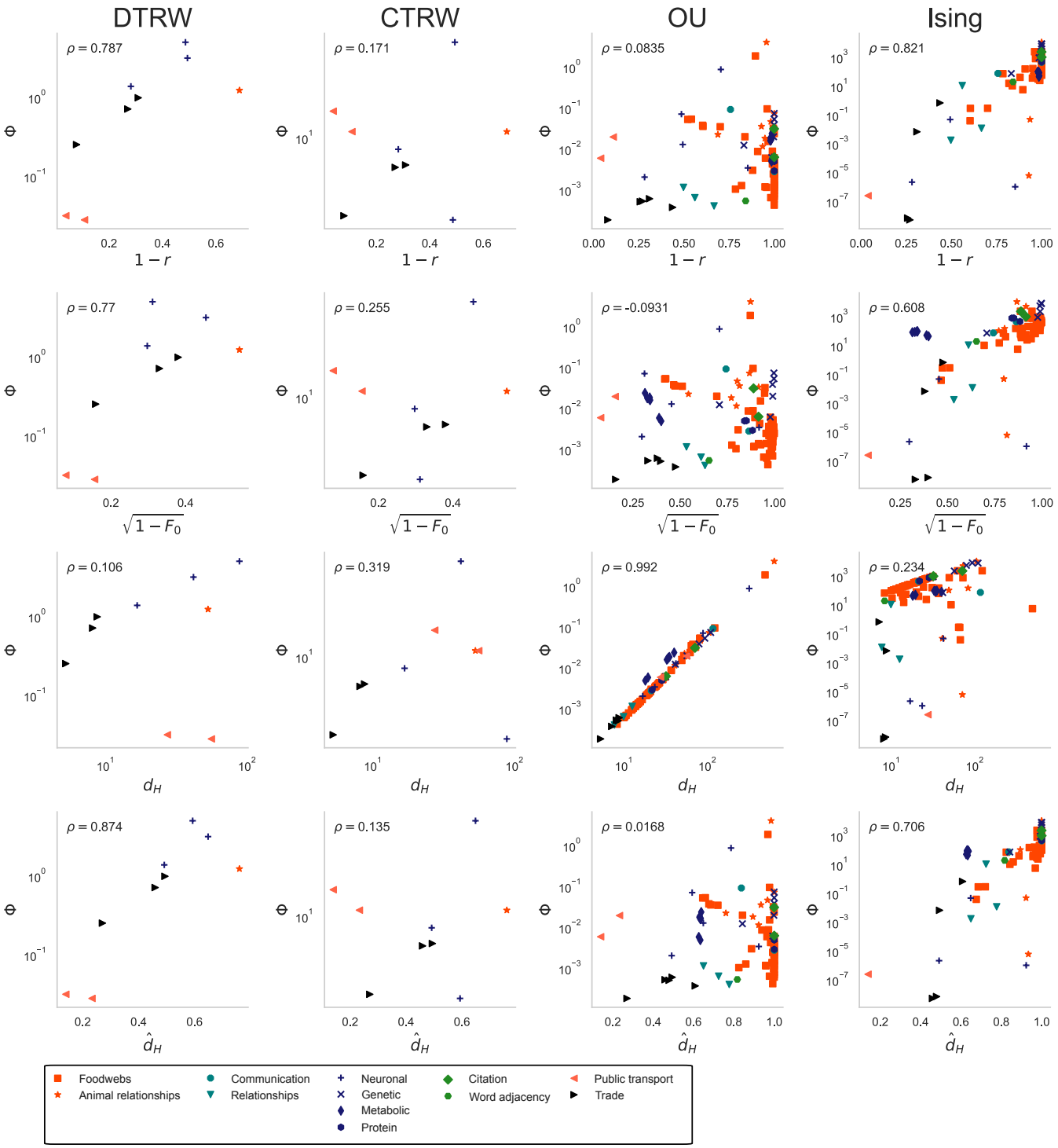

FIG. 6. **Entropy production rate and directedness in real-world networks**: We plot the EPR of four dynamical systems against four directedness measures on real world networks from twelve fields. The first column shows the EPR of the DTRW and the second the EPR of the CTRW. These are restricted to nine strongly connected networks. The third column shows the EPR of the OU and the fourth the EPR of the Ising model. We can that the normalised directedness measures are strongly correlated with the DTRW and Ising model. The Henrici index is extremely strongly correlated with the EPR of the OU process. None of the measures are correlated with the EPR of the CTRW. Ecological networks are in orange, social in teal, biological in blue, language in green, transport in red and economic in black.

or dynamical organisation of a system from the complex patterns in a MVTS is a ubiquitous problem in complex systems science [92, 93]. A range of computational methods exist to quantify broken detailed balance in a MVTS [25, 61, 94–96]. Concurrently, a range of methods for inferring directed network structure from data have been developed [93, 97–103]. In this section, we will present a simple, intuitive method for both estimating the EPR and inferring a directed network structure from a MVTS using linear auto-regression [60]. We fit a first-order multivariate autoregressive model and then associate this model to a corresponding OU process where the EPR can be calculated explicitly [57]. Furthermore, this fitting will infer a network of interactions between the variables, under the assumed model, whose directedness can be measured. In addition to calculating the overall directedness, we can extract the trophic levels as described in Section II in order to unravel the hierarchical organisation of the system. We will apply this approach to MVTS from human neuroimaging data from the Human Connectome Project (HCP) and stock-prices from the New York Stock Exchange (NYSE) to investigate the relation between broken detailed balance and structural directedness in dynamic data.

Whilst the assumption of a linear model may seem presumptuous, previous studies have found that even highly non-linear systems, such as large-scale brain dynamics, are well, or even best, described by linear models [104]. Furthermore, autoregressive models have previously found success in a spectrum of areas including finance and economics [60], neuroscience [105] and beyond. Finally we note that three recent studies in neuroimaging have considered related approaches, fitting linear models to neural recordings in order to quantify the EPR [106] or the asymmetry of interactions [62, 107].

A. Linear auto-regression of multivariate time-series

Consider an $N-$dimensional MVTS of signals of the form $\boldsymbol{X}(t_i) = \{X_1(t_i), ..., X_N(t_i)\}$ recorded at equispaced time-points $t_i \in \{t_0, t_1, ..., t_T\}$. We assume such signals are discrete, finite observations of either a generalised or network-based OU process,

$$\frac{d\boldsymbol{x}}{dt} = -\boldsymbol{B}\boldsymbol{x}(t) + \boldsymbol{\xi}(t), \tag{56}$$

$$\frac{d\boldsymbol{x}}{dt} = (\boldsymbol{W} - \boldsymbol{I})\boldsymbol{x}(t) + \boldsymbol{\nu}(t), \tag{57}$$

with additive noise satisfying,

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^\top(t')\rangle = 2\boldsymbol{D}\delta(t-t'), \tag{58}$$

$$\langle \boldsymbol{\nu}(t)\boldsymbol{\nu}^\top(t')\rangle = 2\sigma\boldsymbol{I}\delta(t-t'), \tag{59}$$

respectively, as previously defined in Section III.

In order to find the parameters, $\boldsymbol{B}, \boldsymbol{D}$ or $\boldsymbol{W}, \sigma$, that best explain our observed data, we fit a first-order linear multivariate auto-regressive (MAR) model of the form,

$$\boldsymbol{X}(t_{i+1}) = \boldsymbol{A}\boldsymbol{X}(t_i) + \boldsymbol{\chi}(t_i), \tag{60}$$

where $\boldsymbol{A}$ is calculated using least-squares auto-regression and $\boldsymbol{\chi}$ is a MVTS of **residuals**. In order to associate the MAR to an OU process, we discretise the continuous-time process with a one-step scheme. Whilst any such discretisation can be applied, we proceed with a Euler-Maruyama discretisation [108], with time-step Δt and obtain,

$$\boldsymbol{x}(t_{i+1}) = [\boldsymbol{I} - \Delta t\boldsymbol{B}]\boldsymbol{x}(t_i) + \boldsymbol{\Lambda}\boldsymbol{\eta}_i, \tag{61}$$

$$\boldsymbol{x}(t_{i+1}) = [\boldsymbol{I} + \Delta t(\boldsymbol{W} - \boldsymbol{I})]\boldsymbol{x}(t_i) + \boldsymbol{\Gamma}\boldsymbol{\zeta}_i, \tag{62}$$

where $\boldsymbol{\Lambda}\boldsymbol{\Lambda}^\top = 2\boldsymbol{D}$, $\boldsymbol{\Gamma}\boldsymbol{\Gamma}^\top = 2\sigma\boldsymbol{I}$ and $\boldsymbol{\eta}_i, \boldsymbol{\zeta}_i$ are $N-$dimensional independent, identically distributed Gaussian vectors with independent components, each with mean 0 and variance Δt. Thus we can associate the discretised OU process to the MAR model using the following relations,

$$\boldsymbol{A} = \boldsymbol{I} - \Delta t(\boldsymbol{B}), \tag{63}$$

$$\boldsymbol{A} = \boldsymbol{I} + \Delta t(\boldsymbol{W} - \boldsymbol{I}). \tag{64}$$

Furthermore, in the case of the network-restricted model, we modify the auto-regressive algorithm to use non-negative least-squares in order to guarantee that $\boldsymbol{W}$ is restricted to non-negative entries (see Appendix I) [109]. In order to estimate $\boldsymbol{D}$ or σ, we take the covariance of the residual time-series $\boldsymbol{\chi}(t)$ and note that,

$$\mathrm{Cov}[\boldsymbol{\chi}] = 2\Delta t\boldsymbol{D}, \tag{65}$$

$$\mathrm{Cov}[\boldsymbol{\chi}] = 2\Delta t\sigma\boldsymbol{I}, \tag{66}$$

depending on the assumed model. We can estimate $\sigma \approx \langle \frac{1}{2\Delta t}\mathrm{diag}\,\mathrm{Cov}[\boldsymbol{\chi}]\rangle$, where $\langle\cdot\rangle$ is the mean. We note that Δt represents the time-scale of the process and is not discernible directly from the time-series data but that $\Delta t << 1$ is an assumption of the discretisation. In the following, we take $\Delta t = 0.1$. Once $\boldsymbol{B}, \boldsymbol{D}$ or $\boldsymbol{W}, \sigma$ have been obtained, if $-\boldsymbol{B}$ or $\boldsymbol{W} - \boldsymbol{I}$ are stable matrices, one can use the analysis presented in Section III to calculate the EPR of the unique steady-state and quantify the degree to which we have broken detailed balance in the time-series. Furthermore, the directedness of the effective network, $\boldsymbol{W}$, can be analysed using the measures presented in Section II.

B. Applications to real-world multivariate time-series

1. *Human neuroimaging at rest and during task*

We first apply this approach to neuroimaging data from the HCP [110]. We consider BOLD fMRI in

the Desikan-Killany (DK80) parcellation [111] with 62 cortical regions and 18 sub-cortical regions taken from the same 100 (unrelated) participants at rest and during a social and motor task. For further details on the experimental paradigms see Ref. [110]. The data was pre-processed following standard HCP protocols and is further described in Ref. [112] and Appendix K. For each participant in each condition, we apply the method described in the previous section to fit both a general linear model and a network-restricted model using auto-regression. We then measure the EPR from the general model and the directedness from the effective network extracted in the restricted model. In addition, we partition the nodes of the network into the 7 canonical Yeo functional sub-networks, each of which is associated with a specific aspect of brain function [113] (see Appendix K).

Figure 7 shows the results of the analysis applied to the neuroimaging data. Panel *a*) shows that the EPR is elevated in task conditions when compared to rest, which coincides with results from previous studies [25, 26, 61, 62]. However, we can also see that the effective network becomes significantly more directed in the task states which drives the increase in the EPR. Panel *b*) shows that the EPR is positively correlated with the directedness using all measures. Going beyond aggregate quantities, panel *c*) shows the trophic organisation of a participant-averaged effective network coloured according to the Yeo partition. This representation allows us to see where each sub-network sits in the overall hierarchy and how the network is reorganised by the task stimulus. Similarly, we can plot the mean trophic level of each sub-network in each state, as shown in panel *d*). We can see that during tasks some sub-networks re-position themselves higher in the hierarchy, such as the dorsal attention network, whilst others are re-positioned lower, such as the sub-cortical regions. The link between the hierarchical organisation and the overall directedness of the underlying network and the EPR of the dynamics is supported by this empirical analysis. Furthermore, our results support previous resulting regarding the hierarchical reorganisation of the brain during tasks [61, 62, 112, 114, 115].

2. *Stock-prices from the New York Stock Exchange*

We consider the daily stock-prices of 119 U.S. companies listed on the NYSE over the period 2000-2021, as previously analysed and published in Ref. [92]. Each company represents a node whose signal is the fluctuating stock price. Therefore we have a single MVTS in this case-study. Again we fit the general linear model and calculate the EPR and the network-restricted model to obtain an effective network of interactions between the prices. Notably, under this model we restrict to positive connections between nodes, a common assumption for whole-brain modelling, but less common in the modelling of financial time-series where variables may have strong negative correlations [60]. However, our notions of directedness and hierarchy assume positive weights, and so this is an important modelling assumption.

Firstly, we note that the financial time-series is out of equilibrium and has EPR $\Phi = 26.2$ (3 s.f.). Furthermore, the effective network of stock interactions is trophically organised with trophic directedness, $\sqrt{1 - F_0} = 0.405$ (3 s.f.), strongly directed with irreciprocity, $1 - r = 0.975$ (3 s.f.) and non-normal ($d_H = 1.53$, $\hat{d}_H = 0.160$ (3 s.f.)). In Figure 8, panel *a*) shows the trophic decomposition of the effective network of interactions between stocks, coloured by industry, and panel *b*) shows the distribution of trophic levels for each industry. We can see that consumer services and goods sit atop the hierarchy, feeding into the dynamics of other stocks, whilst market-sensitive indicators, such as financials and utilities, sit at the bottom, following the trends. The conclusion that the NYSE is operating out of equilibrium can be interpreted both through the lens of thermoeconomics [116, 117] which argues that the law of statistical mechanics can describe economics systems, or more simply through the lens of economic forces creating a hierarchical interaction structure between stocks and industries [118], that results in non-equilibrium dynamics.

VIII. DISCUSSION AND CONCLUSIONS

Bridging the gap between the structure and dynamics of complex networks is a fundamental challenge in the modelling of real-world systems [9]. Here, we presented novel results linking the hierarchical, directed structure of a network with the emergence of non-equilibrium dynamics and broken detailed balance. For a diverse range of dynamical systems and directedness measures, we first showed that the EPR increased with directedness in synthetic hierarchical networks. We then decoupled the notions of directedness using a null model and we were able to show how the nature of the dynamics dictated which directedness measure would predict its EPR, highlighting the difference between locally- and globally-evolving processes. Next we considered a vast range of real-world network topologies and showed that the link between the EPR and directedness held in actual network data. Finally, we applied our theory to multivariate time-series using a simple auto-regressive model to measure the EPR directly from dynamic node time-series, but also to unravel the hierarchical structure of the interactions between variables. Applying this method to human neuroimaging at task and at rest, we found that the brain operates further from equilibrium in task compared to rest indicating that the EPR is a key indicator of cognitive exertion and complexity of neural dynamics which confirms previous results [25, 26, 61, 62, 114]. Additionally, our approach extracts

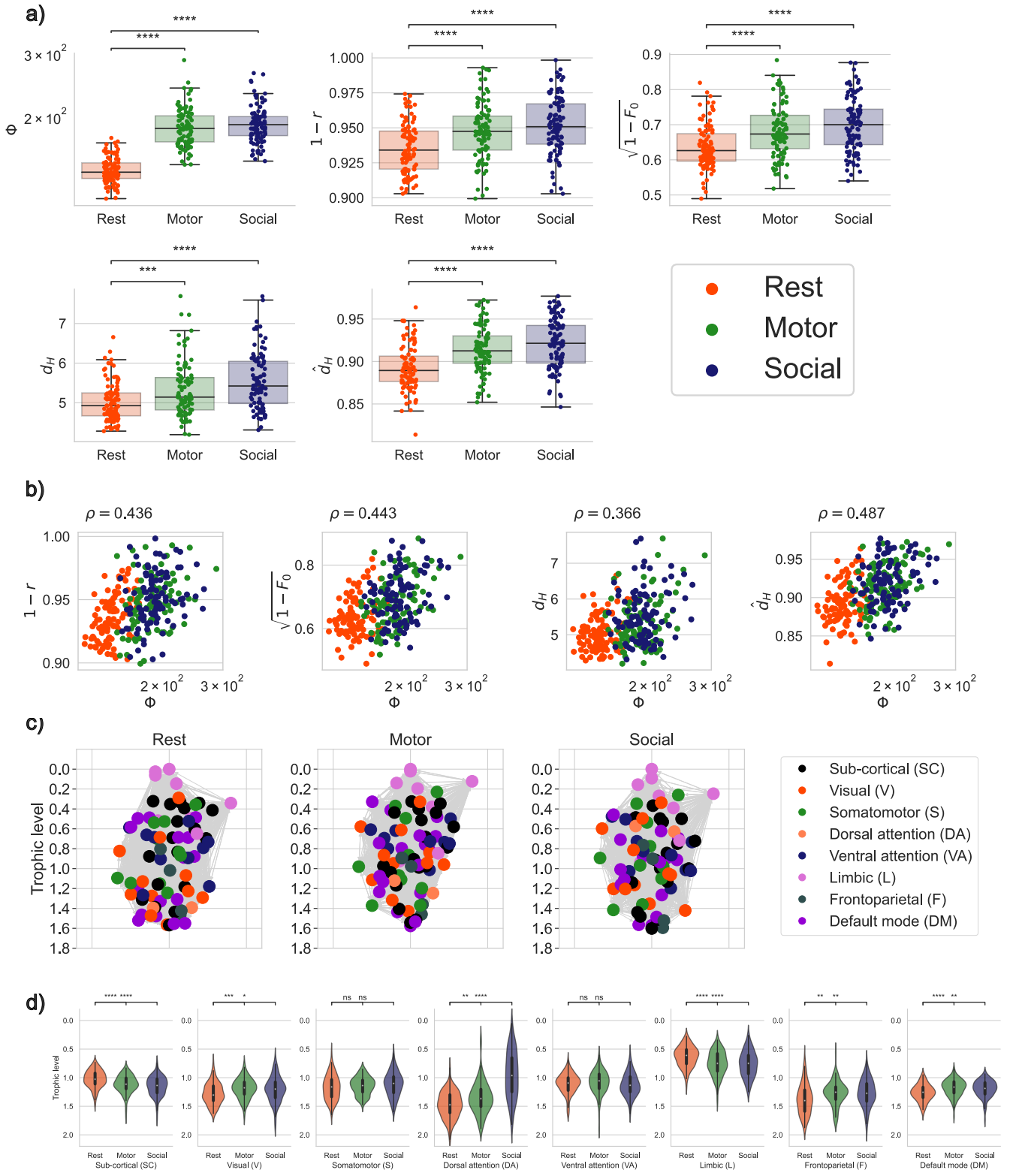

FIG. 7. **Entropy production rate and directedness of effective brain networks:** *a*) The EPR of the linear model and the directedness of the effective network are significantly elevated during task when compared to rest. *b*) EPR is positively correlated with each of the directedness measures. *c*) This panel shows the trophic levels of a participant-averaged effective network in each condition. The hierarchy of brain regions reorganises during tasks. Different functional networks occupy different positions within the processing hierarchy and this position changes depending on the task. *d*) The distribution among participants of trophic levels for each functional network in each condition. The significant changes in the distribution suggests the reorganisation of the processing hierarchy during task. (ns) = $p > 0.05$; (*) = $p > 0.01$; (**) = $p > 0.001$; (***) = $p > 0.0001$; (****) = $p \leq 0.0001$.

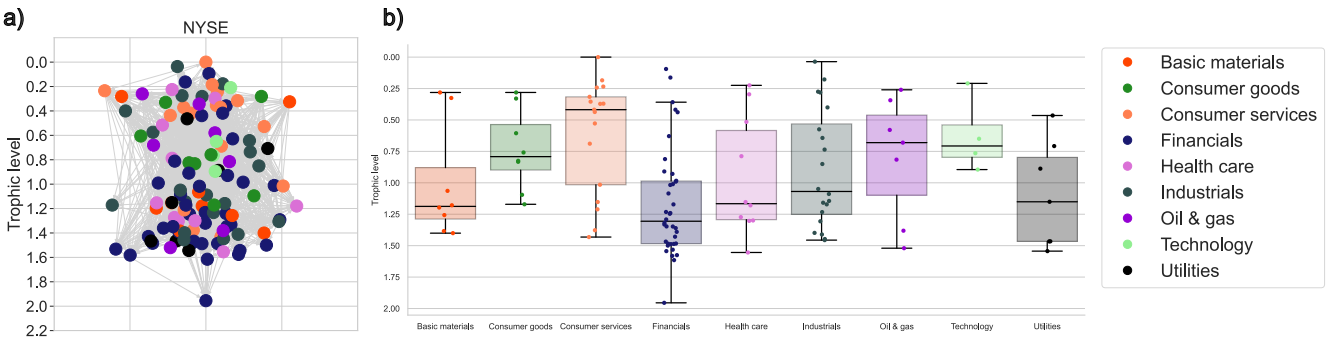

FIG. 8. **Trophic hierarchy in the New York Stock Exchange:** *a*) The effective network of stock interactions inferred from prices on the NYSE and coloured by industry. The network is hierarchically organised with certain industries like 'Consumer services' sitting atop the hierarchy and 'Financials' sitting at the bottom. *b*) Distributions of trophic level by industry. Each industry occupies a different position in the hierarchy with 'Consumer goods/services' at the top and 'Financials' and 'Utilities' at the bottom.

a network representation of regional interactions that demonstrates the hierarchical re-organisation of the brain during tasks [114, 115]. In addition, we analysed stock-prices from the NYSE to identify the directed influence structure between stocks created by the economic forces at play during speculation [118].

Our work represents the first attempt to link the structure of a network to its non-equilibrium thermodynamics. Given that non-equilibrium dynamics and the emergence of broken detailed balance is thought to be a unifying phenomenon characterizing living systems [13], our work indicates that directed, hierarchical structures are imperative to the functioning of biological complex systems. These results place newfound importance on both the structural [41, 49, 53, 63, 91] and the dynamical [40, 42, 47, 50] phenomena that are unique to complex systems with directed connections. In particular, in neuroimaging, where broken detailed balance emerges consistently [25], the traditional assumption of undirected structural connections limits the accuracy of whole-brain models which cannot explain the non-equilibrium nature of the empirical data. Our simple linear modeling approach provides a first-step to understand simultaneously the hierarchical structure and the non-equilibrium thermodynamics of multivariate data. Conversely, our analysis of empirical structural data showed that real-world networks display strong directedness and hierarchical organisation [51, 91]. Therefore, one can expect that the dynamics of such systems will be operating out of equilibrium and thus motivates the use of non-equilibrium thermodynamics to describe their evolution [117].

These results provide a general framework to study a range of applications. For example, hierarchically *modular* network structure has been shown to significantly impact critical dynamics [119], yet its influence on thermodynamics remains to be elucidated. Further, computational techniques for both measuring broken detailed balance [94–96] and inferring interactions [93, 98, 100, 103] directly from data, will allow these results to be extended beyond the limitations of the analytically tractable processes or linear models considered here, to a range of non-linear dynamics and real-world time-series. Finally, our approach provides a clear and important interpretation for empirically observed violations of detailed balance, as found in the brain [25]. As a result, such observations can be reframed from the perspective of hierarchical network organisation which can bring new insight into the workings of complex real-world systems and the relationship between structure and function.

To conclude, we have shown that directedness and non-equilibrium dynamics are intimately linked. The degree to which a system is hierarchically directed in its interactions determines its divergence from thermodynamic equilibrium. As a result, taking into consideration the non-reciprocity of interactions becomes fundamental to understanding the dynamic trajectories of both models and real-world systems.

Code availability

The R and Matlab code used in this project will be made available on publication at `https://github.com/rnartallo/brokendetailedbalance`.

Data availability

The network data used in this project is collated from multiple freely available locations and references are given in Appendix J. The human neuroimaging data used in this project is freely available from the HCP website [110]. The financial time-series used in this project is freely available from the Python package 'yfinance' and

at Ref. [120].

ACKNOWLEDGMENTS

The authors wish to acknowledge M. Aguilera for his assistance understanding the Ising mean-field approximation.

R.N.K was supported by an EPSRC doctoral scholarship from grants EP/T517811/1 and EP/R513295/1. G.D. was supported by the AGAUR research support grant (ref. 2021 SGR 00917) funded by the Department of Research and Universities of the Generalitat of Catalunya, by the project NEurological MEchanismS of Injury, and the project Sleep-like cellular dynamics (NEMESIS) (ref. 101071900) funded by the EU ERC Synergy Horizon Europe and by the project PID2022-136216NB-I00 financed by the MCIN /AEI /10.13039/501100011033 / FEDER, UE., the Ministry of Science and Innovation, the State Research Agency and the European Regional Development Fund. M.L.K was supported by the Centre for Eudaimonia and Human Flourishing funded by the Pettit and Carlsberg Foundations and by the Center for Music in the Brain (MIB), funded by the Danish National Research Foundation (project number DNRF117). R.L. was supported by the EPSRC grants EP/V013068/1 and EP/V03474X/1.

Appendix A: Relationship between trophic directedness and non-normality

In previous studies, the relationship between trophic coherence and non-normality has been discussed [41, 51]. In this appendix, we build on this by illustrating the situations where they are (in)equivalent, which lends insight into what they are measuring. First, we define **perfect trophic directedness** to be $F_0 = 0$ ($\sqrt{1-F_0} = 1$) i.e. maximally trophically directed. This occurs iff the level differences $z_{ji} = h_i - h_j$ are 1 for each pair of levels. Oppositely, we define a network to be **trophically flat** if $F_0 = 1$ ($\sqrt{1-F_0} = 0$) i.e. maximally trophically undirected. This occurs iff the level differences $z_{ji} = h_i - h_j$ are 0 for each pair of levels. Similarly, a network is defined to be **perfectly (maximally) non-normal** if $d_H(\boldsymbol{W}) = ||\boldsymbol{W}||_F$ ($\hat{d}_H(\boldsymbol{W}) = 1$). This means that $\sum_i |\lambda_i|^2 = 0$ where $\{\lambda_i\}$ is the eigenspectrum. Figure 9 shows examples of (un)weighted directed networks in all four categories of trophically flat/directed and non-normal/normal.

MacKay et al showed that a normal unweighted network must be trophically flat [51]. In the unweighted case, they begin by noting that $(\boldsymbol{W}^\top\boldsymbol{W})_{ij}$ is the number of common sources between nodes i and j, whilst $(\boldsymbol{W}\boldsymbol{W}^\top)_{ij}$ is the number of common targets. In particular, $(\boldsymbol{W}^\top\boldsymbol{W})_{ii} = w_i^{\text{in}}$ and $(\boldsymbol{W}\boldsymbol{W}^\top)_{ii} = w_i^{\text{out}}$. If

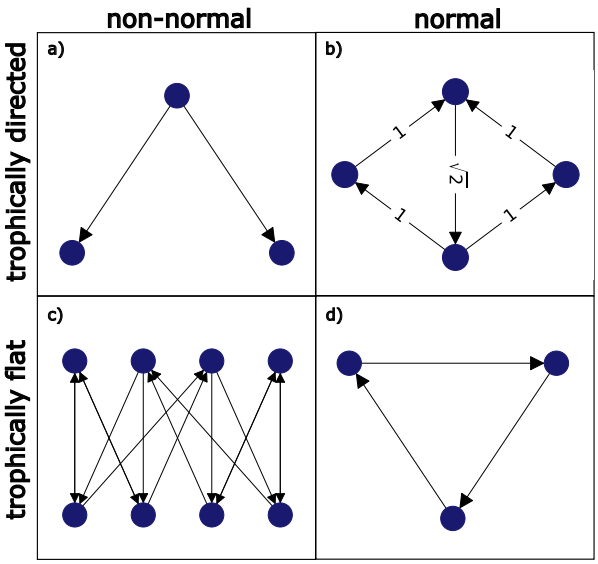

FIG. 9. **Non-normality and trophic directedness:** Non-normality and trophic directedness are in-equivalent. *a*) A hierarchical motif which is both non-normal and trophically directed. Most real-world directed networks are in this category. *b*) A normal but trophically directed network. There are no such unweighted networks. *c*) A trophically flat but non-normal network. Intuitively, this network has no clear hierarchy but is still non-normal. *d*) A directed cycle which is both normal and trophically flat. This network has no hierarchy.

$\boldsymbol{W}$ is normal i.e. $\boldsymbol{W}^\top\boldsymbol{W} = \boldsymbol{W}\boldsymbol{W}^\top$, then $w_i^{\text{in}} = w_i^{\text{out}}$ and $v_i = 0$ for all i i.e. the network is trophically flat.

For a weighted network, the same result does not hold. By following the proof of the unweighted case, we can construct a counter example. In the case of a network with non-negative weights, we now have,

$$(\boldsymbol{W}\boldsymbol{W}^\top)_{ii} = \sum_k W_{ik}^2, \tag{A1}$$

$$(\boldsymbol{W}^\top\boldsymbol{W})_{ii} = \sum_k W_{ki}^2, \tag{A2}$$

which are no longer simply $w_i^{\text{in}}, w_i^{\text{out}}$. If $\boldsymbol{W}$ is normal we therefore have, $\sum_k W_{ik}^2 = \sum_k W_{ki}^2$ for each i. On the other hand, a trophically flat network i.e. $v_i = 0$ implies $\sum_k W_{ik} = \sum_k W_{ki}$ for each i. These two conditions are not equivalent so we can construct a network that is both normal and trophically directed. One such network, shown in panel *b*) of Fig. 9, is given by,

$$\boldsymbol{W} = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ \sqrt{2} & 0 & 0 & 0 \end{pmatrix}, \tag{A3}$$

which satisfies that $\sum_k W_{ik}^2 = \sum_k W_{ki}^2$ for each i whilst $\sum_k W_{ik} \neq \sum_k W_{ki}$. Therefore, it is normal yet trophically directed ($F_0 = 0.9737$).

On the other hand, trophic flatness does not imply normality, even in the unweighted case. One can consider the network,

$$\boldsymbol{W} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{A4}$$

shown in panel c) of Fig. 9, which is non-normal ($d_H(\boldsymbol{W}) = 2.8284$) yet trophically flat. MacKay et al present a simpler example of a trophically flat, non-normal network, ($d_H(\boldsymbol{W}) = 0.6982$), but including a self loop,

$$\boldsymbol{W} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{A5}$$

MacKay et al further showed that perfect trophic directedness implies implies all eigenvalues are 0 and, consequently, perfect non-normality [51]. On the other hand, perfect non-normality does not imply perfect trophic directedness, as evidenced by the so-called 'feed-forward' motif,

$$\boldsymbol{W} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \tag{A6}$$

which has $\hat{d}_H(\boldsymbol{W}) = 1$ ($d_H(\boldsymbol{W}) = 1.7321$) but $F_0 = 0.1111$.

Finally, in real-world directed networks, strong correlations have been found between trophic directedness and non-normality, but the relationship is non-linear [41, 51]. This relationship can be partially explained by introducing the 'loop exponent' of a network which bridges the spectral properties of a network and the trophic decomposition [41, 91]. Figure 10 shows the correlations between the four different measures in the 97 real-world networks considered in Section VI.

Appendix B: A preferential attachment scheme for generating hierarchical networks

The seminal contributions of de Solla Price [121, 122] and the subsequent development of the preferential attachment (PA) algorithm for generating random networks [123], has long been studied for its power-law degree distribution. Modifications to the PA algorithm can cause it to generate non-normal, irreciprocal, trophically directed networks with a strong hierarchy [50]. For the networks considered in Sections II and III, we used the following generative algorithm.

We consider a growing network where, at each step, a new node is added to the network. The new node j connects to an exisiting node i with a probability proportional to d_i, the degree of node i. We randomly sample the weight of this connection W_{ji} from a uniform distirbution $U(0,1)$, and then introduce a weak reciprocal link $W_{ij} = \frac{W_{ji}}{\gamma}$ for $\gamma \gg 1$ such that $0 \leq W_{ij} \ll W_{ji}$.

As the network begins with no edges and a single node, we must initialise the network in some way. When a new node j is added to the network, if an exisiting node i has $d_i = 0$, they are connected with probability $0 < p_0 \ll 1$. If the exisiting node i has $d_i > 0$, then they are connected with probability $\min(1, \frac{d_i}{d_{\text{tot}}} + \mu)$ where d_{tot} is the total degree of the network. This approach produces a strongly connected graph where each edge is reciprocated, which may not be true for all real-world networks.

Appendix C: Linear interpolation parameterises trophic directedness and irreciprocity

In Section II, we varied the directedness of networks by first generating a hierarchically directed network using the PA algorithm, as described above, and then linearly interpolating between this network and its Hermitian,

$$\hat{\boldsymbol{W}}(\epsilon) = (1-\epsilon)\tilde{\boldsymbol{W}} + \epsilon \boldsymbol{W}, \tag{C1}$$

for $\epsilon \in [0,1]$ where the Hermitian network is given by $\tilde{\boldsymbol{W}} = \frac{1}{2}\left(\boldsymbol{W} + \boldsymbol{W}^\top\right)$. Figure 2 showed that each of the four measures increases almost linearly with ϵ. In this section, we show that the irreciprocity and trophic directedness increase exactly linearly i.e.

$$1 - r(\hat{\boldsymbol{W}}(\epsilon)) = \epsilon(1 - r(\boldsymbol{W})), \tag{C2}$$

$$\sqrt{1 - F_0(\hat{\boldsymbol{W}}(\epsilon)} = \epsilon\sqrt{1 - F_0(\boldsymbol{W})}. \tag{C3}$$

Unfortunately, non-normality is a spectral measure and the eigenvalues of the matrix $\hat{\boldsymbol{W}}(\epsilon)$ are not calculable from $\tilde{\boldsymbol{W}}$ and $\boldsymbol{W}$.

We begin with the irreciprocity. We denote the reciprocity of $\boldsymbol{W}$ by r ($= r(1)$). The irreciprocity of the interpolated network is given by,

$$1 - r(\epsilon) = 1 - \frac{\sum_i \sum_{j \neq i} \min(\hat{W}_{ij}(\epsilon), \hat{W}_{ji}(\epsilon))}{\sum_i \sum_{j \neq i} \hat{W}_{ij}(\epsilon)}, \tag{C4}$$

$$= \frac{\sum_i \sum_{j \neq i} \hat{W}_{ij}(\epsilon) - \min(\hat{W}_{ij}(\epsilon), \hat{W}_{ji}(\epsilon))}{\sum_i \sum_{j \neq i} \hat{W}_{ij}(\epsilon)}. \tag{C5}$$

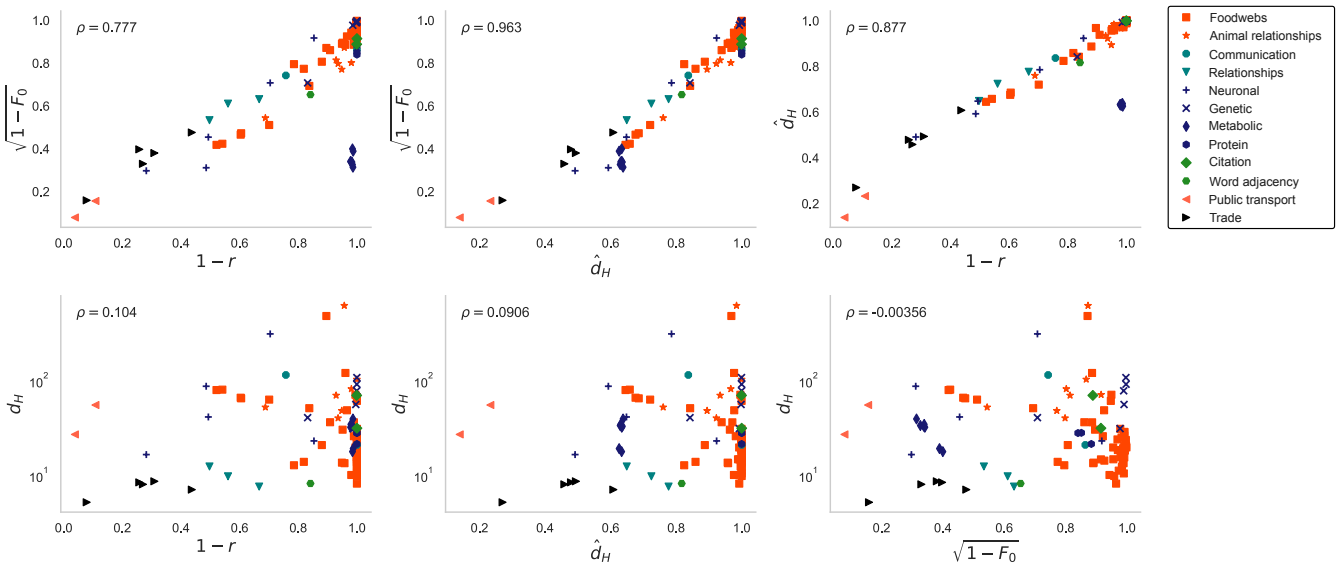

FIG. 10. **Directedness of real-world networks**: A comparison between the directedness measures on real-world networks. Each pane shows a pairwise comparison between two of the four directedness measures on 97 real world networks from different domains. We can see strong correlations between the three normalised measures irreciprocity, trophic directedness and normalised Henrici index with a much weaker correlation to the unbounded Henrici index. We can also identify which fields the strongly, or weakly, directed networks belong to. Ecological networks are in orange, social in teal, biological in blue, language in green, transport in red and economic in black.

Firstly, we note that the interpolation preserves the row sums (even excluding the diagonal), so the denominator, which we denote k, is the same as in the irreciprocity of the original network $\boldsymbol{W}$.

$$1 - r(\epsilon) = \frac{\sum_i \sum_{j\neq i} \hat{W}_{ij}(\epsilon) - \min(\hat{W}_{ij}(\epsilon), \hat{W}_{ji}(\epsilon))}{k}, \tag{C6}$$

$$k(1 - r(\epsilon)) = \sum_i \sum_{j\neq i} \hat{W}_{ij}(\epsilon) - \min(\hat{W}_{ij}(\epsilon), \hat{W}_{ji}(\epsilon)). \tag{C7}$$

Then, using the defintion of $\hat{W}_{ij}(\epsilon)$ and cancelling the Hermitian terms, we obtain,

$$k(1 - r(\epsilon)) = \sum_i \sum_{j\neq i} \epsilon W_{ij} - \epsilon \min(W_{ij}, W_{ji}). \tag{C8}$$

Finally, we get,

$$1 - r(\epsilon) = \frac{\epsilon \sum_i \sum_{j\neq i} W_{ij} - \min(W_{ij}, W_{ji})}{k}, \tag{C9}$$

$$1 - r(\epsilon) = \epsilon(1 - r). \tag{C10}$$

□

Next, we consider the trophic directedness. As defined in Section II, for the original network $\boldsymbol{W}$, we have quantities, $w_i^{\text{in}}, w_i^{\text{out}}$. For the interpolated network these become,

$$w_i^{\text{in}}(\epsilon) = \epsilon w_i^{\text{in}} + (1-\epsilon)\frac{1}{2}(w_i^{\text{in}} + w_i^{\text{out}}), \tag{C11}$$

$$w_i^{\text{out}}(\epsilon) = \epsilon w_i^{\text{out}} + (1-\epsilon)\frac{1}{2}(w_i^{\text{in}} + w_i^{\text{out}}). \tag{C12}$$

From this, we obtain,

$$u_i(\epsilon) = w_i^{\text{in}}(\epsilon) + w_i^{\text{out}}(\epsilon) \tag{C13}$$

$$= w_i^{\text{out}} + w_i^{\text{out}} \tag{C14}$$

$$= u_i, \tag{C15}$$

i.e. u_i is preserved under interpolation. On the other hand,

$$v_i(\epsilon) = w_i^{\text{in}}(\epsilon) - w_i^{\text{out}}(\epsilon) \tag{C16}$$

$$= \epsilon(w_i^{\text{in}} - w_i^{\text{out}}) \tag{C17}$$

$$= \epsilon v_i. \tag{C18}$$

Furthermore, the symmetric weighted graph Laplacian of the original network, $\boldsymbol{\Lambda}$, is also preserved under the interpolation,

$$\boldsymbol{\Lambda}(\epsilon) = \text{diag}(\boldsymbol{u}) - \hat{\boldsymbol{W}}(\epsilon) - \hat{\boldsymbol{W}}^\top(\epsilon) \tag{C19}$$

$$= \text{diag}(\boldsymbol{u}) - 2\tilde{\boldsymbol{W}} \tag{C20}$$

$$= \boldsymbol{\Lambda}. \tag{C21}$$

For original network, $\boldsymbol{h}$ is the vector of trophic levels which solves,

$$\boldsymbol{\Lambda h} = \boldsymbol{v}. \tag{C22}$$

The trophic levels of the interpolated network are given by the solution to,

$$\boldsymbol{\Lambda h}(\epsilon) = \epsilon \boldsymbol{v}, \tag{C23}$$

i.e. $\boldsymbol{h}(\epsilon) = \epsilon \boldsymbol{h}$. Finally, we use an alternative, but equivalent, formulation for $1 - F_0$, shown by Mackay et al [51], given by,

$$1 - F_0 = \frac{\sum_{i,j} W_{ij}(h_j - h_i)}{\sum_{i,j} W_{ij}}. \tag{C24}$$

Therefore, for the interpolated network, we have,

$$1 - F_0(\epsilon) = \frac{\sum_{i,j} \hat{W}_{ij}(\epsilon)(h_j - h_i)\epsilon}{\sum_{i,j} W_{ij}}, \tag{C25}$$

as the denominator, which we again denote k, is preserved under the interpolation. Expanding the interpolation, we obtain,

$$\begin{aligned} k(1 - F_0(\epsilon)) = \epsilon(1-\epsilon)\frac{1}{2}\sum_{i,j}(W_{ij} + W_{ji})(h_j - h_i) \\ + \epsilon^2 \sum_{i,j} W_{i,j}(h_j - h_i). \end{aligned} \tag{C26}$$

Notice that the sum $\sum_{i,j}(W_{ij}+W_{ji})(h_j-h_i)$ vanishes as the term for (i,j) cancels with the term for (j,i). Therefore, we get

$$1 - F_0(\epsilon) = \epsilon^2(1 - F_0), \tag{C27}$$

or for the trophic directedness defined in Section II,

$$\sqrt{1 - F_0(\epsilon)} = \epsilon\sqrt{1 - F_0}. \tag{C28}$$

□

Appendix D: Deriving the entropy production rate in the Ornstein-Uhlenbeck process

We recall that the multivariate OU process is given by the Langevin system,

$$\frac{d\boldsymbol{x}}{dt} = -\boldsymbol{B}\boldsymbol{x}(t) + \boldsymbol{\xi}(t), \tag{D1}$$

where $\boldsymbol{x}(t) \in \mathbb{R}^N$ is the time-dependent state vector, $\boldsymbol{B} \in \mathbb{R}^{N\times N}$ is the **friction matrix**, and $\boldsymbol{\xi}(t) \in \mathbb{R}^N$ is additive white noise with covariance given by,

$$\langle \boldsymbol{\xi}(t)\boldsymbol{\xi}^\top(t')\rangle = 2\boldsymbol{D}\delta(t - t'), \tag{D2}$$

where $\boldsymbol{D} \in \mathbb{R}^{N\times N}$ is the *noise covariance matrix* which is symmetric. We follow Godrèche and Luck [57] to derive the EPR rate of the OU in a steady state. Assuming each eigenvalue of the friction matrix, $\boldsymbol{B}$, has positive real part, then the system relaxes exponentially fast to a steady state with Gaussian fluctuations given by,

$$\boldsymbol{x}(t) = e^{-\boldsymbol{B}t}\boldsymbol{x}(0) + \int_0^t e^{-\boldsymbol{B}(t-s)}\boldsymbol{\xi}(s)\, ds, \tag{D3}$$

and covariance,

$$\boldsymbol{S} = \lim_{t\to\infty} \boldsymbol{S}(t), \tag{D4}$$

$$= \lim_{t\to\infty} \langle \boldsymbol{x}(t)\boldsymbol{x}^\top(t)\rangle. \tag{D5}$$

The covariance can be written as,

$$\begin{aligned} \boldsymbol{S} = \lim_{t\to\infty} [e^{-\boldsymbol{B}t}\boldsymbol{S}(0)e^{-\boldsymbol{B}^\top t}, \\ + 2\int_0^t e^{-\boldsymbol{B}(t-s)}\boldsymbol{D}e^{-\boldsymbol{B}^\top(t-s)}\, ds], \end{aligned} \tag{D6}$$

$$= 2\int_0^\infty e^{-\boldsymbol{B}t}\boldsymbol{D}e^{-\boldsymbol{B}^\top t}\, dt. \tag{D7}$$

It can also be shown that $\boldsymbol{S}$ satisfies the following Sylvester equation [81],

$$\boldsymbol{B}\boldsymbol{S} + \boldsymbol{S}\boldsymbol{B}^\top = 2\boldsymbol{D}. \tag{D8}$$

Next, we define the Onsager matrix, $\boldsymbol{L}$, of kinetic coefficients,

$$\boldsymbol{L} = \boldsymbol{B}\boldsymbol{S} = \boldsymbol{D} + \boldsymbol{Q}, \tag{D9}$$

$$\boldsymbol{L}^\top = \boldsymbol{S}\boldsymbol{B}^\top = \boldsymbol{D} - \boldsymbol{Q}, \tag{D10}$$

parameterising the asymmetries through the matrix $\boldsymbol{Q}$, which provides an intuitive measure of the degree of non-equilibrium.

The EPR, Φ, can then be written in the form,

$$\Phi = \langle \boldsymbol{x}^\top(\boldsymbol{D}^{-1}\boldsymbol{B} - \boldsymbol{S}^{-1})^\top \boldsymbol{D}(\boldsymbol{D}^{-1}\boldsymbol{B} - \boldsymbol{S}^{-1})\boldsymbol{x}\rangle, \tag{D11}$$

$$= -\langle \boldsymbol{x}^\top \boldsymbol{S}^{-1}\boldsymbol{Q}\boldsymbol{D}^{-1}\boldsymbol{Q}\boldsymbol{S}^{-1}\boldsymbol{x}\rangle, \tag{D12}$$

with the second equation following from the relations $\boldsymbol{D}^{-1}\boldsymbol{B} - \boldsymbol{S}^{-1} = \boldsymbol{D}^{-1}\boldsymbol{Q}\boldsymbol{S}^{-1}$ and $\boldsymbol{Q} = -\boldsymbol{Q}^\top$. Using that the steady state is Gaussian we have that $\langle \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}\rangle = \mathrm{Tr}(\boldsymbol{S}\boldsymbol{A})$ for a general matrix $\boldsymbol{A}$ and thus we have that,

$$\Phi = -\,\mathrm{Tr}(\boldsymbol{Q}\boldsymbol{D}^{-1}\boldsymbol{Q}\boldsymbol{S}^{-1}). \tag{D13}$$

This can be rewritten in the form,

$$\Phi = -\,\mathrm{Tr}(\boldsymbol{D}^{-1}\boldsymbol{B}\boldsymbol{Q}). \tag{D14}$$

For further details see Ref. [57].

Appendix E: 2-node networks

In Section IV, we considered directed 2-node networks and showed that the EPR of the OU increases with the asymmetry between the two connections. Furthermore,

for larger networks, we consider the OU and calculated the EPR as a function of the interpolation parameter ϵ using the numerical solution of the Sylvester equation (D8). Here we consider the linear interpolation applied to the case of 2 nodes, where the EPR as a function of ϵ can be calculated explicitly [57]. We show, analytically, that the Henrici index increases linearly and the EPR of the OU increases quadratically in ϵ, consistent with the conjectured relationship obtained numerically for hierarchical networks. This varies from the analysis in Section IV, where we varied the weights (W_{12}, W_{21}) as we are now fixing these weights and performing the interpolation between the network and its Hermitian. The 2-node directed network is defined by the weight matrix,

$$\boldsymbol{W} = \begin{bmatrix} 0 & W_{12} \\ W_{21} & 0 \end{bmatrix}. \tag{E1}$$

In the previous section, we showed that the irreciprocity and trophic directedness are linearly interpolated by ϵ. Recalling from Section IV, that the irreciprocity and trophic directedness of the 2-node network coincides and is given by,

$$1 - r = \sqrt{1 - F_0} = \frac{|W_{12} - W_{21}|}{W_{12} + W_{21}}, \tag{E2}$$

then the irreciprocity and trophic directedness of the interpolated network is given by,

$$1 - r(\epsilon) = \sqrt{1 - F_0(\epsilon)} = \epsilon \frac{|W_{12} - W_{21}|}{W_{12} + W_{21}}. \tag{E3}$$

On the other hand, we were not able to show that, for general networks, the Henrici indices scaled linearly with ϵ, but we can do so in the case of the 2-node network. Recall that the Henrici indices of the 2-node network were given by,

$$d_H = |W_{12} - W_{21}|, \tag{E4}$$

$$\hat{d}_H = \frac{|W_{12} - W_{21}|}{\sqrt{W_{12}^2 + W_{21}^2}}. \tag{E5}$$

First we define,

$$\tilde{W}_{12}(\epsilon) = \epsilon W_{12} + \frac{1}{2}(1 - \epsilon)(W_{12} + W_{21}) \tag{E6}$$

$$\tilde{W}_{21}(\epsilon) = \epsilon W_{21} + \frac{1}{2}(1 - \epsilon)(W_{12} + W_{21}) \tag{E7}$$

The eigenvalues of the interpolated network are given by,

$$\lambda_{\pm}(\epsilon) = \pm\sqrt{\tilde{W}_{12}(\epsilon)\tilde{W}_{21}(\epsilon)}, \tag{E8}$$

thus the Henrici index factorises,

$$d_H^2(\epsilon) = [\tilde{W}_{12}(\epsilon)]^2 + [\tilde{W}_{21}(\epsilon)]^2 - 2\tilde{W}_{12}(\epsilon)\tilde{W}_{21}(\epsilon) \tag{E9}$$

$$= \epsilon^2 (W_{12} - W_{21})^2 \tag{E10}$$

$$d_H(\epsilon) = \epsilon |W_{12} - W_{21}| \tag{E11}$$

$$= \epsilon d_H. \tag{E12}$$

Next we consider the OU evolving on an interpolated 2-node network. We recall that for an MOU defined by matrices,

$$\boldsymbol{B} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \boldsymbol{D} = \begin{pmatrix} u & w \\ w & v \end{pmatrix}, \tag{E13}$$

the EPR is given by,

$$\Phi = \frac{(cu - bv + (d - a)w)^2}{(a + d)(uv - w^2)}. \tag{E14}$$

For the 2-node network, this expression becomes,

$$\Phi = \frac{\Theta\gamma^2}{2}(W_{12} - W_{21})^2, \tag{E15}$$

We now fix $W_{12} \neq W_{21}$ which has an associated EPR Φ. Under the interpolation, the network becomes,

$$\hat{\boldsymbol{W}}(\epsilon) = \begin{bmatrix} 0 & \tilde{W}_{12}(\epsilon) \\ \tilde{W}_{21}(\epsilon) & 0 \end{bmatrix} \tag{E16}$$

Thus, we have,

$$\boldsymbol{B} = \begin{pmatrix} \Theta & \Theta(1 - \gamma\tilde{W}_{12}(\epsilon)) \\ \Theta(1 - \gamma\tilde{W}_{21}(\epsilon)) & \Theta \end{pmatrix}, \tag{E17}$$

$$\boldsymbol{D} = \begin{pmatrix} 2\sigma & 0 \\ 0 & 2\sigma \end{pmatrix}, \tag{E18}$$

which, after simplification, yields an expression for the EPR of the interpolated system,

$$\Phi(\epsilon) = \frac{\Theta\gamma^2\epsilon^2(W_{12} - W_{21})^2}{2} \tag{E19}$$

$$= \epsilon^2 \Phi. \tag{E20}$$

□

Linking this to previous result, the EPR in the OU on the interpolated 2-node network is also an exact quadratic function of the unnormalised Henrici index of the corresponding network.

Appendix F: Circulant networks

Next we consider the case of networks with circulant weight matrices. These correspond to k-regular directed cyclic networks. For example, Figure 11 shows two k-regular 4-cycles ($k = 1, 2$). All k-regular N-cycles are special cases of the $(N-1)-$regular N-cycle which can be written as,

$$\boldsymbol{W} = \begin{pmatrix} 0 & w_1 & w_2 & \dots & w_{N-1} \\ w_{N-1} & 0 & w_1 & \dots & w_{N-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ w_1 & w_2 & w_3 & \dots & 0 \end{pmatrix}. \tag{F1}$$

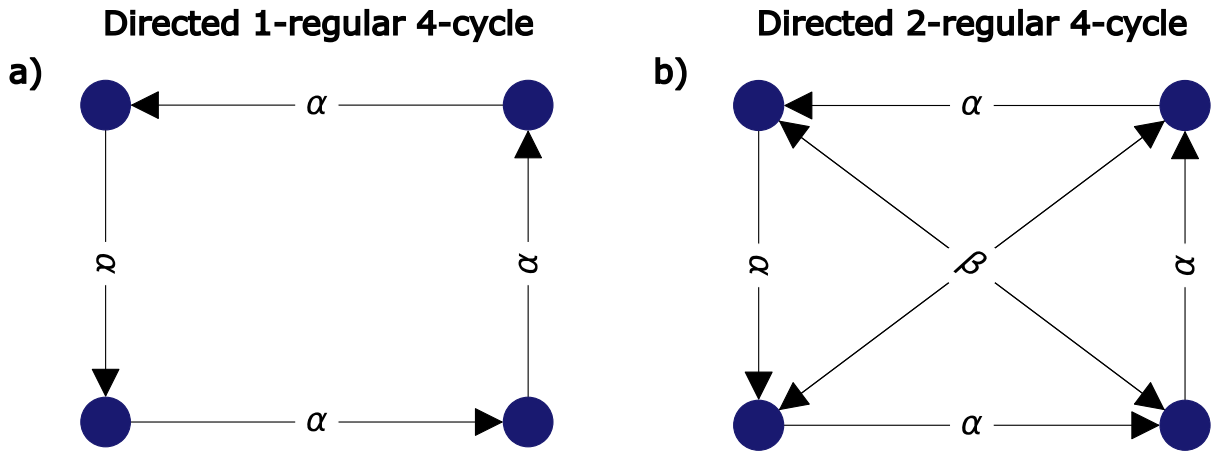

FIG. 11. **Circulant matrices and $k-$regular directed cycles:** *a*) A directed 1-regular 4 cycle defined by the vector $\boldsymbol{w} = (0, \alpha, 0, 0)$. *b*) A directed 2-regular 4 cycle defined by the vector $\boldsymbol{w} = (0, \alpha, \beta, 0)$.

This is uniquely defined by the top row,

$$\boldsymbol{w} = (0, w_1, ..., w_{N-1}). \tag{F2}$$

It is important to note that $\boldsymbol{W}$ is circulant and therefore normal. However, as it is directed, it breaks detailed balance for all the systems considered in this paper (with the exception of the RW for $N = 2$). More importantly, as we will show, applying the interpolation to the cyclic networks, ϵ increases the EPR of the corresponding NOU. We found in Sections V and VI that the non-normality of the underlying network was a extremely strong indicator of the EPR of the NOU, yet in this case, the non-normality is 0 and the EPR still varies with ϵ. This represents an important special case where hierarchical asymmetry or a global direction are not necessary for the OU to progressively break detailed balance. The lack of these cyclic structures in real-world networks [91] explains why, despite this special case, the non-normality correlates so closely with the EPR for the OU on real-world networks, as shown in Section VI.

Firstly, as mentioned, the Henrici indices are 0 for these networks as their weight matrices are normal. Furthermore, for all i

$$w_i^{\text{in}} = w_i^{\text{out}} = \sum_j w_j, \tag{F3}$$

meaning $\boldsymbol{v} = 0$ and therefore $F_0 = 1$ ($\sqrt{1 - F_0} = 0$) so the network is trophically flat. This implies that the network has no global direction or hierarchy as both global measures vanish. On the other hand, it is locally asymmetric and so has non-zero irreciprocity,

$$1 - r = 1 - \frac{\sum_j \min(w_j, w_{N-j})}{\sum_j w_j} \tag{F4}$$

$$= \frac{\sum_j |w_j - w_{N-j}|}{2\sum_j w_j}. \tag{F5}$$

Next, we consider the EPR of the OU on circulant networks. The interpolated network is also circulant and so we can write it in terms of the top row. First, we define,

$$\tilde{w}_i(\epsilon) = \epsilon w_i + \frac{1}{2}(1 - \epsilon)(w_i + w_{N-i}), \tag{F6}$$

then the top row of the interpolated weight matrix is given by

$$\hat{\boldsymbol{w}}(\epsilon) = (0, \tilde{w}_1(\epsilon), \ldots, \tilde{w}_{N-1}(\epsilon)). \tag{F7}$$

Therefore, the friction matrix $\boldsymbol{B} = \Theta(\boldsymbol{I} - \gamma\hat{\boldsymbol{W}}(\epsilon))$ is also circulant and defined by the vector,

$$\boldsymbol{b}(\epsilon) = (\Theta, -\Theta\gamma\tilde{w}_1(\epsilon), \ldots, -\Theta\gamma\tilde{w}_{N-1}(\epsilon)). \tag{F8}$$

The matrix $\boldsymbol{D}$ is diagonal for the OU and therefore circulant. Godrèche and Luck [57] showed that the EPR of the MOU with circulant $\boldsymbol{B}, \boldsymbol{D}$ is given by,

$$\Phi = \sum_{k=0}^{N-1} \frac{(\Im(\tilde{b}_k))^2}{\Re(\tilde{b}_k)}, \tag{F9}$$

where $\tilde{\boldsymbol{b}} = (\tilde{b}_0, ..., \tilde{b}_{N-1})$ is the discrete Fourier transform of the vector $\boldsymbol{b}$, $\Im(\cdot)$ is the imaginary part and $\Re(\cdot)$ is the real part. Furthermore, $\tilde{\boldsymbol{b}}$ is also the vector of eigenvalues of the circulant matrix $\boldsymbol{B}$.

The discrete Fourier transform gives us,

$$\tilde{b}_k = \Theta - \Theta\gamma \sum_{j=1}^{N-1} \tilde{w}_j(\epsilon) \exp \boldsymbol{i} \frac{2\pi jk}{N}. \tag{F10}$$

Using the periodicity and oddness/evenness of sine/cosine we have that,

$$\sin(\frac{2\pi jk}{N}) + \sin(\frac{2\pi(N-j)k}{N}) = 0, \tag{F11}$$

$$\cos(\frac{2\pi jk}{N}) + \cos(\frac{2\pi(N-j)k}{N}) = 2, \tag{F12}$$

which allows us to simplify our expression to,

$$\tilde{b}_k = \Theta - \Theta\gamma \sum_{j=1}^{N-1} (\epsilon w_j)(\boldsymbol{i}\sin(\frac{2\pi jk}{N}) + \cos(\frac{2\pi jk}{N})) - \Theta\gamma \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} (1-\epsilon)(w_j + w_{N-j})\cos(\frac{2\pi jk}{N}), \tag{F13}$$

which we can then substitute into the formula for $\Phi(\epsilon)$. Doing so, we get,

$$\Phi(\epsilon) = \sum_{k=0}^{N-1} \frac{\Theta^2\gamma^2 \left(\sum_{j=1}^{N} \epsilon w_j \sin(\frac{2\pi jk}{N})\right)^2}{\Theta - \Theta\gamma \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} (1-\epsilon)(w_j + w_{N-j})\cos(\frac{2\pi jk}{N}) - \Theta\gamma \sum_{j=1}^{N-1} \epsilon w_j \cos(\frac{2\pi jk}{N})}, \tag{F14}$$

$$= \epsilon^2 \sum_{k=0}^{N-1} \frac{\Theta^2\gamma^2 \left(\sum_{j=1}^{N} w_j \sin(\frac{2\pi jk}{N})\right)^2}{\Theta - \Theta\gamma \sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} (w_j + w_{N-j})\cos(\frac{2\pi jk}{N})}, \tag{F15}$$

$$= \epsilon^2 \Phi, \tag{F16}$$

where Φ is the EPR of the fully directed network ($\Phi(1)$). Again, we can see that the EPR of the OU scales quadratically with ϵ, but in this case all the interpolated networks are normal.

Appendix G: Coarse-graining fails to capture entropy production in small Ising models

In an Ising model with N variables, there are 2^N possible configurations, meaning the state space expands exponentially with the system size. For small Ising networks, using the Glauber dynamics [86], we were able to sample trajectories from the Ising model and estimate the steady state probabilities. Then, with the conditional transition probability defined by the model, we were able to estimate the joint transition probability and the EPR. However, for $N > 10$, the state space is so large that estimating the steady state probabilities required too many samples and sorting samples into distinct states became computational infeasible. An alternative approach that has been applied for the Ising model and in empirical time-series [25], is to measure the EPR in a coarse-grained state-space [87, 88]. Coarse-graining the state-space is achieved by mapping a number of **microstates**, in the original high-dimensional state-space (micro-space), to a lower number of **macrostates** in a lower dimensional space (macro-space). This can be achieved by clustering, in particular hierarchical clustering, which maps 'similar' microstates into a single macrostate, as shown by the schematic in Figure 12. Then, the EPR can be estimated by measuring the divergence between forward and backward joint transition probabilities in the coarse-grained space [88]. This measurement is a lower-bound on the 'true' entropy production in micro-space meaning that broken detailed balance at the coarse-grained level implies non-equilibrium dynamics at the micro-level [25]. However, as we show here, this lower bound may be inaccurate as a relative estimate of the EPR.

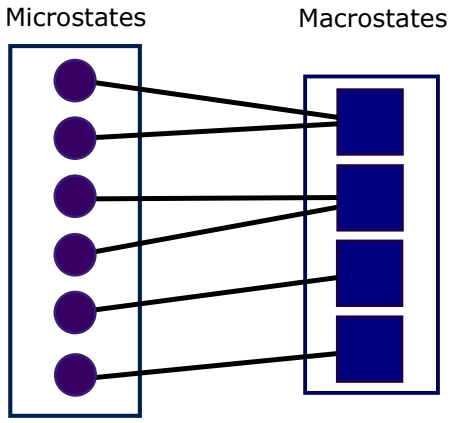

FIG. 12. **Coarse-graining of state space:** High dimensional state space can be coarse-grained into a lower one by mapping multiple microstates into a lower number of macrostates. This by done via a clustering approach. The EPR in the coarse-grained state space is a lower bound on the entropy production in the original state-space.

In order to illustrate this, we consider an Ising model with 10 spins. This system has a state space of size $2^{10} = 1024$ which is close to the size-limit where we can estimate the steady-state probabilities, yet large enough to cluster states sensibly, meaning we can compare the results in micro- and macro-space. We generate a hierarchical 10-node network and apply the interpolation described previously to vary the directedness of the network. At each value of ϵ, we sample from the Ising model using Glauber dynamics. We then estimate the steady-state probabilities in micro-space and use the conditional transition probability that defines the dynamics to calculate the joint transition probabilities and the EPR, as described in Section III. Concurrently, at each ϵ, we apply bisecting, hierarchical $k-$means clustering [124] as applied by Lynn et al [25] to the samples to coarsen the state-space into $k = 10$ macrostates. We note that, while we present the results for $k = 10$, this result remained consistent over a reasonable number of macro-states. In the macro-space, we no longer have the conditional transition probability and so we directly estimate the joint transition probability by counting the occurrences of each transition, following Lynn et al [25], which can then be used to estimate the EPR. Figure 13 shows the normalised EPR in both state-spaces as a function of ϵ. Panel *a*) shows that the EPR in micro-space follows the expected behaviour, increasing non-linearly with ϵ, as was found in Sections IV, V and VI. On the other hand, Panel *b*) shows that the coarse-graining procedure causes an inaccurate relative measurement of the EPR and there is no correlation between the directedness of the network and the EPR.

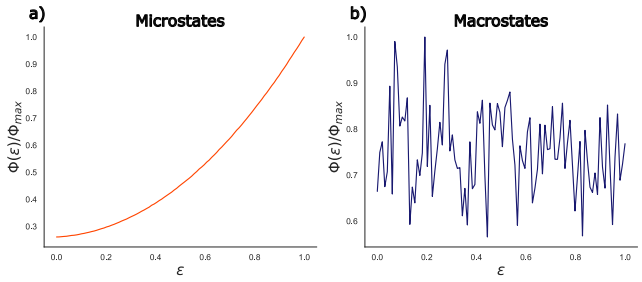

FIG. 13. **Entropy production in micro- and macro-space:** Normalised EPR in the asymmetric Ising model with 10 nodes, calculated from samples, as a function of the interpolation parameter ϵ. *a*) The normalised EPR calculated in micro-space which follows the expected behaviour found in Sections III, V and VI. *b*) The normalised EPR calculated in a coarse-grained macro-space. There is no correlation between the directedness of the network, ϵ, and the EPR $\frac{\Phi(\epsilon)}{\Phi_{\max}}$.

Appendix H: Derivation of the naive mean field for the asymmetric Ising model

The naive mean-field (NMF) and the Thouless-Anderson-Palmer (TAP) mean-field are standard approaches to solving the so-called 'inverse Ising problem' in equilibrium Ising models [85]. For non-equilibrium (asymmetric) Ising systems, an information-geometric approach can be used to approximate the mean-field solution [125]. Here, we will derive the NMF for the asymmetric Ising model following a recent framework that unifies a number of mean-field approaches to the Ising model [85].

We recall that the Ising model is defined by a discrete time Markov chain where the spins at time $t+1$ are updated according to,

$$P(\boldsymbol{x}(t+1)|\boldsymbol{x}(t)) = \prod_i \frac{e^{x_i(t+1)h_i(t+1)}}{2\cosh h_i(t+1)}, \tag{H1}$$

$$h_i(t+1) = \frac{1}{T}(H_i + \sum_j W_{ij}x_j(t)), \tag{H2}$$

where T is the thermodynamic temperature, H_i are external fields and $\boldsymbol{W} = (W_{ij})$ are the pairwise coupling strengths defined by a weighted network.

Furthermore, we recall that the sufficient thermodynamic quantities we aim to approximate are,

$$m_i(t) = \sum_{\boldsymbol{x}(t)} x_i(t) P(\boldsymbol{x}(t)), \tag{H3}$$

$$D_{il}(t) = \sum_{\boldsymbol{x}(t),\boldsymbol{x}(t-1)} x_i(t) x_l(t-1) P(\boldsymbol{x}(t), \boldsymbol{x}(t-1)) - m_i(t) m_l(t-1). \tag{H4}$$

where $\boldsymbol{m}(t) = (m_1(t), ..., m_N(t))$ is the average activation rate of the system and $\boldsymbol{D}(t)$ is the delayed correlation matrix.

Using the language of information geometry [126, 127], we define $\mathcal{P}(t)$ to be the manifold of $P(\boldsymbol{x}(t))$ where each point on the manifold corresponds to a set of parameter values. Within $\mathcal{P}(t)$, there are sub-manifolds $\mathcal{Q}(t)$ that are 'analytically tractable'. The simplest such manifold $\mathcal{Q}(t)$ is the manifold of models where each spin is independent. Each point on the sub-manifold is defined by a vector of parameters $\boldsymbol{\Theta}(t) = \{\Theta_i(t)\}$, and the distribution on this sub-manifold is given by,

$$P(\boldsymbol{x}(t)|\boldsymbol{\Theta}(t)) = \prod_i \frac{e^{x_i(t)\Theta_i(t)}}{2\cosh\Theta_i(t)}. \tag{H5}$$

The average activation rate of the spins is therefore given by,

$$m_i(t) = \tanh\Theta_i(t). \tag{H6}$$

Given local fields $\boldsymbol{H}$ and a network $\boldsymbol{W}$, we aim to approximate the thermodynamic quantities of the intractable target distribution $P(\boldsymbol{x}(t)|\boldsymbol{H}, \boldsymbol{W}) \in \mathcal{P}(t)$ with a tractable distribution from $\mathcal{Q}(t)$. To do this, we aim to find a distribution $Q(\boldsymbol{x}(t)) \in \mathcal{Q}(t)$ that minimises the KL-divergence to $P(\boldsymbol{x}(t)|\boldsymbol{H}, \boldsymbol{W})$. An important result is that the independent model that minimises the KL divergence, which we denote $Q(\boldsymbol{x}(t)|\boldsymbol{\Theta}^*(t))$, has identical activation rates to the target [125].

Next, we perform the so-called $\alpha-$projection and the Plefka expansion [127]. We parameterise a curve between a tractable distribution $Q(\boldsymbol{x}(t)|\boldsymbol{\Theta}(t))$ and the target distribution $P(\boldsymbol{x}(t)|\boldsymbol{H}, \boldsymbol{W})$ with a parameter $\alpha \in [0, 1]$ such that we have a family of distributions,

$$P_\alpha(x_i(t+1)|\boldsymbol{x}(t)) = \prod_i \frac{e^{x_i(t+1)h_i^\alpha(t+1)}}{2\cosh h_i^\alpha(t+1)}, \tag{H7}$$

$$h_i^\alpha(t+1) = (1-\alpha)\Theta_i(t) + \alpha\left(\frac{1}{T}(H_i + \sum_j W_{ij}x_j(t))\right). \tag{H8}$$

Therefore, at $\alpha = 0$, we have $P_0(\boldsymbol{x}(t+1)|\boldsymbol{x}(t)) = Q(\boldsymbol{x}(t+1)|\boldsymbol{\Theta}(t+1))$ and at $\alpha = 1$, we have $P_1(\boldsymbol{x}(t+1)|\boldsymbol{x}(t)) = P(\boldsymbol{x}(t+1)|\boldsymbol{x}(t))$. We can write the thermodynamic quantities of the distribution at each value of α as $\boldsymbol{m}^\alpha(t), \boldsymbol{D}^\alpha(t)$, functions of α.

The Plefka expansion is a Taylor expansion around $\alpha = 0$,

$$\boldsymbol{m}^\alpha(t) = \boldsymbol{m}^0(t) + \sum_{k=1}^{n} \frac{\alpha^k}{k!} \frac{\partial^k \boldsymbol{m}^0(t)}{\partial \alpha^k} + \mathcal{O}(\alpha^{(n+1)}), \tag{H9}$$

but we note that $\boldsymbol{m}^1(t) = \boldsymbol{m}^0(t)$ as mentioned earlier [125]. Thus, the optimal tractable distribution, $Q(\boldsymbol{x}(t)|\boldsymbol{\Theta}^*(t))$ satisfies,

$$\sum_{k=1}^{n} \frac{\alpha^k}{k!} \frac{\partial^k \boldsymbol{m}^0(t)}{\partial \alpha^k} = 0, \tag{H10}$$

which should be solved with respect to $\boldsymbol{\Theta}(t)$. The approximation is defined first by the number of terms in this sum that we set equal to 0 and then solve for, but also by the time-points at which the model is assumed to have independent units [85]. The NMF is obtained by setting only the first derivative to 0 and by assuming that the model at both t and $t-1$ have independent spins. The first derivative at 0 is given by,

$$\frac{\partial m_i^{\alpha=0}(t)}{\partial \alpha} = (1 - m_i(t)^2)\left(-\Theta_i(t) + H_i + \sum_j W_{ij} m_j(t-1)\right), \tag{H11}$$

where we direct the reader to the appendices of [85, 125] for further details. This gives an approximation of the optimal parameter setting,

$$\boldsymbol{\Theta}^*(t) \approx H_i + \sum_j W_{ij} m_j(t-1), \tag{H12}$$

and the NMF,

$$m_i(t) \approx \tanh\left(H_i + \sum_j W_{ij} m_j(t-1)\right). \tag{H13}$$

Similarly, one can expand $D_{il}^{\alpha}(t)$ around $\alpha = 0$ and set $\Theta_i(t) = \Theta_i^*(t)$ to obtain the approximation,

$$D_{il}(t) \approx W_{il}(1 - m_i^2(t))((1 - m_l^2(t-1)), \tag{H14}$$

where we direct again to the appendices of [85, 125] for further details.

Appendix I: Auto-regression with (un)constrained least-squares

In Section VII, we defined a method to quantify the EPR and infer an interaction network directly from a MVTS using a linear model. This model is either an unconstrained

$$\frac{d\boldsymbol{x}}{dt} = -\boldsymbol{B}\boldsymbol{x}(t) + \boldsymbol{\nu}(t), \tag{I1}$$

multivariate OU process, or one that was constrained to the edges of a network,

$$\frac{d\boldsymbol{x}}{dt} = (\boldsymbol{W} - \boldsymbol{I})\boldsymbol{x}(t) + \boldsymbol{\nu}(t). \tag{I2}$$

Using the time-discretisation presented in Section VII, we associate either model with an auto-regressive process of the form [60],

$$\boldsymbol{X}(t_{i+1}) = \boldsymbol{A}\boldsymbol{X}(t_i) + \boldsymbol{\chi}(t_i), \tag{I3}$$

where we find the coefficient matrix $\boldsymbol{A}$ using least-square regression. In the case of the unconstrained model, we simply solve the convex optimisation problem,

$$\min_{\boldsymbol{A}\in\mathbb{R}^N\times\mathbb{R}^N} ||\boldsymbol{X}_{1:T} - \boldsymbol{A}\boldsymbol{X}_{0:T-1}||^2, \tag{I4}$$

where $\boldsymbol{X}_{i_1:i_T}$ is an $N \times T$ matrix of data-points where each column is the multivariate observation taken from a time-point $t = i_1, ..., i_T$. Such a problem can be solved using any standard convex optimisation solver but is also available as a stand alone function in most scientific programming languages.

From $\boldsymbol{A}$, we can calculate the residuals at each time-step,

$$\boldsymbol{\chi}(t_{i-1}) = \boldsymbol{X}(t_i) - \boldsymbol{A}\boldsymbol{X}(t_{i-1}). \tag{I5}$$

We then estimate $\boldsymbol{B}$ and the noise covariance $\boldsymbol{D}$ by performing,

$$\boldsymbol{B} = \frac{1}{\Delta t}(\boldsymbol{I} - \boldsymbol{A}), \quad \boldsymbol{D} = \frac{1}{2\Delta t}\operatorname{Cov}[\boldsymbol{\chi}] \tag{I6}$$

for some choice of $\Delta t << 1$. Clearly, Δt simply scales the process and as such can be chosen arbitrarily to be 0.1.

In the case of the network-constrained model, we assume $\boldsymbol{W}$ has non-negative entries which means we must restrict our solution space. Using the relation,

$$\boldsymbol{W} = \frac{1}{\Delta t}(\boldsymbol{A} - \boldsymbol{I}) + \boldsymbol{I}, \tag{I7}$$

we solve the following optimisation problem,

$$\min_{\boldsymbol{W}:W_{ij}\geq 0} ||\boldsymbol{X}_{1:T} - ((1 - \Delta t)\boldsymbol{I} + \Delta t\boldsymbol{W})\boldsymbol{X}_{0:T-1}||^2, \tag{I8}$$

which remains convex. Additional constraints such as no self-loops or restriction to the existence of particular edges can be added without breaking the convexity of the problem. In addition, this particular problem can also be solved using built in non-negative least-squares algorithms in most programming languages subject to appropriate modification. We estimate the noise intensity using,

$$\sigma = \langle \frac{1}{2\Delta t}\operatorname{diag}\operatorname{Cov}[\boldsymbol{\chi}]\rangle, \tag{I9}$$

where $\langle\cdot\rangle$ is the mean. Again, the Δt simply scales the process and only affects the diagonal entries (self-loop) of the network, but does not affect the asymmetries.

Finally, we note that in order to define the EPR, we require that the process converges to a steady-state. This requires that $-\boldsymbol{B}$ or $\boldsymbol{W} - \boldsymbol{I}$ is a stable matrix i.e. all eigenvalues have negative real part. This constraint is non-convex and so cannot be enforced as part of the algorithm without resorting to more heuristic and

complex optimisation methods. In this study, we found that all time-series considered, were best fit with an unconstrained model that converged to a stationary state, which did not always hold for the constrained model. As a result the EPR estimates are obtained with the unconstrained model, whilst the network inference is done with the constrained model.

Appendix J: Empirical network data

In this study we consider 97 real-world directed networks from a range of different fields including ecology, sociology, biology, language, transport and economics. These network were compiled from a range of different sources and are reported in Table I.

TABLE I: Real-world directed network data from various sources.

Name	Ref.	Nodes	Edges
Ecological			
Marine Foodweb in Bahia Falsa, Mexico	[128]	166	9576
Marine Foodweb in Estero de Punta Banda, Mexico	[128]	143	3696
Marine Foodweb in Flensburg Fjord, Germany/Denmark	[128]	77	576
Marine Foodweb 1 in Ythan Estuary, Scotland	[128]	166	9029
Marine Foodweb in Carpinteria Salt Marsh Reserve, USA	[128]	166	7682
Marine Foodweb in Sylt Tidal Basin, Germany	[128]	215	14963
Marine Foodweb in Otago Harbour, New Zealand	[128]	215	15266
River Foodweb in Berwick Stream, New Zealand	[129]	77	240
River Foodweb 1 in Coweeta, USA	[129]	58	126
River Foodweb 2 in Coweeta, USA	[129]	71	148
River Foodweb in Martins Stream, USA	[129]	105	343
River Foodweb in Powder Stream, New Zealand	[129]	78	268
River Foodweb in Troy Stream, USA	[129]	77	181
River Foodweb in Venlaw Stream, New Zealand	[129]	66	187
River Foodweb in Black Rock Stream, New Zealand	[130]	86	375
River Foodweb in Broad Stream, New Zealand	[130]	94	564
River Foodweb in Dempsters Stream during summer, New Zealand	[130]	107	965
River Foodweb in German Creek, New Zealand	[130]	84	352
River Foodweb in Healy Creek, New Zealand	[130]	96	634
River Foodweb in Kye Burn, New Zealand	[130]	98	629
River Foodweb in Little Kye Burn, New Zealand	[130]	78	375
River Foodweb in Stony Stream, New Zealand	[130]	109	827
River Foodweb in Sutton Stream during summer, New Zealand	[130]	87	424
River Foodweb in Canton Creek, New Zealand	[130]	102	696
River Foodweb in Catlins Stream, New Zealand	[131]	48	110
River Foodweb in Dempsters Stream during autumn, New Zealand	[131]	83	414
River Foodweb in Dempsters Stream during spring, New Zealand	[131]	93	538
River Foodweb in Sutton Stream during autumn, New Zealand	[131]	80	335
River Foodweb in Sutton Stream during spring, New Zealand	[131]	74	391
River Foodweb in Narrowdale Stream, New Zealand	[132]	71	154
River Foodweb in North Col Stream, New Zealand	[132]	78	241
Terrestrial Foodweb in Scotch Broom, England	[133]	86	219
Marine Foodweb in Cayman Islands	[134]	242	3764
Marine Foodweb in Chesapeake Bay, USA	[135]	31	67
Dominance amongst ants	[136]	16	36
Dominance amongst kangaroos	[137]	17	91
Marine Foodweb in St. Marks Estuary, US	[138]	48	218
Terrestrial Foodweb in Saint-Martin Island, Lesser Antilles	[139]	42	205
Marine Foodweb 2 in Ythan Estuary, Scotland	[140]	82	391
Lake Foodweb in Lough Hyne, Ireland	[141]	349	5102
Marine Foodweb in Weddel Sea, Antarctica	[141]	483	15317
Fossil Assemblage Foodweb from Chengjiang Shale, China	[142]	33	90
Fossil Assemblage Foodweb from Burgess Shale, Canada	[142]	48	243
Lake Foodweb in Bridge Broom Lake	[143]	25	104

Dominance amongst wolves	[144]	16	148
Lake Foodweb in Little Rock Lake, USA 1	[145]	183	2476
Lake Foodweb in Little Rock Lake, USA 2	[146]	92	997
Marine Foodweb in Northeast United States Shelf	[147]	79	1378
Lake Foodweb in Skipwith Common, England	[148]	25	189
Marine Foodweb in Benguela Current, South Africa	[149]	29	196
Marine Foodweb in Florida Bay during dry season	[150]	128	2137
Dominance among ponies	[151]	17	146
Dominance among cattle	[152]	28	217
Dominance among sheep	[153]	28	250
Dominance among bison	[154]	26	314
Dominance among macaques	[155]	62	1187
Terrestrial Foodweb in grasslands of the United Kingdom	[146]	61	97
Terrestrial Foodweb in El Verde Field Station, Puerto Rico	[146]	155	1507
Terrestrial Foodweb in Coachella Valley, USA	[156]	29	262
Marine Foodweb in the Caribbean	[157]	155	1507
Sociological			
Political Blogs Network	[158]	1224	18957
Friendship among college students in a course about leadership	[159]	32	96
Friendship among highschool students	[146]	70	366
Co-purchased political books on Amazon	[160]	105	441
Social interactions between inmates in prison	[161]	67	182
Social interactions between inmates in prison	[161]	67	182
Biological			
Protein network for 1A4J	[159]	95	404
Protein network for 1AOR	[159]	96	406
Protein network for 1EAW	[159]	53	236
Gene regulatory network for *Saccharomyces cerevisiae*	[162]	2933	6152
Human gene regulatory network for a healthy person	[163]	4071	8466
Human gene regulatory network for a person with cancer	[163]	4049	11707
Gene regulatory network for *Pseudomonas aeruginosa*	[164]	691	991
Gene regulatory network for Mycobacterium tuberculosis	[165]	1624	3169
Neuronal network for a mouse brain	[146]	213	21654
Connectome of the Rhesus brain, extracted from tract tracing	[166]	242	4090
Connectome of the Rhesus brain via retrograde tracer	[167]	91	628
Neuronal network for *Caenorhabditis elegans*	[168]	297	2345
Connectome of the cat brain	[169]	65	1139
Connectome of the rat brain	[170]	503	47329
Metabolic network of *Archaeoglobus fulgidus*	[171]	1267	3011
Metabolic network of *Caenorhabditis elegans*	[171]	1172	2864
Metabolic network of *Chlamydia pneumoniae*	[171]	386	792
Metabolic network of *Chlamydia trachomatis*	[171]	446	941
Metabolic network of *Methanococcus jannaschii*	[171]	1081	2589
Metabolic network of *Saccharomyces cerevisiae*	[171]	1510	3833
Metabolic network of *Methanobacterium thermoautotrophicum*	[171]	1111	2705
Language			
Citations from papers that cite "Small World Problem"	[172]	233	994
Citations to Small, Griffith and descendants	[172]	1024	4918
Word adjacency network for Dr. Seuss's Green Eggs and Ham book	[146]	50	101
Trade			
International trade network of minerals	[173]	24	135
International trade network of manufactured food products	[173]	24	307
International trade network of manufactured goods	[173]	24	310
International trade network of crude animal and vegetable material	[173]	24	307
International trade network of diplomatic exchanges	[173]	24	369
Transport			
London tube network	[174]	270	628
Paris metropolitan train grid	[174]	302	705

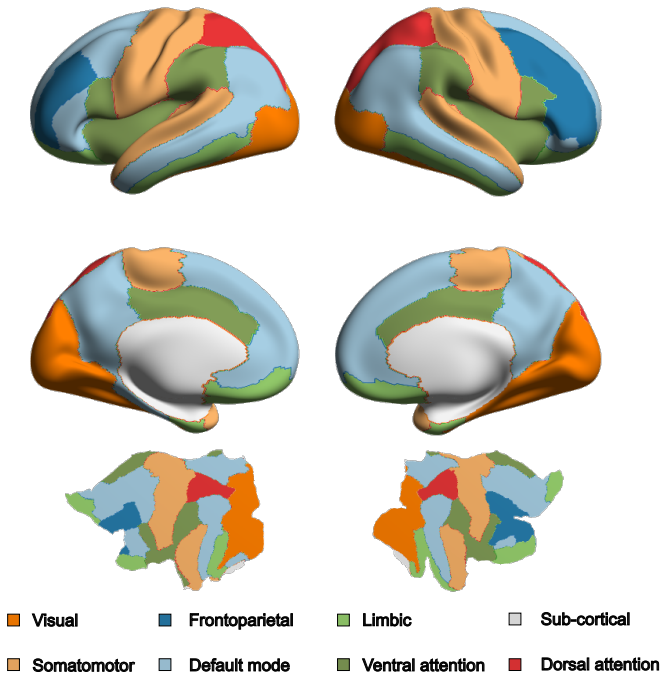

FIG. 14. **Yeo parcellation projected onto the connectome:** The 7 canonical resting state networks are show projected onto the DK80 parcellation with the sub-cortical regions also labelled.

Appendix K: Empirical time-series data

1. Human neuroimaging

In Section VII we analysed functional magnetic resonance imaging (fMRI) from 100 unrelated participants at rest and during task. This data is freely available as part of the Human Connectome Project HCP1003 release. The data used here is the same as that analysed previously in Ref. [112]. In brief, we used the Desikan-Killany parcellation [111] made up of 62 cortical regions and 18 sub-cortical regions for a total of 80 regions of interest (DK80). The data was pre-processed using the HCP pipeline using standard software packages from the FMRIB Software Library, FreeSurfer and the Connectome Workbench [175, 176]. This included correcting for spatial and gradient distortions, head motions and further included bias-field removal, intensity normalisation, registration to a T1-structural image, transformation to 2mm MNI (Montreal Neurological Institute) space and application of the FIX artefact removal procedure [176, 177]. Head motion was regressed out and, using independent component analysis, artefacts were removed using ICA+FIX processing [178, 179]. Using the Fieldtrip toolbox [180], the average time-series of the grayordinates in each region of the DK80 parcellation was extracted and the resulting BOLD signal was filtered in the range of 0.008-0.08 Hz using a second order Butterworth filter.

We further parcellate the DK80 into 8 sub-networks, the 7 Yeo resting-state networks [113], and the sub-cortical regions. The projection of the 8 sub-networks onto the DK80 parcellation is shown in Figure 14.

2. Stock-prices from the New York Stock Exchange

In Section VII, we analysed stock prices from 119 U.S. companies from the New York Stock Exchange (NYSE) in the period 1 January 2000 to 17 June 2021. This is the same financial dataset previously studied in Ref. [92]. It was obtained from the Yahoo! finance historical data application programming interface ('yfinance' Python library) but is freely deposited online in Ref. [120].

[1] J. A. Dunne, R. J. Williams, and N. D. Martinez, Food-web structure and network theory: The role of connectance and size, Proceedings of the National Academy of Sciences of the United States of America **99**, 12917 (2002).
[2] M. O. Jackson, *Social and Economic Networks* (Princeton University Press, 2010).
[3] S. Wasserman, *Social Network Analysis* (Cambridge University Press, 1994).
[4] A.-L. Barabási, N. Gulbahce, , and J. Loscalzo, Network medicine: A network-based approach to human disease, Nature Reviews Genetics **12**, 56 (2011).
[5] E. Bullmore and O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, Nature Reviews Neuroscience **10**, 186 (2009).
[6] D. S. Bassett and O. Sporns, Network neuroscience, Nature Neuroscience **20**, 353–364 (2017).
[7] M. Newman, *Networks* (Oxford University Press, 2018).
[8] M. E. J. Newman, The structure and function of complex networks, SIAM Review **45**, 167 (2003).
[9] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, Complex networks: Structure and dynamics, Physics Reports **424**, 175 (2006).
[10] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).
[11] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes* (John Wiley & Sons, California, United States, 1968).
[12] S. Carnot, *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance* (Bachelier, Paris, 1824).
[13] E. Schrödinger, *What is Life? The Physical Aspect of the Living Cell and Mind* (Cambridge University Press, Cambridge, United Kingdom, 1944).

[14] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Physical Review E **60** (1999).

[15] C. Jarzynski, Nonequilibrium equality for free energy differences, Physical Review Letters **78** (1997).

[16] E. H. Feng and G. E. Crooks, Length of time's arrow, Physical Review Letters **101** (2008).

[17] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Broken detailed balance and non-equilibrium dynamics in living systems: a review, Reports on Progress in Physics **81** (2018).

[18] C. P. Brangwynne, G. H. Koenderink, F. C. MacKintosh, and D. A. Weitz, Cytoplasmic diffusion: molecular motors mix it up, Journal of Cell Biology **183**, 583 (2008).

[19] H. Yin, I. Artsimovitch, R. Landick, and J. Gelles, Nonequilibrium mechanism of transcription termination from observations of single RNA polymerase molecules, Proceedings of the National Academy of Sciences of the United States of America **96**, 13124 (1999).

[20] K. C. Huang, Y. Meir, and N. S. Wingreen, Dynamic structures in escherichia coli: Spontaneous formation of MinE rings and MinD polar zones, Proceedings of the National Academy of Sciences of the United States of America **100**, 12724 (2003).

[21] P. Mehta and D. J. Schwab, Energetic costs of cellular computation, Proceedings of the National Academy of Sciences of the United States of America **109**, 17978 (2012).

[22] B. Stuhrmann, M. S. E. Silva, M. Depken, F. C. Mackintosh, and G. H. Koenderink, Nonequilibrium fluctuations of a remodeling in vitro cytoskeleton, Physical Review E **86** (2012).

[23] C. Battle, C. P. Broedersz, N. Fakhri, V. F. Geyer, J. Howard, C. F. Schmidt, and F. C. Mackintosh, Broken detailed balance at mesoscopic scales in active biological systems, Science **352**, 604 (2016).

[24] J. L. England, Statistical physics of self-replication, Journal of Chemical Physics **139** (2013).

[25] C. W. Lynn, E. J. Cornblath, L. Papadopoulos, and D. S. Bassett, Broken detailed balance and entropy production in the human brain, Proceedings of the National Academy of Sciences of the United States of America **118** (2021).

[26] G. Deco, Y. Sanz-Perl, L. de la Fuente, J. D. Sitt, B. T. T. Yeo, E. Tagliazucchi, and M. L. Kringelbach, The arrow of time of brain signals in cognition: Potential intriguing role of parts of the default mode network, Network Neuroscience **7**, 966–998 (2023).

[27] T. Bolton, D. Van De Ville, E. Amico, M. Preti, and R. Liégeois, The arrow-of-time in neuroimaging time series identifies causal triggers of brain function, Human Brain Mapping **44**, 4077 (2023).

[28] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Reports of Progress in Physics **475** (2012).

[29] M. Aguilera, M. Igarashi, and H. Shimazaki, Nonequilibrium thermodynamics of the asymmetric Sherrington-Kirkpatrick model, Nature Communications **14** (2023).

[30] T. Herpich, J. Thingna, and M. Esposito, Collective power: minimal model for thermodynamics of nonequilibrium phase transitions, Physical Review X **8** (2018).

[31] T. Herpich, T. Cossetto, G. Falasco, and M. Esposito, Stochastic thermodynamics of all-to-all interacting many-body systems, New Journal of Physics **22** (2020).

[32] M. Suñé and A. Imparato, Out-of-equilibrium clock model at the verge of criticality, Physical Review Letters **123** (2019).

[33] J.Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation systems, Reviews of Modern Physics **48** (1976).

[34] G. Oster, A. Perelson, and A. Katchalsky, Network thermodynamics, Nature **234**, 393 (1971).

[35] R. Rao and M. Esposito, Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics, Physical Review X **6** (2016).

[36] M. Polettini, A. Wachtel, and M. Esposito, Dissipation in noisy chemical networks: The role of deficiency, Journal of Chemical Physics **143** (2015).

[37] S. D. Cengio, V. Lecomte, and M. Polettini, Geometry of nonequilibrium reaction networks, Physical Review X **13** (2023).

[38] C. W. Lynn, C. M. Holmes, W. Bialek, and D. J. Schwab, Decomposing the local arrow of time in interacting systems, Physical Review Letters **129** (2022).

[39] D. Papo and J. Buldú, Does the brain behave like a (complex) network? I. Dynamics, Physics of Life Reviews **48**, 47 (2024).

[40] D. C. Krakauer, Symmetry–simplicity, broken symmetry–complexity, Interface Focus **13** (2023).

[41] S. Johnson, Digraphs are different: why directionality matters in complex systems, Journal of Physics: Complexity **1** (2020).

[42] M. Fruchart, R. Hanai, P. B. Littlewood, and V. Vitelli, Non-reciprocal phase transitions, Nature **592**, 363 (2021).

[43] R. Muolo, T. Carletti, J. P. Gleeson, and M. Asllani, Synchronization dynamics in non-normal networks: The trade-off for optimality, Entropy **23** (2020).

[44] R. Muolo, J. D. O'Brien, T. Carletti, and M. Asllani, Persistence of chimera states and the challenge for synchronization in real-world networks, The European Physical Journal B **97**, 6 (2024).

[45] M. Asllani and T. Carletti, Topological resilience in non-normal networked systems, Physical Review E **97** (2018).

[46] S. Nicoletti, D. Fanelli, N. Zagli, M. Asllani, G. Battistelli, T. Carletti, L. Chisci, G. Innocenti, and R. Livi, Resilience for stochastic systems interacting via a quasi-degenerate network, Chaos: An Interdisciplinary Journal of Nonlinear Science **29**, 083123 (2019).

[47] M. Asllani, J. D. Challenger, F. S. Pavone, L. Sacconi, and D. Fanelli, The theory of pattern formation on directed networks, Nature Communications **5** (2014).

[48] R. Muolo, M. Asllani, D. Fanelli, P. K. Maini, and T. Carletti, Patterns of non-normality in networked systems, Journal of Theoretical Biology **480**, 81 (2019).

[49] S. Johnson, V. Domínguez-García, L. Donetti, and M. A. Muñoz, Trophic coherence determines food-web stability, Proceedings of the National Academy of Sciences **111**, 17923 (2014).

[50] M. Asllani, R. Lambiotte, and T. Carletti, Structure and dynamical behavior of non-normal networks, Science Advances **4** (2018).

[51] R. MacKay, S. Johnson, and B. Sansom, How directed is a directed network?, Royal Society Open Science **7** (2020).

[52] J. O'Brien, K. Oliveira, J. Gleeson, and M. Asllani, Hierarchical route to the emergence of leader nodes in real-world networks, Physical Review Research **3** (2021).

[53] N. Rodgers, P. Tiňo, and S. Johnson, Strong connectivity in real directed networks, Proceedings of the National Academy of Sciences of the United States of America **120** (2023).

[54] L. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behaviour of Non-Normal Matrices and Operators* (Princeton University Press, 2005).

[55] N. Masuda, M. Porter, and R. Lambiotte, Random walks and diffusion on networks, Physics Reports **716-717**, 1 (2017).

[56] R. Lambiotte, J.-C. Delvenne, and M. Barahona, Random walks, Markov processes and the multiscale modular organization of complex networks, IEEE Transactions on Network Science and Engineering **1**, 76 (2014).

[57] C. Godrèche and J. Luck, Characterising the nonequilibrium stationary states of Ornstein–Uhlenbeck processes, Journal of Physics A: Mathematical and Theoretical **52** (2018).

[58] H. Nishimori, *Statistical physics of spin glasses and information processing: an introduction* (Clarendon Press, 2001).

[59] P. Erdös and A. Rényi, On random graphs I, Publicationes Mathematicae Debrecen **6**, 290 (1959).

[60] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications* (Springer, 2017).

[61] G. Deco, Y. Sanz-Perl, H. Bocaccio, E. Tagliazucchi, and M. L. Kringelbach, The INSIDEOUT framework provides precise signatures of the balance of intrinsic and extrinsic dynamics in brain states, Communications Biology **5** (2022).

[62] G. Deco, C. Lynn, Y. Sanz-Perl, and M. L. Kringelbach, Violations of the fluctuation-dissipation theorem reveal distinct non-equilibrium dynamics of brain states, Physical Review E **108** (2023).

[63] P. Kale, A. Zalesky, and L. Gollo, Estimating the impact of structural directionality: How reliable are undirected connectomes?, Network Neuroscience **2**, 259 (2018).

[64] K. J. Friston, E. D. Fagerholm, T. S. Zarghami, T. Parr, I. Hipólito, L. Magrou, and A. Razi, Parcels and particles: Markov blankets in the brain, Network Neuroscience **5**, 211 (2021).

[65] B. Corominas-Murtra, J. Goñi, R. V. Solé, and C. Rodríguez-Caso, On the origins of hierarchy in complex networks, Proceedings of the National Academy of Sciences of the United States of America **110**, 13316 (2013).

[66] We consider the (ir)reciprocity of weighted networks as defined in Ref. [68]. Alternative definitions of reciprocity for unweighted graphs are given in Ref. [7, 67].

[67] D. Garlaschelli and M. I. Loffredo, Patterns of link reciprocity in directed networks, Physical Review Letters **93** (2004).

[68] T. Squartini, F. Picciolo, F. Ruzzenenti, and D. Garlaschelli, Reciprocity of weighted networks, Scientific Reports **3** (2013).

[69] C. Voisin, *Hodge Theory and Complex Algebraic Geometry* (Cambridge University Press, 2010).

[70] Equivalently, the 'SpringRank' formulation considers directed springs between nodes and aims to find a ranking that minimizes the total energy of these springs [181].

[71] G. Hennequin, T. P. Vogels, and W. Gerstner, Non-normal amplification in random balanced neuronal networks, Physical Review E **86** (2012).

[72] G. Baggio, V. Rutten, G. Hennequin, and S. Zampieri, Efficient communication over complex dynamical networks: The role of matrix non-normality, Science Advances **6** (2020).

[73] G. Lindmark and C. Altafini, Centrality measures and the role of non-normality for network control energy reduction, IEEE Control Systems Letters **5**, 1013 (2021).

[74] C. Duan, T. Nishikawa, D. Eroglu, and A. E. Motter, Network structural origin of instabilities in large complex systems, Science Advances **8** (2022).

[75] R. Lambiotte, Continuous-time random walks and temporal networks, in *Temporal Network Theory* (Springer, 2023) pp. 225–239.

[76] A. Schwarze and M. Porter, Motifs for processes on networks, SIAM Journal on Applied Dynamical Systems **20** (2021).

[77] U. Seifert, Entropy production along a stochastic trajectory and an integral fluctuation theorem, Physical Review Letters **95** (2005).

[78] F. Nielsen, On the Jensen–Shannon symmetrization of distances relying on abstract means, Entropy **21** (2019).

[79] A. Dechant, Minimum entropy production, detailed balance and Wasserstein distance for continuous-time markov processes, Journal of Physics A: Mathematical and Theoretical **55** (2022).

[80] G. Uhlenbeck and L. Ornstein, On the theory of the Brownian motion, Physical Review **36** (1930).

[81] J. Zabczyk, *Mathematical Control Theory: An Introduction* (Springer, 2020).

[82] M. Lax, Fluctuations from the nonequilibrium steady state, Reviews of Modern Physics **32** (1960).

[83] V. Simoncini, Computational methods for linear matrix equations, SIAM Review **58** (2016).

[84] H. Huang and Y. Kabashima, Dynamics of asymmetric kinetic Ising systems revisited, Journal of Statistical Mechanics: Theory and Experiment (2014).

[85] M. Aguilera, S. A. Moosavi, and H. Shimazaki, A unifying framework for mean-field theories of asymmetric kinetic Ising systems, Nature Communications **58** (2021).

[86] M. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Oxford University Press, 1999).

[87] M. Esposito, Stochastic thermodynamics under coarse graining, Physical Review E **85** (2012).

[88] E. Roldán and J. M. R. Parrondo, Estimating dissipation from single stationary trajectories, Physical Review Letters **105** (2010).

[89] T. Martynec, S. H. L. Klapp1, and S. A. M. Loos, Entropy production at criticality in a nonequilibrium Potts model, New Journal of Physics **22** (2020).

[90] G. Bianconi, N. Gulbahce, and A. E. Motter, Local structure of directed networks, Physical Review Letters **100** (2008).

[91] S. Johnson and N. S. Jones, Looplessness in networks is linked to trophic coherence, Proceedings of the National Academy of Sciences of the United States of America **114**, 5618 (2017).

[92] A. Santoro, F. Battiston, G. Petri, and E. Amico, Higher-order organization of multivariate time series,

Nature Physics **19**, 221–229 (2023).
[93] O. M. Cliff, A. G. Bryant, J. T. Lizier, N. Tsuchiya, and B. D. Fulcher, Unifying pairwise interactions in complex dynamics, Nature Computational Science **3**, 883 (2023).
[94] A. Seif, M. Hafezi, and C. Jarzynski, Machine learning the thermodynamic arrow of time, Nature Physics **17**, 105 (2021).
[95] W. Gilpin, Generative learning for nonlinear dynamics, Nature Reviews Physics (2024).
[96] A. Frishman and P. Ronceray, Learning force fields from stochastic trajectories, Physical Review X **10** (2020).
[97] M. Timme and J. Casadiego, Revealing networks from dynamics: an introduction, Journal of Physics A: Mathematical and Theoretical **47** (2014).
[98] L. Novelli, P. Wollstadt, P. Mediano, M. Wibral, and J. T. Lizier, Large-scale directed network inference with multivariate transfer entropy and hierarchical statistical testing, Network Neuroscience **3**, 827 (2019).
[99] C. Zou and J. Feng, Granger causality vs. dynamic Bayesian network inference: a comparative study, BMC Bioinformatics **10** (2009).
[100] C. Merger, A. René, K. Fischer, P. Bouss, S. Nestler, D. Dahmen, C. Honerkamp, and M. Helias, Learning interacting theories from data, Physical Review X **13** (2023).
[101] M. Gilson, R. Moreno-Bote, A. Ponce-Alvarez, P. Ritter, and G. Deco, Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome, PLOS Computational Biology **12** (2016).
[102] M. Gilson, G. Zamora-López, V. Pallarés, M. H. Adhikari, M. Senden, A. T. Campo, D. Mantini, M. Corbetta, G. Deco, and A. Insabato, Model-based whole-brain effective connectivity to study distributed cognition in health and disease, Network Neuroscience **4**, 338 (2020).
[103] A. F. Villaverde, J. Ross, F. Morán, and J. R. Banga, Mider: Network inference with mutual information distance and entropy reduction, PLOS ONE **9** (2014).
[104] E. Nozari, M. A. Bertolero, J. Stiso, L. Caciagli, E. J. Cornblath, X. He, A. S. Mahadevan, G. J. Pappas, and D. S. Bassett, Macroscopic resting-state brain dynamics are best described by linear models, Nature Biomedical Engineering (2023).
[105] L. Harrison, W. Penny, and K. Friston, Multivariate autoregressive modeling of fMRI time series, NeuroImage **19**, 1477 (2003).
[106] M. Gilson, E. Tagliazucchi, and R. Cofré, Entropy production of multivariate Ornstein-Uhlenbeck processes correlates with consciousness levels in the human brain, Physical Review E **107** (2023).
[107] D. Benozzo, G. Baggio, G. Baron, A. Chiuso, S. Zampieri, and A. Bertoldo, Analyzing asymmetry in brain hierarchies with a linear state-space model of resting-state fMRI data, bioRxiv (2023).
[108] P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, 1992).
[109] D. Chen and R. Plemmons, Nonnegativity constraints in numerical analysis, in *The Birth of Numerical Analysis* (World Scientific, 2009) pp. 109–139.
[110] D. Van Essen, S. Smith, D. Barch, T. Behrens, E. Yacoub, K. Ugurbil, and for the WU-Minn HCP Consortium, The WU-Minn Human Connectome Project: An overview, NeuroImage **80**, 62 (2013).
[111] R. Desikan, F. Ségonne, B. Fischl, B. Quinn, B. Dickerson, D. Blacker, R. Buckner, A. Dale, R. Maguire, B. Hyman, M. Albert, and R. Killiany, An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest, Neuroimage **31**, 968 (2006).
[112] G. Deco, D. Vidaurre, and M. L. Kringelbach, Revisiting the global workspace orchestrating the hierarchical organization of the human brain, Nature Human Behaviour **5**, 497 (2021).
[113] B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner, The organization of the human cerebral cortex estimated by intrinsic functional connectivity, Journal of Neurophysiology **106**, 1125 (2011).
[114] M. L. Kringelbach, Y. Sanz-Perl, E. Tagliazucchi, and G. Deco, Toward naturalistic neuroscience: Mechanisms underlying the flattening of brain hierarchy in movie-watching compared to rest and task, Science Advances **9** (2023).
[115] S. Dehaene, M. Kerszberg, and J.-P. Changeux, A neuronal model of a global workspace in effortful cognitive tasks, Proceedings of the National Academy of Sciences of the United States of America **95**, 14529 (1998).
[116] S. Sieniutycz and P. Salamon, *Finite-time thermodynamics and thermoeconomics* (Taylor & Francis, 1990).
[117] V. N. Pokrovskii, *Thermodynamics of Complex Systems: Principles and applications* (IOP Publishing, 2020).
[118] N.-F. Chen, R. Roll, and S. A. Ross, Economic forces and the stock market, The Journal of Business **59**, 383 (1986).
[119] P. Moretti and M. A. Muñoz, Griffiths phases and the stretching of criticality in brain networks, Nature Communications **4** (2013).
[120] A. Santoro, F. Battiston, G. Petri, and E. Amico, Higher-order organization of multivariate time series, Zenodo https://doi.org/10.5281/zenodo.7210076 (2022).
[121] D. J. de Solla Price, Networks of scientific papers, Science **149**, 510 (1965).
[122] D. J. de Solla Price, A general theory of bibliometric and other cumulative advantage processes, Journal of the American Society for Information Science , 292 (1976).
[123] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, Science **286**, 509 (1999).
[124] S. Lamrous and M. Taileb, Divisive hierarchical k-means, IEEE International Conference on Computational Inteligence for Modelling Control and Automation (2006).
[125] H. J. Kappen and J. J. Spanjers, Mean field theory for asymmetric neural networks, Physical Review E **61** (2000).
[126] T. Tanaka, Information geometry of mean-field approximation, in *Advanced mean field methods: Theory and practice* (MIT Press, 2001) pp. 351–360.
[127] S. Amari, S. Ikeda, and H. Shimokawa, Information geometry of alpha-projection in mean field approximation, in *Advanced mean field methods: Theory and practice* (MIT Press, 2001) pp. 241–257.
[128] J. A. Dunne, K. D. Lafferty, A. P. Dobson, R. F. Hechinger, A. M. Kuris, N. D. Martinez, J. P. McLaughlin, K. N. Mouritsen, R. Poulin, K. Reise, D. B. Stouf-

fer, D. W. Thieltges, R. J. Williams, and C. D. Zander, Parasites affect food web structure primarily through increased diversity and complexity, PLOS Biology **11** (2013).

[129] R. Thompson and C. Townsend, Impacts on stream food webs of native and exotic forest: an intercontinental comparison, Ecology **84**, 145 (2003).

[130] C. R. Townsend, R. Thompson, A. McIntosh, C. Kilroy, E. Edwards, and M. R. Scarsbrook, Disturbance, resource supply, and food-web architecture in streams, Ecology Letters **1**, 200 (1998).

[131] J. Klaise and S. Johnson, The origin of motif families in food webs, Scientific Reports **7** (2017).

[132] R. M. Thompson and C. R. Townsend, Energy availability, spatial heterogeneity and ecosystem size predict food-web structure in streams, OIKOS **108**, 137 (2004).

[133] J. Memmott, N. Martinez, and J. Cohen, Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web, Journal of Animal Ecology **69**, 1 (2001).

[134] J. Bascompte, C. J. Melián, and E. Sala, Interaction strength combinations and the overfishing of a marine food web, Proceedings of the National Academy of Sciences of the United States of America **102**, 5443 (2005).

[135] R. E. Ulanowicz and D. Baird, Nutrient controls on ecosystem dynamics: the chesapeake mesohaline community, Journal of Marine Systems **19**, 159 (1999).

[136] B. J. Cole, Dominance hierarchies in leptothorax ants, Science **212**, 83 (1981).

[137] T. Grant, Dominance and association among members of a captive and a free-ranging group of grey kangaroos (macropus giganteus), Animal Behaviour **21**, 449 (1973).

[138] R. R. Christian and J. J. Luczkovich, Organizing and understanding a winter's seagrass foodweb network through effective trophic levels, Ecological Modelling **117**, 99 (1999).

[139] L. Goldwasser and J. Roughgarden, Construction and analysis of a large Caribbean food web, Ecology **74**, 1216 (1993).

[140] M. Huxham, S. Beaney, and D. Raffaelli, Do parasites reduce the chances of triangulation in a real food web?, OIKOS **76**, 284 (1996).

[141] A. Eklöf, U. Jacob, J. Kopp, J. Bosch, R. Castro-Urgal, N. P. Chacoff, B. Dalsgaard, C. de Sassi, M. Galetti, P. R. Guimarães, S. B. Lomáscolo, A. M. M. González, M. A. Pizo, R. Rader, A. Rodrigo, J. M. Tylianakis, D. P. Vázquez, and S. Allesina, The dimensionality of ecological networks, Ecology Letters **16**, 577 (2013).

[142] J. A. Dunne, R. J. Williams, N. D. Martinez, R. A. Wood, and D. H. Erwin, Compilation and network analyses of Cambrian food webs, PLOS Biology **6** (2008).

[143] K. Havens, Scale and structure in natural food webs, Science **257**, 1107 (1992).

[144] J. A. van Hooff and J. A. B. Wensing, in *Man and Wolf: Advances, Issues, and Problems in Captive Wolf Research* (Springer Science & Business Media, 1987).

[145] N. D. Martinez, Artifacts or attributes? Effects of resolution on the Little Rock Lake food web, Ecological Monographs **61**, 367 (1991).

[146] S. Johnson, Network data repository from various sources.

[147] J. Link, Does food web theory work for marine ecosystems?, Marine Ecology Progress Series **230**, 1 (2002).

[148] P. H. Warren, Spatial and temporal variation in the structure of a freshwater food web, OIKOS **55**, 299 (1989).

[149] P. Yodzis, Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem, Journal of Animal Ecology **67**, 635 (1998).

[150] R. Ulanowicz and D. L. DeAngelis, Network analysis of trophic dynamics in South Florida ecosystems, in *U.S. Geological Survey Programon the South Florida Ecosystem; Proceedings of South Florida Restoration Science-Forum* (BiblioGov, 1999).

[151] T. H. Clutton-Brock, P. J. Greenwood, and R. P. Powell, Ranks and relationships in highland ponies and highland cows, Zeitschrift für Tierpsychologie **41**, 202 (1976).

[152] M. W. Schein and M. H. Fohrman, Social dominance relationships in a herd of dairy cattle, The British Journal of Animal Behaviour **3**, 45 (1955).

[153] C. C. Hass, Social status in female bighorn sheep (Ovis canadensis): expression, development and reproductive correlates, Journal of Zoology **225**, 509 (1991).

[154] D. F. Lott, Dominance relations and breeding rate in mature male American bison, Zeitschrift für Tierpsychologie **49**, 418 (1979).

[155] F. F. Strayer and M. S. Cummins, *Dominance Relations: An Ethological View of Human Conflict and Social Interaction* (Livingstone, 1980).

[156] G. A. Polis, Complex trophic interactions in deserts: An empirical critique of food-web theory, The American Naturalist **138**, 123 (1991).

[157] S. Opitz, Trophic interactions in Caribbean coral reefs, International Center for Living Aquatic Resources Management (1996).

[158] L. A. Adamic and N. Glance, The political blogosphere and the 2004 U.S. election: divided they blog, LinkKDD '05: Proceedings of the 3rd international workshop on Link discovery , 36 (2005).

[159] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Superfamilies of evolved and designed networks, Science **303**, 1538 (2004).

[160] A. Clauset, E. Tucker, and M. Sainz, The Colorado index of complex networks, (2016).

[161] J. Duncan MacRae, Direct factor analysis of sociometric data, Sociometry **23**, 360 (1960).

[162] C. T. Harbison, D. B. Gordon, T. I. Lee, N. J. Rinaldi, K. D. Macisaac, T. W. Danford, N. M. Hannett, J.-B. Tagne, D. B. Reynolds, J. Yoo, E. G. Jennings, J. Zeitlinger, D. K. Pokholok, M. Kellis, P. A. Rolfe, K. T. Takusagawa, E. S. Lander, D. K. Gifford, E. Fraenkel, and R. A. Young, Transcriptional regulatory code of a eukaryotic genome, Nature **431**, 99– (2004).

[163] M. B. Gerstein, A. Kundaje, M. Hariharan, S. G. Landt, K.-K. Yan, C. Cheng, X. J. Mu, E. Khurana, J. Rozowsky, R. Alexander, R. Min, P. Alves, A. Abyzov, N. Addleman, N. Bhardwaj, A. P. Boyle, P. Cayting, D. Z. C. Alexandra Charos, Y. Cheng, D. Clarke, C. Eastman, G. Euskirchen, S. Frietze, Y. Fu, J. Gertz, F. Grubert, A. Harmanci, P. Jain, M. Kasowski, P. Lacroute, J. Leng, J. Lian, H. Monahan, H. O'Geen, Z. Ouyang, E. C. Partridge, D. Patacsil, F. Pauli, D. Raha, L. Ramirez, T. E. Reddy, B. Reed, M. Shi, T. Slifer, J. Wang, L. Wu, X. Yang, K. Y. Yip,

G. Zilberman-Schapira, S. Batzoglou, A. Sidow, P. J. Farnham, R. M. Myers, S. M. Weissman, and M. Snyder, Architecture of the human regulatory network derived from ENCODE data, Nature **489**, 91 (2012).

[164] E. Galán-Vásquez, B. Luna, and A. Martínez-Antonio, The regulatory network of Pseudomonas aeruginosa, Microbial Informatics and Experimentation **1** (2011).

[165] J. Sanz, J. Navarro, A. Arbués, C. Martín, P. C. Marijuán, and Y. Moreno, The transcriptional regulatory network of Mycobacterium tuberculosis, PLOS One **6** (2011).

[166] L. Harriger, M. P. van den Heuvel, and O. Sporns, Rich club organization of macaque cerebral cortex and its role in network communication, PLOS One **7** (2012).

[167] N. T. Markov, M. Ercsey-Ravasz, C. Lamy, A. R. R. Gomes, L. Magrou, P. Misery, P. Giroud, P. Barone, C. Dehay, Z. Toroczkai, K. Knoblauch, D. C. V. Essen, , and H. Kennedy, The role of long-range connections on the specificity of the macaque interareal cortical network, Proceedings of the National Academy of Sciences of the United States of America **110**, 5187 (2013).

[168] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature **393**, 440 (1998).

[169] M. A. de Reus and M. P. van den Heuvel, Rich club organization and intermodule communication in the cat connectome, Journal of Neuroscience **33**, 12929 (2013).

[170] M. Bota and L. W. Swanson, Online workbenches for neural network connections, Journal of Comparative Neurology **500**, 807 (2006).

[171] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. Barabási, The large-scale organization of metabolic networks, Nature **407**, 651 (2000).

[172] E. Garfield, Index of citation networks produced by analyses from the software HistCite.

[173] W. D. Nooy, A. Mrvar, and V. Batagelj, *Exploratory Social Network Analysis with Pajek* (Cambridge University Press, 2018).

[174] M. J. Williams and M. Musolesi, Spatio-temporal networks: reachability, centrality and robustness, Royal Society Open Science **3** (2016).

[175] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. V. Essen, M. Jenkinson, and for the WU-Minn HCP Consortium, The minimal preprocessing pipelines for the Human Connectome Project, NeuroImage **18**, 105 (2013).

[176] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, M. Kelly, T. Laumann, K. L. Miller, S. Moeller, S. Petersen, J. Power, G. Salimi-Khorshidi, A. Z. Snyder, A. T. Vu, M. W. Woolrich, J. Xu, E. Yacoub, K. Uğurbil, D. C. V. Essen, M. F. Glasser, and for the WU-Minn HCP Consortium, Resting-state fMRI in the Human Connectome Project, NeuroImage **80**, 144 (2013).

[177] T. N. Schröder, K. V. Haak, N. I. Z. Jimenez, C. F. Beckmann, and C. F. Doeller, Functional topography of the human entorhinal cortex, eLife **4** (2015).

[178] G. Salimi-Khorshidi, G. Douaud, C. F. Beckmann, M. F. Glasser, L. Griffanti, and S. M. Smith, Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers, NeuroImage **90**, 449 (2014).

[179] L. Griffanti, G. Salimi-Khorshidi, C. F. Beckmann, E. J. Auerbach, G. Douaud, C. E. Sexton, E. Zsoldos, K. P. Ebmeier, N. Filippini, C. E. Mackay, S. Moeller, J. Xu, E. Yacoub, G. Baselli, K. Ugurbil, K. L. Miller, and S. M. Smith, ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging, NeuroImage **95**, 232 (2014).

[180] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data, Computational Intelligence and Neuroscience **2011** (2011).

[181] C. D. Bacco, D. B. Larremore, and C. Moore, A physical model for efficient ranking in networks, Science Advances **4** (2018).