

Duality between predictability and reconstructability in complex systems

Charles Murphy,^{1,2,*} Vincent Thibeault,^{1,2} Antoine Allard,^{1,2} and Patrick Desrosiers^{1,2,3}

¹*Département de physique, de génie physique et d'optique,
Université Laval, Québec (Québec), Canada G1V 0A6*

²*Centre interdisciplinaire en modélisation mathématique,
Université Laval, Québec (Québec), Canada G1V 0A6*

³*Centre de recherche CERVO, Québec (Québec), Canada G1J 2G3*

(Dated: June 20, 2022)

Predicting the evolution of a large system of units using its structure of interaction is a fundamental problem in complex system theory. And so is the problem of reconstructing the structure of interaction from temporal observations. Here, we find an intricate relationship between predictability and reconstructability using an information-theoretical point of view. We use the mutual information between a random graph and a stochastic process evolving on this random graph to quantify their codependence. Then, we show how the uncertainty coefficients, which are intimately related to that mutual information, quantify our ability to reconstruct a graph from an observed time series, and our ability to predict the evolution of a process from the structure of its interactions. Interestingly, we find that predictability and reconstructability, even though closely connected by the mutual information, can behave differently, even in a dual manner. We prove how such duality universally emerges when changing the number of steps in the process, and provide numerical evidence of other dualities occurring near the criticality of multiple different processes evolving on different types of structures.

I. INTRODUCTION

The relationship between *structure* and *function* is fundamental in complex systems [1–3], and important efforts have been invested in developing network models to better understand it. In particular, models of dynamics on networks [4–7] have been proposed to assess the influence of network structure over the temporal evolution of the activity in the system. In turn, data-driven models [8, 9], dimension-reduction techniques [10–13] and mean-field frameworks [14–18] have deepened our predictive capabilities. Among other things, these mathematical tools helped shed light on the relationship between dynamics criticality and many network properties such as the degree distribution [14, 16], the eigenvalue spectrum [19–21] and their group structure [17, 22, 23]. Fundamentally, these contributions justify our inclination for measuring and using real-world networks as a proxy to predict how their internal dynamics behave.

Models of dynamics on networks have also been—sometimes implicitly—used as reverse engineering tools for network reconstruction [24], when the networks of interactions are unavailable, noisy [25–27] or faulty [28]. The network reconstruction problem has stimulated many technical contributions [29]: Thresholding matrices built from correlation [30] or other more sophisticated measures [31, 32] of time series, bayesian inference of graphical models [33–38] and models of dynamics on networks [39], among others. These techniques are commonly used in neurosciences [40–42], genetics [43], epidemiology [39, 44] and finance [45], to name a few, in

order to build proxy networks on which the network science framework can be applied.

Interestingly, dynamics prediction and network reconstruction are usually considered separately, even though they are related to one another. The emergent field of the network neuroscience [46, 47] is perhaps the most actively using both notions: Network reconstruction for building brain connectomics from functional time series, then dynamics prediction for inferring various brain disorders from these connectomes [48, 49]. Recent theoretical works have also taken advantage of these notions to show that dynamics hardly depend on the structure [50], but their claim lacks a theoretical bedrock which would carefully define and relate predictability and reconstructability. Furthermore, recent breakthroughs in deep learning on graphs have benefited from proxy network substrates to enhance the predictive power of their models [51–53], with applications in epidemiology [9, 54], and pharmaceuticals [55, 56]. However, they omitted to provide theoretical justifications—other than numerical evidence of the performance—supporting their use of graph neural networks and those proxy network substrates. As a result, their assessment of enhanced predictability remains to be fully corroborated. In retrospect, there is a need for a solid, theoretical foundation of reconstructability, predictability and their relationship in networked systems.

In this work, we establish a rigorous framework that lays such a foundation based on information theory. We use the mutual information as a measure of codependence between structure and dynamics, and demonstrate its natural connection to predictability and reconstructability in dynamics on networks. Of course, the use of information theory in networks and dynamics is not new. In the context of network science, it has been used to characterize random graph ensembles [57–59]—e.g. the

* charles.murphy.1@ulaval.ca

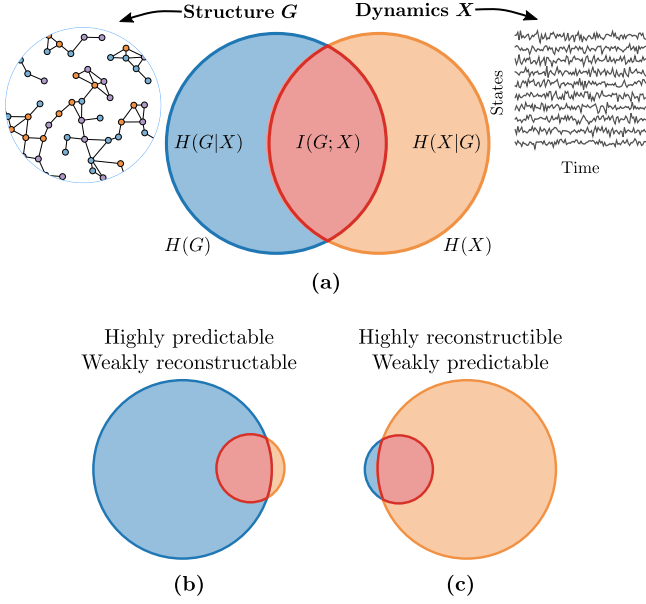


FIG. 1. **Information diagram of dynamics on random graphs.** (a) Areas represent amounts of information: The entropies related to G are shown on the left in blue and those related to X are on the right in orange. Mutual information, in red, corresponds to the information shared by both G and X . (b) The highly predictable / weakly reconstructable scenario, where $H(X) \gg H(G)$ meaning that $I(X; G)$ contains most of the information related to the dynamics, but only a small fraction of the information related to the graph. (c) The reverse scenario, i.e., highly reconstructible / weakly predictable, where $H(G) \gg H(X)$ meaning that $I(X; G)$ contains most of the information related to the graph, but only a small fraction of the information related to the dynamics..

configuration model [60, 61] and stochastic block models [62, 63]—, develop network null models [64] and perform community detection [65, 66]. In stochastic dynamical systems, information-theoretical measures have been proposed to quantify their predictability [67–69] and complexity [70, 71]. Moreover, in statistical mechanics, information transmission has been shown to reach a maximum value near the critical point of spin systems in equilibrium [72, 73]. Here, we aim to combine these ideas in a single framework, building on recent works [74, 75] involving spin dynamics on lattices. Our work extends to a formulation of networked systems beyond equilibrium spin dynamics on lattice for quantifying the structure-function relationship in complex systems, which in addition establishes a quantifiable relationship between predictability and reconstructability in dynamics on networks.

II. INFORMATION THEORY OF DYNAMICS ON RANDOM GRAPHS

Let us consider a random graph G whose support, \mathcal{G}_N , consists in the set of all graphs of N vertices, each of

which having its respective non-zero prior probability $P(G^*)$ with $G^* \in \mathcal{G}_N$. From the Bayesian perspective, the random graph G represents our prior knowledge on the structure of the system of interest. We also consider a stochastic process (also called a dynamics hereafter) of length T , noted X , evolving on a realization of G and representing the possible dynamic states of the system. We note $P(X|G)$ the probability of a random time series $X = (X_{i,t})_{i,t}$ conditioned on G , where $X_{i,t}$ is the random state, with support Ω , of vertex i at time t . Together, X and G form a Bayesian chain $G \rightarrow X$, where the arrow indicates conditional dependence [76].

We are interested in the mutual information between G and X —denoted $I(X; G)$ —which is a symmetric measure that quantifies the codependence between G and X [77], with $I(X; G) = 0$ corresponding to statistical independence. It is equivalently given by [77]

$$\begin{aligned} I(X; G) &= H(G) - H(G | X) \\ &= H(X) - H(X | G), \end{aligned} \quad (1)$$

where $H(G) = -\langle \log P(G) \rangle$ and $H(X) = -\langle \log P(X) \rangle$ are respectively the marginal entropies of G and X , and $H(G|X) = -\langle \log P(G|X) \rangle$ and $H(X|G) = -\langle \log P(X|G) \rangle$ are their corresponding conditional entropies. In the previous equations, the marginal distribution for X , the *evidence*, is defined as $P(X) = \sum_{G^* \in \mathcal{G}_N} P(G^*)P(X|G^*)$, and the posterior distribution is obtained from Bayes' theorem as $P(G|X) = P(G)P(X|G)/P(X)$. In the case where Ω is a countable set (i.e. vertices have discrete dynamical states), $I(X; G)$ is a non-negative measure bounded by $0 \leq I(X; G) \leq \min\{H(G), H(X)\}$. Figure 1(a) provides an illustration of Eq. (1).

The mutual information $I(X; G)$ is related to the *reconstructability* of a random graph and to the *predictability* of a dynamics. Intuitively, the predictability measures our ability to predict the outcome of X ahead of time simply by knowing G and conversely, the reconstructability measures our ability to reconstruct G by observing X only. In other words, $I(X; G)$ represents the knowledge gained about G when X is known or, by symmetry, the knowledge gained about X when G is known. When measured in bits, $I(X; G)$ can be interpreted as the average reduction of the number of binary questions that needs to be answered: Knowing X , is there an edge between vertices i and j ? Knowing G , was vertex i active at time t ? We say that a system is perfectly predictable when the mutual information contains all the information about X , that is when $I(X; G) = H(X)$ [see Fig. 1(b)]. Likewise, we say that it is perfectly reconstructable when $I(X; G) = H(G)$ [see Fig. 1(c)]. Yet, the magnitude of the mutual information is not a good measure of predictability and reconstructability in itself. Indeed, a specific value of $I(G; X)$ may correspond to opposing scenarios when it comes to predictability and reconstructability, as shown in Fig. 1(b-c). We thus in-

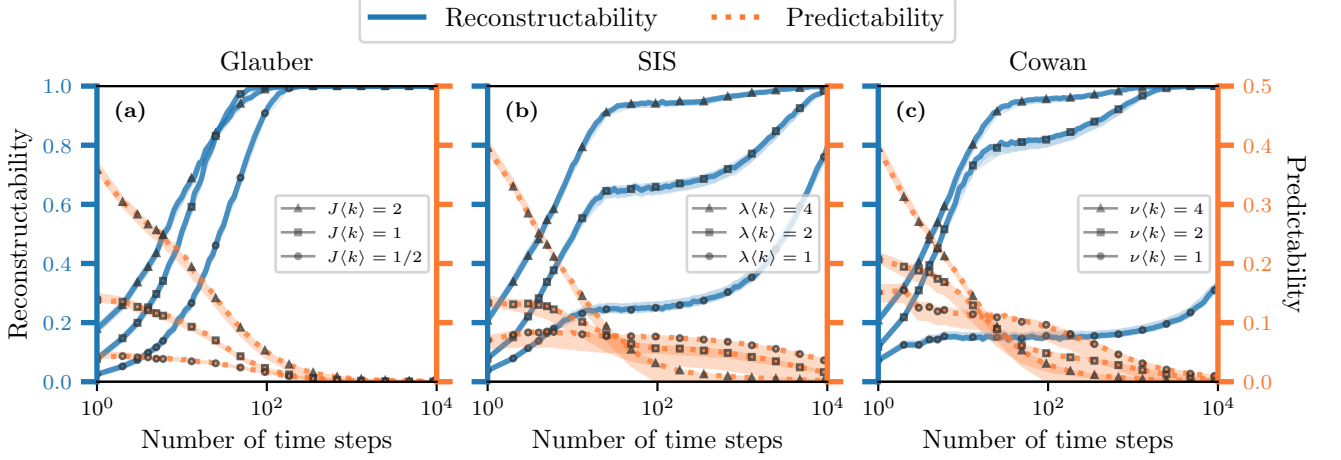


FIG. 2. *T*-duality in binary dynamics evolving on small Erdős-Rényi random graphs: (a) Glauber dynamics, (b) SIS dynamics and (c) Cowan dynamics. Each panel shows the reconstructability coefficient $U(G|X) \in [0, 1]$ (blue) and the predictability coefficient $U(X|G) \in [0, 1]$ (orange) as a function of the number of time steps T . We used graphs of $N = 5$ vertices and $E = 5$ edges, meaning an average degree of $\langle k \rangle = 2$. Each symbol corresponds to the average value measured over 1000 samples. We also show different values of the coupling parameters—normalized by the average degree—using different symbols: (a) $J\langle k \rangle \in \{\frac{1}{2}, 1, 2\}$ for Glauber, (b) $\lambda\langle k \rangle \in \{1, 2, 4\}$ for SIS and (c) $\nu\langle k \rangle \in \{1, 2, 4\}$ for Cowan.

roduce the uncertainty coefficients

$$U(X|G) = \frac{I(X;G)}{H(X)}, \quad (2a)$$

$$U(G|X) = \frac{I(X;G)}{H(G)}, \quad (2b)$$

as measures, bounded between 0 and 1, of the *reconstructability* and the *predictability*, respectively.

III. θ -DUALITY BETWEEN PREDICTABILITY AND RECONSTRUCTABILITY

Predictability and reconstructability in dynamics on random graphs offer two perspectives of the same information shared by G and X —two sides of the same coin. However, it does not mean that predictability and reconstructability go hand in hand even though they are related: A high value of $U(G|X)$ does not necessarily imply a high value of $U(X|G)$, which can somewhat be counterintuitive. This observation is well illustrated by Figs. 1(b)–(c), where $U(G|X)$ and $U(X|G)$ can take op-

posing values, depending on $H(G)$ and $H(X)$, for a same value of $I(X;G)$. As an example, let us consider X to be a Markov chain evolving on a random graph G for different values of the number of time steps, T . Theorem 1 (see App. B) states that, for any Markov chain whose entropy rate is non-zero and for sufficiently large T , $U(G|X)$ is an increasing function of T , while $U(X|G)$ is a decreasing one. This is a consequence of the fact that the mutual information is strictly increasing with T , and so is $U(G|X)$ whenever $H(G)$ is independent of T . Yet, we show in App. B that $I(X;G)$ increases more slowly than $H(X)$ with T , which results in a decreasing $U(X|G)$. We refer to this opposing behavior as a *duality* between $U(G|X)$ and $U(X|G)$ with respect to T , or a *T*-duality for short [78].

Figure 2 illustrates the universality of the *T*-duality using different binary Markov chains (i.e. $\Omega = \{0, 1\}$). In each of these chains, the probability $P(X|G)$ is

$$P(X|G) = \prod_{t=1}^T P(X_{t+1} | X_t, G), \quad (3)$$

where

$$P(X_{t+1} | X_t, G) = \prod_{i=1}^N \left\{ [\alpha(n_{i,t}, m_{i,t})]^{(1-X_{i,t})X_{i,t+1}} [1 - \alpha(n_{i,t}, m_{i,t})]^{(1-X_{i,t})(1-X_{i,t+1})} \right. \\ \left. [\beta(n_{i,t}, m_{i,t})]^{X_{i,t}(1-X_{i,t+1})} [1 - \beta(n_{i,t}, m_{i,t})]^{X_{i,t}X_{i,t+1}} \right\} \quad (4)$$

is the transition probability from state X_t to state X_{t+1} .

We also denote the activation ($0 \rightarrow 1$) and the deacti-

vation ($1 \rightarrow 0$) probability functions with $\alpha(n_{i,t}, m_{i,t})$ and $\beta(n_{i,t}, m_{i,t})$, respectively, where $n_{i,t}$ and $m_{i,t}$ denote the number of active and inactive neighbors of vertex i at time t . We consider three well known Markov chain models of different origins: The Glauber dynamics, the Susceptible-Infectious-Susceptible (SIS) dynamics and the Cowan dynamics. The Glauber dynamics [79], which have been used to describe the time-reversible evolution of magnetic spins aligning in a crystal, have been tremendously studied because of its critical behavior and its phase transition. Its stationary distribution is given by the Ising model which have found many applications in condensed-matter physics [80] and statistical machine learning [76, 81]. The SIS dynamics represents a canonical model of simple contagion describing how infectious diseases spread into a population of individuals, where susceptible individuals can get infected by their infectious contacts. Finally, the Cowan dynamics [82] has been proposed to model the neuronal activity in the brain. In this model, quiescent neurons fire if their received action potential, coming from their firing neighbors, is above a given threshold. For each model, we can identify an inactive state—down, susceptible or quiescent—and an active one—up, infectious or firing. The corresponding activation and deactivation probabilities are given in Table 1.

Figure 2 validates Theorem 1 and clearly illustrates the T -duality for each dynamics and with different values of their parameters. We used the Erdős-Rényi model as the random graph on which these dynamics evolve. The support \mathcal{G}_N is the set of all simple graphs of N vertices with E edges, and

$$P(G) = \left(\frac{\binom{N}{2}}{E} \right)^{-1}. \quad (5)$$

The observation of the T -duality begs for a more general definition of duality for any arbitrary parameter θ (see Appendix A). In fact, we say that $U(G|X)$ and $U(X|G)$ are *dual* with respect to θ , or θ -*dual*, in an interval Θ if and only if the signs of their derivative with respect to θ are different for every $\theta^* \in \Theta$:

$$\left[\frac{\partial U(G|X)}{\partial \theta} \frac{\partial U(X|G)}{\partial \theta} \right]_{\theta=\theta^*} < 0. \quad (6)$$

This criterion formally relies on the existence of regions Θ where the variations of $U(G|X)$ and $U(X|G)$ with respect to θ are contradictory, regardless of their amplitude. We use this criterion to relate the existence of extrema of $U(G|X)$ and $U(X|G)$ with that of regions of θ -duality (see Lemma 1 in App. A), and to prove Theorem 1.

Knowing the existence of the T -duality and having a general definition of θ -duality, it is now natural to ask if there exist other types of θ -dualities in dynamics on random graphs. A large variety of parameters could lead to interesting θ -dualities—some controlling the general behavior of the dynamics, and others controlling some

Dynamics	$\alpha(n, m)$	$\beta(n, m)$	Coupling
Glauber [79]	$\sigma[2J(n - m)]$	$\sigma[2J(m - n)]$	J
SIS [5]	$1 - \left(1 - \frac{\lambda}{\beta}\right)^m$	β	λ
Cowan [82]	$\sigma[a(\nu m - \mu)]$	β	ν

TABLE I. Activation and deactivation probability functions, $\alpha(n, m)$ and $\beta(n, m)$, respectively, for the binary dynamics considered in this study, where n corresponds to the number of inactive neighbors whose states are 0, and m corresponds to the number of active neighbors whose states are 1. We define $\sigma(x) = [\exp(-x) + 1]^{-1}$ as the logistic function. Some of these parameters are fixed throughout the paper: $\beta = 0.5$ for SIS and Cowan, and $a = 7$ and $\mu = 1$ for Cowan. The coupling parameters (J for Glauber, λ for SIS and ν for Cowan) are specified in each figure. Also, to prevent the SIS dynamics from being completely inactive, we allow the inactive vertices to spontaneously activate with probability $\epsilon = 10^{-3}$ [83].

structural properties of the random graph which, in turn, also impact the dynamics. Most of them require the system to be larger, if the effects over X and G of varying θ are to be significant (e.g. phase transitions). However, in high-dimensional systems, theoretical and numerical challenges arise in the evaluation of the reconstructability and the predictability, which complicate the search for dualities. We address this problem in the next section.

IV. EVALUATION OF THE MUTUAL INFORMATION IN LARGE SYSTEMS

The exact evaluation of the mutual information, $H(X)$ and $H(G|X)$ for large systems becomes rapidly tedious as the size N of the system increases. This challenge arises because of the graph enumeration needed to evaluate the *log-evidence* probability, namely

$$\log P(X) = \log \left[\sum_{G^* \in \mathcal{G}_N} P(G^*) P(X | G^*) \right], \quad (7)$$

becomes quickly intractable with N . This difficulty frequently arises in Bayesian model selection [84–86] and machine learning [34, 36], and is known as the marginal likelihood, partition function or evidence probability estimation problem. Fortunately, several approaches exist to circumvent this problem: We considered two different estimators of the log-evidence probability to compute the mutual information (see App. C). These two estimators are biased, but we show that they in fact bound the mutual information below and above, which will be useful to ensure their numerical validity (see App. C).

The first estimator (noted MF) is based on a variational mean-field (MF) approximation [87], where the

posterior probability is bounded such that

$$P(G | X) \geq P_{\text{MF}}(G | X) = \prod_{i < j} [\pi_{ij}(X)]^{A_{ij}} [1 - \pi_{ij}(X)]^{1 - A_{ij}}, \quad (8)$$

where A_{ij} is the element (i, j) of the adjacency matrix of the undirected graph G ($A_{ij} = 1$ if there exists an edge between vertices i and j , $A_{ij} = 0$ otherwise), and $\pi_{ij}(X)$ is the marginal posterior probability that edge (i, j) exists given X . The MF estimator is biased and converges to a lower bound of the posterior probability, hence yielding a lower bound for $I(X; G)$, $U(G | X)$ and $U(X | G)$ when used in Eqs. (1), (2a) and (2b).

The second estimator (noted AIS) was presented in Ref. [86] as a stepping-stone algorithm for computing the evidence probability—a version of annealed importance sampling (AIS) [84]. This estimator converges to a lower bound of the evidence probability, but its bias can be arbitrarily reduced by increasing the number K of temperature steps. Therefore, when used to estimate Eq. (1), we show that this biased estimator of mutual information converges to an upper bound of $I(X; G)$, $U(G | X)$ and $U(X | G)$ (see App. C). Altogether, these two estimators allow us to bound the mutual information even when its exact evaluation is intractable.

Figure 3(a) shows the behavior of $I(X; G)$ in the Glauber dynamics on a small Erdős-Rényi random graph as approximated using the MF and AIS estimators, and compares them to an exact evaluation based on an explicit graph enumeration used in Fig. 2. As expected the two estimators provide a lower and an upper bound for $I(X; G)$, and these bounds are fairly tight.

Several caveats are in order. On the one hand, the bias of the AIS estimator can, in principle, be reduced arbitrarily by increasing the number K of temperature steps, but its evaluation becomes quickly computationally costly. On the other hand, the evaluation of MF estimator is comparatively quicker, but cannot be improved by further sampling. The AIS estimator is accordingly closer to the exact value throughout, but it can sometimes overestimate the mutual information above its upper bound since $H(X)$ is overestimated while $H(X|G)$ is not. The MF estimator can also yield negative values of $I(X; G)$ for small values of J —i.e. regimes where $H(G|X) \simeq H(G)$ —due to an overestimated $H(G|X)$ becoming larger than $H(G)$.

Figure 3(b) shows the same experiment as in Fig. 3(a) but with larger graphs of $N = 100$ vertices and leads to similar observations: the AIS estimator is always greater than the MF estimator, and both estimators sometimes yields approximated values for $I(X; G)$ outside of the valid range $[0, \max\{H(G), H(X)\}]$. Interestingly, these bounds are nevertheless fairly close to one another, as in the case $N = 5$.

With reliable estimations of the mutual information in hand, it is now possible to investigate the existence of other θ -dualities in more complicated systems, which is the subject of the next section.

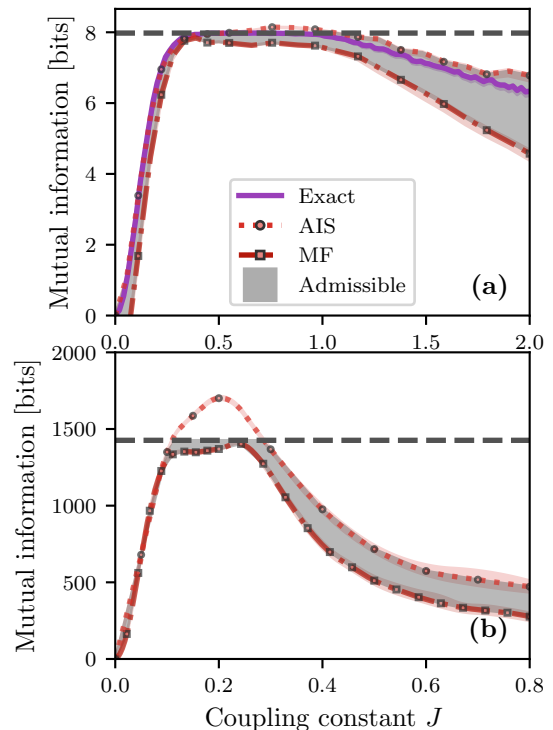


FIG. 3. **Estimators of the mutual information in the Glauber dynamics on Erdős-Rényi graphs as a function of the normalized coupling parameter $J \langle k \rangle$:** (a) $N = 5$, $E = 5$ and $T = 100$ (b) $N = 100$, $E = 250$ and $T = 1000$. The solid line in (a) corresponds to the exact evaluation of $I(X; G)$ and is the same line as the one in Fig. 2(a). The circles and square in both (a) and (b) represent the values of $I(X; G)$ computed using the AIS and the MF estimators, respectively. The dashed line indicates the upper bound of $I(X; G)$, i.e., $\max\{H(G), H(X)\}$. We also show with a gray area the admissible values of $I(X; G)$ bounded by the biased MF and AIS estimators (see App. C).

V. DUALITY AND CRITICALITY

Despite their different nature and range of applications, the three models presented in Table I share several properties of interest. For instance, each model has a coupling parameter that controls the influence of the state of the first neighbors on the transition probabilities. They also all feature a phase transition in the infinite size limit whose position is determined by the coupling parameter (see Fig. 5 and App. E). We now investigate the influence of criticality over the existence of θ -dualities, where θ is a coupling parameter.

For the Glauber dynamics, this parameter is the coupling constant J , which dictates the reduction (increase) in the total energy of a spin configuration when two neighboring spins are parallel (antiparallel). The Glauber dynamics features a continuous phase transition at a critical point J_c between a disordered and an ordered phase, where for $J < J_c$ the spins are disordered resulting

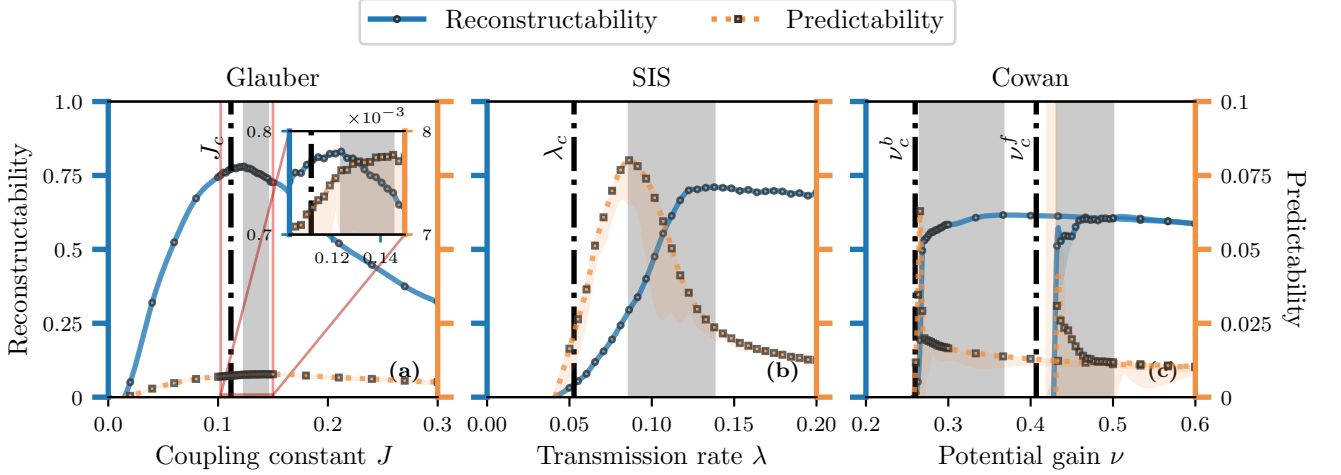


FIG. 4. **Dynamics evolving on configuration model graphs with negative binomial degree distribution:** (a) Glauber dynamics, (b) SIS dynamics and (c) Cowan dynamics. We generated graphs of $N = 1000$ vertices, where the degree distribution $\rho(k) = (1-p)p^k$ is a geometric distribution with $p = \langle k \rangle / (1 + \langle k \rangle)$ and $\langle k \rangle = 5$, and time series of length $T = 2000$. See Table I for the remaining parameters. Similar to Fig. 2, $U(G|X)$ is shown in blue (left axis) and $U(X|G)$ is shown in orange (right axis). We show, for each dynamics, the uncertainty coefficients as a function of the—normalized—coupling parameter, $J\langle k \rangle$ for Glauber, $\lambda\langle k \rangle$ for SIS and $\nu\langle k \rangle$ for Cowan. The vertical dotted-dashed lines correspond to the phase transition thresholds of each dynamics, which are estimated from Monte Carlo simulations (see Appendix E). For the Cowan dynamics, the forward and backward branches are shown with their corresponding thresholds and dual regions (see main text).

in a vanishing magnetization, and for which this magnetization is non-zero when $J > J_c$. For the SIS dynamics, it is the transmission rate λ that acts as a coupling parameter. Like the Glauber dynamics, the SIS dynamics possesses a continuous phase transition where, when $\lambda < \lambda_c$, the system reaches an absorbing—or inactive—state from which it cannot escape, and an active state, when $\lambda > \lambda_c$, where a non-zero fraction of the vertices remain active over time [88]. The Cowan dynamics can both feature a continuous or a first-order phase transition between an inactive and an active phase depending on the value of slope a , for which the coupling parameter is ν , i.e. the potential gain for each firing neighbors. The continuous and first-order phase transitions of the Cowan dynamics are quite different in that the latter is characterized by two thresholds, namely the forward and backward thresholds $\nu_c^b < \nu_c^f$, respectively (see Appendix E for further details). Hence, the Cowan dynamics has a first-order phase transition that exhibits a bistable region $\nu \in (\nu_c^b, \nu_c^f)$, where both the inactive and active phases are reachable depending on the initial conditions.

To account for the heterogeneous network structure observed in a wide range of complex systems [1], we simulate the dynamics on the configuration model, a random graph whose—potentially heterogeneous—degree sequence \mathbf{k} is fixed and whose support \mathcal{G}_N corresponds to the set of all loopy multigraphs of degree sequence \mathbf{k} . The probability of a graph G^* in this ensemble is

$$P(G^*) = \frac{(2E)!!}{(2E)!} \frac{\prod_i k_i!}{\prod_{i < j} M_{ij}! \prod_i M_{ii}!!}, \quad (9)$$

where M_{ij} counts the number of edges connecting vertices i and j in the multigraph G^* and $2E = \sum_i k_i$ is the number of half-edges in G^* . Like the Erdős-Rényi model, the configuration model fixes the number of edges, but also fixes the degree distribution $\rho(k)$.

Figure 4 shows the predictability and reconstructability of the three dynamics evolving on graphs drawn from the configuration model whose distribution, $\rho(k)$, is geometric, as estimated by the MF estimator. First, these results allow us to compare the dynamics with one another. For example, on the one hand, the Glauber dynamics is globally less predictable than the other two, since its predictability coefficient is overall smaller. In other words, the knowledge of a graph G^* provides less information about X^* in the Glauber dynamics in comparison with the others, relatively to the total amount of information needed to reconstruct X^* . This is related to the time reversibility of the Glauber dynamics, which allows any vertex to transition from the inactive to the active state (and vice versa) with non-zero probability, at any time, effectively making the Glauber dynamics more random than the others. On the other hand, the SIS and Cowan dynamics are portrayed by the MF estimator as practically unpredictable and unreconstructable when their coupling parameter is below their respective critical point. This precisely occurs in the inactive phase, where no mutual information can be generated after a short time, when the system reaches the inactive state. By contrast, the Glauber dynamics does not reach an inactive state below its critical point, which explains the gradual increase in predictability and reconstructability

in that region.

Several additional observations are worth making. All dynamics exhibit maxima for $U(X|G)$ and $U(G|X)$ which delineate a region of duality illustrated by the shaded areas (two for Cowan, that is one for each branch). These regions are close to, but systematically above, their respective phase transition thresholds. A similar phenomenon in spin dynamics on non-random lattices has been reported by previous works [74, 75], in which the information transmission rate between spins—a measure akin to $I(X;G)$ —is maximized above the critical point. Our numerical results are consistent with theirs, and suggest that their findings regarding near-critical systems even apply beyond spin dynamics on fixed lattices, to other types of processes on more heterogeneous and random structures.

VI. CONCLUSION

In this work, we used information theory to characterize the structure-function relationship with mutual information. We showed how mutual information is a natural starting point to define both predictability and reconstructability in dynamics on networks, in turn showing how they are intrinsically related. Our approach is quite general allowing the exploration of different configurations of dynamics on networks of the form $G \rightarrow X$, thus varying the nature of the process itself as well as the random graph on which it evolves. Our framework could be extended to adaptive systems [89–92] where both X and G influence each other (i.e. $X \leftrightarrow G$). The relationship between X and G could also go the other way around: A system in which X generates a graph G (i.e., $X \rightarrow G$). Hyperbolic graphs [93, 94] falls into this category, where X represents a set of coordinates, and our framework could be extended to quantifying the feasibility of network geometry inference [95–97].

We found efficient ways to estimate the mutual information numerically, thus allowing us to investigate relatively large systems. More work on this front is required, however, since the evaluation of these estimators remains quite computationally costly. It would be worth investigating simpler models, for which it is possible to analytically—or at least approximately—evaluate $U(X|G)$ and $U(G|X)$. Approximate mean-field models of the sort described in Refs. [11, 13, 98] come to mind and could potentially provide reliable approximations of $I(X;G)$, $U(X|G)$ and $U(G|X)$.

Central to our findings is the peculiar discovery that predictability and reconstructability are not only related, but sometimes dual to one another. We proved that such θ -duality appears when the length of the processes changes and presented numerical evidence of duality near the criticality in three different dynamics on random heterogeneous networks. These findings generalize and formalize—while being consistent with—previous works [74, 75] and suggest that criticality in these systems

is intrinsically related to the duality between predictability and reconstructability in complex systems.

ACKNOWLEDGMENTS

We are grateful to Guillaume St-Onge and Vincent Painchaud for useful comments, and to Simon Lizotte and François Thibault for their help in designing the software. This work was supported by the Fonds de recherche du Québec – Nature et technologies (VT, PD), the Conseil de recherches en sciences naturelles et en génie du Canada (CM, VT, AA, PD), and the Sentinelle Nord program of Université Laval, funded by the Fonds d’excellence en recherche Apogée Canada (CM, VT, AA, PD). We acknowledge Calcul Québec and Compute Canada for their technical support and computing infrastructures.

Appendix A: Formal definition of θ -duality

In what follows, we define the duality between predictability and reconstructability by taking a more general stance: Instead of considering a stochastic process X evolving on a random graph G , we let X be conditioned on an arbitrary discrete random variable Y . First, we define the local duality of the uncertainty coefficients. The latter are considered as continuously differentiable functions with respect to a parameter θ whose domain is some non-empty interval of the real line.

Definition 1 (Local duality). The uncertainty coefficients $U(X|Y)$ and $U(Y|X)$ are *locally dual* with respect to θ at $\theta = \theta^*$ if and only if

$$\left[\frac{\partial U(X|Y)}{\partial \theta} \frac{\partial U(Y|X)}{\partial \theta} \right]_{\theta=\theta^*} < 0. \quad (\text{A1})$$

The definition of the θ -duality, a global property, follows that of the local duality.

Definition 2 (θ -Duality). The uncertainty coefficients $U(X|Y)$ and $U(Y|X)$ are *dual* with respect to θ , or θ -*dual*, in the interval Θ if and only if they are locally dual for all values of θ^* in Θ .

From these definitions, we relate the presence of extrema of $U(X|Y)$ and $U(Y|X)$ with the existence of a θ -duality.

Lemma 1. *Let Θ be a non-empty subinterval of the variable θ whose one endpoint is a local extremum of $U(X|Y)$ and the other, a local extremum of $U(Y|X)$. Moreover, suppose that $U(X|Y)$ and $U(Y|X)$ do not have critical points in Θ . Then the extrema points delineate a region of θ -duality if and only if they are both maxima (or both minima).*

Proof. Let θ_R and θ_P be the extrema points of $U(Y|X)$ and $U(X|Y)$, respectively. Thus

$$\left. \frac{\partial U(Y|X)}{\partial \theta} \right|_{\theta=\theta_R} = \left. \frac{\partial U(X|Y)}{\partial \theta} \right|_{\theta=\theta_P} = 0. \quad (\text{A2})$$

Suppose for a moment that $\theta_R < \theta_P$ and let $\Theta = (\theta_R, \theta_P)$. This implies that $\frac{\partial U(Y|X)}{\partial \theta}$ changes sign at θ_R , before $\frac{\partial U(X|Y)}{\partial \theta}$, for which the sign change happens at θ_P .

On the one hand, if the extrema points θ_R and θ_P are both maxima (or minima), then $\frac{\partial U(Y|X)}{\partial \theta}$ and $\frac{\partial U(X|Y)}{\partial \theta}$ have different signs in Θ . Hence, inequality (A1) is verified in this region. The uncertainty coefficients are therefore θ -dual in Θ .

On the other hand, if the uncertainty coefficients are θ -dual in Θ , then inequality (A1) is satisfied in this interval. This in turn implies that either $U(Y|X)$ decreases in Θ while $U(X|Y)$ increases or $U(Y|X)$ increases in Θ while $U(X|Y)$ decreases. Therefore, the endpoints of Θ are either both maximum points or both minimum points.

Finally, repeating the same arguments with $\theta_R > \theta_P$ and $\Theta = (\theta_P, \theta_R)$ leads to the same conclusions about θ -duality of $U(X|Y)$ and $U(Y|X)$ in Θ . \square

Appendix B: Universality of the T -duality

We demonstrate the universality of the T -duality, where T is the number of steps in the process X . First, we need to show that the mutual information is a monotonically increasing function of T .

Lemma 2. *Let $X = (X_1, X_2, \dots, X_T)$ be a Markov chain of length T whose transition probabilities are conditional to some discrete random variable Y that is independent of T and such that $H(X_{t+1}|X_t) > 0$ for all $t \in \{1, \dots, T-1\}$. Suppose moreover that the state spaces of X and Y are finite. Then the mutual information $I(X; Y)$ is nonzero and monotonically increasing with $T \in \mathbb{Z}_+$.*

Proof. Let us define a Markov chain $X' = (X_1, X_2, \dots, X_{T-1})$ of size $T-1$, such that the concatenation of X' with state variable X_T yields X . Hence, we can express the mutual information between X and Y in terms of X' as $I(X; Y) = I(X', X_T; Y)$. Furthermore, proving the monotonicity of mutual information can be reformulated as proving the following inequality:

$$I(X', X_T; Y) - I(X'; Y) > 0, \quad (\text{B1})$$

for all T . By the chain rule for conditional mutual information, that is $I(X', X_T; Y) = I(X_T; Y|X') + I(X'; Y)$, inequality (B1) becomes

$$I(X_T; Y|X') = H(X_T|X') - H(X_T|X', Y) > 0. \quad (\text{B2})$$

The term $H(X_T|X') - H(X_T|X', Y)$ is always at least non-negative, by virtue of the non-negativity of mutual information [77, Theorem 2.6.5]. Then, to prove

inequality (B2), we must verify that $H(X_T|X')$ never equals $H(X_T|X', Y)$. Recalling that $H(X_T|X') \geq H(X_T|X', Y) \geq 0$, inequality (B2) does not hold if (i) $H(X_T|X') = 0$ or if (ii) X_T is independent of Y (i.e. $I(X_T; Y|X') = 0$). According to the hypothesis $H(X_{t+1}|X_t) > 0$ for all $t \in \{1, \dots, T-1\}$, condition (i) cannot be true. Moreover, condition (ii) implies that $I(X; Y) = I(X_T, X'; Y) = I(X'; Y) = 0$. Therefore, the only instance where Eq. (B1) is not satisfied is when the Markov chain X is independent of Y , i.e. $I(X; Y) = 0$ for all length T . However, this contradicts the assumption about the transition probabilities. Hence, $I(X; Y) > 0$ and monotonically increases with T . \square

Before presenting the main result of this section, let us make a few remarks about the restrictions imposed in the last lemma. The condition $H(X_{t+1}|X_t) > 0$ for all $t \in \{1, \dots, T-1\}$ only asserts that the Markov chain is nondeterministic in the sense that knowing the state of the chain at time t does not completely eliminate the uncertainty about the state at time $t+1$. This condition is satisfied for wide variety of stochastic processes, including the irreducible Markov chains, where there is always a nonzero probability to transition from a state to any other state in a finite number of time steps. Moreover, the finiteness of the state spaces for the chain X and the variable Y is imposed to make $H(X)$, $H(Y)$, and $I(X; Y)$ finite. This in turn ensures that the uncertainty coefficients $U(Y|X)$ and $U(X|Y)$ are well defined for all $T \in \mathbb{Z}_+$, a property that is necessary to prove the next lemma.

Lemma 3. *Let $X = (X_1, X_2, \dots, X_T)$ and Y respectively be a Markov chain and a discrete random variable as in Lemma 2. Then the uncertainty coefficients $U(Y|X)$ and $U(X|Y)$, interpreted as functions of $T \in \mathbb{Z}_+$, can be uniquely generalized to functions, respectively $f(T)$ and $g(T)$, that are real analytic for all $T \in \mathbb{R}_+$. Moreover, $H(X)$ can be extended to a function $h(T)$ that is analytic for all $T \in \mathbb{R}_+$ except where $f(T) = 0$.*

Proof. We first consider $U(X|Y)$ and $U(Y|X)$, which are defined in Eqs. (2a)–(2b). These can be interpreted as functions of $T \in \mathbb{Z}_+$ whose values belong to the interval $[0, 1]$. According to Guichard's Theorem [99, Theorem 5.2.1] (see also [100, Theorem 15.13]), there exist two functions of $z \in \mathbb{C}$, denoted f and g , that are holomorphic in the whole complex plane and whose values at $z = T \in \mathbb{Z}_+$ equal those of $U(X|Y)$ and $U(Y|X)$, respectively.

Now, $U(X|Y)$ and $U(Y|X)$, and consequently $f(z)$ and $g(z)$, have bounded values for all $z = T \in \mathbb{Z}_+$. Moreover, f and g are holomorphic, so their restriction to the axis $z = T \in \mathbb{R}$ is real analytic. Hence, on that axis, f and g are Lipschitz continuous, which means that there are positive and finite constants, a and b , such that $|f(T) - g(T')| \leq a|T - T'|$ and $|g(T) - g(T')| \leq b|T - T'|$ for all $T, T' \in \mathbb{R}$. Choosing $T = T' + \epsilon$ with $T' \in \mathbb{Z}_+$ and $|\epsilon| < 1$, we conclude that $f(T)$ and $g(T)$ have finite values for all $T \in \mathbb{R}_+$.

The functions f and g are thus holomorphic in the whole complex plane and bounded on the positive real axis. This allows to use a special case of Carlson's Theorem [101, Theorem 2.8.1] according to which holomorphic functions that are bounded on the positive real axis are uniquely defined by their values on the set \mathbb{Z}_+ . Therefore, f is the unique extension $U(X|Y)$ that is analytic for all $T \in \mathbb{R}_+$. The same conclusion holds for g and $U(Y|X)$.

To finish the proof, we need to tackle $H(X)$. We cannot use the same strategy as above because $H(X)$ is not a bounded function of $T \in \mathbb{Z}_+$. However, by definition, the identity

$$H(X) = \frac{H(Y)U(Y|X)}{U(X|Y)}. \quad (\text{B3})$$

is valid whenever $U(X|Y) > 0$. Now, according to Lemma 2, $I(X;Y) > 0$ and hence $U(X|Y) > 0$ for all $T \in \mathbb{Z}_+$. This means that Eq. (B3) is well defined for all $T \in \mathbb{Z}_+$. To extend the domain of validity of the identity, we use the analytic functions f and g introduced above and define a new function h as

$$h(T) = H(Y) \frac{g(T)}{f(T)}. \quad (\text{B4})$$

The values of h coincide with those of $H(X)$ for all $t \in \mathbb{Z}_+$, so that Eq. (B3) defines a unique extension of $H(X)$. Moreover, h is analytic for all $T \in \mathbb{R}_+$ except at the points t where $f(t) = 0$. \square

With Lemmas 2 and 3 in hand, we now proceed to prove our main theoretical result: the universality of the T -duality in Markov chains.

Theorem 1. *Let $X = (X_1, X_2, \dots, X_T)$ and Y respectively be a Markov chain and a discrete random variable as in Lemma 2. Suppose additionally that X has a finite nonzero entropy rate and that Y has a nonzero entropy. Then there exists a positive constant τ such that the uncertainty coefficients $U(Y|X)$ and $U(X|Y)$ are T -dual for all $T \geq \tau$.*

Proof. According to Lemma 3, the quantities $U(X|Y)$, $U(Y|X)$, and $H(X)$, which were originally defined as real functions of $T \in \mathbb{Z}_+$, have unique analytic extensions on the positive real axis, i.e., $T \in \mathbb{R}_+$. This allows us to treat $U(X|Y)$, $U(Y|X)$, and $H(X)$ as continuously differentiable functions with respect to T .

Now, by hypothesis, the entropy rate of the Markov chain X , $R := \lim_{T \rightarrow \infty} \frac{H(X)}{T}$, is well defined and nonzero. Hence, $H(X) \sim RT$, i.e., $H(X)$ is positive and asymptotically linearly increasing with T . Moreover, since Y is independent of T and $I(X;Y) > 0$, it follows that $I(X;Y)$ is monotonically increasing with respect to T by Lemma 2. As a result, $U(Y|X) = \frac{I(X;Y)}{H(Y)}$ is also monotonically increasing, since its denominator is independent of T , by assumption. This translates to the strict inequality $\frac{\partial U(Y|X)}{\partial T} > 0$. If there exists a T -duality, i.e., there is

a domain of T where Eq. (A1) is true, then $U(X|Y)$ must be monotonically decreasing with T —or $\frac{\partial U(X|Y)}{\partial T} < 0$ —in that domain. To prove this, note that we can relate the two uncertainty coefficients using Eq. (B3). This leads to the following differential equation

$$\frac{\partial}{\partial T} [\log U(X|Y)] = \frac{\partial}{\partial T} [\log U(Y|X)] - \frac{\partial}{\partial T} [\log H(X)], \quad (\text{B5})$$

where we used the fact that $\frac{\partial H(Y)}{\partial T} = 0$. Hence, to show that $U(X|Y)$ is monotonically decreasing with T , the following inequality must hold

$$\frac{\partial}{\partial T} [\log U(Y|X)] < \frac{\partial}{\partial T} [\log H(X)]. \quad (\text{B6})$$

Suppose for a moment that $U(X|Y)$ is in fact increasing, such that Eq. (B6) is false. This will eventually give rise to a contradiction. Let $g(T) := U(Y|X)$ and $h(T) := H(X)$ be continuous functions of T such that their derivative with respect to T are respectively given by $g'(\tau) := \left. \frac{\partial f(T)}{\partial T} \right|_{T=\tau}$ and $h'(\tau) := \left. \frac{\partial h(T)}{\partial T} \right|_{T=\tau}$. Note that $0 < f(\tau) \leq 1$ and $h(\tau) > 0$ for all $\tau \in \mathbb{R}_+$. If Eq. (B6) is false, then

$$(\log g(T))' \geq (\log h(T))'. \quad (\text{B7})$$

Using Grönwall's inequality [102, Theorem 1.2.1], we get

$$\frac{g(T)}{g(a)} \geq \frac{h(T)}{h(a)}, \quad 0 < a < T. \quad (\text{B8})$$

So far, we have established that $h(T) = H(X) \sim RT$ and that $U(Y|X)$ is monotonically increasing. We have also proved that if $U(X|Y)$ is not monotonically decreasing with T , then inequality (B8) is satisfied. However, the latter inequality and $h(T) \sim RT$ readily imply that $g(T)$ belongs to the class $\Omega(T)$, which is the set of all $\tilde{g}(T)$ such that there exist positive constants, S and T^* , for which $\tilde{g}(T) \geq ST$ for all $T \geq T^*$ (i.e. Knuth's Big Omega [103]).

Two cases must be considered. First, if $ST^* > 1$, then $\tilde{g}(T) \geq ST^* > 1$, which is in direct contradiction with $g(T) \leq 1$ whenever $T \geq T^*$. Second, if $ST^* \leq 1$, then choose $T^{**} > S^{-1} \geq T^*$, so that $\tilde{g}(T) \geq ST^{**} > 1$ for all $T \geq T^{**}$. This again contradicts the inequality $g(T) \leq 1$ whenever $T \geq T^{**}$. As a result, inequality (B8) cannot be satisfied when $T \geq \tau$, with $\tau = \max\{T^*, T^{**}\}$. We thus conclude that $U(X|Y)$ is monotonically decreasing for all $T \geq \tau$. Therefore, $U(Y|X)$ and $U(X|Y)$ are T -dual in the interval $[\tau, \infty)$. \square

Appendix C: Estimators of the mutual information

The mutual information $I(X;G)$ is generally intractable. Its intractability stems from the evaluation

of the *evidence* probability, which is defined by the following equation:

$$P(X) = \sum_{G \in \mathcal{G}_N} P(G)P(X | G). \quad (\text{C1})$$

Indeed, this sum potentially counts a number of terms which grows exponentially with the number of vertices N in the random graph. More specifically, the evidence probability appears in two entropy terms needed to compute the mutual information, namely the marginal entropy $H(X) = -\langle \log P(X) \rangle$ and the reconstruction entropy $H(G | X) = -\left\langle \log \frac{P(G)P(X|G)}{P(X)} \right\rangle$, where $\langle f(Y) \rangle$ denotes the expectation of $f(Y)$. Fortunately, the evidence probability, and in turn the mutual information, can be estimated efficiently using Monte Carlo techniques, which we present in this section.

1. Graph enumeration approach

For sufficiently small random graphs ($N \leq 5$), the evidence probability can be efficiently computed by enumerating all graphs of \mathcal{G}_N and by adding explicitly each term of Eq. (C1). Then, we can estimate the mutual information by sampling M graph-states pairs, denoted $(G^{*(m)}, X^{*(m)})$, and by computing the following arithmetic average:

$$I(X; G) \simeq \frac{1}{M} \sum_{m=1}^M \log P(X^{*(m)} | G^{*(m)}) - \log P(X^{*(m)}). \quad (\text{C2})$$

The variance of this estimator scales with the inverse of \sqrt{M} . In Fig. 2, we used this estimator to compute the mutual information, where $M = 1000$.

2. Variational mean-field approximation

In this approach, we estimate the posterior probability instead of the evidence probability. According to Bayes' theorem, the posterior probability is

$$P(G | X) = \frac{P(G)P(X | G)}{P(X)}. \quad (\text{C3})$$

Behind this estimator is a variational mean-field (MF) approximation that assumes the conditional independence of the edges. For simple graphs, the MF posterior is

$$P_{\text{MF}}(G | X) = \prod_{i \leq j} [\pi_{ij}(X)]^{A_{ij}} [1 - \pi_{ij}(X)]^{1 - A_{ij}}, \quad (\text{C4})$$

where $\pi_{ij}(X) := P(A_{ij} = 1 | X)$ is the marginal conditional probability of existence of the edge (i, j) given X .

For multigraphs, a similar expression can be obtained, but instead involves a probability $\pi_{ij}(\omega | X) := P(M_{ij} = \omega | X)$ that there are ω multi-edges between i and j . In this case, the MF posterior becomes

$$P_{\text{MF}}(G | X) = \prod_{i < j} \prod_{\omega=0}^{\infty} [\pi_{ij}(\omega | X)]^{\delta_{\omega, M_{ij}}}, \quad (\text{C5})$$

where $\delta_{x,y}$ is the Kronecker delta. The MF approximation allows to compute a lower bound of the true posterior entropy, such that

$$H(G | X) \geq -\langle \log P_{\text{MF}}(G | X) \rangle, \quad (\text{C6})$$

as a consequence of the conditional independent between the edges [77, Theorem 2.6.5]. Using the MF approximation and a strategy similar to the exact estimator, we compute the MF estimator of the mutual information as follows:

$$I(G | X) \geq \frac{1}{M} \sum_{m=1}^M \left[\log P_{\text{MF}}(G^{*(m)} | X^{*(m)}) - \log P(G^{*(m)}) \right]. \quad (\text{C7})$$

To compute $P_{\text{MF}}(G^{*(m)} | X^{*(m)})$, we sample a set $\mathcal{Q}^{(m)} := \{G_1^{*(m)}, \dots, G_Q^{*(m)}\}$ of Q graphs from the posterior distribution $P(G | X^{*(m)})$. Then, we estimate the probabilities $\pi_{ij}(X) \simeq \frac{n_{ij}^{(m)}}{Q}$ using their corresponding maximum likelihood estimate, where $n_{ij}^{(m)}$ is the number of times the edge (i, j) is seen in $\mathcal{Q}^{(m)}$. An analogous maximum likelihood estimate is made in the multigraph case, where $\pi_{ij}(\omega | X) \simeq \frac{n_{ij;\omega}^{(m)}}{K}$ and $n_{ij;\omega}^{(m)}$ counts the number of times there were ω multiedges between i and j in $\mathcal{Q}^{(m)}$. This estimator is a lower bound of the mutual information—a consequence of Eq. (C6). Hence, it is biased, and the extent of this bias is dependent on the quality of the conditional independence assumption with respect to the true random graph. Note that the MF estimator can yield negative estimates of the mutual information (see Fig. 3).

In Figs. 3 and 4, we fix the number of graphs sampled from the posterior distribution to $Q = 1000$, and propose $5N$ moves between each sample, as also mentioned in App. D.

3. Annealed important sampling

Whereas the MF estimator represents a biased estimator of the posterior probability $P(G | X)$, there exists other Markov chain Monte-Carlo (MCMC) techniques that tackle the problem of estimating the evidence probability directly. The one we consider in this paper is obtained from an *annealed importance sampling* (AIS) procedure called the stepping-stone (SS) algorithm [86].

The procedure of the stepping-stone algorithm takes advantage of the fact that it is possible to sample efficiently from the posterior distribution $P(G | X)$ using MCMC (see Section D). In order to compute an accurate estimator of the evidence probability $P(X)$, the procedure samples the space \mathcal{G}_N according to $P_\beta(G | X)$, where $0 \leq \beta \leq 1$ is an inverse temperature parameter that dampens the influence of the likelihood such that

$$P_\beta(G | X) \propto [P(X | G)]^\beta P(G). \quad (\text{C8})$$

The inverse temperature basically allows the Markov chain to navigate \mathcal{G}_N efficiently to construct an accurate estimator of $P(X)$, that is where the graph samples are not all too close or too far from the maximum posterior. More specifically, the AIS estimator is defined by

$$P_{\text{AIS}}(X) = \prod_{k=1}^K \langle [P(X | G_k^*)]^{\beta_k - \beta_{k-1}} \rangle, \quad (\text{C9})$$

where $0 = \beta_0 < \dots < \beta_{K-1} = 1$ and the expectation is evaluated with respect to $G_k^* \sim P_{\beta_k}(G | X^*)$, for each k . Similarly to the mean-field estimator, we estimate this expectation by collect a sample $\mathcal{Q}_k^{(m)}$ of Q graphs distributed according to $P_{\beta_k}(G | X^{*(m)})$, for each k .

Taking the log of this equation gives us an estimator of the log-evidence probability, which we can use to compute the mutual information directly:

$$\log P_{\text{AIS}}(X) = \sum_{k=1}^K \log \langle [P(X | G_k^*)]^{\beta_k - \beta_{k-1}} \rangle. \quad (\text{C10})$$

Although the estimator for P_{AIS} is unbiased, the one for the log-evidence probability introduces a bias:

$$\log P(X) \geq \log P_{\text{AIS}}(X). \quad (\text{C11})$$

This bias can be arbitrarily reduced by increasing K [86], although we found that doing so provides diminishing returns. Using the AIS estimator of the evidence probability, we obtain an AIS estimator of the mutual information such that

$$I(G; X) \leq \frac{1}{M} \sum_{m=1}^M \left[\log P(X^{*(m)} | G^{*(m)}) - \log P_{\text{AIS}}(X^{*(m)}) \right]. \quad (\text{C12})$$

Following Ref. [86], we use values of β_k distributed according to a beta distribution $\text{Beta}(\alpha, 1)$, where $\beta_k = \left(\frac{k}{K}\right)^{1/\alpha}$, such that increasing α controls how skewed around zero the sequence $\{\beta_k\}_k$ is. For Fig. 3, we fix $\alpha = 0.5$ and $K = 20$ and, for each value of β_k , we sample 1000 graphs from $P_{\beta_k}(G | X^*)$, proposing $5N$ moves in-between each sample (see Appendix D).

4. Biases of the uncertainty coefficients

When an estimation of the mutual information is biased, it necessarily follows that an estimation of the resulting uncertainty coefficients will also be biased. Fortunately, we can show that the direction of the bias does not change either for the reconstructability $U(G|X)$ or the predictability $U(X|G)$. Suppose that $\mathcal{I}_\varepsilon = I(X; G)(1 + \varepsilon)$ is an estimator of the mutual information, where $\varepsilon \in \mathbb{R}$ is a small bias which can be either positive or negative. Then, the corresponding estimators of the uncertainty coefficients, that we denote \mathcal{P}_ε and \mathcal{R}_ε for the predictability and the reconstructability, respectively, are

$$\mathcal{P}_\varepsilon = \frac{\mathcal{I}_\varepsilon}{H(X|G) + \mathcal{I}_\varepsilon}. \quad (\text{C13})$$

and

$$\mathcal{R}_\varepsilon = \frac{\mathcal{I}_\varepsilon}{H(G)} = U(G|X)(1 + \varepsilon), \quad (\text{C14})$$

Note that we also suppose that $H(G)$ and $H(X|G)$ are not affected by the bias ε . For the first expression, we consider the first-order development of \mathcal{P}_ε with respect to ε :

$$\mathcal{P}_\varepsilon = U(X|G) \left[1 + \left(1 - U(X|G) \right) \varepsilon - \mathcal{O}(\varepsilon^2) \right]. \quad (\text{C15})$$

Indeed, given that $U(X|G) \geq 0$, the leading biased term $(1 - U(X|G))\varepsilon$ must have the same sign as ε . The second expression clearly shows that the bias of \mathcal{R}_ε is exactly given by ε . Therefore, both \mathcal{P}_ε and \mathcal{R}_ε retain the direction of bias of \mathcal{I}_ε .

Appendix D: Markov chain Monte-Carlo algorithm

To sample from the posterior distribution, we use a Markov chain Monte-Carlo (MCMC) algorithm where, starting from a graph G , we propose a move, denoted $\bar{G}^* \leftarrow G^*$, according to a proposition probability $P(\bar{G}^* | G^*)$, and accept it with the Metropolis-Hastings probability:

$$\min \left(1, e^{-\log \Delta \frac{P(G^* | \bar{G}^*)}{P(\bar{G}^* | G^*)}} \right), \quad (\text{D1})$$

where $\Delta = \frac{P(\bar{G}^*)P(X^* | \bar{G}^*)}{P(G^*)P(X^* | G^*)}$ is the ratio between the joint probability of the two graphs with X^* . This ratio can be computed efficiently in $\mathcal{O}(T)$, by keeping in memory $n_{i,t}$, the number of inactive neighbors, and $m_{i,t}$, a number of active neighbors, for each vertex i at each time t (see Ref. [39]). Equation (D1) allows to sample from the posterior distribution $P(G|X)$ without the requirement to compute the intractable normalization constant $P(X)$. We collect graph samples at every $N\delta$ moves, where we fix $\delta = 5$ in all experiments.

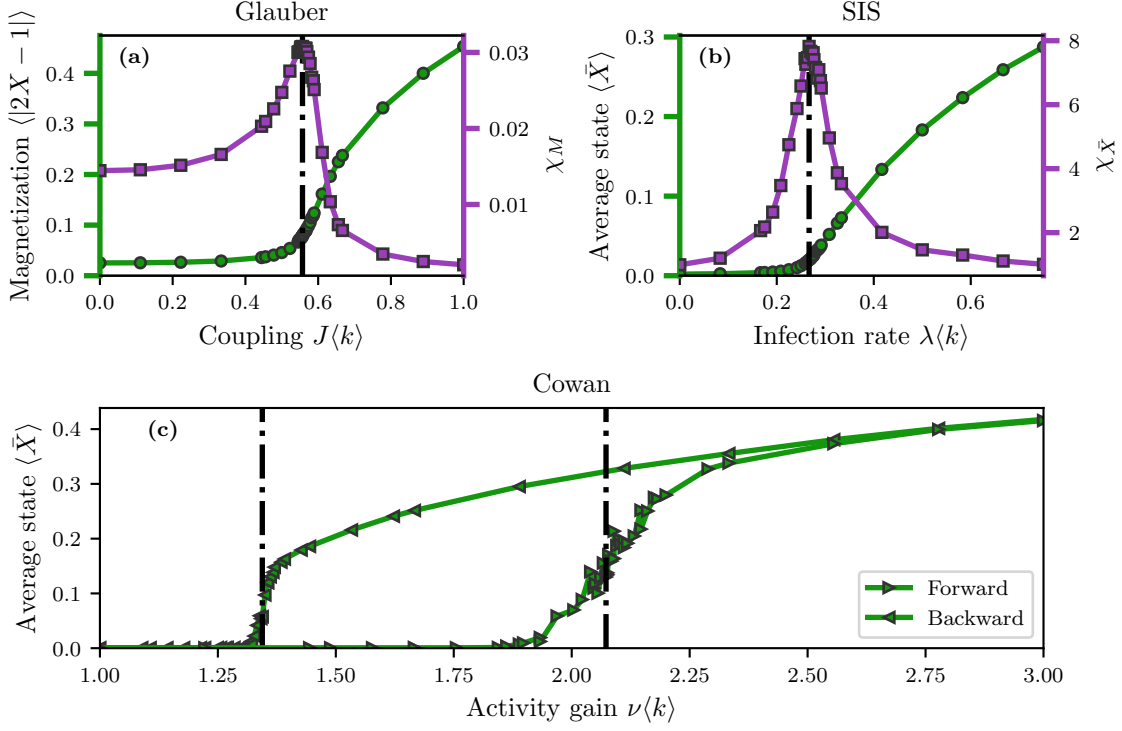


FIG. 5. Numerical evaluation of the phase transition thresholds: (a) Glauber dynamics, (b) SIS dynamics, (c) Cowan dynamics. For panels (a) and (b), the left axis (green) shows the order parameter (green circles), and the right axis (purple) shows the susceptibility (purple squares). For panel (c), only the order parameter is shown but for both the forward (right triangle) and backward (left triangle) branches. The values of the thresholds are indicated by the vertical dashed lines. We used the same parameters as those of Fig. 4, but increased the number of steps $T = 10^4$ to better sample from the dynamics. Each marker has been average over 48 realizations.

We consider two types of random graphs with different constraints: The Erdős-Rényi model and the configuration model. Hence, we need two different sampling propositions to apply our MCMC algorithm, that is one for each model. We assume that the support of the Erdős-Rényi model is the set of all simple graphs of N vertices with E edges. In this case, we consider an *hinge flip* move, where an edge (i, j) is sampled uniformly from the edge set of the graph G and a vertex k is sampled uniformly from its vertex set. Then, with probability $\frac{1}{2}$, we rewire edge (i, j) by either selecting i or j to connect with k . Note that, because we consider the support \mathcal{G}_N of G to be a space of simple graphs, all moves resulting in the addition of a self-loop or a multiedges are rejected with probability 1. As a result, the proposition probability is the same for any move $\bar{G}^* \leftarrow G^*$:

$$P(\bar{G}^*|G^*) = \frac{1}{EN} \Rightarrow \frac{P(G^*|\bar{G}^*)}{P(\bar{G}^*|G^*)} = 1. \quad (\text{D2})$$

For the configuration model, we assume that the support is the set of all loopy multigraphs of N vertices whose degree sequence is \mathbf{k} . In this case, we propose *double-edge swap* moves according to the prescription of Ref. [104]. We refer to it for further details.

Appendix E: Numerical estimation of the phase transition thresholds

We evaluate the phase transition thresholds of each dynamics using standard finite-size scaling techniques and Monte Carlo simulations (see Fig. 5). For Glauber, an adequate order parameter to visualize the phase transition is the magnetization $M := \frac{1}{N} \sum_i |2X_i - 1|$, where the absolute value breaks the spin symmetry [81]. In this process, it is well known that the susceptibility of the order parameter M , given by

$$\chi_M = \frac{\langle M^2 \rangle - \langle M \rangle^2}{\langle M \rangle}, \quad (\text{E1})$$

diverges at the threshold $J = J_c$ of the phase transition for infinite size systems [81]. In finite systems, χ_M instead reaches a maximum at $J = J_c$. We use this fact to locate J_c and show the corresponding results in Fig. 5(a).

For the SIS dynamics, a similar finite-size scaling analysis can be carried out, but a suitable order parameter is rather the average state $\bar{X} := \frac{1}{N} \sum_i X_i$. We also use a definition of the susceptibility that is more convenient

for spreading processes [19], given in terms of \bar{X} :

$$\chi_{\bar{X}} = \frac{\langle \bar{X}^2 \rangle - \langle \bar{X} \rangle^2}{\langle \bar{X} \rangle}, \quad (\text{E2})$$

which also diverges at the phase transition threshold $\lambda = \lambda_c$ for infinite size systems. We show the results for SIS in Fig. 5(b).

Finally, for the Cowan dynamics, we have a first-order phase transition characterized by a discontinuity of the order parameter \bar{X} in the infinite size limit, and a bistable region bounded by two thresholds $\nu_c^b < \nu_c^f$. To find these two thresholds, we evaluate the order parameter \bar{X} for varying values of the parameter ν , and find the location where the discontinuity occurs. We obtain the forward and backward branches by using different initial conditions, where the system is nearly inactive—with one

active vertex—and completely active—with no inactive vertex—, respectively.

For the Cowan dynamics, it is important to mention that since we consider relatively small systems ($N = 1000$ vertices), the bistable region is not clearly defined. Hence, a system starting in the forward branch can jump on the backward branch with a non-zero probability. This is why the expected discontinuity at the threshold is, in fact, populated (see Fig. 5(c)). This finite-size effect should be reduced for considering larger systems, but increasing N is unfortunately too computationally costly at the moment. Hence, to get a reasonable estimation of the thresholds in this scenario, we uniformly sample the set of ν 's, compute $\langle \bar{X} \rangle$ for all values of ν and find the point ν^* corresponding to the maximum gap between two points. Then, to increase the precision of this estimation, we zoom on a region centered at ν^* and do it again, until it converges. This method provides reasonably accurate thresholds for our purposes.

-
- [1] A.-L. Barabási, “Network science,” *Phil. Trans. R. Soc. A* **371**, 20120375 (2013).
 - [2] V. Latora, V. Nicosia, and G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, 2017).
 - [3] M. E. J. Newman, *Networks*, 2nd ed. (Oxford University Press, 2018).
 - [4] B. Barzel and A.-L. Barabási, “Universality in network dynamics,” *Nat. Phys.* **9**, 673–681 (2013).
 - [5] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Rev. Mod. Phys.* **87**, 925 (2015).
 - [6] S. Boccaletti, J. A. Almendral, S. Guan, I. Leyva, Z. Liu, I. Sendiña-Nadal, Z. Wang, and Y. Zou, “Explosive transitions in complex networks’ structure and dynamics: Percolation and synchronization,” *Phys. Rep.* **660**, 1–94 (2016).
 - [7] I. Iacopini, G. Petri, A. Barrat, and V. Latora, “Simplicial models of social contagion,” *Nat. Commun.* **10**, 2485 (2019).
 - [8] L. Hébert-Dufresne, S. V. Scarpino, and J.-G. Young, “Macroscopic patterns of interacting contagions are indistinguishable from social reinforcement,” *Nat. Phys.* **16**, 426–431 (2020).
 - [9] C. Murphy, E. Laurence, and A. Allard, “Deep learning of contagion dynamics on complex networks,” *Nat Commun.* **12**, 4720 (2021).
 - [10] J. Gao, B. Barzel, and A.-L. Barabási, “Universal resilience patterns in complex networks,” *Nature* **530**, 307 (2016).
 - [11] E. Laurence, N. Doyon, L. J. Dubé, and P. Desrosiers, “Spectral dimension reduction of complex dynamical networks,” *Phys. Rev. X* **9**, 011042 (2019).
 - [12] B. Pietras and A. Daffertshofer, “Network dynamics of coupled oscillators and phase reduction techniques,” *Phys. Rep.* **819**, 1–109 (2019).
 - [13] V. Thibeault, G. St-Onge, L. J. Dubé, and P. Desrosiers, “Threefold way to the dimension reduction of dynamics on networks: An application to syn-
 - chronization,” *Phys. Rev. Research* **2**, 043215 (2020).
 - [14] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Phys. Rev. Lett.* **86**, 3200 (2001).
 - [15] L. Hébert-Dufresne and B. M. Althouse, “Complex dynamics of synergistic coinfections on realistically clustered networks,” *Proc. Natl. Acad. Sci. USA* **112**, 10551–10556 (2015).
 - [16] G. St-Onge, J.-G. Young, E. Laurence, C. Murphy, and L. J. Dubé, “Phase transition of the susceptible-infected-susceptible dynamics on time-varying configuration model networks,” *Phys. Rev. E* **97**, 022305 (2018).
 - [17] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, “Master equation analysis of mesoscopic localization in contagion dynamics on higher-order networks,” *Phys. Rev. E* **103**, 032301 (2021).
 - [18] G. St-Onge, H. Sun, A. Allard, L. Hébert-Dufresne, and G. Bianconi, “Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks,” *Phys. Rev. Lett.* **127**, 158301 (2021).
 - [19] S. C. Ferreira, C. Castellano, and R. Pastor-Satorras, “Epidemic thresholds of the susceptible-infected-susceptible model on networks: A comparison of numerical and theoretical results,” *Phys. Rev. E* **86**, 041125 (2012).
 - [20] C. Castellano and R. Pastor-Satorras, “Relating topological determinants of complex networks to their spectral properties: Structural and dynamical effects,” *Phys. Rev. X* **7**, 041024 (2017).
 - [21] R. Pastor-Satorras and C. Castellano, “Eigenvector localization in real networks and its implications for epidemic spreading,” *J. Stat. Phys.* **173**, 1110–1123 (2018).
 - [22] L. Hébert-Dufresne, P.-A. Noël, V. Marceau, A. Allard, and L. J. Dubé, “Propagation dynamics on networks featuring complex topologies,” *Phys. Rev. E* **82**, 036115 (2010).
 - [23] G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, “Social confinement and meso-

- scopic localization of epidemics on networks,” *Phys. Rev. Lett.* **126**, 098301 (2021).
- [24] I. Brugere, B. Gallagher, and T. Y. Berger-Wolf, “Network structure inference, a survey: Motivations, methods, and applications,” *ACM Comput. Surv.* **51**, 1–39 (2018).
- [25] T. P. Peixoto, “Reconstructing networks with unknown and heterogeneous errors,” *Phys. Rev. X* **8**, 041011 (2018).
- [26] J.-G. Young, G. T. Cantwell, and M. E. J. Newman, “Bayesian inference of network structure from unreliable data,” *J. Complex Netw.* **8**, cnaa046 (2020).
- [27] J.-G. Young, F. S. Valdovinos, and M. E. J. Newman, “Reconstruction of plant–pollinator networks from observational data,” *Nat. Commun.* **12**, 3911 (2021).
- [28] E. Laurence, C. Murphy, G. St-Onge, X. Roy-Pomerleau, and V. Thibeault, “Detecting structural perturbations from time series using deep learning,” (2020), [arXiv:2006.05232](#).
- [29] S. McCabe, L. Torres, T. LaRock, S. Haque, C.-H. Yang, H. Hartle, and B. Klein, “netrd: A library for network reconstruction and graph distances,” *J. Open Source Softw.* **6**, 2990 (2021).
- [30] M. A. Kramer, U. T. Eden, S. S. Cash, and E. D. Kolaczyk, “Network inference with confidence from multivariate time series,” *Phys. Rev. E* **79**, 061916 (2009).
- [31] T. Schreiber, “Measuring Information Transfer,” *Phys. Rev. Lett.* **85**, 461–464 (2000).
- [32] A. K. Seth, “Causal connectivity of evolved neural networks during behavior,” *Netw. Comput. Neural Syst.* **16**, 35–54 (2005).
- [33] P. Abbeel, D. Koller, and A. Y. Ng, “Learning factor graphs in polynomial time and sample complexity,” *J. Mach. Learn. Res.* **7**, 1743–1788 (2006).
- [34] R. Salakhutdinov and I. Murray, “On the quantitative analysis of deep belief networks,” in *Proceedings of the 25th international conference on Machine learning* (2008) pp. 872–879.
- [35] J. Bento and A. Montanari, “Which graphical models are difficult to learn?” in *Advances in neural information processing systems* (2009) pp. 1303–1311.
- [36] R. Salakhutdinov and H. Larochelle, “Efficient learning of deep boltzmann machines,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics* (2010) pp. 693–700.
- [37] G. Bresler, E. Mossel, and A. Sly, “Reconstruction of markov random fields from samples: some observations and algorithms,” *SIAM J. Comput.* **42**, 563–578 (2013).
- [38] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchitsky, and R. Melko, “Quantum Boltzmann Machine,” *Phys. Rev. X* **8**, 021050 (2018).
- [39] T. P. Peixoto, “Network reconstruction and community detection from dynamics,” *Phys. Rev. Lett.* **123**, 128301 (2019).
- [40] M. Hinne, T. Heskes, C. F. Beckmann, and M. A. J. Van Gerven, “Bayesian inference of structural brain networks,” *NeuroImage* **66**, 543–552 (2013).
- [41] M. Breakspear, “Dynamic models of large-scale brain activity,” *Nat. Neurosci.* **20**, 340–352 (2017).
- [42] D. S. Bassett, P. Zurn, and J. I. Gold, “On the nature and use of models in network neuroscience,” *Nat. Rev. Neurosci.* **19**, 566 (2018).
- [43] Y. Wang, T. Joshi, X.-S. Zhang, D. Xu, and L. Chen, “Inferring gene regulatory networks from multiple microarray datasets,” *Bioinformatics* **22**, 2413–2420 (2006).
- [44] B. Prasse, M. A. Achterberg, L. Ma, and P. Van Mieghem, “Network-inference-based prediction of the COVID-19 epidemic outbreak in the Chinese province Hubei,” *Appl. Netw. Sci.* **5**, 35 (2020).
- [45] N. Musmeci, S. Battiston, G. Caldarelli, M. Puliga, and A. Gabrielli, “Bootstrapping topological properties and systemic risk of complex networks using the fitness model,” *J. Stat. Phys.* **151**, 720–734 (2013).
- [46] D. S. Bassett and O. Sporns, “Network neuroscience,” *Nat. Neurosci.* **20**, 353 (2017).
- [47] O. Sporns, “Structure and function of complex brain networks,” *Dialogues Clin. Neurosci.* **15**, 247–262 (2013).
- [48] A. Fornito, A. Zalesky, and M. Breakspear, “The connectomics of brain disorders,” *Nat. Rev. Neurosci.* **16**, 159–172 (2015).
- [49] M. P. Van den Heuvel and O. Sporns, “A cross-disorder connectome landscape of brain dysconnectivity,” *Nat. Rev. Neurosci.* **20**, 435–446 (2019).
- [50] B. Prasse and P. Van Mieghem, “Predicting Dynamics on Networks Hardly Depends on the Topology,” (2020), [arXiv:2005.14575](#).
- [51] Z. Zhang, P. Cui, and W. Zhu, “Deep Learning on Graphs: A Survey,” (2018), [arXiv:1812.04202](#).
- [52] J. Zhou, Gé Cui, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, “Graph Neural Networks: A Review of Methods and Applications,” (2018), [arXiv:1812.08434](#).
- [53] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, “How Powerful are Graph Neural Networks?” (2018), [arXiv:1810.00826](#).
- [54] C. Shah, N. Dehmamy, N. Perra, M. Chinazzi, A.-L. Barabási, A. Vespignani, and R. Yu, “Finding Patient Zero: Learning Contagion Source with Graph Neural Networks,” (2020), [arXiv:2006.11913](#).
- [55] A. Fout, J. Byrd, B. Shariat, and A. Ben-Hur, “Protein Interface Prediction using Graph Convolutional Networks,” in *Advances in Neural Information Processing Systems* (2017) pp. 6530–6539.
- [56] M. Zitnik, M. Agrawal, and J. Leskovec, “Modeling polypharmacy side effects with graph convolutional networks,” *Bioinformatics* **34**, i457–i466 (2018).
- [57] G. Bianconi, “Entropy of network ensembles,” *Phys. Rev. E* **79**, 036114 (2009).
- [58] K. Anand and G. Bianconi, “Entropy measures for networks: Toward an information theory of complex topologies,” *Phys. Rev. E* **80**, 045102(R) (2009).
- [59] K. Anand and G. Bianconi, “Gibbs entropy of network ensembles by cavity methods,” *Phys. Rev. E* **82**, 011116 (2010).
- [60] S. Johnson, J. J. Torres, J. Marro, and M. A. Munoz, “Entropic origin of disassortativity in complex networks,” *Phys. Rev. Lett.* **104**, 108702 (2010).
- [61] K. Anand, G. Bianconi, and S. Severini, “Shannon and von Neumann entropy of random networks with heterogeneous expected degree,” *Phys. Rev. E* **83**, 036109 (2011).
- [62] T. P. Peixoto, “Entropy of stochastic blockmodel ensembles,” *Phys. Rev. E* **85**, 056122 (2012).
- [63] J.-G. Young, P. Desrosiers, L. Hébert-Dufresne, E. Laurence, and L. J. Dubé, “Finite-size analysis of the detectability limit of the stochastic block model,” *Phys.*

- Rev. E **95**, 062304 (2017).
- [64] G. Cimini, T. Squartini, F. Saracco, D. Garlaschelli, A. Gabrielli, and G. Caldarelli, “The statistical physics of real-world networks,” *Nat. Rev. Phys.* **1**, 58 (2019).
 - [65] T. P. Peixoto, “Hierarchical block structures and high-resolution model selection in large networks,” *Phys. Rev. X* **4**, 011047 (2014).
 - [66] T. P. Peixoto, “Nonparametric bayesian inference of the microcanonical stochastic block model,” *Phys. Rev. E* **95**, 012317 (2017).
 - [67] T. DelSole and M. K. Tippet, “Predictability: Recent insights from information theory,” *Rev. Geophys.* **45**, RG4002 (2007).
 - [68] R. Kleeman, “Information theory and dynamical system predictability,” *Entropy* **13**, 612 (2011).
 - [69] S. V. Scarpino and G. Petri, “On the predictability of infectious disease outbreaks,” *Nat. Commun.* **10**, 1 (2019).
 - [70] J. P. Crutchfield and K. Young, “Inferring statistical complexity,” *Phys. Rev. Lett.* **63**, 105 (1989).
 - [71] D. P. Feldman and J. P. Crutchfield, “Measures of statistical complexity: Why?” *Phys. Lett. A* **238**, 244 (1998).
 - [72] H. Matsuda, K. Kudo, R. Nakamura, O. Yamakawa, and T. Murata, “Mutual information of ising systems,” *Int. J. Theor. Phys.* **35**, 839–845 (1996).
 - [73] S.-J. Gu, C.-P. Sun, and H.-Q. Lin, “Universal role of correlation entropy in critical phenomena,” *J. Phys. A* **41**, 025002 (2007).
 - [74] L. Barnett, J. T. Lizier, M. Harré, A. K. Seth, and T. Bossomaier, “Information flow in a kinetic Ising model peaks in the disordered phase,” *Phys. Rev. Lett.* **111**, 177203 (2013).
 - [75] M. Meijers, S. Ito, and P. R. ten Wolde, “Behavior of information flow near criticality,” *Phys. Rev. E* **103**, L010102 (2021).
 - [76] D. Edwards, *Introduction to Graphical Modelling* (Springer Science & Business Media, 2012).
 - [77] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. (Wiley-Interscience, 2006).
 - [78] Not to be confused with target space duality in string theory [105].
 - [79] R. J. Glauber, “Time-Dependent Statistics of the Ising Model,” *J. Math. Phys.* **4**, 294–307 (1963).
 - [80] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, 2009).
 - [81] K. Binder and D. Heermann, *Monte Carlo Simulation in Statistical Physics* (Springer, 2010).
 - [82] J. D. Cowan, “Stochastic neurodynamics,” in *Advances in Neural Information Processing Systems*, Vol. 3 (1990) p. 62.
 - [83] P. Van Mieghem and E. Cator, “Epidemics in networks with nodal self-infection and the epidemic threshold,” *Phys. Rev. E* **86**, 016116 (2012).
 - [84] R. M. Neal, “Annealed importance sampling,” *Stat. Comput.* **11**, 125 (2001).
 - [85] M. A. Newton and A. E. Raftery, “Approximate bayesian inference with the weighted likelihood bootstrap,” *J. Roy. Stat. Soc. B* **56**, 3 (1994).
 - [86] W. Xie, P. O. Lewis, Y. Fan, L. Kuo, and M.-H. Chen, “Improving marginal likelihood estimation for bayesian phylogenetic model selection,” *Syst. Biol.* **60**, 150 (2011).
 - [87] M. Opper and D. Saad, eds., *Advanced Mean Field Methods: Theory and Practice* (The MIT Press, 2001).
 - [88] It is not strictly accurate to say that our considered version of the SIS dynamics reaches a true absorbing state, since we allow for self-infection ϵ which allows it to escape the completely inactive state. Instead, it reaches a metastable state where most of the vertices are asymptotically inactive. However, it can be shown that the two phase transitions are quite similar for small ϵ [83].
 - [89] T. Gross and B. Blasius, “Adaptive coevolutionary networks: a review,” *J. R. Soc. Interface* **5**, 259–271 (2008).
 - [90] V. Marceau, P.-A. Noël, L. Hébert-Dufresne, A. Allard, and L. J. Dubé, “Adaptive networks: Coevolution of disease and topology,” *Phys. Rev. E* **82**, 036116 (2010).
 - [91] S. V. Scarpino, A. Allard, and L. Hébert-Dufresne, “The effect of a prudent adaptive behaviour on disease transmission,” *Nat. Phys.* **12**, 1042–1046 (2016).
 - [92] A. Khaledi-Nasab, J. A. Kromer, and P. A. Tass, “Long-Lasting Desynchronization of Plastic Neural Networks by Random Reset Stimulation,” *Front. Physiol.* **11**, 622620 (2021).
 - [93] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, “Hyperbolic geometry of complex networks,” *Phys. Rev. E* **82**, 036106 (2010).
 - [94] M. Boguñá, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. A. Serrano, “Network geometry,” *Nat. Rev. Phys.* **3**, 114–135 (2021).
 - [95] M. Boguñá, F. Papadopoulos, and D. Krioukov, “Sustaining the internet with hyperbolic mapping,” *Nat. Commun.* **1**, 1–8 (2010).
 - [96] F. Papadopoulos, R. Aldecoa, and D. Krioukov, “Network geometry inference using common neighbors,” *Phys. Rev. E* **92**, 022807 (2015).
 - [97] G. García-Pérez, A. Allard, M. A. Serrano, and M. Boguñá, “Mercator: uncovering faithful hyperbolic embeddings of complex networks,” *New J. Phys.* **21**, 123033 (2019).
 - [98] J. P. Gleeson, “High-Accuracy Approximation of Binary-State Dynamics on Networks,” *Phys. Rev. Lett.* **107**, 068701 (2011).
 - [99] P. J. Davis, *Interpolation and approximation* (Dover, 1975).
 - [100] W. Rudin, *Real and complex analysis*, 3rd ed. (McGraw-Hill, 1986).
 - [101] G. E. Andrews, R. Askey, R. Roy, and R. Askey, *Special functions*, Vol. 71 (Cambridge University Press, 1999).
 - [102] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities-Ordinary Differential Equations, vol. I* (Academic Press, 1969).
 - [103] D. E. Knuth, “Big Omicron and Big Omega and Big Theta,” *SIGACT News*, 18 (1976).
 - [104] B. K. Fosdick, D. B. Larremore, J. Nishimura, and J. Ugander, “Configuring random graph models with fixed degree sequences,” *SIAM Rev.* **60**, 315–355 (2018).
 - [105] A. Giveon, M. Porrati, and E. Rabinovici, “Target space duality in string theory,” *Phys. Rep.* **244**, 77–202 (1994).