

PAPER: Disordered systems, classical and quantum

Uncovering the non-equilibrium stationary properties in sparse Boolean networks

Giuseppe Torrisi*, Reimer Kühn and Alessia Annibale

Department of Mathematics, King's College London, Strand, WC2R 2LS,
United Kingdom

E-mail: Giuseppe.torrisi@kcl.ac.uk

Received 15 February 2022

Accepted for publication 7 April 2022

Published 13 May 2022

Online at stacks.iop.org/JSTAT/2022/053303
<https://doi.org/10.1088/1742-5468/ac66d0>



Abstract. Dynamic processes of interacting units on a network are out of equilibrium in general. In the case of a directed tree, the dynamic cavity method provides an efficient tool that characterises the dynamic trajectory of the process for the linear threshold model. However, because of the computational complexity of the method, the analysis has been limited to systems where the largest number of neighbours is small. We devise an efficient implementation of the dynamic cavity method which substantially reduces the computational complexity of the method for systems with discrete couplings. Our approach opens up the possibility to investigate the dynamic properties of networks with fat-tailed degree distribution. We exploit this new implementation to study properties of the non-equilibrium steady-state. We extend the dynamic cavity approach to calculate the pairwise correlations induced by different motifs in the network. Our results suggest that just two basic motifs of the network are able to accurately describe the entire statistics of observed correlations. Finally, we investigate models defined on networks containing bi-directional interactions. We observe that the stationary state associated with networks with symmetric or anti-symmetric interactions is biased towards the active or inactive state respectively, even if independent interaction entries are drawn from a symmetric distribution. This phenomenon, which can be regarded as a form of spontaneous symmetry-breaking, is peculiar to systems formulated in terms of Boolean variables, as opposed to Ising spins.

*Author to whom any correspondence should be addressed.

Keywords: cavity and replica method, message-passing algorithms, dynamical heterogeneities, kinetic Ising models

Contents

1. Introduction	2
2. Derivation of dynamic cavity equation on fully asymmetric networks ...	5
3. Node activation heterogeneity	7
3.1. Comparison between theory and simulation.....	8
3.2. Comparison with the nMF approximation.....	11
4. Pairwise correlation for the linear threshold model on fully asymmetric networks	14
5. Extensions	19
5.1. Multi-node interaction model.....	19
5.2. Any Boolean function can be written in terms of a linear combination of AND gates.....	21
5.3. Spin dynamics.....	22
6. Beyond fully asymmetric networks	24
6.1. The one-time-step approximation	25
6.2. Comparison between OTA and simulations	27
7. Spontaneous symmetry breaking of the linear threshold model for unbiased parameter choices	28
8. Summary and discussion	31
Acknowledgments	32
Appendix A. High noise regime	32
Appendix B. Derivation of the OTA	33
References	38

1. Introduction

A broad range of disordered and complex systems can be modelled in terms of binary units, which update their state according to stochastic Boolean functions of the neighbouring units. They include spin glasses [1–3] in physics, gene-regulatory and immune networks [4–8] in biology, artificial neural networks [9–12] in computer science, agent-based models [13–16], models of operational or credit risk [17, 18] in economics and finance, and a variety of hard combinatorial optimization problems [19–23] to name but

a few. In recent years there was, in particular, a surge of interest in systems where each unit interacts only with a *finite* set of neighbours, and which are thus defined on finitely connected networks.

Statistical mechanics provides a rich arsenal of tools to analyse the equilibrium behaviour of sparse systems [3, 24–26]. However, the assumption of equilibrium is not satisfied in systems that are driven, dissipative or exhibit a degree of asymmetry in the interactions. Here, one has to resort to tools of non-equilibrium statistical mechanics, which are however much more challenging. Non-equilibrium dynamics has been analysed in explicit detail for a class of models where the dynamics of units has an absorbing state [27–30] often combined with the simplifying feature that interactions are of pure product form, e.g. in contagion processes [31, 32].

However, the dynamics of a large class of Boolean models does not have an absorbing state. In the present manuscript, we analyse the stochastic dynamics of linear threshold models, formulated in terms of $\{0, 1\}$ variables, and discuss extensions to threshold models with multi-node interactions and models with Ising spin $\{\pm 1\}$, which do not have absorbing states. In such models, approximation schemes that are successfully used for dense systems, such as the heterogeneous mean-field and TAP approaches [23, 33], have been shown to be ineffective for sparse systems [34, 35]. On the other hand, generating functional analyses [36] can accurately characterise site averaged quantities, such as global magnetization and time-lagged correlations [37, 38], in sparse heterogeneous systems. However they require averaging over graph ensembles, hence they are not able to describe single instances. These can instead be investigated by the dynamic cavity method [29, 32, 38, 39] which is particularly effective in the analysis of sparse systems, whenever short loops are rare. In particular, the dynamic cavity method can potentially investigate dynamic properties at the level of individual nodes [39]. Single-node statistics has, for instance, been shown to be highly heterogeneous for non-equilibrium models with an absorbing state [27, 32, 40]. Little, however, is known for other models, including spin models and linear threshold models.

In particular, explicit analysis has been feasible only for quite a narrow choice of networks, namely directed trees and fully asymmetric random graphs, where bi-directional links are sampled independently [38, 39]. Moreover, both the dynamic cavity and generating functional methods require evaluating averages over a configuration space that grows exponentially with the in-degree of the nodes, which limits the applicability of both methods to networks with small in-degrees. Due to this computational complexity barrier, a large class of networks, in particular those characterised by a fat-tailed degree distribution, are out of reach for analytical characterisation. Networks with fat-tailed degree distribution are, however, known to appear in many real-world problems and represent the outcome of generating network models based on preferential attachment [41–46]. They are characterised by large differences in the local environment surrounding nodes, which are known to lead to heterogeneous node properties at equilibrium [26]. Therefore, the aforementioned exponential complexity in the in-degrees has so far prevented a detailed study of node heterogeneities, precisely in those systems where they play a bigger role.

We have recently proposed a method based on dynamic programming that overcomes this complexity barrier when couplings are chosen randomly from *any* discrete set of

equidistant values, the simplest example being couplings chosen from the set $\{\pm J\}$, see [47]. Our algorithm reduces the computational complexity in the in-degrees of the system from exponential to quadratic, thus allowing us to characterise the probabilistic evolution of individual node states in networks with a broad in-degree distribution.

In this manuscript, we extend our analysis of heterogeneous properties of the linear threshold model in fully asymmetric networks to include pairwise correlations between node states. Our analysis shows that the statistics of pairwise correlation is almost completely captured in terms of simple motifs of length two. This reveals the extent to which, in terms of path distance, node states significantly influence each other.

In addition, we extend our formalism to investigate systems with multi-node interactions, using bipartite (or factor) graphs, which provide a more flexible model for combinatorial control; see [48]. Multi-node interaction models can represent generic Boolean functions, which are of particular interest in the context of gene regulation [4, 49–51]. We present our extended formalism for models of indicator variables 0, 1 as well as models of Ising spins ± 1 , for which our formulation in terms of bipartite graphs leads to (sparse) mixed p -spin models.

Finally, we extend our analysis to systems with bi-directional links. For this class of systems, the cavity method requires following the probability of the full dynamic trajectory, which cannot be reduced to the product of one-time-step terms, because bi-directional links introduce retarded self-interactions of nodes with their past. This leads to an exponential scaling with the time horizon considered, which effectively restricts the application of the method to follow the dynamics over a few time-steps only, thus making the investigation of long time behaviour infeasible. An approximation technique called the one-time approximation (OTA) [34, 35, 39] has been proposed to overcome these restrictions.

The OTA factorises the probabilities of dynamic trajectories into one-time-step probabilities, and it leads to two sets of coupled equations, one for the marginals and one for the cavity marginals, which become closed under further assumptions. This makes the problem solvable also for the stationary state. Both equations are affected by the aforementioned in-degree complexity, hence the OTA greatly benefits from our dynamic programming method. Within the OTA framework, different closure schemes have been proposed in the literature [34, 35]. We evaluate the impact that different closure schemes have on the results, and we use such analysis to motivate our choice. A thorough inspection of the result leads to a surprising finding: networks with the same number of positive and negative interactions may sustain a biased distribution of node activation in presence of bi-directional links. This phenomenon, which can be regarded as a form of spontaneous symmetry breaking, is peculiar of systems with $\{0, 1\}$ representation of states (as opposed to Ising spins with $\{\pm 1\}$ representation).

Our paper is organized as follows. In section 2 we describe the dynamic programming approach to the linear threshold model on fully asymmetric networks with fat-tailed degree distribution. Its non-equilibrium stationary state is characterised in terms of individual nodes statistics in section 3, and pairwise site correlations in section 4. In section 5 we extend our method both to models with multi-node interactions, formulated in terms of bipartite networks, and to networks of Ising spins. In section 6 we include bi-directional links in the analysis of linear threshold models, using the OTA. In section 7

we investigate the unexpected occurrence of a form of spontaneous symmetry breaking in these systems. We summarise and discuss our results in section 8. Finally, we include two appendices: the first containing the high noise limit for the asymmetric network model and our multi-node interactions model, the second containing details concerning the derivation of OTA schemes. The interested reader can find the code to reproduce the results shown in this paper at the following link¹.

2. Derivation of dynamic cavity equation on fully asymmetric networks

We consider a directed network consisting of N nodes labelled $i = 1, \dots, N$. For each edge (i, j) , the edge weight $J_{ij} \in \mathbb{R}$ defines the strength of the interaction carried from node j to node i , while $J_{ij} = 0$ if the link (i, j) does not exist. In this section we will consider fully asymmetric networks, i.e. directed networks such that $\forall i, j$ if the link (i, j) exists, the link in the opposite direction (j, i) does not exist. We denote $\partial_i = \{j : J_{ij} \neq 0\}$ the set of predecessors of node i and $k_i^{\text{in}} = |\partial_i|$ the in-degree of node i . We will drop the superscript ‘in’ in k_i^{in} whenever it does not lead to ambiguity. Every node i can be in one of two states, described by a binary state variable $n_i \in \{0, 1\}$ with $i \in \{1, \dots, N\}$. Several Boolean models can be defined on a given graph depending on the choice of the update rule of node states, a prominent example being the random Boolean network model [4].

In the present section, we consider the linear threshold model defined on a weighted directed network, which is specified by the interaction matrix $\mathbf{J} = (J_{ij})$ and thresholds $\{\vartheta_i\}_{i=1}^N$. A node i evaluates a local field

$$h_i(\mathbf{n}) := \sum_j J_{ij} n_j, \quad (1)$$

and becomes active in the next time-step if the sum of the local field and a random noise contribution z_i exceeds the threshold ϑ_i , i.e.

$$n_i(t+1) = \Theta(h_i(\mathbf{n}(t)) - \vartheta_i - z_i(t)), \quad (2)$$

where the $z_i(t)$ are independent identically distributed random variables, with cumulative distribution function $\text{Prob}[z_i(t) < x] := \Phi_\beta(x)$, for all i, t , and $T = \beta^{-1}$ is a parameter that characterises the noise strength. The local field $h_i(\mathbf{n}_{\partial_i})$ depends on the states of nodes that are predecessors of node i . We use $\mathbf{n}_{\partial_i} = \{n_j, j \in \partial_i\}$ to denote the states of this set. In the following, we explicitly write the argument of the local field using the notation $h_i(\mathbf{n}_{\partial_i})$. Given the states \mathbf{n}_{∂_i} , the probability that node i is active at the next time-step is $\Phi_\beta(h_i(\mathbf{n}_{\partial_i}) - \vartheta_i)$, hence the transition probability to state n_i , for node i , is

$$W[n_i | h_i(\mathbf{n}_{\partial_i}), \vartheta_i] = n_i \Phi_\beta(h_i(\mathbf{n}_{\partial_i}) - \vartheta_i) + (1 - n_i)(1 - \Phi_\beta(h_i(\mathbf{n}_{\partial_i}) - \vartheta_i)). \quad (3)$$

¹<https://doi.org/10.5281/zenodo.5996772>.

Given the joint probability $P(\mathbf{n}_{\partial_i}, t)$ of the states \mathbf{n}_{∂_i} at time t , the probability that node i is in state n_i at time $t + 1$ is given by

$$P(n_i, t + 1) = \sum_{\mathbf{n}_{\partial_i}} W[n_i | h_i(\mathbf{n}_{\partial_i}), \vartheta_i] P(\mathbf{n}_{\partial_i}, t), \quad (4)$$

where $\sum_{\mathbf{n}_{\partial_i}}(\cdot) := \prod_{j \in \partial_i} \sum_{n_j = \{0,1\}}(\cdot)$ indicates the sum over all possible configurations of the microscopic variables. For finitely coordinated random graphs of the type considered here, the cavity method [24] can be used to analyse the dynamics of the system. It is exact on trees and known to become exact for finitely coordinated random graphs in the thermodynamic limit, as the typical length of any loops in such systems diverges logarithmically in system size N . On the cavity graph where node i , along with the edges connected to it, has been removed, the cavity approximation assumes

$$P^{(i)}(\mathbf{n}_{\partial_i}, t) = \prod_{j \in \partial_i} P^{(i)}(n_j, t), \quad (5)$$

which holds on a tree-like structure. The probability $P^{(i)}(\mathbf{n}_{\partial_i}, t)$ of the states of nodes neighbouring i will in general be different from the corresponding probability on the original graph, i.e. $P(\mathbf{n}_{\partial_i}, t) \neq P^{(i)}(\mathbf{n}_{\partial_i}, t)$, but equality does hold on a directed tree and it holds approximately on a random fully asymmetric graph, in which loops are rare [38, 39, 50].² Using the equivalence of the cavity and non-cavity dynamics, from equation (4)

$$P(n_i, t + 1) = \sum_{\mathbf{n}_{\partial_i}} W[n_i | h_i(\mathbf{n}_{\partial_i}), \vartheta_i] \prod_{j \in \partial_i} P(n_j, t). \quad (6)$$

In particular, the probability of node activation becomes

$$P(n_i = 1, t + 1) = \sum_{\mathbf{n}_{\partial_i}} \Phi_\beta(h_i(\mathbf{n}_{\partial_i}) - \vartheta_i) \prod_{j \in \partial_i} P(n_j, t). \quad (7)$$

Let us denote by $P_i(t) := P(n_i = 1, t)$ the activation probability of node i at time t , let the angle bracket $\langle \cdot \rangle_{\mathbf{n}_{\partial_i}, t}$ indicate an average over the states of the predecessors of node i , evaluated using their joint node activation probabilities at time t

$$\langle (\cdot) \rangle_{\mathbf{n}_{\partial_i}, t} := \sum_{\mathbf{n}_{\partial_i}} \prod_{j \in \partial_i} P_j(t)^{n_j} [1 - P_j(t)]^{1-n_j} (\cdot). \quad (8)$$

Combining equations (8) and (7) a concise expression for the activation probability $P_i(t + 1)$ is obtained as

$$P_i(t + 1) = \langle \Phi_\beta(h_i(\mathbf{n}_{\partial_i}) - \vartheta_i) \rangle_{\mathbf{n}_{\partial_i}, t}. \quad (9)$$

The average in equation (9) is evaluated over a configuration space containing 2^{k_i} elements, as shown in equation (8). Thus, the evaluation of the average hits an (exponential) complexity barrier whenever $k_i \gg 1$. In [47] we have proposed a new method

²Cavity method can be applied also in the general case, as discussed in [39] and in appendix B. However the evaluation is more challenging.

that uses dynamic programming to efficiently perform the average in equation (9) using a polynomial number of operations. In the following, we recap the main ideas of the method.

Dynamic programming. Equation (9) depends on the distribution of states \mathbf{n}_{∂_i} through the local field $h_i(\mathbf{n}_{\partial_i})$. If couplings are extracted from a discrete set of values, e.g. $J_{ij} = \{0, \pm J\}$, there is a degeneracy of values of $h_i(\mathbf{n}_{\partial_i})$, i.e. several configurations of k_i states $(n_{j_1}, \dots, n_{j_{k_i}})$ are associated with the same value of local field. In fact, for a given realisation of the interaction terms $(J_{ij_1} \dots J_{ij_{k_i}})$ the number of possible values of $h_i(\mathbf{n}_{\partial_i})$ is $k_i + 1$, while the number of configurations of states is 2^{k_i} . We exploit this degeneracy by averaging directly over the local fields, rather than over the micro-states of the set of predecessors, using a dynamic programming approach. We show below that our approach reduces the computational complexity of the problem from $\mathcal{O}(2^{k_i})$ to $\mathcal{O}(k_i^2)$, [47].

Let us consider a node i with in-degree k_i , and let us use $\{1, \dots, k_i\}$ to label the indices of the predecessors of node i . We first define the sub-problem

$$f_i(\ell, \tilde{h}) = \left\langle \Phi_\beta \left(\tilde{h} + \sum_{j=\ell}^{k_i} J_{ij} n_j - \vartheta_i \right) \right\rangle_{n_\ell, \dots, k_i, t} \quad \text{for } \ell \in \{1, \dots, k_i\}, \quad (10)$$

which consists in performing the average over a sub-set of nodes. We will refer to \tilde{h} as an auxiliary field. At any given ℓ , the set of $f_i(\ell, \tilde{h})$ of interest would in fact correspond to the set of averages representing the sub-problems of equation (9) that remain to be evaluated after the average over the first $\ell - 1$ nodes has been performed. Averages in equation (10) are not performed directly, but through the recursive relationship

$$f_i(\ell, \tilde{h}) = P_\ell(t) f_i(\ell + 1, \tilde{h} + J_{i\ell}) + (1 - P_\ell(t)) f_i(\ell + 1, \tilde{h}), \quad (11)$$

for $1 \leq \ell \leq k_i$, with the *terminal boundary condition*

$$f_i(k_i + 1, \tilde{h}) = \Phi_\beta(\tilde{h} - \vartheta_i). \quad (12)$$

The original average $P_i(t + 1)$ of equation (9) that we are ultimately interested in is, within this backward iteration scheme, obtained as

$$P_i(t + 1) = f_i(1, 0). \quad (13)$$

In our earlier work [47], we have shown that equation (11) leads to a dramatic reduction of the computational complexity for binary interactions $J_{ij} \in \{0, \pm J\}$. The computational and memory requirement for the evaluation of $P_i(t + 1)$ is seen to scale as k_i^2 , for a node with in-degree k_i , rather than 2^{k_i} , as it would in a naive evaluation. This entails that the complexity of our algorithm is $\mathcal{O}(\sum_i k_i^2)$.

3. Node activation heterogeneity

It has been shown that fully asymmetric graphs do not exhibit a spin-glass phase in a variety of dynamical models, and the time-dependent local magnetizations or node activation probabilities are expected to converge to a fixed point [37, 39, 52, 53]. We exploit this fact to investigate the trajectory of nodes' states by iteration of equation (9). Our dynamic programming implementation provides an efficient method to evaluate node activation probabilities. We apply our formalism to inspect the dynamical properties of networks with fat-tailed degree distributions. The theoretical evaluations presented in this section would not be feasible without our dynamic programming implementation.

Model parameters used in this study. In the present section, we consider models of synthetic networks in the configuration model class, i.e. we will study ensembles of maximally random directed graphs \mathbf{J} with constrained in-degree and out-degree sequences of nodes. The in-degree of nodes are drawn from the distribution $\rho^{\text{in}}(k) = N^{-1} \sum_i \delta_{k, k_i^{\text{in}}}$. In this paper, we use the same distribution for the out-degree of nodes k_i^{out} .³ The in-degree and out-degree for the same node are drawn independently. We use $\langle k^{\text{in}} \rangle$ to denote the mean in-degree. Our synthetic model uses parameters observed in real-world networks. Specifically, we use parameters extracted from a gene regulatory network [54], which is characterised by a node in-degree distribution with a power-law tail $P_\gamma(k) \sim \gamma k^{-\gamma-1}$ and $\gamma = 2.81$, which we implement in terms of the discrete fat-tailed distribution $P_\gamma(k) = k^{-\gamma} - (k+1)^{-\gamma}$; it is defined for positive integers $1 \leq k \in \mathbb{N}$ and has the desired power-law behaviour for large k . For every link (i, j) , the corresponding interaction term J_{ij} is sampled from the distribution $P(J_{ij}) = \eta \delta(J_{ij} - J) + (1 - \eta) \delta(J_{ij} + J)$ and we use $\eta = 0.621$ as suggested by the data in [54] and $J = 1/\sqrt{\langle k^{\text{in}} \rangle} \approx 0.72$ unless stated otherwise. We set the value of the threshold to be the same for all sites, $\vartheta_i = \vartheta \forall i$. We assume a 'thermal' noise distribution with $\Phi_\beta(x) = 1/2 (1 + \tanh \frac{\beta x}{2})$.

3.1. Comparison between theory and simulation

The dynamic programming algorithm allows to speed up the evaluation of the node activation probability at time $t+1$ given the knowledge of node activation probabilities at time t . By iterating the procedure over the time-steps of interest, we solve the dynamics for the nodes activation probabilities. In this section, we compare the analytical results against Monte Carlo (MC) simulations of the microscopic dynamics. Several results in the literature indicate that the dynamic cavity method is effective to reproduce the trajectory of fully asymmetric graphs, see [39, 55]. While those results inspect the site average of node activations for Ising spins, we consider the full distribution of node activations for indicator variables. In particular, we compare the stationary values of the node activation probabilities $P_i = \lim_{t \rightarrow \infty} P_i(t) \forall i \in \{1, \dots, N\}$ obtained from simulations with those obtained via the cavity approach. The stationary activation probabilities are obtained by running the iterative procedure equation (9) until convergence.

³If the distributions of in- and out-degree are drawn from different distributions, a minimal requirement is $\sum_i k_i^{\text{in}} = \sum_i k_i^{\text{out}}$ that guarantees the conservation of the number of links.

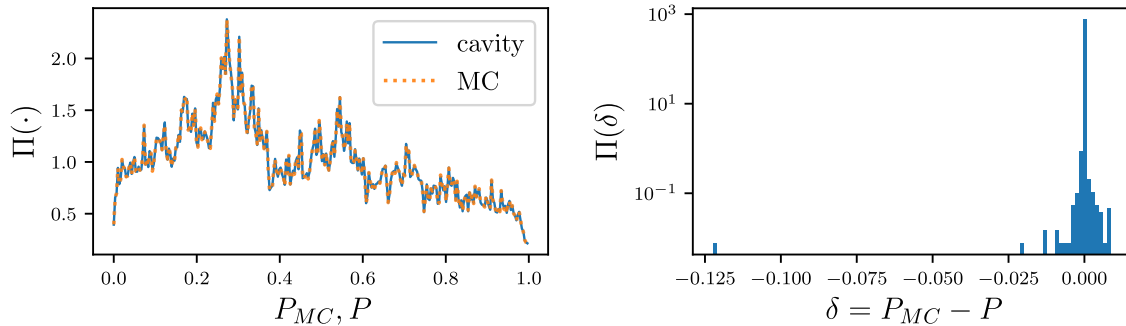


Figure 1. (Left) Distribution of the node activation probability, as computed from cavity theoretical calculation (blue solid line) and from MC direct simulation of the microscopic dynamics (dotted green line). (Right) Distribution of the differences $\delta_i = P_{i,\text{MC}} - P_i$, evaluated for each node i of the network. Parameters: $N = 100\,000$, $T/J = 0.2$, $\gamma = 2.81$, $\eta = 0.5$, $\vartheta/J = 0.1$, $J = 1$.

The convergence is controlled by measuring the difference between activation probabilities $\epsilon_{t+1} = \max_i |P_i(t+1) - P_i(t)|$ in two consecutive steps; this criterion is adopted in the entire manuscript. Let us call $P_i = P_i(t)$ corresponding to the smallest t satisfying the exit condition, which we took to be $\epsilon_t < 10^{-4}$. We test cavity predictions against simulations of the microscopic dynamics equation (2). We evaluate the stationary node activation probability of simulated dynamical trajectories by taking the sample average of node activations

$$P_{i,\text{MC}} = \frac{1}{t_s} \sum_{t=t_0}^{t_0+t_s} n_i(t), \quad (14)$$

for $t_s \gg 1$, and t_0 a sufficiently long time to allow the dynamics to relax to stationarity. In the following, we use the general notation $\Pi(x)$ to denote the probability density function for the realisation of local quantities x_i , e.g. $\Pi(P) = N^{-1} \sum_i \delta(P_i - P)$. In figure 1 we show the distribution of the node activation probability as computed via the cavity method and by estimation from simulations. These are found to be in excellent agreement, confirming the validity of the cavity method to describe the stationary state, see figure 1. We note that, to reach the resolution imposed by cavity $\epsilon_t < 10^{-4}$ through simulations, we need to simulate long trajectories in equation (14) with $t_s \sim 10^8$, which makes the evaluation from simulation computationally far more expensive than an evaluation based on the cavity method. For the same realisation of a network of $N = 100\,000$ nodes and identical parameter choices, the cavity implementation took 9 s, whereas simulations took more than 11 days on the same machine.

Time evolution. We investigate the statistical properties of dynamical trajectories through forward iteration of equation (9). We initialise the system with random initial conditions for the node activation probabilities $P_i(0)$. We stop iterations at convergence using the same criterion as above. The time evolution of the dynamics for each site is shown in figure 2. Node activation probabilities converge after very few time-steps. This property is linked to the absence of feedback signals in the dynamics, thanks to the directed tree-like structure of the network. The probability density function of the node

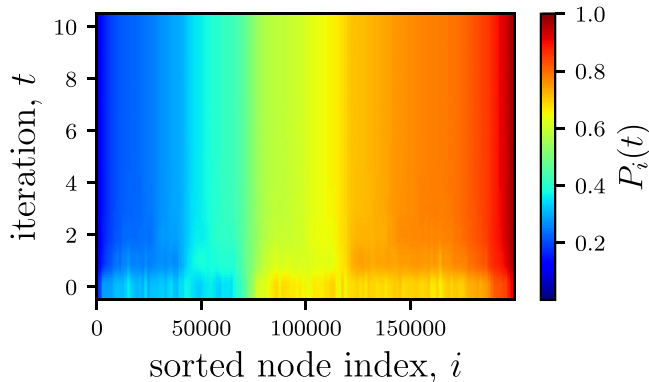


Figure 2. Trajectory of probabilities of node activation up to convergence. The colour of the pixel located at row t and column i maps the value $P_i(t)$. Colour-bar on the right maps the colours to the value of activation probability. For visualisation purposes nodes are sorted by the value of P at the end of iteration, i.e. column locations are sorted according to their value in the top row. Parameters: $N = 200\,000$, $\vartheta/J = 0.2$, $T/J = 0.3$, $J \approx 0.72$, $\eta = 0.621$.

activation probabilities at stationarity is shown in figure 3 for different noise levels T . The distribution of the node activation probabilities is not uni-modal, indicating the need to inspect the full distribution of node activations, as measures of central tendency are not representative. This is also true in the high noise limit, where the problem is directly solvable; see appendix A.

The multi-modal structure of the node activation distribution can be rationalised by simple reasoning. Thanks to the fat tailed property of the network, more than 87% of nodes have in-degree 1. Let us consider a node i with only one predecessor j , let us call $\rho_{\pm}(P)$ the probability that node i is active, given that $J_{ij} = \pm J$, and that the node activation probability of that predecessor is P , i.e. $\rho_{\pm}(P) = \langle \Phi_{\beta}(\pm Jn - \vartheta) \rangle_n$ with $\langle n \rangle_n = P$. The explicit expression of functions $\rho_{\pm}(P)$ becomes

$$\begin{aligned} \rho_{\pm}(P) &= \frac{P}{2} \left[1 + \tanh \frac{\beta}{2} (\pm J - \vartheta) \right] + \frac{1-P}{2} \left[1 - \tanh \frac{\beta \vartheta}{2} \right] \\ &= \frac{1}{2} \left[1 - \tanh \frac{\beta \vartheta}{2} \right] + \frac{P}{2} \left[\tanh \frac{\beta}{2} (\pm J - \vartheta) + \tanh \frac{\beta \vartheta}{2} \right]. \end{aligned} \quad (15)$$

The properties of nodes with in-degree 1 can be captured in terms of the map shown in figure 4. We denote p^* the fixed point defined as $p^* = \rho_{+}(p^*)$. The composition of function ρ_{\pm} is able to describe the trends observed in figure 3. In figure 5, we highlight the contribution to the distribution of node activation probability from nodes with in-degree 1. The values of $\rho_{\pm}(1)$ give the lower and upper bounds of the node activation probability for nodes with in-degree 1. Even though nodes with in-degree 1 account for the vast majority of nodes, the dynamical properties of node activation depends on the contribution of other nodes as well.

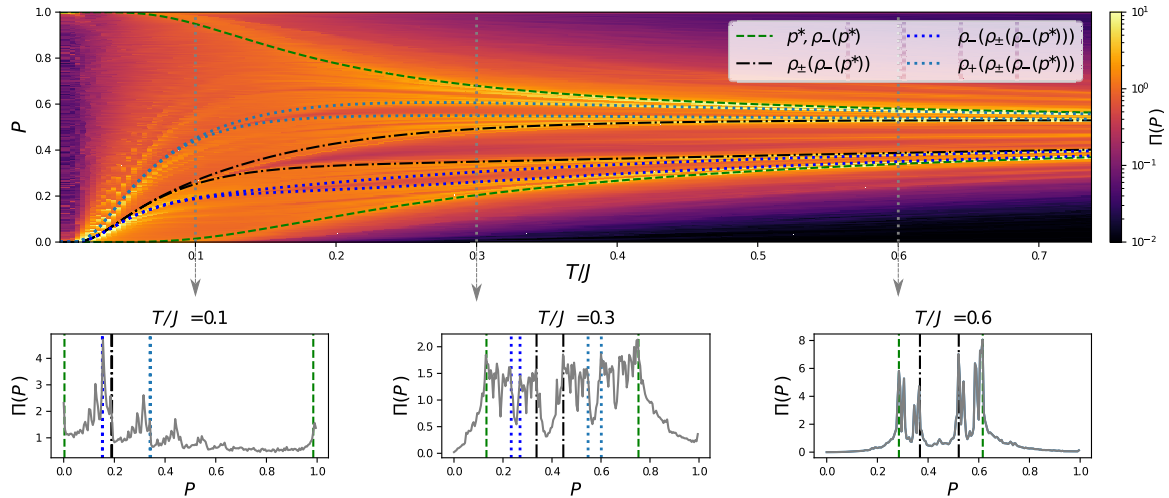


Figure 3. Heat-map of the distribution of node activation probabilities $\Pi(P)$ at different values of noise parameter T/J (top). Dashed, dotted, and dot-dashed lines show the graph of different function compositions of ρ_{\pm} at different values of T , see equation (15). Vertical dotted lines mark the noise levels at which the histograms of $\Pi(P)$ are shown in the lower panels. They probe the regimes at low, intermediate and high noise, corresponding to $T/J = 0.1, 0.3, 0.6$ from left to right. The central lower panel corresponds to the histogram of the $\Pi(P)$ realised in figure 2, once the system has converged. The same parameters as in figure 2 are adopted.

3.2. Comparison with the nMF approximation

Solving the dynamics of the linear threshold model for generic couplings is computationally demanding as already discussed above. In the context of disordered systems, approximate methods have been proposed, which are effective for models with dense interactions, such as the dynamic naive mean-field (nMF) approximation and dynamic TAP equations, for interactions with an arbitrary degree of symmetry [33], and specifically for fully asymmetric interactions [56]. However, for diluted systems, dynamic nMF approximations have been shown to be less effective [35, 57] for systems with both ferromagnetic and spin-glass interactions, especially at low temperature. In this section, we investigate the distribution of node activations and we highlight that even in the region of parameters where the site average of the node activation probability is well captured by the nMF method, the nMF turns out to be incapable of reliably capturing the full distribution of node activation probabilities.

We briefly expose the intuitive idea of the mean-field approximation, and we refer to [35] for more details. For a node i let us consider the local field $h_i = \sum_j J_{ij} n_j$ and the probability associated $\text{Prob}(h_i) = \langle \delta(h_i - \sum_j J_{ij} n_j) \rangle_n$. The core idea of the mean-field approximation is to approximate $\text{Prob}(h_i)$ with a delta-function peaked on the mean value $\langle h_i \rangle$, which implies $\langle \Phi_{\beta}(h_i - \vartheta_i) \rangle \approx \Phi_{\beta}(\langle h_i \rangle - \vartheta_i)$. This assumption is motivated in the limit where the number of neighbours of node i diverges where, thanks to the law of large numbers, one expects $\text{Prob}(h_i)$ to concentrate—for suitably normalized J_{ij} —around the mean value. However, in the case of sparse matrices, the number of

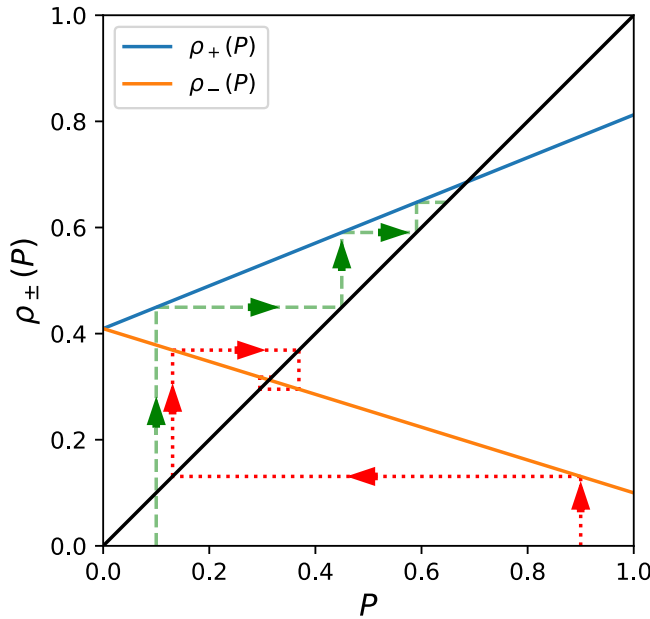


Figure 4. Graphical solution of the equation $\rho_{\pm}(P^*) = P^*$. The black line represents the graph of the function $f(x) = x$. Fixed points p_{\pm}^* are obtained by the intersections of the lines ρ_{\pm} with the bisector. The same parameters as in figure 2.

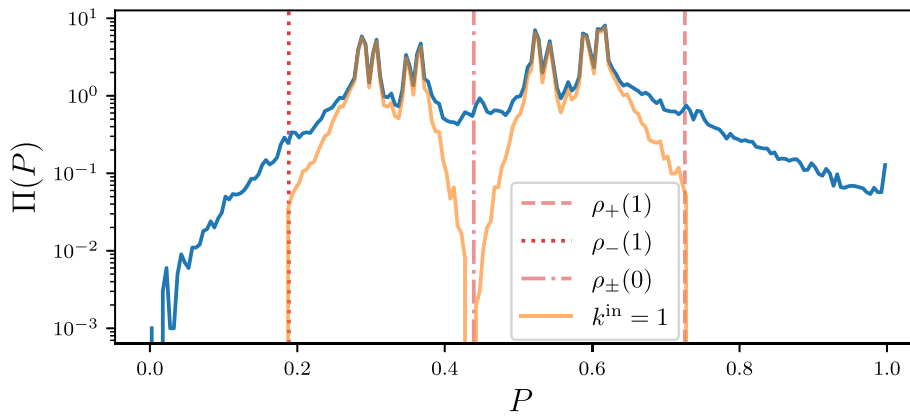


Figure 5. Distribution of the node activation probability and contribution from nodes with in-degree 1 (orange). Vertical lines indicated the range of possible values of P for nodes with in-degree 1, as predicted by figure 4. The same parameters as in figure 2.

neighbours is typically small. Therefore, we expect the nMF to fail. However, our results below indicate that the site averaged node activation is well captured by the mean-field approximation in a variety of conditions.

Starting from equation (2), the associated mean-field equation for fully asymmetric networks is

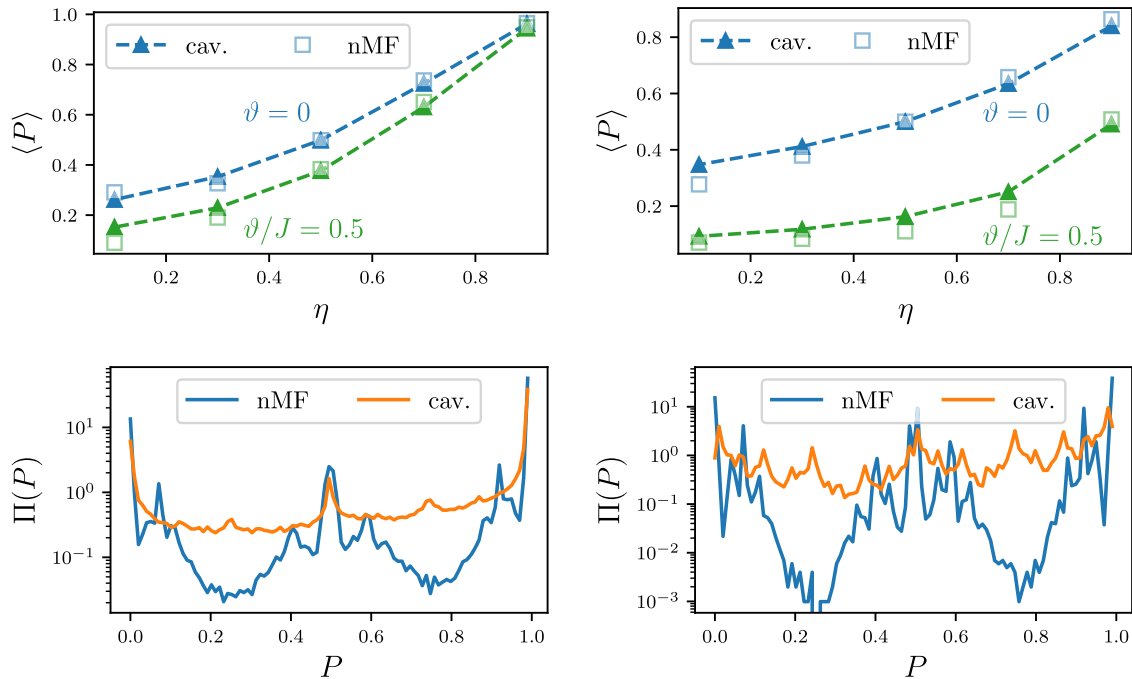


Figure 6. Comparison between nMF and cavity (cav) solutions at $T/J = 0.2$ in random directed networks with $N = 100\,000$ nodes and regular topology with $k^{\text{in}} = k^{\text{out}} = 3$ (left) and heterogeneous in- and out-degrees, drawn from a fat-tailed distribution with $\gamma = 4$ (right). Upper panels show the average activation probability $\langle P \rangle = N^{-1} \sum_{i=1}^N P_i$ versus the bias η , for $\vartheta/J = 0, 0.5$. Results from cavity are shown by triangle markers, while squares show mean-field results. Lower panels show the distribution $\Pi(P)$ of the site-dependent activation probabilities $\{P_i\}_{i=1}^N$, for $\vartheta = 0$, $\eta = 0.7$. Orange lines show results from cavity equation (7), blue lines show results from nMF equation (16).

$$P_i(t+1) = \Phi_\beta \left(\sum_j J_{ij} P_j(t) - \vartheta_i \right), \quad (16)$$

which represents an approximation of equation (7). We run equation (16) for each node i starting from random initial conditions until convergence. We apply the heterogeneous mean-field in the same class of networks as above (fully asymmetric adjacency matrix with discrete coupling strength) and use the cavity result as a ground truth to evaluate the performance of the mean-field approximation. We compute the site average of the node activation probability $\langle P \rangle = N^{-1} \sum_i P_i$ both in the case where the set $\{P_i\}_{i=1}^N$ is computed by using equations (16) and (7). In figure 6 we compare the results for the site average of node activation at different values of bias η and threshold ϑ for two network models, namely the directed regular graphs, and networks with fat-tailed degree distribution. Results indicate the mean-field captures remarkably well the behaviour of the site average for different choices of network models and both in the absence and presence of an external threshold. However, the mean-field approximation captures poorly the heterogeneous

site dependence, and the distribution $\Pi(P)$ presents large differences between the cavity and mean-field results.

4. Pairwise correlation for the linear threshold model on fully asymmetric networks

Temporal and pairwise correlations are of great interest for the analysis of dynamics on networks. For example, correlations are used extensively to discover patterns of relationships in numerous inverse problems [58–61]. In the context of spin models on networks, the correlations have been investigated in the dense regime for asymmetric interactions through mean-field [56] and an analytical expression for the site correlators can be formally solved through TAP equations for generic systems not at equilibrium [33]. However, in the latter case, the analytical equations are difficult to solve. In this section, we introduce a method to evaluate spatio-temporal correlations in the sparse regime using dynamic cavity techniques.

As shown above, in the cavity approach the evaluation of the one-point expectations $\langle n_i \rangle$ relies on the assumption that the states associated with the predecessors of a node are independent of one another, which is true on a tree. Instead, for a given *pair* of nodes, the states of the *joint* predecessors are not independent of one another, even on a tree. The cavity method has recently been applied in the presence of node correlations on loopy graphs to investigate percolation, random matrix spectra [62], and the equilibrium state of Ising spin models [63]. The authors explicitly include the contribution from short loops, e.g. triangles, which break the assumption of independence of states the cavity method relies on. Here, we solve a different problem, namely, we adapt the cavity method to evaluate two-point expectations. In our context, a new set of motifs that break the assumption of independence of states arises, which are present even in directed trees. Therefore, we devise a method that incorporates node correlations inside the cavity equations.

If the graph admits a giant connected component, every set of nodes belonging to the giant out-component shares at least one common ancestor, which would make the evaluation of node correlation infeasible on a large graph. In the following, we assume correlations to be mainly generated by two network motifs, where nodes have common predecessors either at distance one or two, and we test our approximate results against direct simulations.

Let us consider the stationary (time-translation invariant) pairwise connected correlation (also called covariance)

$$C_{ij}^c(\tau) := \lim_{t \rightarrow \infty} \langle n_i(t + \tau) n_j(t) \rangle - P_i P_j, \quad (17)$$

that is defined between any pair of sites $\{i, j\}$ and time-lag τ . The empirical stationary pairwise connected correlation (as computed from MC simulations) is analogously defined as

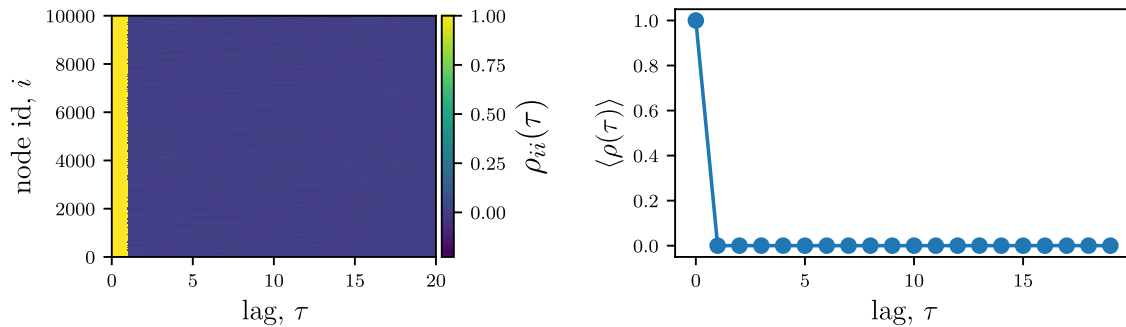


Figure 7. Single-site lagged Pearson autocorrelation $\rho_{ii}(\tau)$ for different nodes i (left) and its site average $\langle \rho(\tau) \rangle$ (right) as a function of the time-lag τ in a directed random regular graph with parameters: $N = 10\,000$, $k^{\text{out}} = k^{\text{in}} = 2$, $T/J = 0.2$, $\vartheta = 0$.

$$\hat{C}_{ij}^c(\tau) := \frac{1}{t_s} \sum_{t=t_0}^{t_0+t_s} n_i(t+\tau)n_j(t) - P_{i,\text{MC}}P_{j,\text{MC}}, \quad (18)$$

for $t_s \gg 1$, and t_0 a sufficiently long time to allow the dynamics to relax to stationarity. The covariance $C_{ij}^c(\tau)$ depends both on the sites and on the lag. In the following we restrict ourselves to two levels of analysis, namely the single-site lagged auto-covariance $C_{ii}^c(\tau)$, and the covariance at zero time-lag, $C_{ij}^c(0)$. We compare analytical results for those two types of covariance against simulations.

Single-site lagged auto-covariance. For the dynamics described by equation (2), we first investigate the lagged auto-covariance $\hat{C}_{ii}^c(\tau)$. We will present and discuss results in term of the time-lagged local Pearson correlation defined as, $\rho_{ii}(\tau) := \hat{C}_{ii}^c(\tau)/\hat{C}_{ii}^c(0)$, as well as its sample average $\langle \rho(\tau) \rangle = N^{-1} \sum_i \rho_{ii}(\tau)$. Simulation results show that $\rho_{ii}(\tau) \approx 0$ for $\forall i, \forall \tau > 0$ (hence trivially $\langle \rho(\tau) \rangle \approx 0$), which means that the dynamics presents no retarded auto-covariance, see figure 7. This is consistent with the theoretical prediction from the cavity formalism on fully asymmetric networks in the absence of loops, $C_{ii}^c(\tau) = 0 \forall i, \forall \tau > 0$. Heuristically, non-zero auto-covariance is associated with feedback loops which send the information back to the starting node with a delay. However, on a directed tree this never happens.

Pairwise connected correlation. Below we compute the analytical pairwise covariance at zero time-lag, $C_{ij}^c(0)$, which we refer to as C_{ij}^c for simplicity. Correlations between two nodes i, j are expected to arise from common ancestors of the two nodes. Given the earlier result of the absence of single-site lagged auto-covariance, we expect that if the common ancestor k is not at the same distance from i and j , signals would need a different amount of time to travel from k to i and j , respectively, entailing that $C_{ij}^c(\tau) \neq 0$ *only* at a non-zero τ corresponding to the difference in distances. Here we are interested in the equal-time connected correlation, $C_{ij}^c(0)$, which is thus non-zero only if the common ancestor k is at the *same* (finite) distance from nodes i and j . Therefore, network motifs that produce non-zero values of the pairwise equal-time covariance are required to have their common root node at *equal* distance from the leaf nodes; see for example the motifs that are shown in figures 8(a) and (b). The (equal-time) connected

correlation is obtained from the joint expectation

$$C_{ij} = \langle n_i n_j \rangle = \langle \Phi_\beta (h_i(\mathbf{n}_{\partial_i}) - \vartheta_i) \Phi_\beta (h_j(\mathbf{n}_{\partial_j}) - \vartheta_j) \rangle_{\mathbf{n}_{\partial_i} \cup \mathbf{n}_{\partial_j}}, \quad (19)$$

via $C_{ij}^c = C_{ij} - P_i P_j$, where $\mathbf{n}_{\partial_i} \cup \mathbf{n}_{\partial_j}$ indicates the set of states associated with the predecessors of nodes i or j . In the following we consider contributions to the correlations coming from the two types of motifs shown in figures 8(a) and (b). They are (a) a pair of nodes sharing one predecessor, and (b) a pair of direct descendants from the nodes in motif (a). We show that those motifs provide the dominant contributions to the C_{ij} distribution. If we assume that the joint probability of the states of the predecessors of nodes i or j factorises over single-node terms⁴, i.e.

$$P(\mathbf{n}_{\partial_i} \cup \mathbf{n}_{\partial_j}, t) \approx \prod_{\ell \in \partial_i \cup \partial_j} P(n_\ell, t), \quad (20)$$

we are able to take into account the effect of motif (a) (but only motif (a)). This is because the nodes belonging to the set of predecessors of both i and j , $\{\ell, \ell \in \partial_i \cap \partial_j\}$, appear only once in the product on the rhs of equation (20). This then explains why $C_{ij}^c \neq 0$ in these cases.

We evaluate equation (19) in the approximation that takes motif (a) into account through dynamic programming in a similar fashion as done above in the context of the node activation probability. Let $K = |\partial_i \cup \partial_j|$ be the number of joint predecessors of nodes i or j and, to ease the notation, let us use $\{1, \dots, K\}$ to enumerate the nodes in this set of predecessors. We define the family of recursive functions $C_{ij}^{(1)}(\ell, \tilde{h}_i, \tilde{h}_j)$ such that ℓ represents the node we perform the average over. If node ℓ is active, it gives a contribution $J_{i\ell}$ and $J_{j\ell}$ to the auxiliary fields \tilde{h}_i and \tilde{h}_j , respectively. Otherwise the auxiliary fields do not change. The dynamic programming evaluation for C_{ij} becomes

$$\begin{aligned} C_{ij} &= C_{ij}^{(1)}(1, 0, 0) \\ C_{ij}^{(1)}(\ell, \tilde{h}_i, \tilde{h}_j) &= P_\ell C_{ij}^{(1)}(\ell + 1, \tilde{h}_i + J_{i\ell}, \tilde{h}_j + J_{j\ell}) \\ &\quad + (1 - P_\ell) C_{ij}^{(1)}(\ell + 1, \tilde{h}_i, \tilde{h}_j) \quad \text{for } 1 \leq \ell \leq K \\ C_{ij}^{(1)}(K + 1, \tilde{h}_i, \tilde{h}_j) &= \Phi_\beta(\tilde{h}_i - \vartheta_i) \Phi_\beta(\tilde{h}_j - \vartheta_j), \end{aligned} \quad (21)$$

where P_ℓ is the stationary probability of node ℓ . From the connected correlation we obtain the Pearson correlation coefficient

$$\rho_{ij} = \frac{C_{ij}^c}{\sqrt{(P_i - P_i^2)(P_j - P_j^2)}}, \quad (22)$$

and we compare the values ρ_{ij} obtained through cavity approximation with that obtained from the direct evaluation of the microscopic dynamics in figure 8(c). For the

⁴Note that while independence of the neighbours of a *single* node, as used in the standard cavity method, e.g. equation (5), is exact on trees and, in the thermodynamic limit, also on locally tree-like graphs, the factorization assumed in (20) is only an approximation, even on trees, due to the presence of common ancestors.

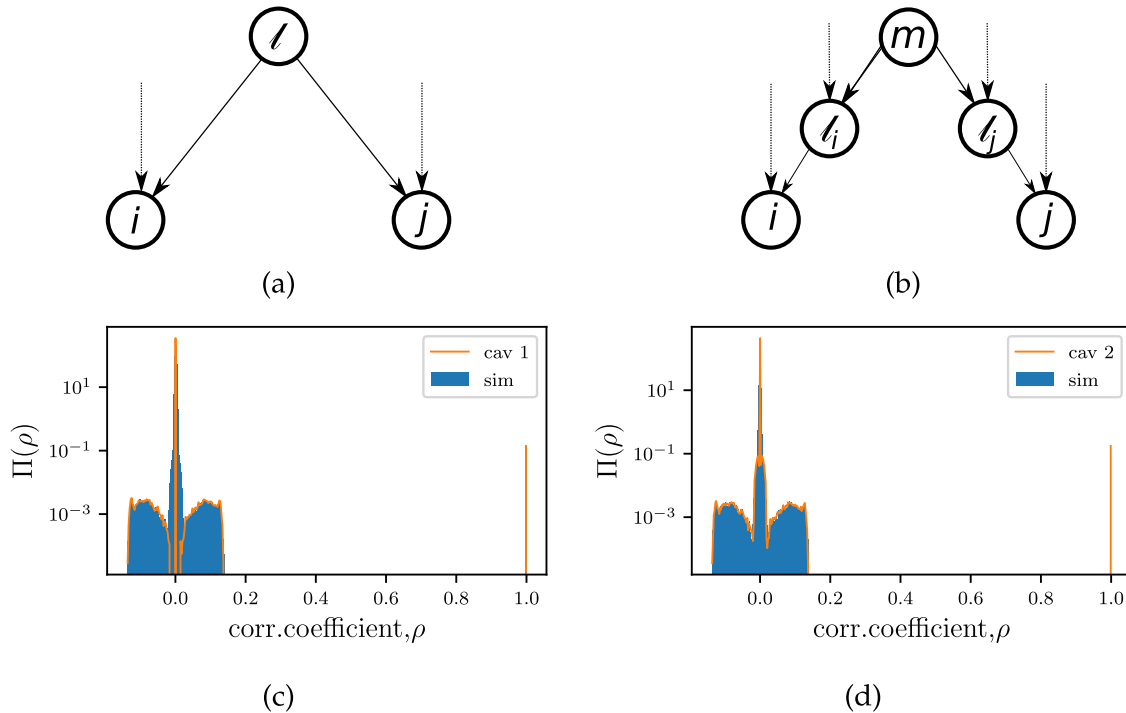


Figure 8. (a), (b) Network motifs that contributes to correlation. (c), (d) Probability density of the Pearson correlation coefficient ρ . Only elements of the upper triangular matrix (including the diagonal) of ρ are taken into account for the histogram. Parameters: $k^{\text{in}} = 2$, $N = 5000$, $T/J = 0.4$, $\vartheta = 0$.

reasons explained above, we obtain non-zero values of ρ_{ij} if and only if nodes i and j have a common predecessor l , as in the motif shown in figure 8(a). Apart from the trivial peak at 1, resulting from the diagonal terms of the distribution, the largest absolute values of ρ are captured nearly perfectly by the motifs illustrated in figure 8(a). On the other hand, such motifs do not capture well the events giving rise to small correlations.

In order to improve the characterisation of these events, for every pair of nodes $\{l_i, l_j\}$ that are described by the motif (a), we investigate the connected correlation associated with the pair of nodes $\{i, j\}$, with i and j a successor of l_i and l_j respectively. This motif is illustrated in figure 8(b) and consists of a pair of nodes $\{i, j\}$ that has at least one common ancestor m at distance two. In the following, we use the estimate of the joint expectation for $\{l_i, l_j\}$ provided by equation (21) to estimate the joint expectation of the pair $\{i, j\}$. For the two nodes denoted l_i and l_j in the figure, there is a non-zero correlation, $\rho_{l_i, l_j} \neq 0$ as shown above, hence the joint probability of the states of the predecessors of i or j does not factorise, and equation (20) does not hold anymore. For a given pair $\{i, j\}$, let us denote the set of pairs of their predecessors $\{\{l_i, l_j\}\}$ which have a predecessor in common by $L_1 = \{\{l_i, l_j\} : l_i \in \partial_i, l_j \in \partial_j \text{ and } \partial_{l_i} \cap \partial_{l_j} \neq \emptyset\}$, and let L_0 be set of predecessors of i or j , which do not appear in L_1 and are treated as independent. Instead of equation (20), the marginal probability of the joint set of neighbours contains

the two-points joint probability of the pairs of states in the set L_1

$$P(\mathbf{n}_{\partial_i} \cup \mathbf{n}_{\partial_j}, t) \approx \prod_{\ell \in L_0} P(n_\ell, t) \prod_{\ell_i, \ell_j \in L_1} P(n_{\ell_i}, n_{\ell_j}, t). \quad (23)$$

For each pair of sites in L_1 , the joint probability $P(n_a, n_b, t)$ is

$$P(n_a, n_b, t) = \langle [n_a \Phi_\beta(h_a(\mathbf{n}_{\partial_a})) + (1 - n_a)(1 - \Phi_\beta(h_a(\mathbf{n}_{\partial_a})))][n_b \Phi_\beta(h_b(\mathbf{n}_{\partial_b})) + (1 - n_b)(1 - \Phi_\beta(h_b(\mathbf{n}_{\partial_b}))) \rangle_{\mathbf{n}_{\partial_a} \cup \mathbf{n}_{\partial_b}, t-1}. \quad (24)$$

To simplify the notation, we drop the explicit time dependence in $P(n_a, n_b, t)$; the result can be written in terms of C_{ab}, P_a, P_b as

$$P(n_a, n_b) = (1 - n_a)(1 - n_b) + P_a(2n_a - 1)(1 - n_b) + P_b(2n_b - 1)(1 - n_a) + C_{ab}(2n_a - 1)(2n_b - 1), \quad (25)$$

and C_{ab} is computed from equation (21). We now evaluate equation (19) using equation (23), which also takes into account motifs (b), in contrast to equation (20) which only accounts for motifs (a). The correction applies to the evaluation of the entries C_{ij} for which a node m exists which is an ancestor of both nodes i and j through the paths (i, ℓ_i, m) and (j, ℓ_j, m) ; see motif in figure 8(b). Let $K_0 = |L_0|$ and $K_1 = |L_1|$, and let $L_0 = \{1, \dots, K_0\}$ and $L_1 = \{\{a_1, b_1\}, \dots, \{a_{K_1}, b_{K_1}\}\}$. Equation (21) is replaced by the dynamic programming relationships

$$\begin{aligned} C_{ij} &= C_{ij}^{(1)}(1, 0, 0) \\ C_{ij}^{(1)}(\ell, \tilde{h}_i, \tilde{h}_j) &= P_\ell C_{ij}^{(1)}(\ell + 1, \tilde{h}_i + J_{i\ell}, \tilde{h}_j + J_{j\ell}) \\ &\quad + (1 - P_\ell) C_{ij}^{(1)}(\ell + 1, \tilde{h}_i, \tilde{h}_j) \quad \text{for } 1 \leq \ell \leq K_0 \\ C_{ij}^{(1)}(K_0 + 1, \tilde{h}_i, \tilde{h}_j) &= C_{ij}^{(2)}(1, 0, 0) \\ C_{ij}^{(2)}(\ell, \tilde{h}_i, \tilde{h}_j) &= P(n_{a_\ell} = 1, n_{b_\ell} = 1) C_{ij}^{(2)}(\ell + 1, \tilde{h}_i + J_{ia_\ell}, \tilde{h}_j + J_{jb_\ell}) \\ &\quad + P(n_{a_\ell} = 1, n_{b_\ell} = 0) C_{ij}^{(2)}(\ell + 1, \tilde{h}_i + J_{ia_\ell}, \tilde{h}_j) \\ &\quad + P(n_{a_\ell} = 0, n_{b_\ell} = 1) C_{ij}^{(2)}(\ell + 1, \tilde{h}_i, \tilde{h}_j + J_{jb_\ell}) \\ &\quad + P(n_{a_\ell} = 0, n_{b_\ell} = 0) C_{ij}^{(2)}(\ell + 1, \tilde{h}_i, \tilde{h}_j) \quad \text{for } 1 \leq \ell \leq K_1 \\ C_{ij}^{(2)}(K_1 + 1, \tilde{h}_i, \tilde{h}_j) &= \Phi_\beta(\tilde{h}_i - \vartheta_i) \Phi_\beta(\tilde{h}_j - \vartheta_j), \end{aligned} \quad (26)$$

where the two-point marginals $P(n_a, n_b)$ are computed from equation (25) using equation (21) to compute C_{ab} . The comparison between simulations and the analytical calculation is shown in figure 8(d). The results indicate that the motifs in figures 8(a) and (b) very accurately capture the salient properties of pairwise correlations in this

system. This finding suggests that methods for network reconstruction from experimentally measured correlations may have limited ability to characterise motifs that involve common ancestors at distance larger than two.

5. Extensions

5.1. Multi-node interaction model

The linear threshold model presented above in equation (9) describes a dynamic model consisting of pairwise interactions between nodes. However, approaches based on pairwise interactions may not always be appropriate [64, 65]. Therefore several models that explicitly consider multi-node interactions (called high-order models) have been proposed, which can be implemented in terms of hypergraphs [66], simplicial complexes [67], and bipartite or factor graphs.

Let us consider a directed bipartite graph $G(V, E)$, where V denotes the set of nodes and E the set of edges. The set of nodes V is composed of two disjoint sets X and Y , with no edges in E having both endpoints in X or Y . Let $|X| = N$, and $|Y| = \alpha N$. We refer to the N elements of X as variable nodes and the αN elements of Y as function nodes. To define a dynamic process on the bipartite network, we associate a dynamic variable with every node of X . We use $\{n_i\}_{i=1}^N$ and $\{\tau_\mu\}_{\mu=1}^{\alpha N}$ to denote the dynamic variables and the function nodes respectively. Every function node μ is associated with a Boolean function $g_\mu(\cdot)$, with the function $g_\mu(\mathbf{n}_{\partial_\mu})$ depending on the states of μ 's predecessors $\mathbf{n}_{\partial_\mu}$. Every variable node i evolves according to a linear threshold function depending on node i 's predecessors τ_{∂_i} . The bipartite graph is defined by the bi-adjacency matrix $\mathbf{A} = (A_{\mu i})$ and the interaction matrix $\mathbf{J} = (J_{i\mu})$, with entries independent of \mathbf{A} . The matrix $\mathbf{A} = (A_{\mu i})$ has dimension $\alpha N \times N$; a value $A_{\mu i} = 1$ indicates that a link from the variable node i to the function node μ exists, signifying the fact that the state n_i is an argument of the function g_μ , and $A_{\mu i} = 0$ otherwise. The interaction matrix $\mathbf{J} = (J_{i\mu})$ has dimension $N \times \alpha N$; a non-zero entry $J_{i\mu}$ represents the strength of the influence of the function μ on the variable node i . For any realisation of the matrices \mathbf{A} and \mathbf{J} , we can define $d_i(\mathbf{J})$ as the number of predecessors of node i and $c_\mu(\mathbf{A})$ the number of predecessors of node μ :

$$d_i(\mathbf{J}) = \sum_{\mu}^{\alpha N} \Theta(|J_{i\mu}|), \quad c_\mu(\mathbf{A}) = \sum_i^N A_{\mu i}. \quad (27)$$

We consider the dynamics where the state of the function node μ , given by the binary variable τ_μ , is determined by $g_\mu(\mathbf{n}_{\partial_\mu})$. Conversely, the binary variables n_i follow a linear threshold dynamics formulated in terms of the τ_μ , i.e.

$$n_i(t+1) = \Theta \left(\sum_{\mu} J_{i\mu} \tau_\mu(t) - \vartheta_i - z_i(t) \right) \quad (28)$$

$$\tau_\mu(t) = g_\mu(\mathbf{n}_{\partial_\mu}(t)),$$

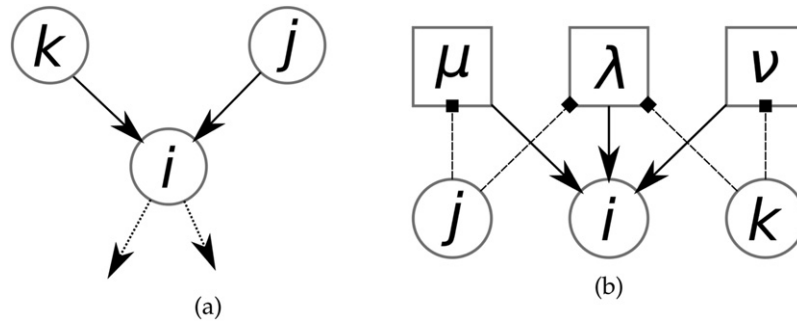


Figure 9. Snapshot of a network describing a pairwise model; each link carries the signal from one node only (left). A bipartite graph is used to represent multi-node interaction, here the node i receives the signal from nodes j, k through function nodes μ, ν and λ , with node λ representing a Boolean function of two variables (right). Arrows indicate a link pointing from a function node to a node, square markers join a function node with the arguments of the function.

where the $z_i(t)$ represent independent identically distributed stochastic noise terms, whose statistics are defined in terms of their cumulative distribution function $\Phi_\beta(z)$ as before. The update rule in equation (28) represents a more flexible model than the pairwise interaction model defined in equation (2). In figure 9(a) we show an example of a small monopartite network that captures pairwise interactions as described by equation (2): a node i receives the signal from the two nodes j and k through distinct links. Instead, the bipartite graph model shown in figure 9(b) is more flexible, as the function nodes return arbitrary Boolean functions $g_\mu(\cdot), g_\nu(\cdot), g_\lambda(\cdot)$ of their respective inputs. These *can* represent the standard pairwise interactions through function nodes such as node μ in figure 9(b), which is a function of a single variable n_j , but also generic multi-node interactions such as through node λ , which depends on two Boolean variables n_j and n_k .

In what follows, we develop the formalism to solve the dynamics of bipartite graphs of the type represented in figure 9(b), and we show that these systems benefit from the speedup of the algorithm discussed above. We will specifically look at an example of a system with multi-node interactions that is motivated by gene regulations [50]. In the following we investigate the case $g_\mu(\mathbf{n}) = \prod_{i \in \partial_\mu} n_i$, which represents an ‘AND’ gate logic,

$$n_i(t+1) = \Theta \left(\sum_{\mu} J_{i\mu} \tau_{\mu}(t) - \vartheta_i - z_i(t) \right) \quad (29)$$

$$\tau_{\mu}(t) = \prod_{j \in \partial_{\mu}} n_j(t).$$

Let $P_i(t)$ be the probability that state $n_i(t) = 1$. Let $\mathcal{P}_{\mu}(t) := \text{Prob}[\tau_{\mu}(t) = 1]$ be the probability that state $\tau_{\mu}(t) = 1$. If the directed bipartite graph does not contain short loops, the dynamic cavity method provides an efficient approximate solution of the dynamics. In particular the node activation probabilities are given by the set of

expressions

$$P_i(t+1) = \langle \Phi_\beta (h_i(\boldsymbol{\tau}_{\partial_i}) - \vartheta_i) \rangle_{\boldsymbol{\tau}_{\partial_i}}, \quad (30)$$

$$\mathcal{P}_\mu(t) = \prod_{j \in \partial_\mu} P_j(t), \quad (31)$$

with $h_i(\boldsymbol{\tau}_{\partial_i}) = \sum_\mu J_{i\mu} \tau_\mu$. Equation (30) is of the same form as equation (9) and thus the dynamic programming approach can speed-up the evaluation of the average. Let us consider a node i with in-degree d_i , and let $\nu \in \{1, \dots, d_i\}$ be the indices of the predecessors of node i . We define a function $f_i(\nu, \tilde{h})$ recursively

$$\begin{aligned} f_i(\nu, \tilde{h}) &= \mathcal{P}_\nu(t) f_i(\nu+1, \tilde{h} + J_{i\nu}) + (1 - \mathcal{P}_\nu(t)) f_i(\nu+1, \tilde{h}) \quad \text{for } 1 \leq \nu \leq d_i \\ f_i(d_i+1, \tilde{h}) &= \Phi_\beta(\tilde{h} - \vartheta). \end{aligned} \quad (32)$$

In terms of this recursion, we have

$$P_i(t+1) = f_i(1, 0). \quad (33)$$

In analogy with the pairwise model, we assume that $J_{i\mu} \in \{-J_i, 0, J_i\}$ with J_i a real positive number that may depend on the variable node. Under this condition, the dynamic programming algorithm provides a speedup of the evaluation of node activation probabilities.

Node activation probability with multi-node interaction. As for the linear threshold model, we inspect the stationary state of the node activation probabilities at different values of the noise intensity; results are shown in figure 10. In our investigation, we use the same degree distribution of the in- and out-degree as in the monopartite network, both for the variable and function nodes. The variable η refers to the fraction of positive terms in the interaction matrix \mathbf{J} in analogy to the monopartite case.

Dominant peaks of the node activation probability can be rationalised in terms of the discrete stochastic map of equation (15). This should not come as a surprise: given our choice of the degree distribution, the majority of nodes only have a pairwise interaction with a single predecessor node. Therefore the same argument used to derive the theoretical lines of equation (15), which describes the noise strength dependence of some of the prominent peaks in the distribution of node activation probabilities, still holds. However, the distribution shown in figure 10 is substantially different from the distribution obtained in the linear threshold model at zero thresholds, which was analysed in [47], highlighting contributions due to multi-node interactions. The effect of multi-node interactions is particularly easy to analyse in the high noise regime, which is discussed in appendix A.

5.2. Any Boolean function can be written in terms of a linear combination of AND gates

We have demonstrated above, how our linear programming method can be extended to analyse the stochastic dynamics of multi-node interaction models, using a formalisation in terms of bipartite graphs in which all function nodes are logical AND gates, i.e. we

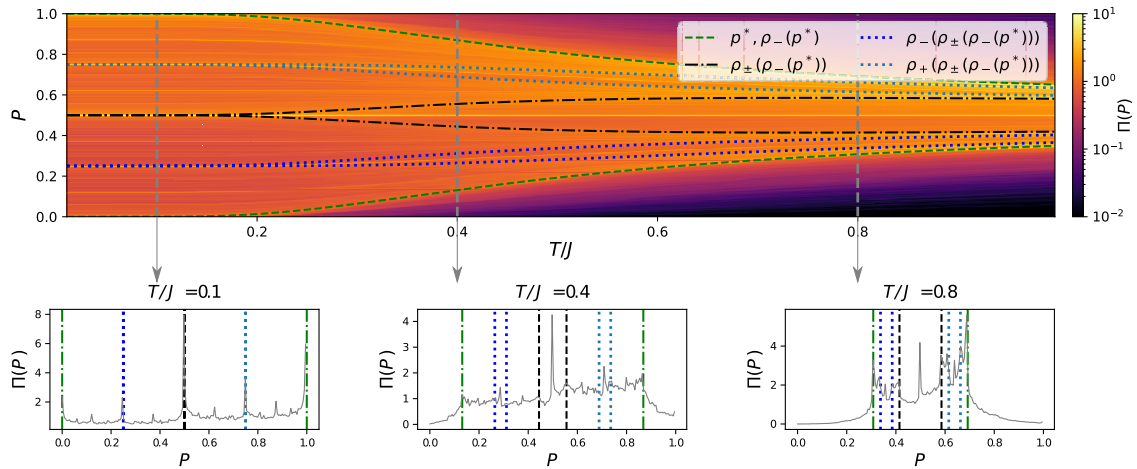


Figure 10. Heat-map of the distribution of node activation probabilities $\Pi(P)$ at different values of noise parameter T resulting from the microscopic dynamics detailed in equation (29) (top). Dashed, dotted, and dot-dashed lines show the graph of different function compositions of ρ_{\pm} at different values of T , see equation (15). Vertical dotted lines mark the noise levels at which the histograms of $\Pi(P)$ are shown in the bottom panels. They probe the regimes at low, intermediate, and high noise corresponding to $T/J = 0.1, 0.4, 0.8$, from left to right. Parameters: $N = 200\,000$, $\alpha = 1$, $\gamma = 2.81$, $\vartheta = 0$.

have $\tau_{\mu} = g_{\mu}(\mathbf{n}_{\partial_{\mu}}) = \prod_{j \in \partial_{\mu}} n_j$ for all μ . We now show that the bipartite model with AND gates is *universal* in the sense that a local field of the form $h_i = \sum_{\mu} J_{i\mu} \tau_{\mu}$ as it appears in equation (29) can represent *arbitrary* functions of $k_i = |\partial_i|$ Boolean variables. To see this, consider an arbitrary real function $g: \{0, 1\}^k \rightarrow \mathbb{R}$ of k Boolean variables. It can always be written using an expansion of the form

$$g(\mathbf{n}) = a^{(0)} + \sum_i a_i^{(1)} n_i + \sum_{i < j} a_{ij}^{(2)} n_i n_j + \sum_{i < j < \ell} a_{ij\ell}^{(3)} n_i n_j n_{\ell} + \cdots + a^{(k)} \prod_{\ell=1}^k n_{\ell}, \quad (34)$$

in which terms are arranged by increasing order, with suitable coefficients $a^0, \mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, a^{(k)}$, which can be found recursively by considering values of $g(\mathbf{n})$ successively for configurations \mathbf{n} , with non-zero components on subsets of $\{1, \dots, k\}$ of increasing size. The local field h_i in equation (29) is exactly of this form if the $\mu \in \partial_i$, i.e. the μ for which $J_{i\mu} \neq 0$, are arranged in order of increasing $|\partial_{\mu}|$.

5.3. Spin dynamics

Let σ_i , $i \in \{1, \dots, N\}$, denote a set of Ising spin variables, taking values in $\{-1, 1\}$, located on the N nodes of a complex network, and let J_{ij} denote the strength of the

interaction from node j to node i . The Ising spin dynamics is defined as

$$\sigma_i(t+1) = \text{sgn} \left[\sum_j^N J_{ij} \sigma_j(t) - \vartheta_i - z_i(t) \right], \quad (35)$$

where ϑ_i is a threshold, and $z_i(t)$ is an independent identically distributed random variable defined as above with $\text{Prob}[z < x] = \Phi_\beta(x)$.⁵ The average of $\text{sgn}(x - z)$ over the noise distribution can be expressed in terms of the cumulative distribution function of the noise as

$$\tilde{\Phi}_\beta(x) := \langle \text{sgn}(x - z) \rangle_z = 2\Phi_\beta(x) - 1. \quad (36)$$

Let us denote by $m_i(t) := \sum_\sigma P_i(\sigma, t) \sigma = \langle \sigma_i \rangle_{\sigma_i, t}$ the time dependent average of the Ising spin at node i . For a fully asymmetric graph, the cavity method provides an expression for the magnetisation in a similar form as above. Performing the average in (35) gives

$$m_i(t+1) = \left\langle \tilde{\Phi}_\beta(h_i(\boldsymbol{\sigma}_{\partial_i}) - \vartheta_i) \right\rangle_{\boldsymbol{\sigma}_{\partial_i, t}}, \quad (37)$$

in which $h_i(\boldsymbol{\sigma}_{\partial_i}) = \sum_j J_{ij} \sigma_j$ and $\boldsymbol{\sigma}_{\partial_i}$ plays a role analogous to equation (8) for spin variables. In analogy to the linear threshold model, equation (35) can be evaluated efficiently through the recursion

$$\begin{aligned} m_i(t+1) &= f_i(1, 0) \\ f_i(\ell, \tilde{h}) &= \frac{1 + m_\ell(t)}{2} f_i(\ell + 1, \tilde{h} + J_{i\ell}) \\ &\quad + \frac{1 - m_\ell(t)}{2} f_i(\ell + 1, \tilde{h} - J_{i\ell}) \quad \text{for } 1 \leq \ell \leq k_i \\ f_i(k_i + 1, \tilde{h}) &= \tilde{\Phi}_\beta(\tilde{h} - \vartheta_i). \end{aligned} \quad (38)$$

Mixed p -spin model. We briefly specify the expression for the p -spin model, whose equilibrium properties have long been investigated in the literature [68]. Let us consider a system defined on a bipartite network with state variables $\sigma_i \in \{\pm 1\}$ and $\tau_\mu \in \{\pm 1\}$. The update rule is

$$\begin{aligned} \sigma_i(t+1) &= \text{sign} \left(\sum_\mu J_{i,\mu} \tau_\mu(t) + z_i(t) - \vartheta_i \right) \\ \tau_\mu(t) &= \prod_{j \in \partial_\mu} \sigma_j(t). \end{aligned} \quad (39)$$

Let $m_\mu(t) = \langle \tau_\mu \rangle_{\tau_\mu, t}$, and $m_i(t) = \langle \sigma_i \rangle_{\sigma_i, t}$ denote the time dependent averages of the τ_μ and the σ_i , respectively. The time evolution of the local magnetizations $m_i(t)$ is given

⁵Note that, in analogy with the rest of the manuscript, the variables ϑ_i represent a threshold and not an external field.

by

$$\begin{aligned}
 m_i(t+1) &= \left\langle \tilde{\Phi}_\beta (h_i(\boldsymbol{\tau}_{\partial_i}) - \vartheta_i) \right\rangle_{\boldsymbol{\tau}_{\partial_i, t}}, \\
 m_\mu(t) &= \prod_{j \in \partial_\mu} m_j(t),
 \end{aligned}
 \tag{40}$$

with $h_i(\boldsymbol{\tau}_{\partial_i}) = \sum_{\mu \in \partial_i} J_{i\mu} \tau_\mu$. Equation (40) can be efficiently evaluated using the recursive procedure below

$$\begin{aligned}
 m_i(t+1) &= f_i(1, 0) \\
 m_i(\nu, \tilde{h}) &= \frac{1 + m_\nu(t)}{2} f_i(\nu + 1, \tilde{h} + J_{i\nu}) \\
 &\quad + \frac{1 - m_\nu(t)}{2} f_i(\nu + 1, \tilde{h} - J_{i\nu}) \quad \text{for } 1 \leq \nu \leq d_i \\
 m_i(d_i + 1, \tilde{h}) &= \tilde{\Phi}_\beta (\tilde{h} - \vartheta_i).
 \end{aligned}
 \tag{41}$$

6. Beyond fully asymmetric networks

Feedback loops are associated with multi-stability and oscillatory behaviour of dynamical systems [69, 70]. These two phenomena are often observed also in biological systems. Indeed, feedback mechanisms represent an essential ingredient of biological networks, e.g. in the context of gene regulation [71–73]. However, it is not yet clear what happens in the case where a combination of many of these patterns is present in a network. The presence of bi-directional links with opposite signs of the interactions (such as in predator–prey systems) have been shown to produce oscillatory behaviour in the context of continuous linear dynamics [74], but little is known in the context of discrete state dynamics, e.g. for the linear threshold model. The dynamics investigated in section 3 is defined on tree-like fully asymmetric network models, where feedback loops are absent.

In this section, we consider the linear threshold model of equation (2) in the presence of bi-directional links. Any bi-directional link can be considered as a de-facto short loop of length two of the network. As expected, the presence of these short loops makes the use of the dynamic cavity method more challenging. In the cavity graph, where a node and its links have been removed, the presence of a short loop causes a retarded self-interactions that make the dynamics of the cavity variables non-Markovian; see [38, 39] and appendix B. An approximation scheme, called the OTA, has been proposed with the intent to reduce the computational complexity created by this form of non-Markovianity [39, 75]. The approximation factorises the memory kernel to include a one-time-step memory term. It exactly describes the full dynamics of fully asymmetric networks and the equilibrium steady-state of systems with symmetric interactions in the replica symmetric phase. Furthermore, the OTA has been shown to effectively

capture the stationary state also of systems with partially symmetric links in the high noise regime [55, 75]. However, the literature has so far focussed on the agreement of macroscopic observables, such as the mean activation, while the distribution of individual node activations has generally been overlooked. In this section, we investigate the heterogeneous node activation in networks with bi-directional links using the OTA and compare the results with simulations. Among other things, our investigation leads to an unexpected observation: the stationary state of networks with symmetric and anti-symmetric interactions are biased towards the active and the inactive states, respectively.

6.1. The one-time-step approximation

In this section we will only provide brief account of the OTA, leaving a more extensive discussion to the appendix B. Let us consider the model equation (2) in the case where the network of interactions contains bi-directional links. For any node i in the graph, the probability of the state at time $t + 1$ depends on the states of node i 's neighbours, \mathbf{n}_{∂_i} , at the previous time-step t through the transition probability $W[n_i|h_i(\mathbf{n}_{\partial_i}), \vartheta_i]$, see equation (4). If links are bi-directional, the states of nodes \mathbf{n}_{∂_i} at time t are not independent, because the state of every node $j \in \partial_i$ depends on the past history states of node i through the terms $J_{ji}n_i(s)$ for $s \in \{t - 1, t - 3 \dots\}$.⁶ In order to describe the dynamics of the nodes $j \in \partial_i$ on the cavity graph from which node i and connections to it are removed, it is useful to introduce the time-dependent cavity thresholds $\vartheta_j^{(i)}(s) := \vartheta_j - J_{ji}n_i(s)$ for $s = 0, \dots, t - 1$ and $\vartheta_j^{(i)}(0) := \vartheta_j$, which encodes the trajectory of node i in terms of threshold terms, and the local field in the cavity graph

$$h_j^{(i)}(\mathbf{n}_{\partial_j}^s) := \sum_{\ell \in \partial_j \setminus \{i\}} J_{j\ell} n_\ell^s, \tag{42}$$

where here and in what follows we have denoted $n_j^s = n_j(s)$ to make the notation more compact. Let us call $\mathbf{n}_j^{0, \dots, t}$ the trajectory of node j from time 0 to t . For every bi-directional link (i, j) , the dynamic cavity method provides a recursive expression for the conditional probability $P_j^{(i)}(\mathbf{n}_j^{0, \dots, t} | \vartheta_j^{(i)0, \dots, t-1})$ of the trajectory of node j in terms of the time-dependent thresholds $\vartheta_j^{(i)0, \dots, t} = (\vartheta_j^{(i)}(0), \dots, \vartheta_j^{(i)}(t))$, as discussed in [38, 39] and in appendix B. The OTA assumes that, for every $j \in \partial_i$, the probability of the trajectory factorises into one-time conditional probabilities

$$P_j^{(i)}(\mathbf{n}_j^{0, \dots, t} | \vartheta_j^{(i)0, \dots, t-1}) \approx P_j^{(i)}(n_j^0) \prod_{s=0}^{t-1} P_j^{(i)}(n_j^{s+1} | \vartheta_j^{(i),s}). \tag{43}$$

The factorisation procedure proposed in equation (43) is not sufficient to fully specify the cavity and marginal probabilities, but an additional closure condition is required, as shown in appendix B. In appendix B we re-investigate the derivation of the OTA, we show that different choices of the closure conditions lead to the different versions

⁶Bi-directional links are the only origin of loops present on a tree, thus every node depends on its own history at even time lag.

of the OTA known in the literature [34, 35], and we also propose a new version of the closure condition. It has been shown that, for symmetric couplings, the version of [35] admits the equilibrium state as a solution, while the version of [34] has not been designed with that goal. We compare the performance of OTA under different closure conditions, including the accuracy in characterising the individual node activation probabilities, as well as the computational complexity of their implementation. We show the version of [35] outperforms the version of [34] at equilibrium, but the situation is reversed in the non-equilibrium model we investigate. However, our results indicates the version of [34] is more consistent through the different network symmetries investigated. Based on our assessment, we expect the version of [34] to provide the best trade-off between accuracy and ease of implementation to tackle the non-equilibrium stationary state. The closed-form expressions for the single-time marginal probability of the OTA we will use are⁷

$$P_i(n_i^t|\vartheta_i) = \sum_{n_i^{t-2}} \left\{ \sum_{n_{\partial_i}^{t-1}} W[n_i^t|h_i(\mathbf{n}_{\partial_i}^{t-1}), \vartheta_i] \left[\prod_{j \in \partial_i} P_j^{(i)}(n_j^{t-1}|\vartheta_j - J_{ji}n_i^{t-2}) \right] \right\} P_i(n_i^{t-2}|\vartheta_i), \tag{44}$$

and for the cavity marginal,

$$P_j^{(i)}(n_j^{t-1}|\vartheta_j^{(i)}) = \sum_{n_j^{t-3}} \sum_{n_{\partial_j \setminus i}^{t-2}} W[n_j^{t-1}|h_j^{(i)}(\mathbf{n}_{\partial_j}^{t-2}), \vartheta_j^{(i)}] \times \left[\prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)}(n_\ell^{t-2}|\vartheta_\ell - J_{\ell j}n_j^{t-3}) \right] P_j(n_j^{t-3}|\vartheta_j). \tag{45}$$

The possible values of $\vartheta_j^{(i)}$ that appear in equation (45) are $\vartheta_j^{(i)} = \{\vartheta_j, \vartheta_j - J_{ji}\}$. To solve the dynamics over the trajectory of length t , the initial conditions for the marginal and cavity probabilities $P_j(n_j^0|\vartheta_j)$, $P_j^{(i)}(n_j^0|\vartheta_j)$, $P_j^{(i)}(n_j^0|\vartheta_j - J_{ji})$ are set for $\forall j$ and $\forall i: J_{ij} \neq 0$, and $P_j(n_j^1|\vartheta_j)$ is computed from equation (6). To evaluate the trajectory of node activation, equations (45) and (44) are propagated over the time-steps of interests $0, \dots, t$. At every time iteration, both equations (45) and (44) would face a complexity barrier created by a large in-degree of node i and j respectively, and their evaluation benefit from our dynamic programming implementation. To this end, we note that the term $\sum_{n_{\partial_i}^{t-1}} W[\dots][\prod_{j \in \partial_i} \dots]$ in equation (44) has the same structure as equation (8) and can be evaluated using dynamic programming as explained in section 3.1. This also applies to the corresponding average in equation (45).

In this section, we apply the OTA to networks with fat-tailed degree distributions and bi-directional links. We will investigate three different degrees of symmetry of the interaction matrix, namely antisymmetric, symmetric and uncorrelated interaction matrix

⁷In the following, we drop the time dependence from the cavity thresholds as this simplifies notation and does not lead to any ambiguity.

J , and we highlight how the symmetry affects the performance of the OTA in the regime of low noise strength.

6.2. Comparison between OTA and simulations

Parameters of the model. We first generate an undirected network in the configuration model class consisting of N nodes. In analogy with above, degrees of the nodes are sampled from the fat-tailed distribution $P_\gamma(k) = k^{-\gamma} - (k+1)^{-\gamma}$. We then assign interaction terms to the two directions of bi-directional links. In particular, interaction terms may or may not have opposite sign in the two directions, i.e. $J_{ij}J_{ji} = \pm J^2$. In the following, we investigate the correlation of the sign in bi-directional links and we focus on three major symmetries: symmetric ($J_{ij} = J_{ji}$), uncorrelated ($\langle J_{ij}J_{ji} \rangle = 0$), and antisymmetric ($J_{ij} = -J_{ji}$) interaction matrix J . In all settings, the number of positive and negative interactions are statistically the same. We iterate the OTA equations until convergence using the same criterion we adopt on fully asymmetric networks, as we have discussed in section 3.1.

Results. In order to compare the performance of OTA against MC simulation, we define P_{MC} and P_{OTA} the probability of node activation derived from microscopic MC simulation of the dynamics and from equations (44) and (45) respectively at stationarity. In figure 11 we compare P_{MC} and P_{OTA} for bi-directional interaction matrices with the three different types of symmetries detailed above. We recall that in the symmetric setting, the stationary state of the dynamics is an equilibrium state, and in the absence of replica symmetry breaking the equilibrium state is described by belief propagation [24]. To evaluate the accuracy of the OTA, we consider the mean square difference of P_{OTA} to P_{MC}

$$\text{error} = \sqrt{\frac{1}{N} \sum_i (P_{i,\text{MC}} - P_{i,\text{OTA}})^2}. \quad (46)$$

Our results suggest that the OTA is able to capture the statistics of node activation in the stationary state remarkably well, even in the regime of moderately low noise strength when bi-directional links are present, both for the symmetric and the uncorrelated case. However, at still lower noise strength we expect the results from OTA to be less accurate, as observed in earlier work for spin models [35, 75].

Our earlier analysis on fully-asymmetric networks revealed that spatial correlations are short-ranged and auto-covariance is zero at any finite time lag. On the other hand, the presence of correlations in the weights is associated with time-lagged auto-correlation at non-zero lag, as shown in figure 12. Simulation results indicate that the antisymmetric and symmetric settings exhibit the slowest decrease of the time-lagged auto-covariance, while the uncorrelated setting exhibits the fastest. Nevertheless, the auto-covariance decays quite consistently in few steps even for the symmetric and antisymmetric interaction models, justifying the use of OTA to investigate the stationary state. This observation suggests that a different factorisation that includes more time-steps than in equation (43) may provide a better characterisation of the symmetric and antisymmetric cases.

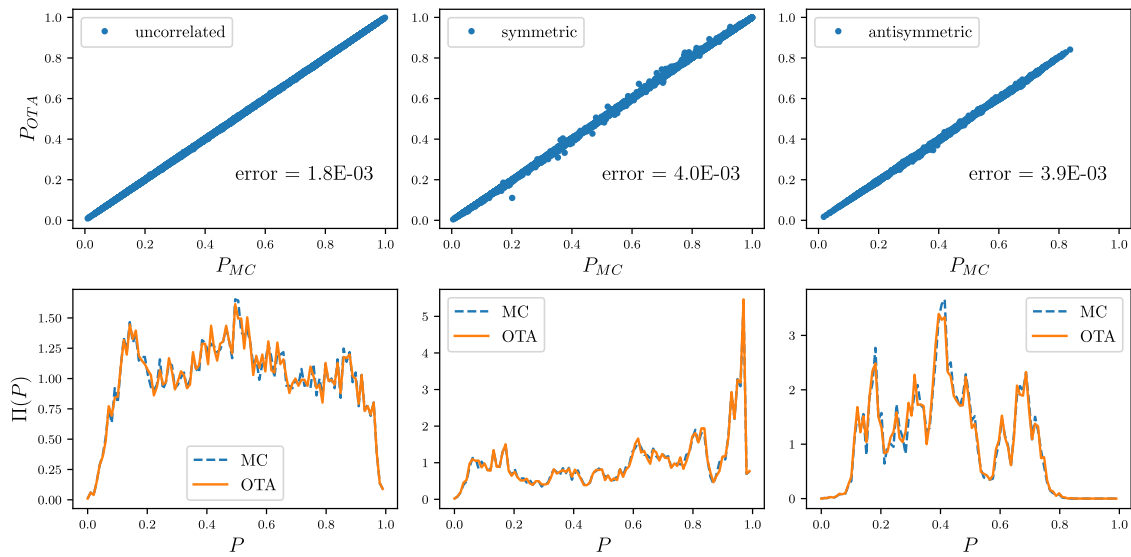


Figure 11. Comparison of the OTA and MC simulations at stationarity in networks with bi-directional links and weights chosen, from left to right, as uncorrelated, symmetric and anti-symmetric. Top panels show scatter plots of single-node activation probabilities computed from OTA versus MC results. Bottom panels show the distribution of node activation probabilities obtained through MC and OTA. Simulations are obtained through an average of over 500 trajectories. Parameters are $N = 10\,000$, $\gamma = 4$, $T/J = 0.5$, $\vartheta = 0$, $t_s = 10^4$ steps.

Finally, figure 11 also reveals that the stationary distribution of node activations for symmetric and antisymmetric interaction is skewed, even though we are in the regime with zero thresholds, $\vartheta_i = 0 \forall i$, and unbiased distribution of couplings, i.e. $\langle J_{ij} \rangle = 0$. We refer to this setting as ‘unbiased parameter choice’ in the following. This behaviour is confirmed in figure 12, where mean node activation are shown at different levels of the noise. In the following, we discuss this unexpected result and we provide an intuitive explanation.

7. Spontaneous symmetry breaking of the linear threshold model for unbiased parameter choices

We have observed that, for an unbiased parameter choice, the linear model defined in terms of $\{0, 1\}$ variables exhibits a skewed distribution of node activations in the presence of bi-directional links. This is somewhat surprising, as in spin systems, the average magnetisation is well known to be zero in the case of an unbiased parameter choice. In this section we compare the two models and we provide a simple example that illustrates how such a phenomenology might arise.

The dynamics described in terms of indicator variables $n_i \in \{0, 1\}$ can be reformulated in terms of spin variables $\sigma_i \in \{\pm 1\}$. Recall that the dynamics of the linear threshold model of equation (2), defined in terms of N indicator variables, is given by

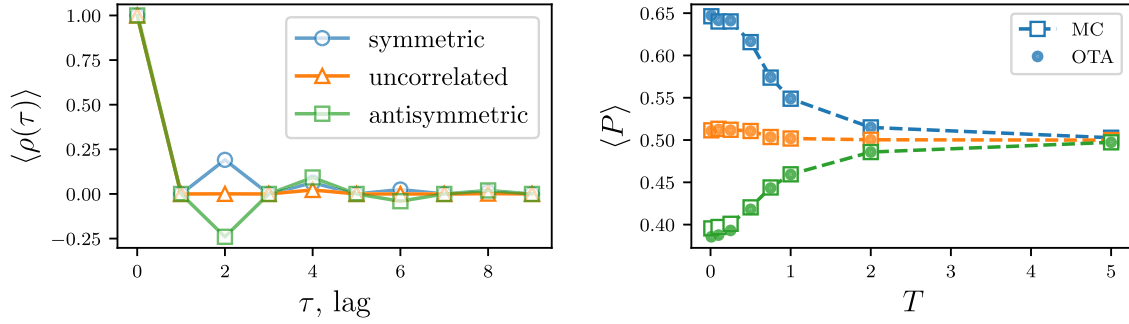


Figure 12. (Left) Site average of the lagged auto-correlation $\langle \rho(\tau) \rangle$ for a network with bi-directional links. Three choices of the interaction matrix J are considered: symmetric (circle), uncorrelated (triangles), and antisymmetric (squares) interaction matrices. The correlation is computed from MC simulated trajectories of the dynamics. (Right) Site average of the node activation at different temperatures $\langle P \rangle = N^{-1} \sum_i P_i$ for symmetric (blue, upper curve), uncorrelated (orange, middle curve), and antisymmetric (green, lower curve) interaction matrices. Squares indicate the result obtained through MC simulations, circles are obtained through the OTA. Network and simulation parameters are the same as for figure 11.

$$n_i(t+1) = \Theta \left(\sum_j J_{ij} n_j(t) - \vartheta_i - z_i(t) \right). \quad (47)$$

Rewriting the analogous dynamics of a model of N interacting spins as defined in equation (35) as

$$\sigma_i(t+1) = \text{sgn} \left(\sum_j J'_{ij} \sigma_j(t) - \vartheta'_i - z_i(t) \right), \quad (48)$$

with interaction matrix J' and thresholds $\{\vartheta'_i\}_{i=1}^N$, one can map one model onto the other by setting

$$J'_{ij} = J_{ij}/2, \quad \vartheta'_i = \vartheta_i - \frac{1}{2} \sum_{\ell} J_{i\ell}, \quad \forall i, j. \quad (49)$$

Unless the mapping of equation (49) is implemented exactly, the trajectories of the Boolean linear threshold model and the spin model will be different. In analogy with the above, we use $m_i = \langle \sigma_i \rangle$ and $P_i = \langle n_i \rangle$ to denote the stationary averages of spins and indicator variables respectively. The phase diagram for spin systems has been extensively studied at equilibrium, and it has been shown [24] that for an unbiased parameter choice the mean magnetisation is zero, i.e. $N^{-1} \sum_i m_i = 0$, even at low temperatures.

In a system with indicator variables instead, both MC simulations and an analysis using OTA show that the system exhibits spontaneous symmetry breaking

$N^{-1} \sum_i P_i \neq 1/2$ when the parameters are unbiased, as shown in figure 12. In particular, at stationarity the average site activation $N^{-1} \sum_i P_i$ is larger than $1/2$ for symmetric interaction matrices and smaller than $1/2$ in the case of antisymmetric interaction matrices.

We present an intuitive argument that explains how the biased distribution of node activations emerges in a model with an unbiased choice of parameters. To this end, let us consider a mini-network composed of just two nodes i, j connected by a bi-directional link, i.e. $J_{ij}J_{ji} \neq 0$. A key mechanism in bi-directional links is the dependence of a node state on its own past, i.e. node i interacts with its neighbours ∂_i which in turn interact with i . The probability of the state of node i at time $t + 2$ can be written in terms of the probability of the state of node i at time t using the Markov property of the microscopic dynamics,

$$\begin{aligned} P_i(n_i^{t+2}) &= \langle W(n_i^{t+2} | J_{ij}n_j) \rangle_{n_j \sim P(n_j^{t+1})} \\ P_j(n_j^{t+1}) &= \langle W(n_j^{t+1} | J_{ji}n_i) \rangle_{n_i \sim P(n_i^t)}. \end{aligned} \tag{50}$$

Considering the cases $J_{ij}, J_{ji} \in \{\pm J\}$, we note that $\langle W(1 | Jn) \rangle_{n \sim P} = \rho_+(P)$ and $\langle W(1 | -Jn) \rangle_{n \sim P} = \rho_-(P)$ with $\rho_{\pm}(P)$ defined earlier in equation (15). In accordance with the unbiased parameters choice, the external threshold is set to zero, $\vartheta_i = \vartheta_j = 0$.

Symmetric network. In a symmetric network one has $J_{ij} = J_{ji}$. Thus only the two following cases can be realised:

- Ferromagnetic coupling $J_{ij} = J_{ji} = J$, leading to $P_i(t + 2) = \rho_+(\rho_+(P_i(t)))$.
- Antiferromagnetic coupling $J_{ij} = J_{ji} = -J$, leading to $P_i(t + 2) = \rho_-(\rho_-(P_i(t)))$.

In the stationary state, the node activation probability is given by the fixed point of the map $p^+ = \rho_+(p^+)$ in the case of ferromagnetic couplings, or $p^- = \rho_-(p^-)$ in the case of antiferromagnetic couplings. We now exploit the above observation by considering an unbiased network of isolated dimers, allowing us to use the above result obtained for individual dimers. The site average of the node activation probability gives

$$\langle P \rangle = \frac{p^+ + p^-}{2} = \left(2 - \frac{1}{2} \tanh^2(\beta J/2) \right)^{-1} > 1/2. \tag{51}$$

While this result is derived for a special class of networks, we expect that for a generic unbiased symmetric network the site average activation is larger than $1/2$, as shown by figure 12 where MC simulation shows that $\langle P \rangle > 1/2$ at stationarity. Indeed, from equation (51) this simple argument predicts that at $T = 0$, $P = 2/3$ which is comparable with the simulation result in figure 12 (see blue markers).

Antisymmetric network. An antisymmetric network defined by $J_{ij} = -J_{ji}$ is always unbiased. Let $J_{ij} = J$ (then $J_{ji} = -J$), equation (50) gives

- $P_i(t + 2) = \rho_+(\rho_-(P_i(t)))$
- $P_j(t + 2) = \rho_-(\rho_+(P_j(t)))$.

In the stationary state, the node activation of node i satisfies $p_i = \rho_+(\rho_-(p_i))$, and the node activation of node j satisfies $p_j = \rho_-(\rho_+(p_j))$. Again, we consider a network of

isolated dimers. The site average of the node activation probability gives

$$\langle P \rangle = \frac{p_i + p_j}{2} = \left(2 + \frac{1}{2} \tanh^2(\beta J/2) \right)^{-1} < 1/2, \quad (52)$$

which suggests that in an unbiased antisymmetric network, the site average activation is smaller than $1/2$ and it is confirmed in figure 12. From equation (52), this simple argument predicts that at $T = 0$, $P = 2/5$ which is comparable with the result in figure 12 (see green markers).

The result of equation (51) is an example where the dynamics of Ising spin and indicator variables leads to macroscopically different stationary states. While a spin model defined for unbiased parameter choices presents $N^{-1} \sum_i m_i = 0$ at equilibrium, for indicator variable we have rationalised that in general $2 \sum_i P_i - 1 \neq 0$.

8. Summary and discussion

In this paper we have solved the stochastic dynamics of Boolean linear threshold models for a class of networks characterised by fat-tailed degree distribution. On fully asymmetric networks, the distribution of the node activation probabilities in the stationary state is not unimodal in general but displays a rich structure. Salient features of the distribution can in part be rationalised in terms of discrete stochastic maps resulting from the underlying network structure and dynamics.

We have adapted the dynamic cavity method to study equal-time pairwise correlations in the stationary state. For every pair of nodes, we have computed the correlation produced by common ancestors at a distance of one or two from the pair. We demonstrate that these two motifs capture most of the statistics of values, indicating that pairwise correlation is effectively a local property of the network.

Finally, we have investigated the individual node activation statistics in the stationary state for two additional classes of models, namely for models with multi-node interactions and for linear threshold models defined on networks with bi-directional links. For models with multi-node interactions, we have adopted a bipartite graph representation in which factor nodes represent ‘AND’ gates, and we have shown that this construction is universal in the sense that it allows representing *any* Boolean function. We have previously used models of this type to describe the regulatory mechanism of transcription factors (TFs), and we have so far focussed on the percolation problem [50, 51] as a way to identify structural properties that gene regulatory networks need to exhibit in order to support multi-cellular life. The present manuscript provides an additional perspective that incorporates the role of noise and of inhibition, which were outside the focus of our previous analysis. In the high noise limit, we observe that TFs having a large in-degree are more robust to spontaneous noise-induced activation, which suggests a potential use as noise filters. On the other hand, results from percolation theory suggest that the average in-degree of TF complexes should be small to allow sustaining a multiplicity of attractors. Hence, the combination of those results may hint at a trade-off between these antagonistic conditions being relevant for gene regulatory networks.

To study systems with bi-directional links, we apply an approximation procedure, called OTA, to analyse the stationary state. The OTA assumes a factorisation of the memory kernel which in some sense amounts to a Markov assumption. The reason why the OTA is found to be effective in describing the non-equilibrium stationary state may thus be very well be related to the fact that the lagged auto-correlation is found to be short-ranged.

The dynamics of the linear threshold model with bi-directional links yields a somewhat surprising result: a network with an unbiased distribution of $\pm J$ interactions sustains a biased distribution of node activation probability, both for symmetric and anti-symmetric interaction matrices. We provide a heuristic argument that helps to explain this phenomenon. Symmetric and anti-symmetric interactions are known to appear in biological networks as a mechanism for positive and negative feedback mechanisms, respectively [76]. The present finding is of potential interest in the context of biological networks and may help explain sustained activation or quiescence of nodes. Also, it demonstrates once more how the intuition built about the Ising spins model may be misleading if one investigates models defined on indicator variables [77].

Acknowledgments

G T is supported by the EPSRC Centre for Doctoral Training in Cross-Disciplinary Approaches to Non-Equilibrium Systems (CANES EP/L015854/1).

Appendix A. High noise regime

In the high noise ($\beta \rightarrow 0$ limit), the expansion of $\Phi_\beta(x)$ in equation (9) to first order in βx gives⁸

$$P_i(t+1) \approx \frac{1}{2} + \beta \Phi_1'(0) \left(\sum_j J_{ij} P_j(t) - \vartheta_i \right). \quad (\text{A.1})$$

Using $P_j(t) = \frac{1}{2} + \mathcal{O}(\beta)$, we obtain that

$$P_i(t+1) = \frac{1}{2} + \beta \Phi_1'(0) \left(\frac{1}{2} \sum_j J_{ij} - \vartheta_i \right) + \mathcal{O}(\beta^2), \quad (\text{A.2})$$

which is independent of the time-step t . The result of equation (A.2) states that in the large noise limit, the node activation of node i , P_i , is solely determined by the imbalance of the interactions with i 's predecessors, i.e. by $\sum_j J_{ij}$. The mean node activation probability over sites $\langle P \rangle = \sum_i P_i/N$ reduces to

$$\langle P \rangle = \frac{1}{2} + \beta \Phi_1'(0) \left(\frac{1}{2N} \sum_{i,j} J_{ij} - \frac{1}{N} \sum_i \vartheta_i \right) + \mathcal{O}(\beta^2), \quad (\text{A.3})$$

⁸Note that $\Phi_\beta(x) = \Phi_1(\beta x)$, both for thermal noise, and for Gaussian noise, if we equate $\beta = \sigma$ in the Gaussian case.

where the $\mathcal{O}(\beta)$ correction depends only on the macroscopic quantities $\sum_{i,j} J_{ij}/N$ and $\sum_i \vartheta_i/N$. This result is compatible with the expansion around the paramagnetic phase done in [39].

Similarly, the high noise approximation for the AND gate dynamics discussed in equation (31) gives

$$P_i \approx \frac{1}{2} + \beta \Phi_1'(0) \left[\sum_{\mu} J_{i\mu} \mathcal{P}_{\mu} - \vartheta_i \right], \quad (\text{A.4})$$

where

$$\mathcal{P}_{\mu} = \left(\frac{1}{2}\right)^{c_{\mu}} + \beta \Phi_1'(0) \sum_{j \in \partial_{\mu}} \sum_{\nu} J_{j\nu} \left(\frac{1}{2}\right)^{c_{\nu}} + \mathcal{O}(\beta^2). \quad (\text{A.5})$$

In this case, the average node activation probability is therefore given by

$$\langle P \rangle = \frac{1}{2} + \beta \Phi_1'(0) \left[\frac{1}{N} \sum_{i,\mu} J_{i\mu} \left(\frac{1}{2}\right)^{c_{\mu}} - \frac{1}{N} \sum_i \vartheta_i \right] + \mathcal{O}(\beta^2). \quad (\text{A.6})$$

In the high noise limit, the average activation probability of factor nodes converges to

$$\langle P \rangle = \sum_{c_{\mu}} P(c_{\mu}) \left(\frac{1}{2}\right)^{c_{\mu}} := G_0^C(1/2), \quad \text{as } \beta \rightarrow 0. \quad (\text{A.7})$$

with $G_0^C(x)$ the generating function of the in-degree of factor nodes. This demonstrates, among other things, that the signalling through AND gates with large in-degrees is effectively suppressed in the high noise limit. Thus, factor nodes with large in-degree may operate as noise filters.

Appendix B. Derivation of the OTA

Dynamic cavity trajectory with bi-directional links. For completeness, we include in this appendix a brief derivation of the OTA approximation from the dynamic cavity equations, mainly following [35, 39]. We also compare different closure schemes of the OTA which have been proposed in the literature and propose a new version as an alternative.

Let us define $P_i(\mathbf{n}_i^{0,\dots,t} | \vartheta_i)$ the probability of the trajectory $\mathbf{n}_i^{0,\dots,t} := (n_i(0) \dots n_i(t))$ of node i over the time span $0, \dots, t$, and $P_j^{(i)}(\mathbf{n}_j^{0,\dots,t-1} | \vartheta_j^{(i)0,\dots,t-2})$ the conditional probability of the trajectory $\mathbf{n}_j^{0,\dots,t-1} := (n_j(0) \dots n_j(t-1))$ of node j , given the history of node i , in the cavity graph where node i and its links have been removed, so that the past values of node i act as external thresholds. Their contribution can be combined with those of the constant threshold ϑ_j , to define time-dependent thresholds $\vartheta_j^{(i)0,\dots,t-1} := (\vartheta_j^{(i)}(0) \dots \vartheta_j^{(i)}(t-1))$ with

$$\vartheta_j^{(i)}(s) = \vartheta_j - J_{ji} n_i(s), \quad (\text{B.1})$$

where $J_{ji}n_i^s$ is the signal received by j from i as a retarded interaction. Above, we have used the notation with time superscript to indicate the collection of variables over the trajectory. In what follows, to ease the notation, we will also use superscripts for single time quantities, e.g. $n_i(s) = n_i^s$ and $\vartheta_j^{(i)}(s) = \vartheta_j^{(i),s}$.

The probability of a single-node trajectory does not factorise in the time-steps, but one needs to solve the set of equations, see [39, 75]

$$P_i(\mathbf{n}_i^{0,\dots,t}|\vartheta_i) = \sum_{\mathbf{n}_{\partial_i}^0} \dots \sum_{\mathbf{n}_{\partial_i}^{t-1}} \left[\prod_{s=0}^{t-1} W[n_i^{s+1}|h_i(\mathbf{n}_{\partial_i}^s), \vartheta_i] \right] \times \prod_{j \in \partial_i} P_j^{(i)}(\mathbf{n}_j^{0,\dots,t-1}|\vartheta_j - J_{ji}\mathbf{n}_i^{0,\dots,t-2}) P_i(\mathbf{n}_i^0), \quad (\text{B.2})$$

$$P_j^{(i)}(\mathbf{n}_j^{0,\dots,t-1}|\vartheta_j^{(i),0,\dots,t-2}) = \sum_{\mathbf{n}_{\partial_j \setminus i}^0} \dots \sum_{\mathbf{n}_{\partial_j \setminus i}^{t-2}} \left[\prod_{s=0}^{t-2} W[n_j^{s+1}|h_j^{(i)}(\mathbf{n}_{\partial_j}^s), \vartheta_j^{(i),s}] \right] \times \prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)}(\mathbf{n}_\ell^{0,\dots,t-2}|\vartheta_\ell - J_{\ell j}\mathbf{n}_j^{0,\dots,t-3}) P_j(\mathbf{n}_j^0), \quad (\text{B.3})$$

where $\vartheta_j = (\vartheta_j, \dots, \vartheta_j)$ is a constant vector with identical entries ϑ_j . Note that for any node $j \in \partial_i$, i.e. such that $J_{ij} \neq 0$, one has $J_{ji} \neq 0$ only if the link (i, j) is bidirectional, so the retarded self-interaction terms in equation (B.1) are absent for unidirectional links. The transition probability $W[n_j^{s+1}|h_j^{(i)}(\mathbf{n}_{\partial_j}^s), \vartheta_j^{(i),s}]$ indicates the conditional probability of the state n_j^{s+1} at time $s+1$ given the values of the cavity local field $h_j^{(i)}(\mathbf{n}_{\partial_j}^s)$,

$$h_j^{(i)}(\mathbf{n}_{\partial_j}^s) := \sum_{\ell \in \partial_j \setminus i} J_{j\ell} n_\ell^s, \quad (\text{B.4})$$

and the cavity threshold $\vartheta_j^{(i),s}$. The detailed function form of $W[n_j^{s+1}|h_j^{(i)}(\mathbf{n}_{\partial_j}^s), \vartheta_j^{(i),s}]$ depends on the choice of the noise distribution. In analogy with the above we adopt a thermal noise model for which

$$W[n_j^{s+1}|h_j^{(i)}(\mathbf{n}_{\partial_j}^s), \vartheta_j^{(i),s}] = \frac{1}{2} \left\{ 1 + (2n_j^{s+1} - 1) \tanh \left[\frac{\beta}{2} \left(h_j^{(i)}(\mathbf{n}_{\partial_j}^s) - \vartheta_j^{(i),s} \right) \right] \right\}. \quad (\text{B.5})$$

In equations (B.2) and (B.3) the probability of the state of node i depends on the states of neighbours ∂_i , which in turn depend parametrically on the node i through the cavity thresholds. This feedback effect is responsible for the non-Markovian structure of equations (B.2) and (B.3), entailing that the computational complexity for a trajectory of time length t grows as 2^t , making the evaluation feasible only for relatively few time-steps.

From now on, we focus on the single site cavity probability and we discuss the quantity $P_j^{(i)}(\mathbf{n}_j^{0,\dots,t}|\vartheta_j^{(i),0,\dots,t-1})$ for the trajectory up to time t and not up to time $t -$

1 as done in equation (B.3) to simplify the notation of the lagged time dependence. A pragmatic approach is to assume that the single site cavity probability trajectory $P_j^{(i)}(\mathbf{n}_j^{0,\dots,t} | \vartheta_j^{(i),0,\dots,t-1})$ factorises in a Markovian fashion under the OTA [39]

$$P_j^{(i)}(\mathbf{n}_j^{0,\dots,t} | \vartheta_j^{(i),0,\dots,t-1}) \approx P_j^{(i)}(n_j^0) \prod_{s=0}^{t-1} P_j^{(i)}(n_j^{s+1} | \vartheta_j^{(i),s}). \quad (\text{B.6})$$

Substituting equation (B.6) into equation (B.3)

$$P_j^{(i)}(\mathbf{n}_j^{0,\dots,t} | \vartheta_j^{(i),0,\dots,t-1}) = \left[\sum_{\mathbf{n}_{\partial_j \setminus i}^{t-1}} W \left[n_j^t | h_j^{(i)}(\mathbf{n}_{\partial_j}^{t-1}), \vartheta_j^{(i),t-1} \right] \prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)}(n_\ell^{t-1} | \vartheta_\ell - J_{\ell j} n_j^{t-2}) \right] \times P_j^{(i)}(\mathbf{n}_j^{0,\dots,t-1} | \vartheta_j^{(i),0,\dots,t-2}), \quad (\text{B.7})$$

where we have grouped together the terms of equation (B.6) at the previous time-steps

$$P_j^{(i)}(\mathbf{n}_j^{0,\dots,t-1} | \vartheta_j^{(i),0,\dots,t-2}) = \sum_{\mathbf{n}_{\partial_j \setminus i}^0} \dots \sum_{\mathbf{n}_{\partial_j \setminus i}^{t-2}} \left[\prod_{s=0}^{t-2} W \left[n_j^{s+1} | h_j^{(i)}(\mathbf{n}_{\partial_j}^s), \vartheta_j^{(i),s} \right] \right] \times \prod_{\ell \in \partial_j \setminus i} \left[\prod_{s=0}^{t-3} P_\ell^{(j)}(n_\ell^{s+1} | \vartheta_\ell - J_{\ell j} n_j^s) P_\ell(n_\ell^0) \right] P_j(n_j^0). \quad (\text{B.8})$$

Applying the OTA factorization to the rightmost factor in equation (B.7), i.e. equation (B.8), one obtains

$$P_j^{(i)}(\mathbf{n}_j^{0,\dots,t} | \vartheta_j^{(i),0,\dots,t-1}) = \left[\sum_{\mathbf{n}_{\partial_j \setminus i}^{t-1}} W \left[n_j^t | h_j^{(i)}(\mathbf{n}_{\partial_j}^{t-1}), \vartheta_j^{(i),t-1} \right] \prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)}(n_\ell^{t-1} | \vartheta_\ell - J_{\ell j} n_j^{t-2}) \right] \times \left[\prod_{s=0}^{t-2} P_j^{(i)}(n_j^{s+1} | \vartheta_j^{(i),s}) \right] P_j(n_j^0). \quad (\text{B.9})$$

This expression is particularly suitable to compute the marginal over the trajectory of node j , i.e.

$$P_j^{(i)}(n_j^t | \vartheta_j^{(i),0,\dots,t-1}) = \sum_{\mathbf{n}_j^{0,\dots,t-1}} P_j^{(i)}(\mathbf{n}_j^{0,\dots,t} | \vartheta_j^{(i),0,\dots,t-1}). \quad (\text{B.10})$$

Thanks to the factorization of the last term in equation (B.9), the sum over $\mathbf{n}_j^{0,\dots,t-3,t-1}$ is trivial, as the term in the first square brackets does not depend on these variables, and

conditional probabilities are normalised to one.

$$\begin{aligned}
 P_j^{(i)} \left(n_j^t | \vartheta_j^{(i),t-1}, \vartheta_j^{(i),t-3} \right) &= \sum_{n_j^{t-2}} \left[\sum_{\mathbf{n}_{\partial_j \setminus i}^{t-1}} W \left[n_j^t | h_j^{(i)}(\mathbf{n}_{\partial_j}^{t-1}), \vartheta_j^{(i),t-1} \right] \right. \\
 &\quad \left. \times \prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)} \left(n_\ell^{t-1} | \vartheta_\ell - J_{\ell j} n_j^{t-2} \right) \right] P_j^{(i)}(n_j^{t-2} | \vartheta_j^{(i),t-3}).
 \end{aligned} \tag{B.11}$$

In equation (B.11) the probability of node j depends on the two cavity thresholds $\vartheta_j^{(i),t-1}, \vartheta_j^{(i),t-3}$. This expression is directly derived from the OTA factorisation equation (B.6). The one-step object $P_j^{(i)} \left(n_j^t | \vartheta_j^{(i),t-1} \right)$ can be obtained imposing a ‘closure’ condition that enforces the Markovian behaviour, and different choices could be made. In [39, 75] the authors assume that, in the stationary state, the cavity threshold does not have an explicit time dependence $\vartheta_j^{(i),s} \approx \vartheta_j^{(i)} \forall s$. In [34] instead, the cavity term is approximated by the non-cavity marginal $P_j^{(i)}(n_j^{t-2} | \vartheta_j^{(i),t-3}) \approx P_j(n_j^{t-2} | \vartheta_j)$. Another approach is to truncate the retarded dependence by taking an average over the state n_i^{t-3} , which we detail below. Remembering from equation (B.1) that the $\vartheta_j^{(i),s} = \vartheta_j - J_{ji} n_i^s$, then

$$P(\vartheta_j^{(i),t-3}) = \sum_{n_i^{t-3}} \delta_{\vartheta_j^{(i),t-3}, \vartheta_j - J_{ji} n_i^{t-3}} P(n_i^{t-3} | \vartheta_i), \tag{B.12}$$

with δ indicating the Kronecker delta. The one-time cavity marginal appearing in equation (B.6) is given by

$$P_j^{(i)} \left(n_j^t | \vartheta_j^{(i),t-1} \right) = \sum_{n_i^{t-3}} P_j^{(i)} \left(n_j^t | \vartheta_j^{(i),t-1}, \vartheta_j - J_{ji} n_i^{t-3} \right) P_i(n_i^{t-3} | \vartheta_i). \tag{B.13}$$

Substituting equation (B.11) in equation (B.13), the expression for the one-step probability becomes

$$\begin{aligned}
 P_j^{(i)} \left(n_j^t | \vartheta_j^{(i),t-1} \right) &= \sum_{n_j^{t-2}} \left[\sum_{\mathbf{n}_{\partial_j \setminus i}^{t-1}} W \left[n_j^t | h_j^{(i)}(\mathbf{n}_{\partial_j}^{t-1}), \vartheta_j^{(i),t-1} \right] \right. \\
 &\quad \left. \times \prod_{\ell \in \partial_j \setminus i} P_\ell^{(j)} \left(n_\ell^{t-1} | \vartheta_\ell - J_{\ell j} n_j^{t-2} \right) \right] P_j^{(i)}(n_j^{t-2} | \vartheta_j),
 \end{aligned} \tag{B.14}$$

with

$$P_j^{(i)}(n_j^{t-2} | \vartheta_j) = \sum_{n_i^{t-3}} P_j^{(i)}(n_j^{t-2} | \vartheta_j - J_{ji} n_i^{t-3}) P_i(n_i^{t-3} | \vartheta_i). \tag{B.15}$$

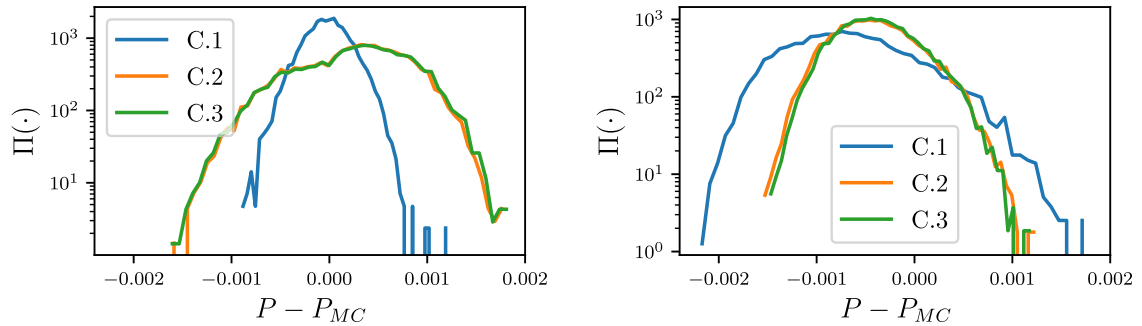


Figure B1. Comparison of different closures in the OTA. C.1 is the closure of [35]. C.2 is the closure of [34], C.3 is the closure of equation (B.15). The distribution of the difference between the probability obtained using OTA and the probability obtained using MC simulations. Network with symmetric (left) and antisymmetric interactions (right). The network belongs to the random regular graph class with $k^{\text{in}} = k^{\text{out}} = 3$. Simulations are obtained through an average of over 500 trajectories. Parameters are $N = 10\,000$, $T/J = 1$, and $\vartheta = 0$, $t_s = 10^4$ steps.

We compare the different OTA closure conditions in equation (B.11), which we denote:

- C.1 corresponding to the assumption $P_j^{(i)}(n_j^{t-2}|\vartheta_j^{(i),t-3}) \approx P_j^{(i)}(n_j^{t-2}|\vartheta_j^{(i),t-1})$, from [35],
- C.2 corresponding to the assumption $P_j^{(i)}(n_j^{t-2}|\vartheta_j^{(i),t-3}) \approx P_j(n_j^{t-2}|\vartheta_j)$, from [34],
- C.3 corresponding to the assumption $P_j^{(i)}(n_j^{t-2}|\vartheta_j^{(i),t-3}) \approx P_j^{(i)}(n_j^{t-2}|\vartheta_j)$ using the definition in equation (B.15).

We evaluate the distribution of the stationary single-node activation probability for the OTA expressions corresponding to closures C.1, C.2, and C.3. We benchmark the theory with MC simulations for symmetric and antisymmetric networks, and we show in figure B1 the distribution of the distance between theory and simulation for the single-site activation probabilities. Our results indicate that in the symmetric case closure C.1 outperforms both C.2 and C.3, which is expected since the expression associated with C.1 admits the equilibrium solution of belief propagation, as discussed in [39]. In the antisymmetric case instead, the relative performances of the methods are changed, with closure C.1 providing the worst result roughly by a factor of two compared to closures C.2 and C.3. We use the mean square distance between OTA closures and simulation to quantify the error in table B1. Our results indicate that closures C.2 and C.3 are less dependent on the network symmetry, while C.1 presents important variations between the two cases examined. From a computational point of view, closures C.1 and C.2 are comparable and they are simpler to perform compared to C.3, since in closure C.3, the equation (B.15) needs to be performed for every link and for every time-step of interest. Hence we adopt closure C.2 since it is simpler to implement than C.3, and our results suggest it gives the same performances as closure C.3.

Table B1. Error, as defined in equation (46), of different closure schemes for symmetric, antisymmetric and uncorrelated interactions. Same parameters as in figure B1.

	C.1	C.2	C.3
Symmetric	2.30×10^{-4}	6.0×10^{-4}	6.2×10^{-4}
Antisymmetric	8.6×10^{-4}	5.5×10^{-4}	5.3×10^{-4}
Uncorrelated	3.8×10^{-4}	3.4×10^{-4}	3.4×10^{-4}

References

- [1] Edwards S F and Anderson P W 1975 Theory of spin glasses *J. Phys. F: Met. Phys.* **5** 965–74
- [2] Sherrington D and Kirkpatrick S 1975 Solvable model of a spin-glass *Phys. Rev. Lett.* **35** 1792–6
- [3] Viana L and Bray A J 1985 Phase diagrams for dilute spin glasses *J. Phys. C: Solid State Phys.* **18** 3037
- [4] Kauffman S A 1969 Metabolic stability and epigenesis in randomly constructed genetic nets *J. Theor. Biol.* **22** 437–67
- [5] Derrida B and Pomeau Y 1986 Random networks of automata: a simple annealed approximation *Europhys. Lett.* **1** 45
- [6] Parisi G 1990 A simple model for the immune network *Proc. Natl Acad. Sci. USA* **87** 429–33
- [7] Agliari E, Barra A, Galluzzi A, Guerra F and Moauro F 2012 Multitasking associative networks *Phys. Rev. Lett.* **109** 268101
- [8] Agliari E, Annibale A, Barra A, Coolen A C C and Tantari D 2013 Immune networks: multitasking capabilities near saturation *J. Phys. A: Math. Theor.* **46** 415003
- [9] Hopfield J J 1982 Neural networks and physical systems with emergent collective computational abilities *Proc. Natl Acad. Sci. USA* **79** 2554–8
- [10] Hertz J, Krogh A and Palmer R G 2018 *Introduction to the Theory of Neural Computation* (Boca Raton, FL: CRC Press)
- [11] Amit D J, Gutfreund H and Sompolinsky H 1987 Statistical mechanics of neural networks near saturation *Ann. Phys., NY* **173** 30–67
- [12] Sollich P, Tantari D, Annibale A and Barra A 2014 Extensive parallel processing on scale-free networks *Phys. Rev. Lett.* **113** 238106
- [13] Challet D and Zhang Y-C 1997 Emergence of cooperation and organization in an evolutionary game *Physica A* **246** 407–18
- [14] Coolen A C C 2005 *The Mathematical Theory of Minority Games: Statistical Mechanics of Interacting Agents* (Oxford: Oxford University Press)
- [15] Iori G 1999 Avalanche dynamics and trading friction effects on stock market returns *Int. J. Mod. Phys. C* **10** 1149–62
- [16] Bornholdt S 2001 Expectation bubbles in a spin model of markets: intermittency from frustration across scales *Int. J. Mod. Phys. C* **12** 667–74
- [17] Anand K and Kühn R 2007 Phase transitions in operational risk *Phys. Rev. E* **75** 016111
- [18] Hachtet J P L and Kühn R 2006 Effects of economic interactions on credit risk *J. Phys. A: Math. Gen.* **39** 2231–51
- [19] Weigt M and Hartmann A K 2000 Number of guards needed by a museum: a phase transition in vertex covering of random graphs *Phys. Rev. Lett.* **84** 6118
- [20] Cocco S and Monasson R 2001 Statistical physics analysis of the computational complexity of solving random satisfiability problems using backtrack algorithms *Eur. Phys. J. B* **22** 505–31
- [21] Martin O C, Monasson R and Zecchina R 2001 Statistical mechanics methods and phase transitions in optimization problems *Theor. Comput. Sci.* **265** 3–67
- [22] Franz S, Leone M, Ricci-Tersenghi F and Zecchina R 2001 Exact solutions for diluted spin glasses and optimization problems *Phys. Rev. Lett.* **87** 127209
- [23] Mézard M, Parisi G and Zecchina R 2002 Analytic and algorithmic solution of random satisfiability problems *Science* **297** 812–5

- [24] Mézard M and Parisi G 2001 The Bethe lattice spin glass revisited *Eur. Phys. J. B* **20** 217–33
- [25] Coolen A 2016 Replica methods for loopy sparse random graphs *J. Phys.: Conf. Ser.* **699** 012022
- [26] Annibale A, Coolen A C C and Bianconi G 2010 Network resilience against intelligent attacks constrained by the degree-dependent node removal cost *J. Phys. A: Math. Theor.* **43** 395001
- [27] Altarelli F, Braunstein A, Dall’Asta L and Zecchina R 2013 Optimizing spread dynamics on graphs by message passing *J. Stat. Mech.* **P09011**
- [28] Lokhov A Y, Mézard M and Zdeborová L 2015 Dynamic message-passing equations for models with unidirectional dynamics *Phys. Rev. E* **91** 012811
- [29] Paga P and Kühn R 2015 Contagion in an interacting economy *J. Stat. Mech.* **P03008**
- [30] Li B and Saad D 2021 Impact of presymptomatic transmission on epidemic spreading in contact networks: a dynamic message-passing analysis *Phys. Rev. E* **103** 052303
- [31] Karrer B and Newman M E J 2010 Message passing approach for general epidemic models *Phys. Rev. E* **82** 016101
- [32] Lokhov A Y, Mézard M, Ohta H and Zdeborová L 2014 Inferring the origin of an epidemic with a dynamic message-passing algorithm *Phys. Rev. E* **90** 012801
- [33] Roudi Y and Hertz J 2011 Dynamical tap equations for non-equilibrium Ising spin glasses *J. Stat. Mech.* **P03031**
- [34] Zhang P 2012 Inference of kinetic Ising model on sparse graphs *J. Stat. Phys.* **148** 502–12
- [35] Aurell E and Mahmoudi H 2012 Dynamic mean-field and cavity methods for diluted Ising systems *Phys. Rev. E* **85** 031119
- [36] De Dominicis C 1978 Dynamics as a substitute for replicas in systems with quenched random impurities *Phys. Rev. B* **18** 4913–9
- [37] Hatchett J P L, Wemmenhove B, Castillo I P, Nikolettopoulos T, Skantzos N S and Coolen A C C 2004 Parallel dynamics of disordered Ising spin systems on finitely connected random graphs *J. Phys. A: Math. Gen.* **37** 6201–20
- [38] Mimura K and Coolen A C C 2009 Parallel dynamics of disordered Ising spin systems on finitely connected directed random graphs with arbitrary degree distributions *J. Phys. A: Math. Theor.* **42** 415001
- [39] Neri I and Bollé D 2009 The cavity approach to parallel dynamics of Ising spins on a graph *J. Stat. Mech.* **P08009**
- [40] Kühn R and Rogers T 2017 Heterogeneous micro-structure of percolation in sparse networks *Europhys. Lett.* **118** 68003
- [41] Leskovec J and Krevl A 2014 SNAP datasets: stanford large network dataset collection <http://snap.stanford.edu/data>
- [42] Albert R and Barabási A-L 2002 Statistical mechanics of complex networks *Rev. Mod. Phys.* **74** 47–97
- [43] Dorogovtsev S N and Mendes J F 2003 *Evolution of Networks: From Biological Networks to the Internet and WWW* (Oxford: Oxford University Press)
- [44] Newman M E J 2018 *Networks: An Introduction* 2nd edn (Oxford: Oxford University Press)
- [45] Estrada E 2011 *The Structure of Complex Networks: Theory and Applications* (Oxford: Oxford University Press)
- [46] Latora V, Nicosia V and Russo G 2017 *Complex Networks: Principles, Methods and Applications* (Cambridge: Cambridge University Press)
- [47] Torrisi G, Annibale A and Kühn R 2021 Overcoming the complexity barrier of the dynamic message-passing method in networks with fat-tailed degree distributions *Phys. Rev. E* **104** 045313
- [48] Fink T and Hannam R 2021 Boolean composition restricts biological logics (arXiv:2109.12551)
- [49] Gates A J, Correia R B, Wang X and Rocha L M 2021 The effective graph reveals redundancy, canalization, and control pathways in biochemical regulation and signaling *Proc. Natl Acad. Sci.* **118** e2022598118
- [50] Torrisi G, Kühn R and Annibale A 2020 Percolation in the gene regulatory network *J. Stat. Mech.* **083501**
- [51] Hannam R, Kühn R and Annibale A 2019 Percolation in bipartite Boolean networks and its role in sustaining life *J. Phys. A: Math. Theor.* **52** 334002
- [52] Hertz J A, Grinstein G and Solla S A 1986 Memory networks with asymmetric bonds *AIP Conf. Proc.* **151** 212–8
- [53] Gutfreund H, Reger J D and Young A P 1988 The nature of attractors in an asymmetric spin glass with deterministic dynamics *J. Phys. A: Math. Gen.* **21** 2775
- [54] Han H *et al* 2018 TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions *Nucleic Acids Res.* **46** D380–6
- [55] Del Ferraro G and Aurell E 2015 Dynamic message-passing approach for kinetic spin models with reversible dynamics *Phys. Rev. E* **92** 010102
- [56] Mézard M and Sakellariou J 2011 Exact mean-field inference in asymmetric kinetic Ising systems *J. Stat. Mech.* **L07001**

- [57] Gleeson J P 2011 High-accuracy approximation of binary-state dynamics on networks *Phys. Rev. Lett.* **107** 068701
- [58] Deplancke B and Gheldof N 2012 *Gene Regulatory Networks: Methods and Protocols* (Berlin: Springer)
- [59] Bitbol A-F, Dwyer R S, Colwell L J and Wingreen N S 2016 Inferring interaction partners from protein sequences *Proc. Natl Acad. Sci. USA* **113** 12180–5
- [60] Morcos F *et al* 2011 Direct-coupling analysis of residue coevolution captures native contacts across many protein families *Proc. Natl Acad. Sci. USA* **108** E1293–301
- [61] Cavagna A, Giardina I, Ginelli F, Mora T, Piovani D, Tavarone R and Walczak A M 2014 Dynamical maximum entropy approach to flocking *Phys. Rev. E* **89** 042707
- [62] Cantwell G T and Newman M E J 2019 Message passing on networks with loops *Proc. Natl Acad. Sci. USA* **116** 23398–403
- [63] Kirkley A, Cantwell G T and Newman M E J 2021 Belief propagation for networks with loops *Sci. Adv.* **7** eabf1211
- [64] Lambiotte R, Rosvall M and Scholtes I 2019 From networks to optimal higher-order models of complex systems *Nat. Phys.* **15** 313–20
- [65] Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A, Young J-G and Petri G 2020 Networks beyond pairwise interactions: structure and dynamics *Phys. Rep.* **874** 1–92
- [66] Ghoshal G, Zlatić V, Caldarelli G and Newman M E J 2009 Random hypergraphs and their applications *Phys. Rev. E* **79** 066118
- [67] Petri G and Barrat A 2018 Simplicial activity driven model *Phys. Rev. Lett.* **121** 228301
- [68] Kirkpatrick T R and Thirumalai D 1987 Dynamics of the structural glass transition and the p -spin-interaction spin-glass model *Phys. Rev. Lett.* **58** 2091–4
- [69] Thomas R 1981 On the relation between the logical structure of systems and their ability to generate multiple steady states or sustained oscillations *Numerical Methods in the Study of Critical Phenomena* (Berlin: Springer) pp 180–93
- [70] Plahte E, Mestl T and Omholt S W 1995 Feedback loops, stability and multistationarity in dynamical systems *J. Biol. Syst.* **03** 409–13
- [71] Freeman M 2000 Feedback control of intercellular signalling in development *Nature* **408** 313–9
- [72] Zhang H, Chen Y and Chen Y 2012 Noise propagation in gene regulation networks involving interlinked positive and negative feedback loops *PLoS One* **7** e51840
- [73] Chakravarty S and Csikász-Nagy A 2021 Systematic analysis of noise reduction properties of coupled and isolated feed-forward loops *PLoS Comput. Biol.* **17** e1009622
- [74] Mambuca A M, Cammarota C and Neri I 2022 Dynamical systems on large networks with predator–prey interactions are stable and exhibit oscillations *Phys. Rev. E* **105** 014305
- [75] Aurell E and Mahmoudi H 2011 A message-passing scheme for non-equilibrium stationary states *J. Stat. Mech.* **P04014**
- [76] Cardelli L, Csikász-Nagy A, Dalchau N, Tribastone M and Tschaikowski M 2016 Noise reduction in complex biological switches *Sci. Rep.* **6** 1–12
- [77] Campajola C, Lillo F, Mazzarisi P and Tantari D 2021 On the equivalence between the kinetic Ising model and discrete autoregressive processes *J. Stat. Mech.* **033412**
- [78] Anderson J A and Rosenfeld E (ed) 1988 *Neurocomputing: Foundations of Research* (Cambridge, MA: MIT Press)