CrossMark

On the Sufficiency of Pairwise Interactions in Maximum Entropy Models of Networks

Lina Merchan[1,2] · Ilya Nemenman[3]

Abstract Biological information processing networks consist of many components, which are coupled by an even larger number of complex multivariate interactions. However, analyses of data sets from fields as diverse as neuroscience, molecular biology, and behavior have reported that observed statistics of states of some biological networks can be approximated well by maximum entropy models with only pairwise interactions among the components. Based on simulations of random Ising spin networks with p-spin ($p > 2$) interactions, here we argue that this reduction in complexity can be thought of as a natural property of densely interacting networks in certain regimes, and not necessarily as a special property of living systems. By connecting our analysis to the theory of random constraint satisfaction problems, we suggest a reason for why some biological systems may operate in this regime.

Keywords Collective dynamics · p-spin models · Numerical simulations

1 Introduction

The increased throughput of biological experiments now allows joint measurements of activities of many basic components underlying collective dynamics in biological systems. Such multivariate data must be interpreted within models. Within this context, Maximum Entropy (MaxEnt) models [1] have been some of the most successful. The logic of such models is that, ultimately, one wants to find an approximation $Q(\mathbf{x})$ to the joint probability distribution $P(\mathbf{x})$ of the observed multivariate data $\{x_i\} = \mathbf{x}$, $i = 1, \ldots, N$. Unfortunately, for a large number of components, N, the datasets can never be large enough to estimate P directly

✉ Ilya Nemenman
  ilya.nemenman@emory.edu

1   Present Address: Department of Engineering Technology, Savannah State University, Savannah, GA 31404, USA

2   Department of Physics, Emory University, Atlanta, GA 30322, USA

3   Departments of Physics and Biology, Emory University, Atlanta, GA 30322, USA

from data. One may only be able to estimate various expectation values of functions of the data, $\langle f_\kappa(\mathbf{x}) \rangle_P = \tilde{f}_\kappa$, $\kappa = 1, \ldots, K$. Then one can search for Q that matches the reliable estimates. If additionally one requests that Q has no structure beyond that required by the matching, then this is equivalent to asking for Q with the maximum entropy, subject to the constraints imposed by the matching,

$$Q = \arg \max S(Q) - \sum_\kappa \lambda_\kappa \left(\langle f_\kappa \rangle_Q - \tilde{f}_\kappa \right), \tag{1}$$

where the entropy S is defined as

$$S(Q) \equiv S(\mathbf{x}) = -\sum_\mathbf{x} Q(\mathbf{x}) \log_2 Q(\mathbf{x}). \tag{2}$$

A common special case of this general formulation is when the variables are binary, which we will denote as $x_i = \sigma_i \in \{-1, 1\}$, and the data constrain their various low-order correlation functions, such as $\langle \sigma_i \rangle$ or $\langle \sigma_i \sigma_j \rangle$. In this case, the MaxEnt approximation Q is [2]:

$$Q(\boldsymbol{\sigma}) = \frac{1}{Z} \exp \left(\sum_i h_i \sigma_i + \sum_{ij} J_{ij} \sigma_i \sigma_j + \sum_{ijk} K_{ijk} \sigma_i \sigma_j \sigma_k + \cdots \right). \tag{3}$$

Here every constrained correlation function gets a term in the exponent, Z is the partition function, and the Lagrange multipliers h_i, J_{ij}, K_{ijk}, ... must be chosen to satisfy the constraints. This is generally not an analytically solvable problem, and even numerics are hard [3–9].

Equation (3) has the form of the Ising spin problem, allowing a wholesale import of intuition from statistical physics to MaxEnt data analysis. Correspondingly, these ideas have been applied to many biological systems in the last decade [10], starting with neurophysiological recordings from salamander retina [11]. There N was a few dozen neurons, and $\sigma_i = \pm 1$ corresponded to the i'th neuron spiking/not spiking at a certain time. A surprising result was that truncating Eq. (3) at the quadratic order in σ_i (or, in other words, constraining Q up to pairwise correlations) provided a good fit to P. We will refer to this finding as *pairwise sufficiency* from now on.

The pairwise sufficiency was later found in other neural systems [12–14] (though it is violated at larger N [15]). It was observed further for natural images [16]; for discrete, yet non-binary $\mathbf{x}$ in sequencing data [17,18]; and for real-valued velocities of birds in flocking experiments [19]. Even for some non-MaxEnt approaches, similar findings were also reported [20,21]. One can interpret these observations in the context of biological systems operating in a special regime [22,23]. However, the wide applicability of the findings suggests an alternative: pairwise sufficiency may emerge for a wide class of biological and non-biological networks *generically*. Indeed, sparse sampling of variables in experiments is similar to decimation in statistical physics, and the resulting renormalization group-like flow may decrease the importance of the higher order couplings [24]. Further, in a perturbative regime, where fluctuations away from the independence are small, the pairwise sufficiency also appears [25]. Here we propose one more possibility, arguing that the pairwise sufficiency arises naturally in *strongly coupled* multivariate systems.

In what follows, we first introduce the idea in an intuitive toy model, and then develop it numerically by analyzing randomly generated networks. We show that the pairwise MaxEnt models approximate such random networks surprisingly well. Further, we explore distributions of states of these networks and their models, leading to an explanation of the pairwise sufficiency. Finally, we discuss why diverse biological systems may find themselves in the

pairwise-sufficient regime, but we leave it for the future to investigate if this mechanism is responsible for the sufficiency in experimental networks.

2 Results

2.1 Building Intuition: Networks of XORs

For a tractable example of emergence of the pairwise sufficiency, we focus on Boolean gates. These are the limit of Ising spin networks in the low temperature (strong coupling) regime [2,26]. For example, $\sigma_3 = \sigma_1$ OR σ_2 can be written as $P(\sigma_3|\sigma_2, \sigma_1) = \frac{1}{Z} \exp[J(\sigma_1\sigma_3 + \sigma_2\sigma_3 + \sigma_3)]$ with $J \rightarrow \infty$. Indeed, $\log[P(\sigma_3 = 1|\sigma_1, \sigma_2)/P(\sigma_3 = -1|\sigma_1, \sigma_2)] = 2J(\sigma_1 + \sigma_2 + 1)$, which gives the OR truth table, as can be readily verified. If also $P(\sigma_1 = \pm 1) = P(\sigma_2 \pm 1) = 1/2$, then $1/4P(\sigma_3|\sigma_2, \sigma_1) = P(\sigma_1, \sigma_2, \sigma_3)$. Thus the joint probability distribution for OR has the pairwise MaxEnt form, Eq. (3). Similarly, for $J \rightarrow \infty$, $\sigma_3 = \sigma_1$ AND σ_2 is equivalent to $P(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{Z} \exp[J(\sigma_1\sigma_3 + \sigma_2\sigma_3 - \sigma_3)]$, or, alternatively, $\log[P(\sigma_3 = 1|\sigma_1, \sigma_2)/P(\sigma_3 = -1|\sigma_1, \sigma_2)] = 2J(\sigma_1 + \sigma_2 - 1)$. This is again a pairwise MaxEnt distribution. However, $\sigma_3 = \sigma_1$ XOR σ_2, is equivalent to $P(\sigma_1, \sigma_2, \sigma_3) = \frac{1}{Z} \exp(-K\sigma_1\sigma_2\sigma_3)$, $K \rightarrow \infty$. This is an example of a purely third-order gate, with no pairwise interactions.

In Fig. 1, we now couple of a few such third-order gates to each other. The spins $\sigma_1, \sigma_2, \sigma_3$ are connected by an XOR (left column, first row), and there is no simpler equivalent represen-
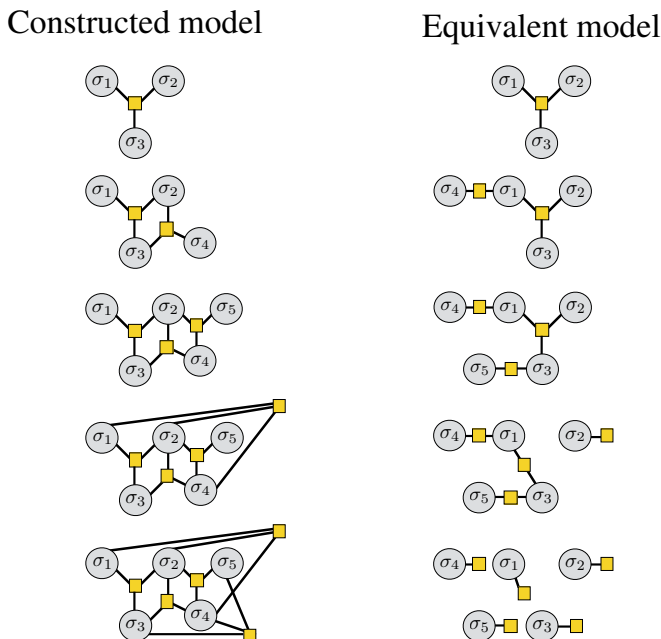

Fig. 1 Emergence of pairwise interactions in a network of XOR gates. On the *left*, we show small networks of spins σ_i (*grey circles*). The spins interact (*yellow squares*) by means of third order XOR interactions. On the *right*, an equivalent network is shown, where some of the XORs get replaced by EQUAL and assignment operations, which are the second and the first order interactions, respectively (Color figure online)

tation of the network (right column, first row). We then add the fourth spin, $\sigma_4 = \sigma_2 \, \text{XOR} \, \sigma_3$ (left column, second row). However, then $\sigma_4 = \sigma_1$. This can be represented as an effective model $P(\sigma_4|\sigma_1, \ldots, \sigma_3) = P(\sigma_4|\sigma_1) = \frac{1}{Z} \exp(J\sigma_1\sigma_4)$, $J = K \to \infty$. Thus the third order XOR interaction is equivalent to a pairwise EQUAL interaction (right column, second row). The latter is effective and nonlocal, in the sense that σ_4 is coupled to σ_1, with which it does not interact in the original construction. Note that this equivalence is true only for the zero temperature, deterministic case ($J = K \to \infty$). For nonzero temperature, probabilities of excited states are not equal in the models on the left and on the right. For example, the state $\boldsymbol{\sigma} = \{+1, +1, +1, +1\}$ has the probability e^{-2K}/Z for the model on the left, and $1/Z$ for the model on the right. Similarly, the state $\boldsymbol{\sigma} = \{-1, +1, -1, +1\}$ has the probabilities $1/Z$ and e^{-2K}/Z, respectively.

We further add $\sigma_5 = \sigma_2 \, \text{XOR} \, \sigma_4$ (third row), and this is equivalent to an effective model $\sigma_5 = \sigma_3$. In short, of the three third order interactions, each constraining one spin and hence "carrying" 1 bit of information, two can be represented without any error as pairwise interactions. Now this deterministic network can exist in four distinct global states out of $2^5 = 32$, determined by $\sigma_{1,2} = \pm 1$ (namely, $\boldsymbol{\sigma}^{(1)} = \{-1, -1, -1, -1, -1\}$, $\boldsymbol{\sigma}^{(2)} = \{-1, +1, +1, -1, +1\}$, $\boldsymbol{\sigma}^{(3)} = \{+1, -1, +1, +1, +1\}$, and $\boldsymbol{\sigma}^{(4)} = \{+1, +1, -1, +1, -1\}$). Thus it is far from the perturbative regime of Ref. [25]. We can grow the network further so that each new spin is coupled by a third order interaction to two existing spins. Then the number of spins, N, and the number of interactions, M, are related as $N = M + 2$, and all but one third order interaction can be represented as a second order one. In other words, a fully pairwise model of such deterministic networks has a relative error of only $1/(N-2)$ when accounting for the statistics of the network states, even if the network was originally constructed with third order gates only.

Alternatively, we can add more XORs without adding new nodes. This may be inconsistent or redundant with already existing couplings. Or in a case such as $\sigma_1 = \sigma_2 \, \text{XOR} \, \sigma_4$ (fourth row), this sets $\sigma_2 = -1$ (thus adding the bias, or the first order term), and all other spins are equal to each other, so that the pairwise effective model is exact. Finally, adding $\sigma_3 = \sigma_4 \, \text{XOR} \, \sigma_5$ sets every spin to -1, and makes even the first-order model exact (bottom row).

We see that networks of deterministic XORs can exhibit the pairwise sufficiency nonperturbatively for their ground states because of emergence of chains of effective pairwise interactions. Formation of such chains and singleton frozen spins, see e. g., Fig. 1 (third row), has been explored previously in hSAT and related Boolean (deterministic gates) problems, under the names of hyperloops, backbone, total constraining, and freezing [27,28]. However, the relation between hyperloop formation and pairwise sufficiency has not been explicitly articulated. Formation of hyperloops is a generic phenomenon in such networks when the density of constraints (gates) is high. Of course, more realistic physical or biological systems are non-deterministic ($J, K < \infty$), and these arguments do not immediately apply. However, the examples above and the previous work for the deterministic case [27,28] give hope. Indeed, one may expect effective pairwise models to approximate more complex networks well when nodes in the network interact strongly and densely, and the space of network states is sufficiently constrained. In such cases, even for networks with $J, K < \infty$, there might be many pairs of nodes that are relatively strongly correlated simply by chance, thus forming a nondeterministic equivalent of a hyperloop. This would still allow replacement of higher order interactions with pairwise ones. In what follows, we will develop this intuition further using numerical studies.

2.2 Pairwise Approximations to Random Networks with Higher Order Interactions

To verify our intuition, we proceed by generating random networks that have *only* higher order nondeterministic interactions among spins (p-spin models, $p > 2$ [29]). We then quantify the accuracy with which lower order MaxEnt models approximate these networks. We explore networks with $p = 3, 4$ to ensure that our findings do not depend on the exact structure of the true higher order interactions. Further, systems with only fourth order couplings have the Z_2 symmetry, and thus cannot include any first order terms in their MaxEnt approximations, Eq. (3). Studying them will allow us to understand if the eventual freezing to a single well-defined state, as in the last row of Fig. 1, is crucial for the pairwise sufficiency, or if it emerges even for nonperturbative networks with more than one highly probable state.

To generate the random networks, we first specify N, the number of nodes, and M the number of interactions. Then for each interaction $\mu = 1, \ldots, M$, we generate its coupling constant K_μ from a zero-mean Gaussian distribution with a certain variance s^2. We then choose three or four spins at random to couple. The overall probability of states for these networks is

$$P_3(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left(\sum_{\mu=1}^{M} K_\mu \sigma_{\mu_1} \sigma_{\mu_2} \sigma_{\mu_3} \right), \quad \text{3-spin model,} \tag{4}$$

$$P_4(\boldsymbol{\sigma}) = \frac{1}{Z} \exp\left(\sum_{\mu=1}^{M} K_\mu \sigma_{\mu_1} \sigma_{\mu_2} \sigma_{\mu_3} \sigma_{\mu_4} \right), \quad \text{4-spin model,} \tag{5}$$

where $\mu_1 < \mu_2 < \mu_3 < \mu_4$, so that the spins do not self-interact. To specify these distributions (and later calculate various errors of approximations), we need to know Z. To decouple studying the problem of the pairwise sufficiency from a hard problem of efficient sampling, we focus on $N \leq 22$, which allows us to estimate Z by direct summation fast enough to do it many times and collect statistics. We generate many such distributions P_3 and P_4, every time picking random $N \in [10, 22]$, $M \in [1, 250]$, and $s \in [0.2, 2.0]$.

For each generated distribution, we estimate its individual and pairwise marginals $P(\sigma_i)$, $P(\sigma_i, \sigma_j)$ for all $i, j = 1, \ldots, N$ by direct marginalization (hereafter we drop subscripts 3 or 4 for P if it does not cause confusion). We then calculate the first order (or independent) MaxEnt approximation

$$Q^{(1)}(\boldsymbol{\sigma}) = \prod_{i=1}^{N} P(\sigma_i). \tag{6}$$

Next we fit the pairwise MaxEnt model $Q^{(2)}$ to P. While good algorithms exist for this purpose (see e.g., [3–9]), it is unclear if their assumptions are satisfied by our networks. Trying again to decouple the problems of efficient inference and the pairwise sufficiency, we choose a classic, well understood Iterative Proportional Fitting Procedure (IPFP) algorithm [30,31]. That is, we start with $Q^{(1)}$ as a guess for $Q^{(2)}$, calculate $Q^{(2)}(\sigma_i, \sigma_j)$, and redefine

$$Q^{(2)}(\boldsymbol{\sigma}) \rightarrow Q^{(2)}(\boldsymbol{\sigma}) \frac{P(\sigma_i, \sigma_j)}{Q^{(2)}(\sigma_i, \sigma_j)}. \tag{7}$$

We cycle through all pairs i, j in order, starting from $\{i, j\} = \{1, 2\}$ through $\{i, j\} = \{N - 1, N\}$, and iterate until $Q^{(2)}(\sigma_i, \sigma_j) \approx P(\sigma_i, \sigma_j)$ for all i and j, up to a relative error of 10^{-5}. This is achieved within $\sim 10^0 \ldots 10^5$ iterations (or more for larger N) depending on how close the final $Q^{(2)}$ is to $Q^{(1)}$, and each iteration itself involves $N(N-1)/2$ applications of
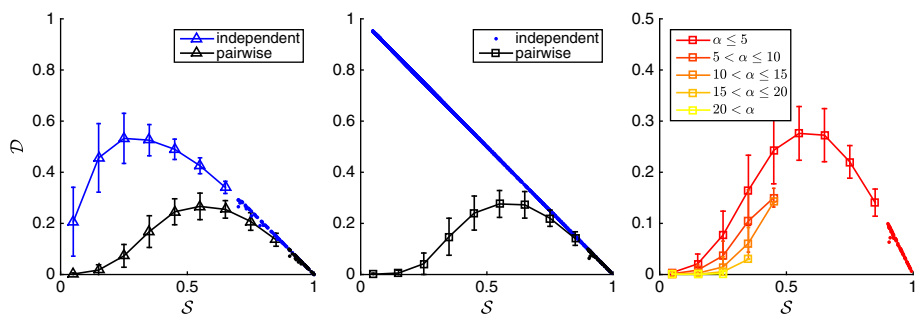
Fig. 2 Error of the MaxEnt fits versus the normalized entropy of the network state space, $\mathcal{S}$. The *left panel* shows errors of the independent, $\mathcal{D}^{(1)}$, and the pairwise, $\mathcal{D}^{(2)}$, approximations for 3-spin networks. We used ~800 random networks with $11 \leq N \leq 20$ spins and with a varying number of interactions, M. We partitioned all the networks by their $\mathcal{S}$ in bins of width of 0.1 and calculated the mean and the standard deviation of $\mathcal{D}$ for each bin. These are indicated by *triangles* and the *error bars*. Wherever the data points for individual networks showed little scatter, we plotted these points instead of the bin averages. The *middle panel* presents similar data, $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$, for 4-spin networks. Here over 4000 random networks were generated with $11 \leq N \leq 22$. $\mathcal{D}^{(2)}$ was again averaged within ten bins, and the means and the standard deviations are plotted. For $\mathcal{D}^{(1)}$, data for individual networks are presented. These merge into a perfect straight line due to the Z_2 symmetry of 4-spin distributions. For both the 3- and the 4-spin cases, the pairwise sufficiency is clear at low $\mathcal{S}$. The *right panel* replots the $\mathcal{D}^{(2)}$ data for the 4-spin networks, but splits them according to α, which measures the average strength of interactions per spin within a network. Large α *curves* are significantly below their small α counterparts, indicating that, other things being equal, densely and strongly interacting networks are more likely to be pairwise sufficient. Notice that large α *curves* end abruptly since such networks cannot have large $\mathcal{S}$ (Color figure online)

Eq. (7) for different pairs of the variables. The algorithm is known to converge to the solution of the MaxEnt problem (at least when there are no zeros in the pairwise marginals) largely irrespective of the specific order in which pairs $\{i, j\}$ are taken, or the initial condition [31]. However, the speed of convergence can strongly depend on these choices. Thus to ensure that the algorithm does not get trapped by very slow convergence rates, we also verify that starting with a random initial condition, or using a random choice of $\{i, j\}$ at every step results in the same solution.

To measure the quality of the MaxEnt models, we calculate the Kullback-Leibler (KL) divergence between the true distribution $P_{3,4}$ and each approximation, normalized by the number of spins in the system:

$$\mathcal{D}^{(1)} = \frac{D_{KL}^{(1)}}{N} \equiv \frac{1}{N} \sum_{\sigma} P(\sigma) \log_2 \frac{P(\sigma)}{Q^{(1)}(\sigma)}, \tag{8}$$

$$\mathcal{D}^{(2)} = \frac{D_{KL}^{(2)}}{N} \equiv \frac{1}{N} \sum_{\sigma} P(\sigma) \log_2 \frac{P(\sigma)}{Q^{(2)}(\sigma)}. \tag{9}$$

Since our maximum N is rather small, this is done by direct summation. Notice that both $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$ are between zero (perfect fit) and one (the worst fit) if single-spin marginals of Q and P are equal.

In Fig. 2, we plot the values of $\mathcal{D}^{(1)}$ and $\mathcal{D}^{(2)}$ measured over different ensembles of random networks versus the normalized entropy of the network's state space $\mathcal{S} = S(\sigma)/N$, which also varies between 0 and 1. For all types of networks and approximations, the quality of fit is high ($\mathcal{D}^{(\cdot)}$ is low) when $\mathcal{S} \sim 1$, so that the networks are unconstrained, and nearly all states are possible. This is trivial since even the zeroth order approximation (each spin up or down with 50% probability) would work well here.

As $\mathcal{S}$ decreases, the fit errors increase. When $\mathcal{S}$ reaches small values, the independent approximation, $\mathcal{D}^{(1)}$, starts behaving differently for the different network types. In the 4-spin case, by construction, $P(\boldsymbol{\sigma}) = P(-\boldsymbol{\sigma})$. Thus $\langle \sigma_i \rangle = 0$ for any i, and the best independent approximation is the uniform distribution. For this construction, the smallest possible entropy is $\mathcal{S} = 1/N$, where the network exists in two mirror states, and there the error of $Q^{(1)}$ is $\mathcal{D}_4^{(1)} = 1 - 1/N$. In contrast, a 3-spin network freezes at $\mathcal{S} = 0$, and each spin is strongly biased (as in our XOR networks above). Thus the independent approximation provides a perfect fit in this case.

The distinction between P_3 and P_4 decreases for the pairwise MaxEnt approximation. Here, for both 3- and 4-spin networks, the fit errors behave similarly: for $\mathcal{S}$ decreasing from 1, $\mathcal{D}^{(2)}$ grows from 0 and reaches its peak at about $\mathcal{D}^{(2)} \approx 0.25 \ldots 0.3$ near $\mathcal{S} \approx 0.5 \ldots 0.6$. This is already interesting: $\mathcal{D}^{(2)}$ almost never goes above 0.3 for all networks we tried. Thus even in some of the worst cases, pairwise approximation is quite good! Further, for even smaller entropies, $\mathcal{D}^{(2)}$ rapidly drops, approaching zero faster than linearly in $\mathcal{S}$. For $\mathcal{S} \approx 0.25$, $\mathcal{D}_3^{(2)} \approx 0.07$. It is even smaller, $\mathcal{D}_4^{(2)} \approx 0.04$, for the quartic case. This is because $\min \mathcal{S}_4 = 1/N$, so that the whole $\mathcal{D}_4^{(2)}$ curve is slightly shifted compared to $\mathcal{D}_3^{(2)}$ at low $\mathcal{S}$.

In summary, for all the networks we have considered, pairwise sufficiency emerges robustly at low (but not too low) entropy. In fact, at $\mathcal{S} \approx 0.25$, our networks can be in more than $2^{\mathcal{S}} = 2^{N\mathcal{S}} \approx 2^5 = 32$ highly probable states. Thus the networks are not totally frozen, and yet the pairwise approximation is nearly sufficient! Crucially, this finding is robust to the changes in the network size: $\mathcal{D}$ versus $\mathcal{S}$ curves are stable over the entire range of N we explored, though if the results hold at even larger N is unclear.

Within a single narrow bin of $\mathcal{S}$, $\mathcal{D}^{(2)}$ may still have a rather large range. We explore this variability in the rightmost panel of Fig. 2. For this, we define $\alpha = sM/N$ (recall that s is the standard deviation of the random couplings used to generate the networks). α measures the strength of interactions (or constraints) per spin, analogously to a similar parameter in the random constraint satisfaction problems [32,33]. For quartic networks, where we have enough samples, we then plot $\mathcal{D}^{(2)}$ versus $\mathcal{S}$ for different ranges of α. Crucially, we find that, for the same $\mathcal{S}$, a larger α results in better pairwise fits. In other words, a denser and stronger interacting network is more likely to be pairwise sufficient. This is potentially a good news for MaxEnt approaches to biological systems, which are known for the immense complexity of the underlying biophysical interactions.

Here we would like to stress that partitioning by α can potentially mix together dense networks with weak interactions (small s, large M) and sparse networks with nearly deterministic interactions (large s, small M). The latter will have high entropies and are unlikely to result in the pairwise sufficiency. In other words, α cannot be the sole control parameter responsible for the sufficiency. This precipitated the use both $\mathcal{S}$ and α in Fig. 2 (right), and, indeed, $\mathcal{D}^{(2)}$ varies a lot for the same α depending on $\mathcal{S}$. Additional analytical investigations are required to verify if there are better choices of control parameters, or if a single control parameter combining $\mathcal{S}$ and α can be formed to analyze the pairwise sufficiency.

We conclude this section by stressing that high probability states of pairwise sufficient p-spin networks are not close to each other. To illustrate this, we focus on the 4-spin case with $N \geq 20$, and on small but not negligible $\mathcal{S}$. We then evaluate the magnitude of the overlap, $|\boldsymbol{\sigma}^\mu \cdot \boldsymbol{\sigma}^\nu|/N$, among all highly probable network states and plot the distribution of the overlaps in Fig. 3. For purely randomly distributed states, we would expect the standard deviation of overlaps to be ~ 0.22, and a peak near zero. And we would expect magnitudes of overlaps near 1 if all highly probable states were clustered near a dominant one. Instead, the distribution in Fig. 3 is not concentrated near 1, and the standard deviation is ≈ 0.39.
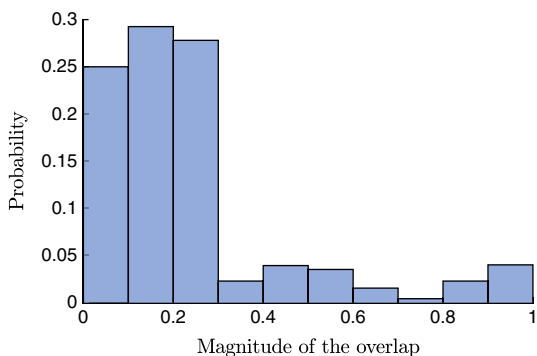
Fig. 3 The pairwise sufficiency emerges in the nonperturbative regime. We select all highly probable states σ^μ, defined somewhat arbitrary as $P(\sigma^\mu) > 0.001$. We then calculate the overlap for all pairs of such states. Finally, we plot the histogram of the magnitudes of the overlaps from all 4-spin networks with $N \geq 20$, and with $0.1 \leq S \leq 0.3$. Such networks can exist in many states, but still very few compared to 2^N. Most of the overlap magnitudes are away from 1, indicating small to moderate similarity among the highly probable states. Thus these states are broadly dispersed and do not cluster together (Color figure online)

Therefore, there is some clustering of probable states, but certainly not strong clustering. Thus the state space of our networks cannot be described as small fluctuations around a dominant state, and the pairwise sufficiency here is not perturbative [25]. It likely emerges due to a previously not investigated mechanism.

2.3 The Structure of the State Space of the Pairwise Sufficient Networks

The toy example of the XOR network suggests that the pairwise sufficiency may emerge when the network "freezes" to a few (but not necessarily just one) highly probable states, and different relatively tightly coupled clusters of spins (hyperloops) decouple from each other. Is this also true for our networks with a nonzero temperature? How do energy landscapes of the sufficient and the insufficient networks differ from each other? And are the MaxEnt fits for both cases structurally different?

To explore this, we extract h_i and J_{ij} from $Q^{(2)}$ (recall that IPFP inferes the distribution, and not the coupling constants). The easiest way to do this is to start with $\geq N(N+1)/2$ states with the highest probabilities from $Q^{(2)}$ found by IPFP and to get the energy of each such state as $E(\sigma) = -\log Q^{(2)}(\sigma)$. We then solve the system of linear equations that relate E to h_i and J_{ij}. Starting with $> N(N+1)/2$ states creates an overdetermined system, which will have a solution only if $Q^{(2)}$ has been determined accurately. This acts as an additional consistency check. Finally, we calculate the eigenvalue spectra of the inferred J_{ij}, having set $J_{ii} = 0$. The averaged spectra are shown in Fig. 4 for different combinations of S and $\mathcal{D}^{(2)}$. We see that the success of $Q^{(2)}$ is correlated with the magnitude of the eigenvalues of J_{ij}—larger magnitudes, which correspond to stronger interactions and more constrained distributions, give the pairwise sufficiency. This is true irrespective of S (though S and α are somewhat dependent, as we have discussed).

Further, we observe that rescaling the spectra by their largest eigenvalue (Fig. 4, inset), makes them collapse within the (rather large) error bars. Thus the J_{ij} (or the energy landscapes) of the pairwise sufficient and the pairwise insufficient fits in this Figure are not intrinsically different: a rescaling (change in temperature) can morph one into the other. This is an interesting observation, but it should not be overinterpreted. Indeed, in this Figure, we
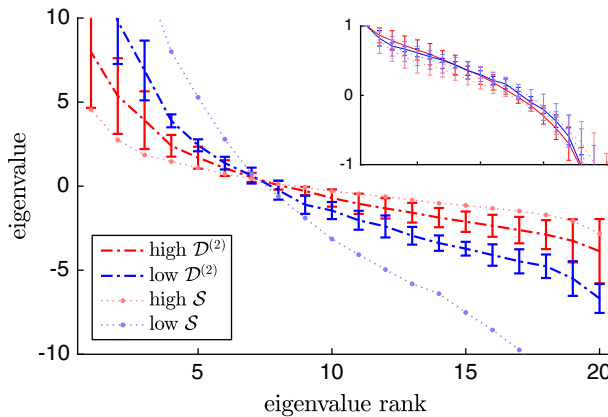
Fig. 4 Spectra of the pairwise coupling matrices J_{ij} for MaxEnt approximations to random 4-spin networks, $N \geq 20$. We order eigenvalues from the largest (lowest energy) to the smallest (highest energy). We then plot the mean spectra (with standard deviations, where it does not obstruct the figures), averaged over different subsets of 4000 networks. Low $\mathcal{S}$ subset corresponds to $0.1 \leq \mathcal{S} < 0.2$. Such networks are fit extremely well by pairwise models, with mean $\mathcal{D}^{(2)} \approx 6 \times 10^{-3}$. High $\mathcal{S}$ corresponds to $0.3 \leq \mathcal{S} < 0.4$. Here the pairwise fits are bad, so that the mean $\mathcal{D}^{(2)} \approx 0.15$. Finally, for the intermediate range $0.2 \leq \mathcal{S} < 0.3$, the quality of fits is diverse. We further partition this range into well fitted, $\mathcal{D}^{(2)} \leq 0.06$, and badly fitted, $\mathcal{D}^{(2)} \geq 0.12$, subsets, leaving intermediate fits off the plot. The four average spectra show that the pairwise sufficiency is directly correlated with the scale of the spectra, with larger magnitude eigenvalues resulting in smaller $\mathcal{D}^{(2)}$. The *inset* shows the averages for each of the four ranges, where each spectrum is normalized by its largest magnitude eigenvalue. The *four curves* are very close for much of their range. However, as discussed in the text, we do not expect a similar collapse for arbitrary networks (Color figure online)

only look at $\mathcal{S} < 0.4$, and we always have $s < 2.0$, so that strong individual interactions are unlikely. Thus to have low entropies, all networks in Fig. 4 must be densely coupled. Then the rescaling property is not that surprising, and it certainly will not generalize to arbitrary networks.

Having analyzed the pairwise sufficient and insufficient solutions, $Q^{(2)}$, we now focus on the landscapes of the p-spin networks themselves. The freezing that results in the decrease of $\mathcal{D}^{(2)}$ and the growth of the eigenvalues of J can create the landscapes of different types. For example, the highly probable states may be essentially uncorrelated, reminiscent of the landscapes of the classical Hopfield network in the ferromagnetic phase [34]. This has been discussed as a possible picture for neural data [11]. Alternatively, as in our XOR network, entire groups of spins can form strongly correlated clusters (hyperloops [27]), which then (nearly) decouple from each other. Then the low energy network states would be direct products of the states of the clusters, which would be far from independence.

To disambiguate the two scenarios, we calculate pairwise spin-spin correlations $c_{ij} = \frac{\text{cov}(\sigma_i, \sigma_j)}{\text{std}\,\sigma_i\,\text{std}\,\sigma_j}$ by direct summation of P (note that, for 4-spin networks, the correlation is equal to the covariance since $\langle \sigma_i \rangle = 0$). We then (hierarchically) cluster the spins based on the absolute value of their correlations, thus putting highly correlated and anti-correlated spins together. We emphasize here that interactions in the underlying "true" model couple spins completely at random. However, due to the random nature of these interactions, certain spins become correlated with some spins stronger than with the others, resulting in some clustering within the co-variance matrix. The question of whether the low-energy states are not independent is equivalent to the question whether the observed clusters in the spin co-variance are statistically significant.
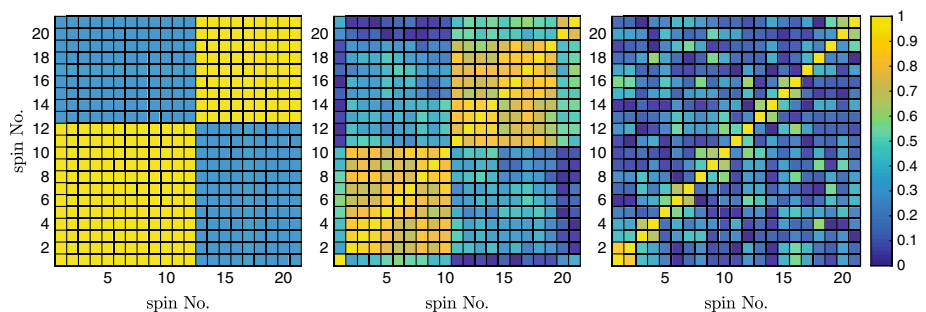
Fig. 5 4-Spin networks decouple into spin clusters in the pairwise sufficient regime. The *three panels* show typical values of $|c_{ij}|$ for $N = 21$, sorted into correlated clusters. *Left panel* a nearly perfectly pairwise sufficient network with $S = 2.05$ bits and $\mathcal{D}^{(2)} = 8.0 \times 10^{-5}$. In accord with the entropy value, the network splits into two clusters, with spins nearly perfectly correlated within each, but almost independent across. *Middle panel* a good pairwise fit with $S = 3.9$ bits and $\mathcal{D}^{(2)} = 0.032$. Correspondingly, four clusters of different sizes are seen. However, now the spins also exhibit some correlations across clusters, which presumably leads to the increase in $\mathcal{D}^{(2)}$. *Right panel* a network with $S = 5.7$ bits and $\mathcal{D}^{(2)} = 0.16$—a bad (though not disastrous) pairwise fit. There are now many small clusters, but correlations within and across the clusters are not very different (Color figure online)

Figure 5 shows the clusters for 4-spin networks (note that since the number, the size, and the spin assignment for clusters are different for each network, we only show typical cases). A network with a near-zero $\mathcal{D}^{(2)}$ (left panel) shows a perfect partitioning into two clusters; $S(\boldsymbol{\sigma}) \approx 2$ bits is a result of this partitioning. As networks with larger entropies are considered, the number of clusters increases, and their boundaries become fuzzy, leading to worse MaxEnt fits. When the definite cluster structure disappears, $\mathcal{D}^{(2)}$ grows dramatically. In other words, the existence of well-defined spin clusters co-occurs with the pairwise sufficiency. Thus the emergence of hyperloops [27] is probably not unique to networks of Boolean gates, and the landscape of our networks looks qualitatively similar.

For 3-spin networks, in addition to the pairwise interactions, there are also nonzero single spin biases in the MaxEnt fits. Thus the entropy and correlations among spins are generally smaller for the same $\mathcal{D}^{(2)}$. Nonetheless, as seen in Fig. 6, the (fuzzy) cluster structure for these networks is not that much different from the 4-spin case.

To further explore the network landscapes, we point out that an inferred symmetric J_{ij} can be rewritten as

$$J_{ij} = \sum_{\nu=1}^{N} \lambda^{(\nu)} \xi_i^{(\nu)} \xi_j^{(\nu)}, \tag{10}$$

where $\lambda^{(\nu)}$ and $\xi_i^{(\nu)}$ are the eigenvalues and the eigenvectors, correspondingly. The eigenvalues take both large positive and large negative values for the pairwise sufficient networks (cf. Fig. 4). The positives correspond to wells in the landscape, and the negatives correspond to peaks. In other words, our networks are not well approximated by the classical Hopfield network [34] with attractive eigenvalues only, but are generalized Hopfield networks, such as discussed in, for example, Refs. [35,36]. If the wells and the peaks were clearly separated, then spin configurations in the vicinity of the well center, distinct from it by just a single spin flip, would have similar high probabilities (this is what allows an eigenvector to act as a broad attractor in the classical Hopfield network [34]). The repulsive peaks far away from the wells would have little effect on $\mathcal{D}^{(2)}$ since the probability of states away from the wells is small in the low temperature regime even without the peaks. In contrast, were spins to form

Fig. 6 A 3-spin network also decouples into spin clusters in the pairwise sufficient regime. For brevity, here we show $|c_{ij}|$ only for one network similar to the *middle panel* in Fig. 5: $S \approx 3.5$, $\mathcal{D}^{(1)} \approx 0.58$, and $\mathcal{D}^{(2)} \approx 0.038$. Four to seven partially overlapping clusters can be seen (Color figure online)
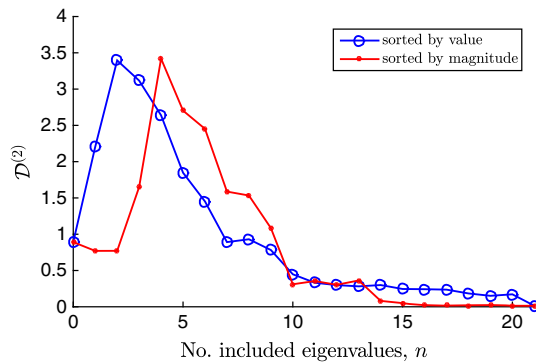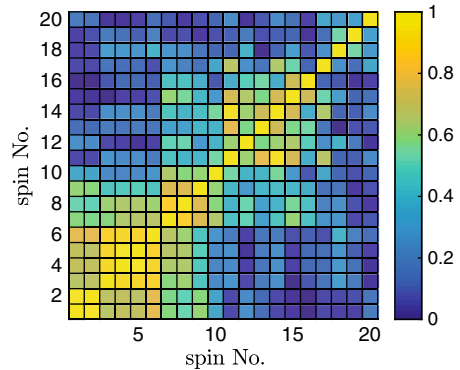

Fig. 7 Repulsive eigenvalues of the MaxEnt coupling matrix J_{ij} contribute to the approximation error. We plot the fit error, $\mathcal{D}^{(2)}$, as a function of the number of the eigenvalues of J_{ij} included in the fit for a 4-spin distribution with $N = 21$, $S \approx 0.1$, and $\mathcal{D}^{(2)} \approx 0.014$. The *blue line* includes eigenvalues in the order from the most attractive to the most repulsive, and the *red one* includes them in the order of their absolute values. The *red line* reaches the limiting value of $\mathcal{D}^{(2)}$ quicker, while the *blue* one requires inclusion of all eigenvalues for this to happen. As explained in the text, this is consistent with emergence of spin clusters (Color figure online)

strongly interacting clusters, flipping just a single spin would lead to a large energy increase. To achieve this, peaks (negative eigenvalues) would be needed to repel the system away from such states, which are near a well, but break the coupling among the in-cluster spins. In other words, negative eigenvalues would decrease the probability of otherwise likely states, and hence would affect $\mathcal{D}^{(2)}$ strongly.

To verify which of the two scenarios holds, for 4-spin networks, we construct the coupling matrix and the pairwise MaxEnt distribution from only $n \leq N$ eigenvalues,

$$J_{ij}(n) = \sum_{v=1}^{n} \lambda^{(v)} \xi_i^{(v)} \xi_j^{(v)}, \tag{11}$$

$$Q^{(2)}(n) = \frac{1}{Z} \exp\left(\sum_{i,j} J_{ij}(n)\sigma_i\sigma_j\right). \tag{12}$$

We then evaluate $\mathcal{D}^{(2)}$ between P and $Q^{(2)}(n)$ as a function of n. Figure 7 shows this dependence for a typical pairwise-sufficient distribution and for two different ways of including

eigenvalues into J. In the first, we proceed from the largest eigenvalue to the smallest one. In the second, we proceed from the largest magnitude eigenvalue to the smallest magnitude one. Since sorting by magnitude (which includes strong repulsive eigenvalues earlier) approaches the terminal $\mathcal{D}^{(2)}$ faster, wells and peaks must both affect probabilities of high-probability spin configurations. As discussed above, this is consistent with the clustering picture, though by itself is certainly not a sufficient condition for clustering.

2.4 The Mechanism of Emergence of the Pairwise Sufficiency

The clustered structure of the network landscapes allows us to propose a hypothesis for why densely coupled p-spin networks exhibit the pairwise sufficiency. We re-group terms in the energies, which define P_3 and P_4 in Eqs. (4, 5). For example, all terms that couple σ_i and σ_j for the 4-spin network can be rewritten as

$$\sigma_i \sigma_j \sum_\mu K_\mu \sigma_{\mu_3} \sigma_{\mu_4} \delta_{i,\mu_1} \delta_{j,\mu_2} \equiv \sigma_i \sigma_j \mathcal{J}_{ij}, \tag{13}$$

where $\delta_{\cdot,\cdot}$ is a Kronecker delta. (Here we slightly abused the notation and imposed that i, j only occur in μ_1 and μ_2.) This equation defines a random coupling $\mathcal{J}_{ij}$, which depends both on the current network state and on the quenched randomness that went into building the network. For a large number of couplings, fluctuations in $\mathcal{J}_{ij}$ will be large enough so that, averaged over the accessible network states, $\mathcal{J}_{ij}$ stays far from zero compared to its standard deviation $\sigma_{\mathcal{J}_{ij}}$. This creates large effective pairwise coupling among spins, so that clusters of spins start behaving coherently. Then the state of every spin in the cluster can be defined by choosing a cluster representative, setting its value, and then coupling each cluster member to the representative through a pairwise interaction—higher order couplings are not needed! The pairwise MaxEnt fit is nearly exact, even though the network is far from frozen since values of the cluster representatives are not necessarily constrained. We illustrate this in Fig. 8, which shows that higher order couplings average to produce large effective pairwise interactions for
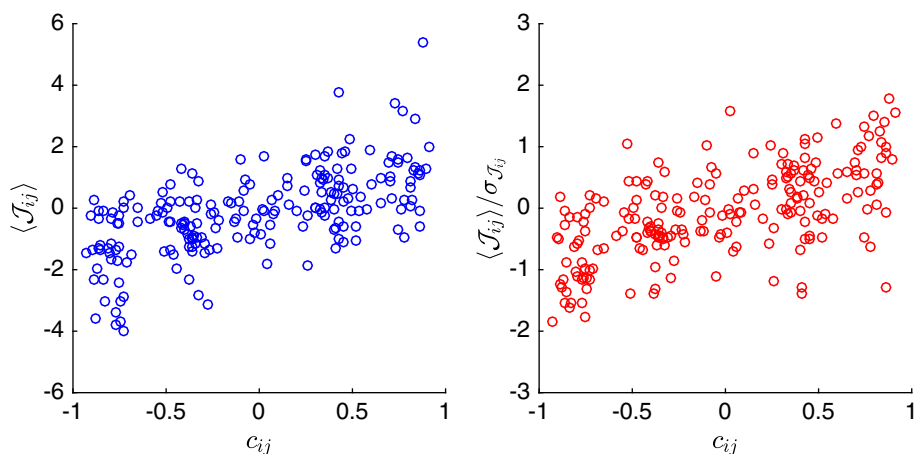

Fig. 8 Formation of effective couplings in 4-spin networks. For the same network as in Fig. 5, *middle*, we plot $\langle \mathcal{J}_{ij} \rangle$ (*left*) and $\langle \mathcal{J}_{ij} \rangle$ in units of the standard deviation (*right*) for each pair of spins versus c_{ij}. For correlated spins, higher order coupling terms add up, on average, to strong pairwise couplings. Typically, the signs of $\mathcal{J}_{ij}$'s and c_{ij} are the same, as excepted (Color figure online)

correlated spins. In other words, the intuition about the structure of the landscape of random networks of deterministic Boolean gates [27] qualitatively holds for networks in Eqs. (4, 5) as well.

3 Discussion

In this numerical study, we showed that pairwise MaxEnt models are more effective in approximating random p-spin networks ($p = 3, 4$) than one would naively expect. Even in the worst cases, the error of such models was rarely above $\mathcal{D}^{(2)} \sim 0.3$, and it was much lower for densely coupled networks, with lower entropy per spin. In particular, this is consistent with the findings in Ref. [37] that the number of constraints required for a MaxEnt model to accurately approximate the underlying probability distribution is an increasing function of the entropy of the system.

We traced the emerging pairwise sufficiency to formation of coherent clusters of spins, largely decoupled among themselves, resulting in a multitude of dependent attractors for the system. Formation of these clusters is reminiscent of hyperloops in the context of random hSAT networks [27]. This is not a perturbative effect and is a new proposal for explaining pairwise sufficiency. Such collective behavior introduces substantial redundancy, and would allow error correction. However, this error correction is of a very different nature compared to, for example, the Hopfield network [34].

Does the mechanism presented here explain the pairwise sufficiency in any *real* biological system? This is unclear since our analysis was limited to specific simulated networks, which may or may not be good models of real biology. Specifically, the network in the original paper that observed the pairwise sufficiency [11] had much smaller entropy per spin (neurons rarely fired), and correlations among spins rarely exceed 0.2. In contrast, while 3-spin networks in our simulations had smaller $\mathcal{S}$ and smaller spin-spin correlations than their 4-spin counterparts, these numbers were still larger than those in the experiments. At the same time, pairwise MaxEnt models do not fit experimental data perfectly (certainly worse than some of our nearly perfect fits) [15]. It may be that some structural features of real systems allow them to operate at higher $\mathcal{D}^{(2)}$ for smaller $\mathcal{S}$ compared to the simple models we investigated here—and exploring a wider class of networks for signatures of behaviors that we observed would be the next step. This is especially important since large coherent deviations from the most probable state into $10 \ldots 100$ metastable states seems to be a crucial feature of many experimental systems (such as bursts of neural activity in the retina that predominantly stays quiet [15]). Such metastable states far away from the ground state at least resemble the models that we studied. In addition, MaxEnt models in other fields may have very different properties compared to those in neuroscience, including different typical entropies and correlation strengths. Therefore, we hope that our models and their generalizations will be able to inform interpretation of experimental data, even if they do not match the experiments in some important properties.

With (approximate) pairwise sufficiency seen in many collective biological phenomena, it is important to ask why these systems operate in the regime that allows it to hold. Indeed, within our model, the pairwise sufficiency is not generic: low $\mathcal{D}^{(2)}$ happens only for small $\mathcal{S}$, and preferentially when the strength of the interactions is high, $\alpha = Ms/N \gg 1$ (cf. Fig. 2). The need for redundancy and error correction is a potential explanation – but there is no obvious reason why the redundancy must result in the pairwise sufficiency (indeed, simple parity-based codes probably do not). Taking the improvement in $\mathcal{D}^{(2)}$ with the increase in

α seriously, we propose a different explanation (a similar argument was first suggested in Ref. [38]).

One can view evolution as trying to satisfy a growing list of constraints imposed upon a biological network by its interactions with the environment. These constraints can include efficient information processing, low energy consumption, robustness to perturbations, fitting within a certain physical size, responding quickly enough so that actions are relevant in the changing world, etc. Some of these global constraints may be equivalent to a large number of local constraints. For example, efficient information transmission in the visual system typically includes removal of redundancy present in the natural stimuli [39], which is equivalent to a multitude of constraints on activities of nearby neurons. When contraints are added, fewer and fewer states of the network remain accessible. Importantly, at least for certain abstract constraint satisfaction problems [27,33,40], before there are no more states left, the accessible states organize themselves in a handful of small, well-separated groups. Whether these states are uncorrelated, or consist of collective flipping of clusters of spins, they can be well represented by pairwise MaxEnt models. (In the former case, such MaxEnt model would have a Hopfield network structure with attractive eigenvalues [34]; in the latter, pairwise interactions would determine cluster assignments.) Therefore, it can be that the pairwise sufficiency is a signature of a biological network nearing the unsatisfiability threshold, being pushed towards it by evolution. Exploring landscapes of satisfiability problems with more realistic ensembles of constraints (or interactions) and comparing them to the landscapes observed in experiments would address this hypothesis.

References

1. Jaynes, E.T.: Information theory and statistical mechanics. Phys. Rev. **106**, 620 (1957)
2. Schneidman, E., Still, S., Berry, M.J., Bialek, W.: Network information and connected correlations. Phys. Rev. Lett. **91**, 238701 (2003)
3. Ackley, D., Hinton, G., Sejnowski, T.: A learning algorithm for boltzmann machines. Cogn. Sci. **9**, 147 (1985)
4. Broderick, T., Dudik, M., Tkacik, G., Schapire, R.E., Bialek, W.: Faster solutions of the inverse pairwise Ising problem. (2007). arXiv:0712.2437v2
5. Weigt, M., White, R.A., Szurmant, H., Hoch, J.A., Hwa, T.: Identification of direct residue contacts in protein-protein interaction by message passing. Proc. Natl. Acad. Sci. (USA) **106**, 67 (2009)
6. Mézard, M., Mora, T.: Constraint satisfaction problems and neural networks: a statistical physics perspective. J. Physiol. (Paris) **103**, 107 (2009)
7. Sessak, V., Monasson, R.: Small-correlation expansions for the inverse Ising problem. J. Phys. A **42**, 055001 (2009)
8. Cocco, S., Monasson, R.: Adaptive cluster expansion for inferring Boltzmann machines with noisy data. Phys. Rev. Lett. **106**, 090601 (2011)
9. Nguyen, H.C., Berg, J.: Mean-field theory for the inverse Ising problem at low temperatures. Phys. Rev. Lett. **109**, 050602 (2012)
10. Tkacik, G., Bialek, W. Information processing in living systems. (2014). arXiv:1412.8752v1
11. Schneidman, E., Berry, M.J., Segev, R., Bialek, W.: Weak pairwise correlations imply strongly correlated network states in a neural population. Nature **440**, 1007 (2006)

12. Tang, A., Jackson, D., Hobbs, J., Chen, W., Smith, J.L., Patel, H., Prieto, A., Petrusca, D., Grivich, M.I., Sher, A., Hottowy, P., Dabrowski, W., Litke, A.M., Beggs, J.M.: A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. J. Neurosci. **28**, 505 (2008)

13. Ohiorhenuan, I., Mechler, F., Purpura, K., Schmid, A., Hu, Q., Victor, J.: Sparse coding and high-order correlations in fine-scale cortical networks. Nature **466**, 617 (2010)

14. Field, G., Gauthier, J., Sher, A., Greschner, M., Machado, T., Jepson, L., Shlens, J., Gunning, D., Mathieson, K., Dabrowski, W., Paninski, L., Litke, A., Chichilnisky, E.J.: Functional connectivity in the retina at the resolution of photoreceptors. Nature **467**, 673 (2010)

15. Tkacik, G., Marre, O., Amodei, D., Schneidman, E., Bialek, W., Berry, M.J.: Searching for collective behavior in a large network of sensory neurons. PLoS Comput. Biol. **10**, e1003408 (2014)

16. Bethge, M., Berens, P.: Near-maximum entropy models for binary neural representations of natural images. In: Platt, J., Koller, D., Singer, Y., Roweis, S. (eds.) Advances in Neural Information Processing Systems, p. 97. MIT Press, Cambridge (2008)

17. Halabi, N., Rivoire, O., Leibler, S., Ranganathan, R.: Protein sectors: evolutionary units of three-dimensional structure. Cell **138**, 774 (2009)

18. Mora, T., Walczak, A., Bialek, W., Callan, C.: Maximum entropy models for antibody diversity. Proc. Natl. Acad. Sci. (USA) **107**, 5405 (2010)

19. Bialek, W., Cavagna, A., Giardina, I., Mora, T., Silvestri, E., Viale, M., Walczak, A.: Statistical mechanics for natural flocks of birds. Proc. Natl. Acad. Sci. (USA) **109**, 4786 (2012)

20. Margolin, A.A., Wang, K., Califano, A., Nemenman, I.: Multivariate dependence and genetic networks inference. IET Syst. Biol. **4**, 428 (2010)

21. Otwinowski, J., Nemenman, I.: Genotype to phenotype mapping and the fitness landscape of the *E. coli* lac promoter. PLoS One **8**, e61570 (2013)

22. Mora, T., Bialek, W.: Are biological systems poised at criticality? J. Stat. Phys. **144**, 268 (2011)

23. Schwab, D.J., Nemenman, I., Mehta, P.: Zipf's law and criticality in multivariate data without fine-tuning. Phys. Rev. Lett. **113**, 068102 (2014)

24. Tkacik, G., Schneidman, E., Berry, M., Bialek, W.: Ising models for networks of real neurons. (2006). arXiv:org/q-bio/0611072v1

25. Roudi, Y., Nirenberg, S., Latham, P.: Pairwise maximum entropy models for studying large biological systems: when they can work and when they can't. PLoS Comput. Biol. **5**, e1000380 (2009)

26. Mézard, M., Ricci-Tersenghi, F., Zecchina, R.: Two solutions to diluted p-spin models and XORSAT problems. J. Stat. Phys. **111**, 505 (2003)

27. Ricci-Tersenghi, F., Weigt, M., Zecchina, R.: Simplest random k-satisfiability problem. Phys. Rev. E **63**, 026702 (2001)

28. Semerjian, G.: On the freezing of variables in random constraint satisfaction problems. J. Stat. Phys. **130**(2), 251–293 (2007)

29. Kirkpatrick, T., Thirumalai, D.: p-Spin-interaction spin-glass models: connections with the structural glass problem. Phys. Rev. B **36**, 5388 (1987)

30. Deming, W.E., Stephan, F.F.: On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. Ann. Math. Stat. **11**, 427 (1940)

31. Csiszár, I.: I-divergence geometry of probability distributions and minimization problems. Ann. Probab. **3**, 146–158 (1975)

32. Mézard, M., Parisi, G., Zecchina, R.: Analytic and algorithmic solution of random satisfiability problems. Science **297**, 812 (2002)

33. Krzakala, F., Montanari, A., Ricci-Tersenghi, F., Semerjian, G., Zdeborova, L.: Gibbs states and the set of solutions of random constraint satisfaction problems. Proc. Natl. Acad. Sci. (USA) **104**, 10318 (2007)

34. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proc. Natl. Acad. Sci. (USA) **79**, 2554 (1982)

35. Cocco, S., Monasson, R., Sessak, V.: High-dimensional inference with the generalized hopfield model: principal component analysis and corrections. Phys. Rev. E **83**, 051123 (2011)

36. Cocco, S., Monasson, R., Weigt, M.: From principal component to direct coupling analysis of coevolution in proteins: low-eigenvalue modes are needed for structure prediction. PLoS Comput. Biol. **9**, e1003176 (2013)

37. Obuchi, T., Cocco, S., Monasson, R.: Learning probabilities from random observables in high dimensions: the maximum entropy distribution and others. J. Stat. Phys. **161**, 598–632 (2015)

38. Lander, A. Beyond q-bio? The closing talk at The Sixth q-bio Conference, Santa Fe, NM. (2012)

39. Barlow, H.B.: Possible principles underlying the transformation of sensory messages. In: Rosenblith, W. (ed.) Sensory Communication, p. 217. MIT Press, Cambridge, MA (1961)

40. Myers, C.R.: Satisfiability, sequence niches and molecular codes in cellular signalling. IET Syst. Biol. **2**, 304 (2008)